

# Development and Validation of the Nonrestorative Sleep Scale (NRSS)

Kate Wilkinson, M.Sc.<sup>1</sup>; Colin Shapiro, M.D.<sup>2</sup>

<sup>1</sup>University of Toronto, Toronto, ON, Canada; <sup>2</sup>Toronto Western Hospital, Toronto, ON, Canada

**Study Objectives:** Nonrestorative sleep (NRS) is defined as the subjective feeling that sleep has been insufficiently refreshing, often despite the appearance of physiologically normal sleep. While NRS has been shown to be associated with a variety of cognitive, affective, and medical complaints, there is currently no valid, reliable instrument available in the public domain for its assessment. The purpose of this study was to develop and validate the Nonrestorative Sleep Scale (NRSS).

**Participants:** The scale was administered to a sample of 226 (age: 46.7 ± 14.9 years; gender: 48% female) consecutive sleep clinic patients and to 30 control participants (age: 36.9 ± 12.5; gender: 53% female).

**Results:** Data screening led to a final instrument of 12 items, and factor analysis resulted in 4 factors accounting for 73.2% of total variance. The scale demonstrated excellent internal

reliability ( $\alpha = 0.88$ ) and good test-retest reliability ( $r = 0.72$ ). Preliminary evaluations of construct validity found that certain subscales correlated reasonably well with previously validated sleep, alertness, and affective scales. Comparisons between global NRSS scores and objective polysomnographic variables revealed a few very small but significant correlations.

**Conclusions:** Based on these findings, the NRSS was confirmed to be a valid and reliable tool for the assessment of nonrestorative sleep.

**Keywords:** Nonrestