Original Investigation

Nocturnal Rapid Eye Movement Sleep Latency for Identifying Patients With Narcolepsy/Hypocretin Deficiency

Olivier Andlauer, MD; Hyatt Moore, IV, MSc; Laura Jouhier, MD; Christopher Drake, PhD; Paul E. Peppard, PhD; Fang Han, MD; Seung-Chul Hong, MD, PhD; Francesca Poli, MD; Giuseppe Plazzi, MD; Ruth O'Hara, PhD; Emmanuel Haffen, MD, PhD; Thomas Roth, PhD; Terry Young, PhD; Emmanuel Mignot, MD, PhD

IMPORTANCE Narcolepsy, a disorder associated with HLA-DQB1*06:02 and caused by hypocretin (orexin) deficiency, is diagnosed using the Multiple Sleep Latency Test (MSLT) following nocturnal polysomnography (NPSG). In many patients, a short rapid eye movement sleep latency (REML) during the NPSG is also observed but not used diagnostically.

OBJECTIVE To determine diagnostic accuracy and clinical utility of nocturnal REML measures in narcolepsy/hypocretin deficiency.

DESIGN, SETTING, AND PARTICIPANTS Observational study using receiver operating characteristic curves for NPSG REML and MSLT findings (sleep studies performed between May 1976 and September 2011 at university medical centers in the United States, China, Korea, and Europe) to determine optimal diagnostic cutoffs for narcolepsy/hypocretin deficiency compared with different samples: controls, patients with other sleep disorders, patients with other hypersomnias, and patients with narcolepsy with normal hypocretin levels. Increasingly stringent comparisons were made. In a first comparison, 516 age- and sex-matched patients with narcolepsy/hypocretin deficiency were selected from 1749 patients and compared with 516 controls. In a second comparison, 749 successive patients undergoing sleep evaluation for any sleep disorders (low pretest probability for narcolepsy) were compared within groups by final diagnosis of narcolepsy/hypocretin deficiency. In the third comparison, 254 patients with a high pretest probability of having narcolepsy were compared within group by their final diagnosis. Finally, 118 patients with narcolepsy/hypocretin deficiency were compared with 118 age- and sex-matched patients with a diagnosis of narcolepsy but with normal hypocretin levels.

MAIN OUTCOME AND MEASURES Sensitivity and specificity of NPSG REML and MSLT as diagnostic tests for narcolepsy/hypocretin deficiency. This diagnosis was defined as narcolepsy associated with cataplexy plus HLA-DQB1*06:02 positivity (no cerebrospinal fluid hypocretin-1 results available) or narcolepsy with documented low ($\leq$ 110 pg/mL) cerebrospinal fluid hypocretin-1 level.

RESULTS Short REML ($\leq$15 minutes) during NPSG was highly specific (99.2% [95% CI, 98.5%-100.0%] of 516 and 99.6% [95% CI, 99.1%-100.0%] of 735) but not sensitive (50.6% [95% CI, 46.3%-54.9%] of 516 and 35.7% [95% CI, 10.6%-60.8%] of 14) for patients with narcolepsy/hypocretin deficiency vs population-based controls or all patients with sleep disorders undergoing a nocturnal sleep study (area under the curve, 0.799 [95% CI, 0.771-0.826] and 0.704 [95% CI, 0.524-0.907], respectively). In patients with central hypersomnia and thus a high pretest probability for narcolepsy, short REML remained highly specific (95.4% [95% CI, 90.4%-98.3%] of 132) and similarly sensitive (57.4% [95% CI, 48.1%-66.3%] of 122) for narcolepsy/hypocretin deficiency (area under the curve, 0.765 [95% CI, 0.707-0.831]). Positive predictive value in this high pretest probability sample was 92.1% (95% CI, 83.6%-97.0%).

CONCLUSIONS AND RELEVANCE Among patients being evaluated for possible narcolepsy, short REML ($\leq$15 minutes) at NPSG had high specificity and positive predictive value and may be considered diagnostic without the use of an MSLT; absence of short REML, however, requires a subsequent MSLT.

JAMA Neurol. 2013;70(7):891-902. doi:10.1001/jamaneurol.2013.1589
Published online May 6, 2013.

The association of excessive daytime sleepiness with cataplexy (episodes of muscle weakness triggered by emotions) has traditionally been used to define narcolepsy.[1,2] In 1960, it was found that during nocturnal polysomnography (NPSG), 50% of cases have rapid transitions from wake to rapid eye movement (REM) sleep (sleep-onset REM periods [SOREMPs]).[3,4] The Multiple Sleep Latency Test (MSLT), a test that allows 4 to 5 daytime nap opportunities for sleep onset and SOREMPs, was subsequently developed to increase sensitivity. A mean sleep latency (MSL) of 8 minutes or less and 2 or more SOREMPs is considered diagnostic for narcolepsy.[2] Positive MSLT results have been reported in 2% to 4% of the population.[5,6]

The opportunity for a new gold standard, low concentration of hypocretin-1 (≤110 pg/mL; hypocretin deficiency) in the cerebrospinal fluid (CSF), has recently arisen with the finding of an autoimmune basis of narcolepsy-cataplexy.[7,8] Almost all cases with low CSF hypocretin-1 levels are HLA-DQB1*06:02 positive (vs 25% in the US general population), making it one of the most tightly HLA antigen–associated disorders known.[9] In a recent meta-analysis, sensitivity and specificity for low CSF hypocretin-1 level were 93% and 100%, respectively, in cases with cataplexy, as defined per the *International Classification of Sleep Disorders*, second edition (*ICSD-2*).[2,10]

Because sleep apnea and sleep deprivation can cause SOREMPs and short MSL during naps,[11-13] NPSG is systematically performed prior to MSLT.[14] Considering the MSLT takes a whole day in a sleep laboratory, we aimed to determine if short REM sleep latency (REML) during NPSG could identify narcolepsy cases associated with hypocretin deficiency, saving time and money. In this study, the value of a short REML during NPSG was compared with MSLT results and HLA-DQB1*06:02 typing findings, using hypocretin deficiency as a gold standard (as defined by the *ICSD-2*).[2] The value of combining measures of short REML at night with MSLT results and HLA-DQB1*06:02 typing findings was also explored.

Methods

Participants

Descriptions of the different samples are summarized in **Table 1**. All participants gave written informed consent. The Stanford institutional review board approved the study for pa-

tients included in the Stanford and Chinese databases and the controls collected from the Wisconsin Sleep Cohort; the Detroit, Michigan, tricounty area; and China.

Patients With Narcolepsy

Eight hundred eight patients with narcolepsy were retrospectively identified using the Stanford Center for Narcolepsy Research database (85.0%; n = 687) and a similar database available in China (15.0%; n = 121). In accordance with *ICSD-2*, narcolepsy diagnosis was based on excessive daytime sleepiness associated with either clear-cut cataplexy or a positive MSLT result (defined as MSL ≤8 minutes associated with ≥2 SOREMPs during 4-5 naps) or a low CSF hypocretin-1 level (≤110 pg/mL).[2] All cases had to have NPSG data available, and patients with a diagnosis of secondary narcolepsy (hypothalamic lesion [n = 1], multiple sclerosis [n = 2], and Turner syndrome [n = 1]) were excluded.[8,15,16] Patients had been recruited between May 1976 and January 2011 at 4 sites: United States (67.1%; n = 542), China (15.0%; n = 121), Korea (10.5%; n = 85), and Italy (7.4%; n = 60). Self-identified ethnicity was 63.5% white (n = 513), 26.9% Asian (n = 217), 7.7% African American (n = 62), and 2.0% other (n = 16). Three samples of patients with narcolepsy were defined as follows:

Sample A: 169 patients with a diagnosis of narcolepsy and documented low CSF hypocretin-1 level (≤110 pg/mL) with or without cataplexy or HLA-DQB1*06:02 positivity.

Sample B: 508 patients with a diagnosis of narcolepsy-cataplexy and HLA-DQB1*06:02 positivity but without CSF hypocretin-1 measurement available. Prior studies have estimated that 98% of such cases have a low CSF hypocretin-1 level (≤110 pg/mL).[10]

Sample C: 131 patients with a diagnosis of narcolepsy and normal CSF hypocretin-1 level (>200 pg/mL). Criteria for a normal CSF hypocretin-1 designation were either CSF hypocretin-1 values of more than 200 pg/mL (n = 60; 45.8% of the sample) or HLA-DQB1*06:02 negativity (n = 71; 54.2% of the sample). Prior studies have estimated that 99.8% of subjects with HLA-DQB1*06:02 negativity have a normal CSF hypocretin-1 level.[10]

Controls

One thousand twenty-six random population controls (sample D) were retrospectively identified from 3 sources: a sample drawn from the Wisconsin Sleep Cohort (36.2%; n = 371; data collection between December 1997 and May 2008), a sample

Table 1. Description of the Samples

Sample	Narcolepsy	Cataplexy	HLA-DQB1*06:02	Low CSF Hypocretin-1 Level	MSLT MSL ≤8 min and ≥2 SOREMPs	Sample Size, No. Before Matching	Sample Size, No. After Matching
A	+	+/−	+/−	+	+/−	169	89
B	+	+	+	Unknown	+/−	508	427
C	+	+/−	−	−	+	131	118
D	−	−	+/−	Unknown	+/−	1026	516 and 118
E	+/−	+/−	+/−	+/−	NA	749	NA
F	Suspected	+/−	+/−	+/−	+/−	254	NA

Abbreviations: CSF, cerebrospinal fluid; MSL, mean sleep latency; MSLT, Multiple Sleep Latency Test; NA, not available or not applicable; SOREMPs, sleep-onset rapid eye movement periods; +, present; −, absent.

drawn from the general population of the Detroit tricounty area (58.3%; n = 598; data collection between 1999 and 2003), and a sample from Beijing University People's Hospital in China (5.5%; n = 57; data collection between 2006 and 2007). Self-identified ethnicity was 72.9% white (n = 744), 17.8% African American (n = 182), 6.8% Asian (n = 69), and 2.5% other (n = 26) (missing data for 5 controls).

The Wisconsin Sleep Cohort is a longitudinal study of sleep habits and disorders in the general population.[17] The cohort was established in 1988 from a sample of employees of 4 state agencies in south central Wisconsin, aged 30 to 60 years. Beginning in 2000, participants enrolled in the Wisconsin Sleep Cohort were recruited for MSLT following an NPSG study, and only patients with MSLT results were drawn from this cohort for this study. Based on survey data, participants showed a typical healthy volunteer bias, with less self-reported hypertension and slightly higher education. They were not paid volunteers.

Participants from the second control sample came from the Detroit tricounty metropolitan area and were aged 18 to 65 years.[18] The sample was generated using a random-digit–dialed, computer-assisted telephone survey and included a laboratory-based (NPSG and MSLT) evaluation in a subset of the full sample. The sample was slightly enriched in individuals with excessive sleepiness, based on a daytime sleepiness scale, although this was shown not to affect the MSLT results.[19]

The Chinese control sample consisted of children seen for a routine dental examination at the dental clinic, healthy local employees, and college students, all from the Beijing University People's Hospital, who volunteered to undergo an NPSG followed by an MSLT.

Cases were retrospectively drawn from sleep centers, and therefore, it is likely that almost all eligible cases were included in the database. For controls, participation for the baseline overnight protocol in the Wisconsin Sleep Cohort rose from 50% to 54% by completion. For the Detroit tricounty area, the participation rate was 33% for the laboratory study. However, there were no differences between those electing to participate in the laboratory study and those who declined participation for age, sex, race, income, employment, marital status, or reported total sleep time.

Naturalistic Samples of Successive Patients

A naturalistic sample of 749 successive patients with a wide range of sleep disorders (sample E) was recruited between October 1999 and March 2007 at the Stanford Sleep Clinic and underwent NPSG. The only exclusion criterion was the use of a continuous positive airway pressure device for already documented sleep apnea prior to inclusion. Final primary diagnoses were sleep apnea (93.1%), narcolepsy-cataplexy (1.9%), insomnia (1.7%), narcolepsy without cataplexy (0.9%), restless legs syndrome (0.5%), REM behavior disorder (0.4%), delayed sleep phase syndrome (0.3%), and other (1.1%). This population included 14 patients with narcolepsy/hypocretin deficiency (1.9%) (all with cataplexy; 5 [35.7%] with low CSF hypocretin-1 levels; 9 with HLA-DQB1*06:02 positivity). Self-identified ethnicity was 84.2% white (n = 631), 10.0% Asian (n = 75), 2.7% African American (n = 20), and 3.1% other (n = 23).

Two hundred fifty-four successive patients from the Beijing University People's Hospital (36.6%; n = 93) and St Vincent's Hospital in Suwon, Korea (63.4%; n = 161) were recruited as a high pretest probability for narcolepsy sample (sample F) between 2005 and September 2011. High pretest probability was defined as clinical excessive daytime sleepiness that subsequently led to the *ICSD-2* diagnosis of either narcolepsy/hypocretin deficiency or narcolepsy without hypocretin deficiency, idiopathic hypersomnia, and Kleine-Levin syndrome (central hypersomnia syndromes). Final primary diagnoses were narcolepsy-cataplexy (53.5%; n = 136), narcolepsy without cataplexy (14.2%; n = 36), idiopathic hypersomnia (31.5%; n = 80), and Kleine-Levin syndrome (0.8%; n = 2). This population included 122 patients with narcolepsy/hypocretin deficiency (48.0%) (118 [96.7%] with cataplexy; 4 without cataplexy; 60 [49.2%] with low CSF hypocretin-1 levels; and 62 with cataplexy and HLA-DQB1*06:02 positivity). Self-identified ethnicity was 100% Asian.

All the described samples are independent, except for 6 patients with low CSF hypocretin-1 levels used in comparison 1a and comparison 4.

Clinical and Polysomnographic Data

Demographic data (age, sex, and self-identified ethnic group, based on predefined categories) were collected through the Stanford Sleep Inventory, a standardized questionnaire for sleep disorders.[20] Self-identified ethnic group is reported for descriptive purpose, because of the multiple sites of inclusion, and the fact that ethnicity has an effect on HLA-DQB1*06:02 prevalence (eg, higher prevalence of HLA-DQB1*06:02 positivity in African American individuals).[21] Body mass index was calculated as weight in kilograms divided by height in meters squared. Clinical features were recorded through clinical interviews, notes, and the Stanford Sleep Inventory. Rapid eye movement sleep latency, total sleep time, and apnea-hypopnea index were measured using a standard overnight NPSG, performed for all participants; an epoch was considered stage R or REM sleep if it contained low-amplitude, mixed-frequency electroencephalography activity and low chin electromyography tone that was the lowest level in the study or at least no higher than the other sleep stages and either had REMs or was preceded by stage R sleep.[22,23] Almost all patients and controls had MSLT following NPSG according to standard protocol[24] (missing data: 7 of the 1026 controls, 40 of the 808 patients; no missing data for the 254 successive patients), except for the naturalistic sample E. Mean sleep latency and the number of SOREMPs (occurrence of REM sleep in ≤15 minutes after sleep onset, if sleep onset occurs within 20 minutes after dark onset) were recorded. During MSLT, patients from the United States, China, and Italy and controls from the Detroit tricounty area and China always had 5 nap opportunities. Patients from Korea and controls from the Wisconsin Sleep Cohort only had 4 nap opportunities, a fifth nap being added only if REM sleep was detected during 1 of the first 4 naps.

HLA-DQB1*06:02 Typing and CSF Hypocretin-1 Measurement

Genetic typing of HLA-DQB1*06:02 was performed between October 1999 and September 2011 using a sequence-specific poly-

merase chain reaction, as described by Hallmayer et al.[25] The CSF hypocretin-1 concentrations were measured between March 2000 and September 2011 as previously described[26] and, in accordance with *ICSD-2* criteria, used as the reference (gold) standard.

Although the retrospective design prevented us from assuring that readers of the index tests (NPSG REML) and the reference standard (CSF hypocretin-1 levels) were universally blind to the result of the other test, CSF hypocretin-1 levels were measured in the laboratory facility, usually without the knowledge of any NPSG result. Technicians scoring REML are usually blind to the diagnosis of the patient under study. Further, given that lumbar puncture and HLA allele typing are typically done after the sleep study, REML measures were mostly done without knowledge of hypocretin deficiency. All participants underwent NPSG and MSLT. The only group that did not undergo systematic MSLT was the group of successive patients coming for a sleep disorder evaluation of any type (sample E), where only patients with a high probability of having narcolepsy underwent the MSLT. Although NPSG and MSLT REM onset parameters were evaluated locally using variable polysomnography equipment, multiple technicians, and bedroom settings, these measures are known to be highly reliable.[6,27-29]

Statistical Analyses

Statistical analyses were performed using the program R version 2.13.2 (The R Foundation). To control for the effect of age and sex on REM sleep propensity, we used the nearest-neighbor matching function from the MatchIt package for R,[30] with a case/control ratio of 1. The number of standard deviations of the distance measure within which to draw control units (caliper) was set to 0.25. To compare means or medians between samples in the 4 comparison groups, we used the Mann-Whitney *U* test or *t* test when relevant and used binomial logistic regression to adjust the results for sex and age (multivariate analysis). Significance level was set at 5% ($P < .05$, 2-sided).

To determine the REML associated with MSLT values providing optimal specificity for narcolepsy, we developed a novel graphical program aimed to assist clinical researchers and provide optimal diagnostic configurations and cutoff values according to test and quality of the receiver operating characteristic (ROC). Interactive plots make it easy to quickly select and compare diagnostic configurations and verify the quality of each selection. Bootstrapping and training/test split validation methods are available to measure 95% confidence intervals of test cutoffs and obtain unbiased ROC values. Our tool allows the user to evaluate a selected model's accuracy via a training/test set validation split and stability by a bootstrap procedure. Because there are 2 criteria of interest, sensitivity and specificity, there is no uniformly best model. To select and validate a single model, these criteria must be aggregated. The software aggregates these criteria by using a convex combination of the 2: $(1-α)$ sensitivity + α specificity, for a user selected α $\in [0, 1]$. Users specify α either explicitly by entering a value or implicitly by choosing what they would like to be the "optimal" point on the full data ROC curve.

Results for sensitivity, specificity, and 95% confidence intervals have all been secondarily confirmed with the epiR package for R.[31] Definitions and short explanations of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) are shown in the eTable in Supplement.[32]

Analytical Plan

Our goal was to determine if short REML during the NPSG could be used as a screening test to predict narcolepsy cases associated with hypocretin deficiency. Age and sex have already been shown to influence NPSG REML in patients with narcolepsy and controls[5,33-35]; we therefore matched patients and controls for these variables before comparisons. As a consequence, the total number of patients and controls was reduced to 516 pairs (some patients did not have a matching control). Optimal REML and MSLT finding cutoffs were first estimated by ROC curves in case vs control comparisons.

In comparison 1a, we used 89 narcolepsy cases with documented low CSF hypocretin-1 levels (sample A) and 89 random population controls (sample D) matched for age and sex. Only 89 cases of 169 (52.7%) could be matched because of a higher number of young adults in sample A vs sample D.

In comparison 1b, we used 427 narcolepsy cases with clear cataplexy and HLA-DQB1*06:02 positivity (sample B) and 427 random population controls (sample D) matched for age and sex. Not all (427 of 508; 84.1%) cases could be matched because of a higher number of young adults in sample B vs sample D.

Because studies have shown that 98% of comparison 1b cases have low CSF hypocretin-1 levels,[10] we then merged comparisons 1a and 1b into comparison 1. Therefore, we used 516 narcolepsy cases either with documented low CSF hypocretin-1 levels (sample A) or with clear cataplexy and HLA-DQB1*06:02 positivity (sample B) and 516 random population controls (sample D) matched for age and sex. Not all (516 of 677; 76.2%) cases could be matched because of a higher number of young adults in samples A and B vs sample D.

To extend our observation, we next performed comparisons in clinical samples independent from each other.

In comparison 2, we used a clinical naturalistic sample of 749 patients (sample E) referred to the sleep clinic and having undergone NPSG. From these, we selected narcolepsy cases either with documented low CSF hypocretin-1 levels or with clear cataplexy and HLA-DQB1*06:02 positivity, and we used all other patients as controls. This sample was used to assess characteristics of our REML criterion in clinical practice in sleep centers. In this sample, narcolepsy cases were not matched with patients with other sleep disorders to maintain the prevalence of the disease (to calculate PPV) and because this sample consists of successive patients. In this sample, only patients suspected for narcolepsy had MSLT findings; therefore, ROC analysis was only performed for optimal NPSG REML cutoffs, not MSLT findings.

In comparison 3, we used a sample of 254 successive patients with hypersomnia with high pretest probability for narcolepsy (sample F) referred to the sleep clinic for daytime sleepiness not due to sleep apnea and having undergone NPSG and MSLT. Among them, we selected narcolepsy cases either with documented low CSF hypocretin-1 levels or with clear

cataplexy and HLA-DQB1*06:02 positivity, and we used all other patients as controls. This sample was used to assess characteristics of our REML criterion, and to compare it with the MSLT findings, in clinical practice among patients with hypersomnia. Cases were not matched with controls for the same reasons noted in comparison 2.

In comparison 4, we used 118 patients with narcolepsy with low CSF hypocretin-1 levels (sample A) and 118 patients with narcolepsy with normal CSF hypocretin-1 levels (sample C) matched for age and sex as controls. Therefore, the patients with narcolepsy with low CSF hypocretin-1 levels in this comparison partially overlap with the patients with narcolepsy with low CSF hypocretin-1 levels from comparison 1. Seventy-three (61.9%) of the 118 controls with normal CSF hypocretin-1 levels had narcolepsy without cataplexy; for these patients, the main diagnostic criterion was not the presence of cataplexy or low CSF hypocretin-1 level, but MSLT positivity. This induces an inclusion bias, and therefore, ROC analysis was only performed for optimal NPSG REML cutoffs, not MSLT results (mandated to be positive for diagnostic inclusion).

Therefore, the total numbers of participants included in the analysis were 1749 different patients (6 patients were used twice, in comparisons 1 and 4) and 516 controls. Most patients (99.0%; n = 1731, see **Table 2** for details) and 17.2% (n = 89) of controls had genetic typing for HLA-DQB1*06:02.

Knowing that at best about 40% to 60% of patients with narcolepsy would have a pathologically abnormal REML during NPSG,[11] we focused on establishing an optimal cutoff with high specificity to use diagnostically, with the aim of avoiding conducting the MSLT the following day. We therefore determined inflection points of the ROC curves for best specificity.[32]

Results

Short REML During NPSG Is a Specific Marker of Narcolepsy
The NPSG REML distribution in narcolepsy/hypocretin deficiency was more skewed to the left than in age- and sex-matched controls (median NPSG REML = 15 minutes and 93 minutes, respectively) (**Figure 1**). Specificity and sensitivity for different cutoffs, with 95% confidence intervals and numbers of patients who tested positive over the total, in the various comparisons are reported in **Table 3**.

Figure 2A shows ROC curves in comparisons 1a, 1b, and 1, with optimal cutoffs of REML of 17 minutes or less, 21 minutes or less, and 18 minutes or less, respectively. These results were virtually identical. Figure 2B shows ROC curves in comparisons 2, 3, and 4, with optimal cutoffs of REML of 16 minutes or less, 17 minutes or less, and 8 minutes or less, respectively. Based on these findings, and considering that SOREMPs on the MSLT are defined as REML of 15 minutes or less, we opted for the same criterion for NPSG REML, ie, 15 minutes or less, as a screening tool for hypocretin deficiency prior to the MSLT. Using the 15-minute cutoff (Table 3), specificity was high (82.2%-99.6%) and sensitivity was low (35.7%-57.4%) in all comparisons. Positive predictive value was high (92.1%) in the relevant high-probability sample (comparison 3).

Short REML During NPSG Plus HLA-DQB1*06:02 Positivity In Narcolepsy
Because almost all patients with narcolepsy and low CSF hypocretin-1 levels (approximately 99%) were HLA-DQB1*06:02 positive, we also assessed whether combining HLA-DQB1*06:02 positivity with short REML on the NPSG improved specificity. This analysis could only be performed in comparison 1a (only the patients with documented low CSF hypocretin-1 levels), since at least some of the patients in all the other comparisons were selected based on HLA allele positivity or negativity (see the Methods section). Combining REML of 15 minutes or less and HLA-DQB1*06:02 positivity gave a specificity of 98.9% and a sensitivity of 43.8% (Table 3). This result indicates that using HLA-DQB1*06:02 positivity further increases specificity from 97.8% to 98.9%.

Short REML During NPSG Plus MSLT Findings In Narcolepsy
The MSLT specificity and sensitivity for narcolepsy/hypocretin deficiency were, respectively, 98.6% and 92.9% in comparison 1 and 71.2% and 93.4% in comparison 3, where PPV and NPV were 75.0% and 92.2%, respectively. Optimal MSLT parameters were determined by multiple ROC curves: MSL of 8 minutes or less associated with 1 or more SOREMPs (specificity, 96.1% and sensitivity, 95.6%) in comparison 1 and MSL of 8 minutes or less associated with 3 or more SOREMPs in comparison 3 (specificity, 85.6%; sensitivity, 86.1%; PPV, 84.7%; and NPV, 86.9%) (Table 3). Combining NPSG REML of 15 minutes or less as a screening test, and MSLT positivity if the screening test is negative, resulted in 97.8% specificity and 94.4% sensitivity in comparison 1 and 67.4% specificity, 94.2% sensitivity, 72.8% PPV, and 92.7% NPV in comparison 3. Combining these results with NPSG REML of 15 minutes or less as a complementary test to the MSLT (ie, using an NPSG REML ≤15 minutes as a supplementary SOREMP and therefore having a positive MSLT finding with MSL ≤8 minutes and only 1 SOREMP associated with NPSG REML ≤15 minutes) gave similar results (Table 3), indicating that adding 1 SOREMP to the MSLT when an NSPG SOREMP (REML ≤15 minutes) has been observed (about 50% of cases) provides little improvement in diagnostic value. Finally, the combination of an NPSG REML of 15 minutes or less associated with HLA-DQB1*06:02 positivity as a screening test and MSLT positivity if the screening test is negative resulted in 98.8% specificity and 90.1% sensitivity in comparison 1a (Table 3), the only setting in which the testing of this combination was possible because of sample selection (see the Methods section).

Discussion

We found the MSLT to be an excellent diagnostic tool for narcolepsy/hypocretin deficiency (cases with low CSF hypocretin-1 levels, plus patients with cataplexy and HLA-DQB1*06:02 positivity), reaching a sensitivity of 92.9% and a specificity of 98.6% vs age- and sex-matched random controls (Table 3). In a sample of 170 narcolepsy cases compared with 1913 patients with other sleep disorders, using clinical interview and MSLT as the gold standard, Aldrich et al[11] had found

Table 2. Demographic, Clinical, Polysomnographic, and Biological Characteristics of Comparisons 1a, 1b, 1, 2, 3, and 4[a]

	Comparison 1a[b]		Comparison 1b[c]		Comparison 1[d]	
	Patients (n = 89)	Controls (n = 89)	Patients (n = 427)	Controls (n = 427)	Patients (n = 516)	Controls (n = 516)
Age, mean (SD), y	30.9 (17.5)	33.9 (19.0)	39.8 (16.8)	41.1 (15.2)	38.2 (17.2)	39.9 (16.1)
Men, No. (%)	46 (51.7)	50 (56.2)	209 (48.9)	218 (51.1)	255 (49.4)	268 (51.9)
Narcolepsy, No. (%)						
With cataplexy	87 (97.8)	NA	427 (100.0)	NA	514 (99.6)	NA
Without cataplexy	2 (2.2)	NA	0	NA	2 (0.4)	NA
BMI, mean (SD)	27.1 (4.7) [85][e]	25.0 (6.6)	28.4 (6.2) [372]	28.4 (6.6)	28.1 (6.0) [457]	27.8 (6.7)
MSLT MSL, mean (SD), min	2.5 (2.3) [82][e]	14.6 (4.2) [82]	2.5 (2.2) [399][e]	11.0 (4.3)	2.5 (2.2) [481][e]	11.6 (4.5) [509]
MSLT SOREMPs, mean (SD)	3.2 (1.3) [81][e]	0.4 (0.8) [83]	3.5 (1.1) [399][e]	0.2 (0.6)	3.4 (1.2) [480][e]	0.2 (0.6) [510]
NPSG REML, mean (SD), min	72.4 (88.3)[e]	110.2 (58.0)	47.9 (68.1)[e]	116.4 (69.2)	52.1 (72.5)[e]	115.3 (67.4)
NPSG REML ≤15 min, No. (%)	39 (43.8)[e]	2 (2.2)	222 (52.0)[e]	2 (0.4)	261 (50.6)[e]	4 (0.8)
NPSG TST, mean (SD), min	386.4 (78.5) [34]	407.4 (69.9)	405.2 (69.3) [311][e]	426.4 (60.4)	403.4 (70.4) [345][e]	423.1 (62.5)
NPSG AHI, mean (SD)	5.5 (9.0) [64]	5.9 (6.4) [32]	6.1 (10.1) [304][e]	2.9 (8.8) [413]	6.0 (9.9) [368][e]	3.1 (8.6) [445]
HLA-DQB1*06:02 positive, No. (%)	87 (97.8)[e]	16 (18.0)	427 (100.0)	NA	514 (99.6)[e]	16 (18.0) [89]
CSF Hcrt level, mean (SD), pg/mL	5.0 (13.2)	NA	NA	NA	5.0 (13.2) [89]	NA

	Comparison 2[f]		Comparison 3[g]		Comparison 4[h]	
	Patients With Narcolepsy (n = 14)	Patients with Other Sleep Disorders (n = 735)	Patients With Narcolepsy (n = 122)	Patients With Other Hypersomnias (n = 132)	Patients With Low CSF Hcrt Levels (n = 118)	Patients With Normal CSF Hcrt Levels (n = 118)
Age, mean (SD), y	46.9 (24.3) [13]	45.2 (13.4)	32.3 (12.1) [120]	33.4 (14.1) [129]	28.3 (13.6)	29.4 (11.1)
Men, No. (%)	6 (42.9)	416 (56.6)	80 (67.2)	72 (54.5)	58 (49.2)	56 (47.5)
Narcolepsy, No. (%)						
With cataplexy	14 (100)[e]	1 (0.1)	118 (96.7)[e]	18 (13.6)	108 (91.5)[e]	45 (38.1)
Without cataplexy	0[e]	2 (0.3)	4 (3.3)[e]	32 (24.2)	10 (8.5)[e]	73 (61.9)
BMI, mean (SD)	26.8 (4.8)	27.5 (6.7) [722]	26.2 (7.4) [62][e]	23.9 (4.0) [93]	25.9 (4.7) [114][e]	25.2 (6.3) [113]
MSLT MSL, mean (SD), min	NA	NA	3.1 (2.5)[e]	6.9 (4.6)	2.8 (2.2) [112][e]	4.8 (3.1) [116]
MSLT SOREMPs, mean (SD)	NA	NA	3.6 (1.2)[e]	1.2 (1.3)	3.4 (1.3) [111][e]	2.7 (1.1) [116]
NPSG REML, mean (SD), min	127.0 (168.1)[e]	141.0 (83.6)	51.6 (75.8)[e]	100.8 (63.9)	66.6 (82.2)[e]	85.9 (68.0)
NPSG REML ≤15 min, No. (%)	5 (35.7)[e]	3 (0.4)	70 (57.4)[e]	6 (4.5)	55 (46.6)[e]	21 (17.8)
NPSG TST, mean (SD), min	387.8 (81.9)	357.5 (82.2) [734]	417.1 (58.1) [121][e]	433.1 (62.2)	423.3 (69.0) [55]	420.1 (69.4) [49]
NPSG AHI, mean (SD)	13.9 (12.3) [13]	23.6 (21.3) [517]	5.2 (9.0) [121]	5.3 (10.3) [104]	4.5 (8.2) [74]	5.9 (12.2) [99]
HLA-DQB1*06:02 positive, No. (%)	12 (92.3) [13][e]	179 (24.7) [726]	121 (100) [121][e]	21 (16.2) [130]	117 (100) [117][e]	13 (11.4) [114]
CSF Hcrt level, mean (SD), pg/mL	4.7 (7.4) [5]	NA	24.0 (17.3) [60][e]	228.8 (61.6) [14]	16.7 (24.5)[e]	310.5 (77.5) [52]

Abbreviations: AHI, apnea-hypopnea index; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CSF, cerebrospinal fluid; Hcrt, hypocretin-1; MSL, mean sleep latency; MSLT, Multiple Sleep Latency Test; NA, not available or not applicable; NPSG REML, nocturnal polysomnography rapid eye movement sleep latency; SOREMPs, sleep-onset rapid eye movement periods; TST, total sleep time.

[a] Numbers in brackets are the number of subjects used for calculations when data were partly missing. Results are adjusted for age and sex.

[b] Patients with narcolepsy and low CSF Hcrt levels vs controls.

[c] Patients with narcolepsy-cataplexy and HLA-DQB1*06:02 positivity vs controls.

[d] Patients with narcolepsy with low CSF Hcrt levels, or cataplexy and HLA-DQB1*06:02 positivity, vs controls.

[e] $P < .05$.

[f] Seven hundred forty-nine sleep clinic patients undergoing NPSG (patients with narcolepsy with low CSF Hcrt levels, or cataplexy and HLA-DQB1*06:02 positivity, vs patients with other sleep disorders).

[g] Two hundred fifty-four sleep clinic patients at high risk for narcolepsy (patients with narcolepsy with low CSF Hcrt levels, or cataplexy and HLA-DQB1*06:02 positivity, vs patients with other hypersomnias [narcolepsy, idiopathic hypersomnia, and Kleine-Levin syndrome] and normal CSF Hcrt levels).

[h] Patients with narcolepsy and low CSF Hcrt levels vs patients with narcolepsy and normal CSF Hcrt levels (positive MSLT findings).

a similar specificity (95%) but a lower sensitivity (78%) for the MSLT. Results of the MSLT are correlated with CSF hypocretin-1 levels.[36,37] Therefore, our higher sensitivity can be explained by the use of a biological marker as the gold standard, reducing the heterogeneity of our population sample. However, a 1.4% false-positive rate is still high considering the large number of patients who undergo MSLT for suspicion of nar-

colepsy, leading to a likely overdiagnosis of narcolepsy without cataplexy and making it essential to interpret these test findings within a specific clinical context.[38]

To our knowledge, our study is the first to report on the sensitivity and specificity of NPSG short REML in a large number of narcolepsy/hypocretin deficiency cases in comparison with a random population sample (Table 3). We also studied

Figure 1. Overlaid Histograms of Rapid Eye Movement Sleep Latency (REML) Distributions for Narcoleptic Cases and Controls
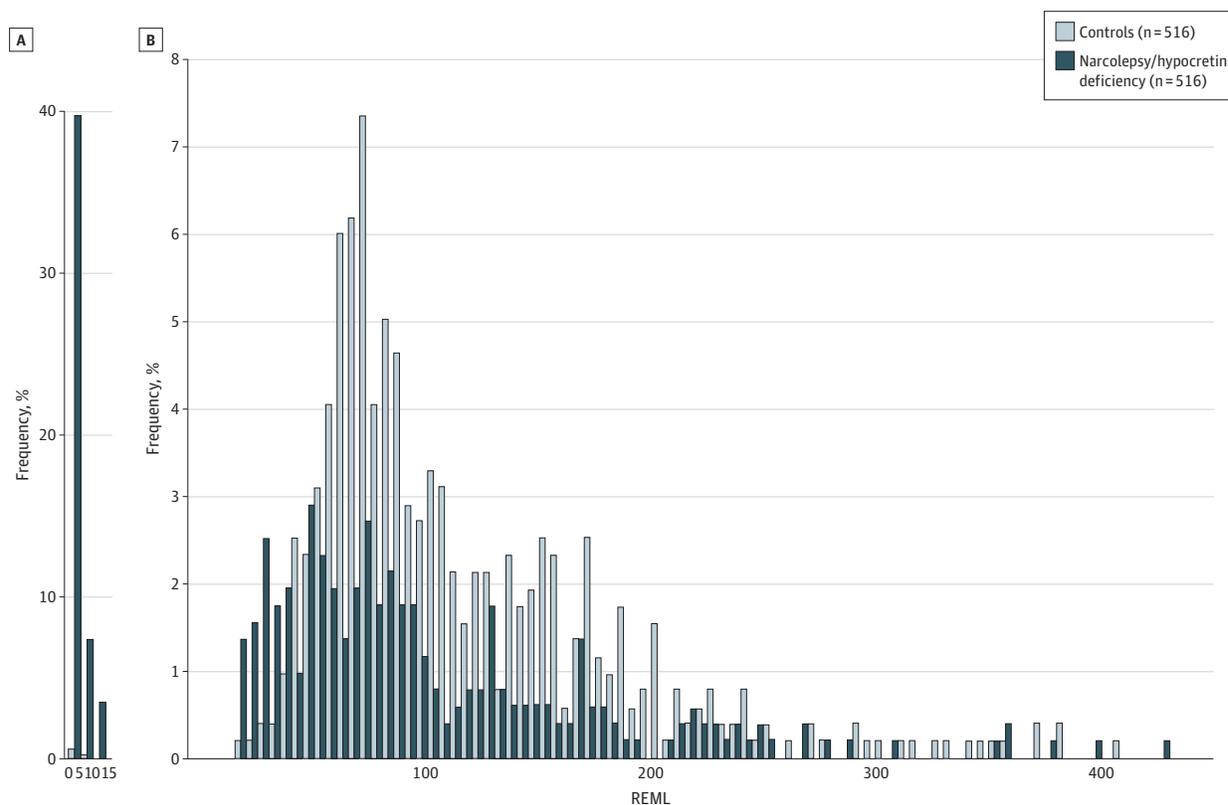
A, Distribution of REML of 15 minutes or less. B, Distribution of REML more than 15 minutes (y-axis has been scaled up). Each bar represents the frequency for a 5-minute REML range.

this biological marker in patients with other sleep disorders, patients with central nervous system hypersomnias who had a high pretest probability for narcolepsy, and patients with narcolepsy based on positive MSLT findings but with normal CSF hypocretin-1 levels (unknown etiology). The choice of using REML of 15 minutes or less as a cutoff was based on our finding that optimal specificity was reached with cutoffs between 8 and 21 minutes, depending on the comparison, and the preexisting definition of an SOREMP on the MSLT as REML of 15 minutes or less. Although longer REML cutoffs may be used to maximize sensitivity, our goal was to select a conservative REML value that would remain robust to changes in clinical context (Table 3). Sensitivity was also less of a concern to us because a negative REML finding was to be followed by a regular MSLT, keeping a similar sensitivity.

The occurrence of NPSG short REML in 30% to 40% of narcolepsy cases has already been described.[11,39] However, in these studies, the gold standard for narcolepsy was based on clinical and MSLT results. Using cases with documented low CSF hypocretin-1 levels (comparison 1a), we observed that an SOREMP at NPSG was highly specific (97.8%) vs controls. This suggests the likelihood of overdiagnosis is low, but sensitivity was poor (43.8%), meaning that half of the patients would be missed with this test. Very similar findings were reported in cases with clear cataplexy and HLA-DQB1*06:02 positivity

(comparison 1b), indicating that these 2 groups are equivalent and that, as expected, almost all cases with cataplexy and HLA-DQB1*06:02 positivity likely have low CSF hypocretin-1 levels if measured.

The comparison with random population controls was helpful but not applicable to the clinical context. We therefore examined how REML behaved diagnostically in successive samples with increasing pretest probability for narcolepsy. In successive patients undergoing a sleep study evaluation (ie, primarily patients with sleep apnea also complaining of sleepiness [comparison 2]), the test was similarly accurate (specificity, 99.6%; sensitivity, 35.7%), consistent with the results of a prior study in a clinic population that did not use CSF hypocretin measurements (specificity, 98.5%; sensitivity, 29%).[11] Because the principal differential diagnoses for narcolepsy/hypocretin deficiency are narcolepsy without hypocretin deficiency, idiopathic hypersomnia, and Kleine-Levin syndrome (all requiring an MSLT for the differential diagnosis), we also tested specificity and sensitivity for a short REML in successive patients with central hypersomnia needing MSLTs for differential diagnosis (comparison 3). In this sample, the most clinically relevant for clinicians, an SOREMP at night was more specific than a positive MSLT finding (95.4% vs 71.2%). It also had a higher PPV (92.1% vs 75.0%) for the diagnosis of narcolepsy with hypocretin deficiency. These findings show

Table 3. Characteristics of Different NPSG REML, MSLT, and HLA-DQB1*06:02 Criteria Combinations to Diagnose Narcolepsy/Hcrt Deficiency in Different Settings[a]

	AUC (95% CI)	Sensitivity		Specificity		PPV		NPV	
		% (95% CI)	No. Positive/ Total No.	% (95% CI)	No. Positive/ Total No.	% (95% CI)	No. Positive/ Total No.	% (95% CI)	No. Positive/ Total No.
Comparison 1a[b]									
NPSG REML ≤15 min	0.704 (0.625-0.786)	43.8 (33.3-54.7)	39/89	97.8 (92.1-99.7)	87/89				
NPSG REML ≤17 min	0.704 (0.625-0.786)	44.9 (34.4-55.9)	40/89	97.8 (92.1-99.7)	87/89				
NPSG REML ≤15 min and HLA-DQB1*06:02	0.933 (0.896-0.968)	43.8 (33.5-54.1)	39/89	98.9 (96.7-100.0)	88/89				
NPSG REML ≤15 min and HLA-DQB1*06:02 or MSLT MSL ≤8 min and MSLT SOREMPs ≥2	0.986 (0.977-0.996)	90.1 (83.6-96.6)	73/81	98.8 (96.4-100.0)	81/82				
Comparison 1b[c]									
NPSG REML ≤15 min	0.820 (0.789-0.850)	52.0 (47.1-56.8)	222/427	99.5 (98.3-99.9)	425/427				
NPSG REML ≤21 min	0.820 (0.789-0.850)	53.9 (49.0-58.7)	230/427	99.5 (98.3-99.9)	425/427				
Comparison 1[d]									
NPSG REML ≤15 min	0.799 (0.771-0.826)	50.6 (46.3-54.9)	261/516	99.2 (98.5-100.0)	512/516				
NPSG REML ≤18 min	0.799 (0.771-0.826)	51.6 (47.2-55.9)	266/516	99.2 (98.5-100.0)	512/516				
MSLT MSL ≤8 min and SOREMPs ≥2	0.928 (0.904-0.946)	92.9 (90.6-95.2)	445/479	98.6 (97.6-99.6)	502/509				
MSLT MSL ≤8 min and SOREMPs ≥1	0.928 (0.904-0.946)	95.6 (93.8-97.4)	458/479	96.1 (94.4-97.8)	489/509				
NPSG REML ≤15 min or MSLT MSL ≤8 min and MSLT SOREMPs ≥2	0.960 (0.946-0.974)	94.4 (92.3-96.4)	452/479	97.8 (96.6-99.1)	498/509				
MSLT MSL ≤8 min and SOREMPs ≥2 (when NPSG REML ≤15 min counts as 1 SOREMP)	0.988 (0.984-0.993)	93.3 (91.0-95.5)	447/479	98.6 (97.5-99.5)	502/509				
Comparison 2[e]									
NPSG REML ≤15 min	0.704 (0.524-0.907)	35.7 (10.6-60.8)	5/14	99.6 (99.1-100.0)	732/735	62.5 (30.6-86.3)	5/8	98.8 (97.7-99.4)	732/741
NPSG REML ≤16 min	0.704 (0.524-0.907)	42.9 (16.9-68.8)	6/14	99.6 (99.1-100.0)	732/735	66.7 (35.4-87.9)	6/9	98.9 (97.9-99.5)	732/740
Comparison 3[f]									
NPSG REML ≤15 min	0.765 (0.707-0.831)	57.4 (48.1-66.3)	70/122	95.4 (90.4-98.3)	126/132	92.1 (83.6-97.0)	70/76	70.8 (63.5-77.3)	126/178
NPSG REML ≤17 min	0.765 (0.707-0.831)	58.2 (48.9-67.1)	71/122	95.5 (90.4-98.3)	126/132	92.2 (83.8-97.1)	71/77	71.1 (63.9-77.7)	126/177
MSLT MSL ≤8 min and SOREMPs ≥2	0.904 (0.868-0.904)	93.4 (87.5-97.1)	114/122	71.2 (62.7-78.8)	94/132	75.0 (67.3-81.7)	114/152	92.2 (85.1-96.6)	94/102
MSLT MSL ≤8 min and SOREMPs ≥3	0.904 (0.868-0.940)	86.1 (78.6-91.7)	105/122	85.6 (78.4-91.1)	113/132	84.7 (77.1-90.5)	105/124	86.9 (79.9-92.2)	113/130
NPSG REML ≤15 min or MSLT MSL ≤8 min and MSLT SOREMPs ≥2	0.931 (0.896-0.957)	94.2 (88.5-97.7)	115/122	67.4 (58.7-75.3)	89/132	72.8 (65.1-79.6)	115/158	92.7 (85.6-97.0)	89/96
MSLT MSL ≤8 min and SOREMPs ≥2 (when NPSG REML ≤15 min counts as 1 SOREMP)	0.921 (0.887-0.953)	93.4 (87.5-97.1)	114/122	69.7 (61.1-77.4)	92/132	74.0 (66.4-80.8)	114/154	92.0 (84.8-96.5)	92/100

Table 3. Characteristics of Different NPSG REML, MSLT, and HLA-DQB1*06:02 Criteria Combinations to Diagnose Narcolepsy/Hcrt Deficiency in Different Settings[a] (continued)

	AUC (95% CI)	Sensitivity		Specificity		PPV		NPV	
		% (95% CI)	No. Positive/ Total No.	% (95% CI)	No. Positive/ Total No.	% (95% CI)	No. Positive/ Total No.	% (95% CI)	No. Positive/ Total No.
Comparison 4[g]									
NPSG REML ≤8 min	0.628 (0.558-0.702)	44.1 (35.1-53.0)	52/118	87.3 (81.3-93.3)	103/118				
NPSG REML ≤15 min	0.628 (0.558-0.702)	46.6 (37.6-55.6)	55/118	82.2 (75.3-89.1)	97/118				

Abbreviations: AUC, area under the curve; Hcrt, hypocretin-1; MSL, mean sleep latency; MSLT, Multiple Sleep Latency Test; NPSG REML, nocturnal polysomnography rapid eye movement sleep latency; NPV, negative predictive value; PPV, positive predictive value; SOREMPs, sleep-onset rapid eye movement periods.

[a] The MSLT specificity and sensitivity for narcolepsy/Hcrt deficiency are, respectively, 98.6% and 92.9% in comparison 1 and 71.2% and 93.4% in comparison 3, where PPV and NPV are, respectively, 75.0% and 92.2%. Optimal MSLT parameters were determined by multiple receiver operating characteristic curves: MSL of 8 minutes or less associated with 1 or more SOREMPs (specificity, 96.1%; sensitivity, 95.6%) were optimal in comparison 1, and MSL of 8 minutes or less associated with 3 or more SOREMPs were optimal in comparison 3 (specificity, 85.6%; sensitivity, 86.1%; PPV, 84.7%; NPV, 86.9%). Using the 15-minute cutoff, specificity was high (82.2%-99.6%) and sensitivity was low (35.7%-57.4%) in all comparisons. The PPV was high (92.1%) in the relevant high-probability sample (comparison 3). Using HLA-DQB1*06:02 positivity (comparison 1a) further increases specificity from 97.8% to 98.9%. Combining NPSG REML of 15 minutes or less as a screening test and MSLT positivity if the screening test findings are negative results in 97.8% specificity and 94.4% sensitivity in comparison 1 and 67.4% specificity,

94.2% sensitivity, 72.8% PPV, and 92.7% NPV in comparison 3. Combining NPSG REML of 15 minutes or less with HLA-DQB1*06:02 positivity as a screening test and MSLT positivity if the screening test findings are negative results in 98.8% specificity and 90.1% sensitivity in comparison 1a.

[b] Eighty-nine patients with narcolepsy and low CSF Hcrt levels vs 89 controls.

[c] Four hundred twenty-seven patients with narcolepsy-cataplexy and HLA-DQB1*06:02 positivity vs 427 controls.

[d] Five hundred sixteen patients with narcolepsy with low CSF Hcrt levels, or cataplexy and HLA-DQB1*06:02 positivity, vs 516 controls.

[e] Fourteen patients with narcolepsy with low CSF Hcrt levels, or cataplexy and HLA-DQB1*06:02 positivity, vs 735 patients with other sleep disorders.

[f] One hundred twenty-two patients with narcolepsy with low CSF Hcrt levels, or cataplexy and HLA-DQB1*06:02 positivity, vs 132 patients with other hypersomnias (narcolepsy, idiopathic hypersomnia, and Kleine-Levin syndrome) and normal CSF Hcrt levels.

[g] One hundred eighteen patients with narcolepsy and low CSF Hcrt levels vs 118 patients with narcolepsy and normal CSF Hcrt levels (positive MSLT findings).

that even in a sample of patients being evaluated for possible narcolepsy (high pretest probability sample) vs other hypersomnias, a short REML (≤15 minutes) has high specificity and PPV and may be considered diagnostic without the need for an MSLT in approximately half of the cases. Patients with a strong suspicion for narcolepsy but a normal REML at night would continue with a 4- to 5-nap MSLT according to typical diagnostic recommendations.[38]

Using these combined criteria in the most clinically relevant sample, overall sensitivity and specificity of the combined test (REML ≤15 minutes or MSLT MSL ≤8 minutes and MSLT SOREMPs ≥2) were estimated to be 94.2% and 67.4%, respectively, similar to the MSLT alone (93.4% and 71.2%) (Table 3). Although specificity of NPSG REML of 15 minutes or less was very high (95.4%) in these patients, 4.6% would still receive a false-positive diagnosis of narcolepsy/hypocretin deficiency, a result that has to be put in balance with the potential cost savings of avoiding the MSLT. Although the cause of nighttime SOREMPs in otherwise healthy individuals (true false positives) has not been systematically investigated, laboratory studies suggest that chronic sleep deprivation or shift work/circadian misalignment[6,40-42] can induce SOREMPs in some circumstances.

In a last comparison (comparison 4), we explored whether REML differentiated patients with narcolepsy who have a measurable biochemical defect (low CSF hypocretin-1 level) vs those who did not have abnormal CSF hypocretin-1 levels but still carried a diagnosis of narcolepsy based on MSLT data or symptoms suggestive of cataplexy. This was a very conservative comparison: lumbar punctures were only performed when there was a very high suspicion of narcolepsy and there was a need to clarify the diagnosis. Yet, even in this case, the specificity was 82.2%. The fact that specificity is slightly lower (82.2% vs 99.5%) in other non–hypocretin-deficient samples can be understood in 2 ways. First, these subjects were selected on the basis of having had a positive MSLT finding with SOREMPs; thus, they may simply have a tendency toward rapid REM sleep onset independent of narcolepsy (due to sleep deprivation, circadian misalignment, or a true central nervous system pathology). Second, it is possible that a subset of patients with normal hypocretin levels have a disease with a cause similar to hypocretin deficiency but with normal hypocretin levels, for example, partial hypocretin cell loss, defects in the hypocretin receptor, or signal transduction pathways. This could explain abnormal REM sleep during the MSLT and a few additional positive subjects with short REML during NSPG. To date, we are aware of no evidence showing that patients with narcolepsy and low and normal CSF hypocretin-1 levels should be treated differently, but it may be argued that patients without a clear biological cause should be treated more carefully, avoiding strong stimulants and drugs of abuse, because it is not even known if their problem is lifelong.

The prevalence of severe excessive daytime sleepiness is 5% in the general population and incidence of narcolepsy has been estimated to be 1.37 per 100 000 person-years, corresponding to about 4000 new cases per year in the United States.[43] This implies that tens of thousands of patients likely undergo MSLT annually, either for the diagnosis of narco-

Figure 2. Receiver Operating Characteristic Curves of Comparisons 1a, 1b, 1, 2, 3, and 4
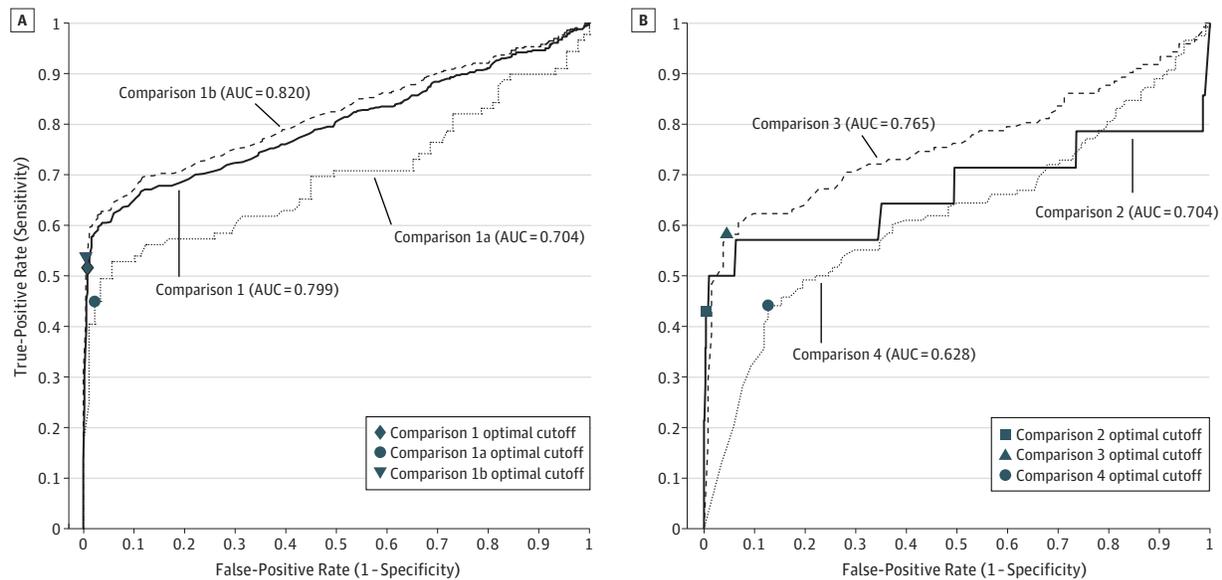
A, Receiver operating characteristic curves in 89 patients with narcolepsy with low cerebrospinal fluid (CSF) hypocretin-1 levels vs 89 controls (comparison 1a), 427 patients with narcolepsy with cataplexy and HLA-DQB1*06:02 positivity vs 427 controls (comparison 1b), and 516 patients with narcolepsy with either low CSF hypocretin-1 levels or cataplexy and HLA-DQB1*06:02 positivity vs 516 controls (comparison 1). In patients with low CSF hypocretin-1 levels (comparison 1a), the optimal cutoff was a rapid eye movement sleep latency (REML) of 17 minutes or less (specificity, 97.8% [95% CI, 92.1-99.7]; sensitivity, 44.9% [95% CI, 34.4-55.9]; area under the curve [AUC], 0.704 [95% CI, 0.625-0.786]); in patients with cataplexy and HLA-DQB1*06:02 positivity (comparison 1b), the optimal cutoff was an REML of 21 minutes or less (specificity, 99.5% [95% CI, 98.3-99.9]; sensitivity, 53.9% [95% CI, 49.0-58.7]; AUC, 0.820 [95% CI, 0.789-0.850]). An REML of 18 minutes or less (specificity, 99.2% [95% CI, 98.5-100.0]; sensitivity, 51.6% [95% CI, 47.2-55.9]) was the best cutoff for optimal specificity in comparison 1 (AUC, 0.799 [95% CI, 0.771-0.826]). These results were virtually identical. B, Receiver operating characteristic curves in 14 patients with narcolepsy/hypocretin deficiency vs 735 patients with other sleep disorders (comparison 2), 122 patients with

narcolepsy/hypocretin deficiency vs 132 patients with narcolepsy without hypocretin deficiency, idiopathic hypersomnia, and Kleine-Levin syndrome (high pretest probability sample, comparison 3), and 118 patients with narcolepsy with low CSF hypocretin-1 levels vs 118 patients with narcolepsy with normal CSF hypocretin-1 levels (comparison 4). In comparison 2, an REML of 16 minutes or less (specificity, 99.6% [95% CI, 99.1-100.0]; sensitivity, 42.9% [95% CI, 16.9-68.8]) was the best cutoff for the diagnosis of narcolepsy/hypocretin deficiency vs other sleep disorders (AUC, 0.704 [95% CI, 0.524-0.907]). In comparison 3, an REML of 17 minutes or less (specificity, 95.5% [95% CI, 90.4-98.3]; sensitivity, 58.2% [95% CI, 48.9-67.1]) was the best cutoff for the diagnosis of narcolepsy with low CSF hypocretin-1 levels vs narcolepsy with normal hypocretin-1 levels, idiopathic hypersomnia, and Kleine-Levin syndrome (AUC, 0.765 [95% CI, 0.707-0.831]). In comparison 4, an REML of 8 minutes or less (specificity, 87.3% [95% CI, 81.3-93.3]; sensitivity, 44.1% [95% CI, 35.1-53.0]) was the best cutoff for the diagnosis of narcolepsy with low CSF hypocretin-1 levels vs narcolepsy with normal CSF hypocretin-1 levels (AUC, 0.628 [95% CI, 0.558-0.702]).

lepsy or for the objective evaluation of hypersomnolence in the absence of a nocturnal sleep pathology, in which case a maintenance of wakefulness test could also be advised.[38] Considering a US Medicare payment of $410.43 per MSLT as of 2011 (conservative estimate, since Medicare payment for MSLT is low relative to other payers), at least $10 million could be saved in direct costs, not including time loss for the patient.

Our results confirm the clinical utility of systematically analyzing nocturnal REML results in patients undergoing NPSG before an MSLT, when narcolepsy is suspected. Indeed, each year, millions of patients undergo polysomnography for the evaluation of sleep apnea, and in all likelihood, patients with unrecognized hypocretin deficiency would be flagged for reevaluation. The prevalence of nocturnal SOREMPs in the general population and patients with other sleep disorders (mostly sleep apnea) is quite low (0.8% and 0.4%, respectively)[6,12]; thus, a nocturnal SOREMP should alert the clinician to the possibility of a missed narcolepsy/hypocretin deficiency diagnosis. Reevaluation of cataplexy (which can be

missed during clinical interviews), confounding factors (sleep deprivation, circadian misalignment, and shift work), and, if needed, NPSG/MSLT testing, HLA-DQB1:0602 typing, and measurement of CSF hypocretin-1 level would be sequentially considered until the situation is clarified. The prevalence of this syndrome is approximately 1 in 3000, and many cases are obese and with a codiagnosis of sleep-disordered breathing (9 of 14 patients with narcolepsy/hypocretin deficiency had comorbid mild/moderate sleep-disordered breathing in our naturalistic sample).[44] Also, a recent study has found that many African American patients have narcolepsy with hypocretin deficiency but no cataplexy, making the diagnosis more difficult in this population.[45]

A last piece of the puzzle in the diagnosis of narcolepsy/hypocretin deficiency is HLA-DQB1*06:02 typing, a marker present in 98% of cases. Because 25% of controls also carry HLA-DQB1:06*02, the test has no specificity by itself but can be useful to exclude narcolepsy/hypocretin deficiency or to further increase specificity in the context of other tests.

Our study has a number of important limitations. First, our control sample was mainly adults, and matching excluded a significant number of pediatric cases. Therefore, the reported findings cannot be extended beyond the adult population. Second, the study is retrospective, and scoring was thus not entirely blind to diagnosis. Technicians may have known when a patient with narcolepsy underwent NPSG and could have been more attentive to NSPG REML. We, however, believe this to be unlikely because short REML is not currently a diagnostic criterion, and scoring often occurs independent of the clinical visit. Third, our financial analysis is approximate, and costs or savings have likely not all been considered. A full 70% of patients suspected for narcolepsy had an NPSG REML of more than 15 minutes and therefore would still have to undergo an MLST, reducing the potential cost saving of avoiding the MSLT in these cases. Savings might also be partially reduced because detecting short REML during NPSG requires night technologists scoring "on the fly" and/or an interpreting physician (familiar with the patient) available to decide on whether to proceed with an MSLT. Finally, health insurance carriers may still require for a time a positive MSLT finding to provide treatment. Nonetheless, in addition to the evidence provided herein, the potential savings will help to recognize a short REML during NPSG as a diagnostic criterion.

Fourth, environmental laboratory conditions, polysomnography equipment, and scorers varied from site to site, and this could have affected some of the comparisons. The fact that REML scoring is known to be highly reliable across scorers[6,27-29] and that optimal REML cutoffs were similar across a diverse set of comparisons argues against clinically meaningful variations. One could also argue that REML can only be optimally measured after a prior night of habituation or in a comfortable environment at home. Indeed, REML is often shorter the second laboratory night after habituation[46] or at home vs in the laboratory.[47] This effect is, however, very small (a few minutes) and not known to apply to the very short REML range (≤15 minutes) observed in patients with narcolepsy/hypocretin deficiency. It is also almost certainly smaller than the effect of an uncomfortable laboratory night on subsequent MSLT findings.[47] Future studies aimed at studying REML at home using ambulatory polysomnography for several successive days may, however, be an interesting avenue to develop cheaper diagnostic alternatives.

Fifth, MSLT was performed according to the standard protocols published in 1986, or in 2005, depending on when the patient had been diagnosed in the sleep clinics.[14,24] The MSLT guidelines of 1986 state that "Drugs known to affect sleep latency...should be withdrawn for 2 weeks before MSLT testing."[24(p520)] These drugs include a number of over-the-counter agents and herbal products with anticholinergic properties, as well as alcohol and prescription drugs. In the community, withdrawing these drugs is the major obstacle to obtaining a valid MSLT result, and these drugs were not withdrawn in the Wisconsin Sleep Cohort sample. The new Practice Parameters for Clinical Use of the Multiple Sleep Latency Test published in 2005 address this difficulty by changing the wording to "REM suppressing medications should ideally be stopped 2 weeks before the MSLT."[14(p119)] Our data on NPSG REML, as well as on MSLT, might therefore be biased for some patients and controls for whom it was impossible to control if the REM-suppressing medication withdrawal recommendation was followed. This would be particularly true when comparing patients with controls (comparison 1), although data from the Wisconsin Sleep Cohort do not indicate a major effect on SOREMP prevalence (unlike shift work) in controls. Further, in our 2 independent naturalistic samples, ie, patients eventually diagnosed with narcolepsy and patients with other diagnoses, both groups were likely to have the same type of medications before NPSG or MSLT. Overall, the stability of the finding across groups makes it very unlikely that medication had a significant effect on the diagnostic validity of short NSPG REML in narcolepsy.

In summary, our findings show that among patients being evaluated for possible narcolepsy, short REML (≤15 minutes) at NPSG had high specificity and PPV and may be considered diagnostic for narcolepsy/hypocretin deficiency without the use of an MSLT; however, the absence of short REML requires a subsequent MSLT.

REFERENCES

1. Schenck CH, Bassetti CL, Arnulf I, Mignot E. English translations of the first clinical reports on narcolepsy and cataplexy by Westphal and Gélineau in the late 19th century, with commentary. *J Clin Sleep Med*. 2007;3(3):301-311.

2. American Academy of Sleep Medicine. *International Classification of Sleep Disorders: Diagnostic and Coding Manual*. 2nd ed. Westchester, IL: American Academy of Sleep Medicine; 2005.

3. Vogel G. Studies in psychophysiology of dreams, III: the dream of narcolepsy. *Arch Gen Psychiatry*. 1960;3:421-428.

4. Rechtschaffen A, Wolpert EA, Dement WC, Mitchell SA, Fisher C. Nocturnal sleep of narcoleptics. *Electroencephalogr Clin Neurophysiol*. 1963;15:599-609.

5. Singh M, Drake CL, Roth T. The prevalence of multiple sleep-onset REM periods in a population-based sample. *Sleep*. 2006;29(7):890-895.

6. 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):1609-1623.

7. Nishino S, Ripley B, Overeem S, Lammers GJ, Mignot E. Hypocretin (orexin) deficiency in human narcolepsy. *Lancet*. 2000;355(9197):39-40.

8. 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):1553-1562.

9. Mignot E, Hayduk R, Black J, Grumet FC, Guilleminault C. HLA DQB1*0602 is associated with cataplexy in 509 narcoleptic patients. *Sleep*. 1997;20(11):1012-1020.

10. Bourgin P, Zeitzer JM, Mignot E. CSF hypocretin-1 assessment in sleep and neurological disorders. *Lancet Neurol*. 2008;7(7):649-662.

11. Aldrich MS, Chervin RD, Malow BA. Value of the multiple sleep latency test (MSLT) for the diagnosis of narcolepsy. *Sleep*. 1997;20(8):620-629.

12. 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):426-431.

13. Marti I, Valko PO, Khatami R, Bassetti CL, Baumann CR. Multiple sleep latency measures in narcolepsy and behaviourally induced insufficient sleep syndrome. *Sleep Med*. 2009;10(10):1146-1150.

14. Littner MR, Kushida C, Wise M, et al; Standards of Practice Committee of the American Academy of Sleep Medicine. Practice parameters for clinical use of the Multiple Sleep Latency Test and the maintenance of wakefulness test. *Sleep*. 2005;28(1):113-121.

15. Kanbayashi T, Shimohata T, Nakashima I, et al. Symptomatic narcolepsy in patients with neuromyelitis optica and multiple sclerosis: new neurochemical and immunological implications. *Arch Neurol*. 2009;66(12):1563-1566.

16. George CF, Singh SM. Juvenile onset narcolepsy in an individual with Turner syndrome: a case report. *Sleep*. 1991;14(3):267-269.

17. Young T. Rationale, design and findings from the Wisconsin Sleep Cohort Study: toward understanding the total societal burden of sleep disordered breathing. *Sleep Med Clin*. 2009;4(1):37-46.

18. Drake C, Roehrs T, Breslau N, et al. The 10-year risk of verified motor vehicle crashes in relation to physiologic sleepiness. *Sleep*. 2010;33(6):745-752.

19. Johnson EO, Breslau N, Roth T, Roehrs T, Rosenthal L. Psychometric evaluation of daytime sleepiness and nocturnal sleep onset scales in a representative community sample. *Biol Psychiatry*. 1999;45(6):764-770.

20. Anic-Labat S, Guilleminault C, Kraemer HC, Meehan J, Arrigoni J, Mignot E. Validation of a cataplexy questionnaire in 983 sleep-disorders patients. *Sleep*. 1999;22(1):77-87.

21. Mignot E, Lin L, Rogers W, et al. Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. *Am J Hum Genet*. 2001;68(3):686-699.

22. Kushida CA, Littner MR, Morgenthaler T, et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep*. 2005;28(4):499-521.

23. Iber C, Ancoli-Israel S, Chesson A. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Westchester, IL: American Academy of Sleep Medicine; 2007.

24. Carskadon MA, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S. Guidelines for the Multiple Sleep Latency Test (MSLT): a standard measure of sleepiness. *Sleep*. 1986;9(4):519-524.

25. Hallmayer J, Faraco J, Lin L, et al. Narcolepsy is strongly associated with the T-cell receptor alpha locus [published correction appears in *Nat Genet*. 2009;41(7):859]. *Nat Genet*. 2009;41(6):708-711.

26. Hong SC, Leen-Kim, Park SA, et al. HLA and hypocretin studies in Korean patients with narcolepsy. *Sleep*. 2002;25(4):440-444.

27. Chen L, Ho CK, Lam VK, et al. Interrater and intrarater reliability in multiple sleep latency test. *J Clin Neurophysiol*. 2008;25(4):218-221.

28. Drake CL, Rice MF, Roehrs TA, Rosenthal L, Guido P, Roth T. Scoring reliability of the multiple sleep latency test in a clinical population. *Sleep*. 2000;23(7):911-913.

29. Folkerts M, Rosenthal L, Roehrs T, et al. The reliability of the diagnostic features in patients with narcolepsy. *Biol Psychiatry*. 1996;40(3):208-214.

30. Ho D, Imai K, King G, Stuart E. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal*. 2007;15(3):199-236. doi:10.1093/pan/mpl013.

31. Carstensen B, Plummer M, Laara E, Hills M. Epi: a package for statistical analysis in epidemiology. R package version 1.1.33. http://cran.r-project.org /web/packages/Epi/index.html. Published 2011. Accessed March 26, 2012.

32. Kraemer HC. *Evaluating Medical Tests: Objective and Quantitative Guidelines*. Newbury Park, CA: SAGE Publication, Inc; 1999.

33. Dauvilliers Y, Gosselin A, Paquet J, Touchon J, Billiard M, Montplaisir J. Effect of age on MSLT results in patients with narcolepsy-cataplexy. *Neurology*. 2004;62(1):46-50.

34. Mignot E, Young T, Lin L, Finn L, Palta M. Reduction of REM sleep latency associated with HLA-DQB1*0602 in normal adults. *Lancet*. 1998;351(9104):727.

35. Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV. 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):1255-1273.

36. Nakamura M, Kanbayashi T, Sugiura T, Inoue Y. Relationship between clinical characteristics of narcolepsy and CSF orexin-A levels. *J Sleep Res*. 2011;20(1, pt 1):45-49.

37. Baumann CR, Khatami R, Werth E, Bassetti CL. Hypocretin (orexin) deficiency predicts severe objective excessive daytime sleepiness in narcolepsy with cataplexy. *J Neurol Neurosurg Psychiatry*. 2006;77(3):402-404.

38. Arand D, Bonnet M, Hurwitz T, Mitler M, Rosa R, Sangal RB. The clinical use of the MSLT and MWT. *Sleep*. 2005;28(1):123-144.

39. Mosko SS, Shampain DS, Sassin JF. Nocturnal REM latency and sleep disturbance in narcolepsy. *Sleep*. 1984;7(2):115-125.

40. Miyasita A, Fukuda K, Inugami M. Effects of sleep interruption on REM-NREM cycle in nocturnal human sleep. *Electroencephalogr Clin Neurophysiol*. 1989;73(2):107-116.

41. Sasaki Y, Fukuda K, Takeuchi T, Inugami M, Miyasita A. Sleep onset REM period appearance rate is affected by REM propensity in circadian rhythm in normal nocturnal sleep. *Clin Neurophysiol*. 2000;111(3):428-433.

42. Khalsa SB, Conroy DA, Duffy JF, Czeisler CA, Dijk DJ. Sleep- and circadian-dependent modulation of REM density. *J Sleep Res*. 2002;11(1):53-59.

43. Ohayon MM. From wakefulness to excessive sleepiness: what we know and still need to know. *Sleep Med Rev*. 2008;12(2):129-141.

44. Sansa G, Iranzo A, Santamaria J. Obstructive sleep apnea in narcolepsy. *Sleep Med*. 2010;11(1):93-95.

45. Andlauer O, Moore H IV, Hong SC, et al. Predictors of hypocretin (orexin) deficiency in narcolepsy without cataplexy. *Sleep*. 2012;35(9):1247-55F.

46. Toussaint M, Luthringer R, Schaltenbrand N, et al. First-night effect in normal subjects and psychiatric inpatients. *Sleep*. 1995;18(6):463-469.

47. Edinger JD, Glenn DM, Bastian LA, et al. Daytime testing after laboratory or home-based polysomnography: comparisons of middle-aged insomnia sufferers and normal sleepers. *J Sleep Res*. 2003;12(1):43-52.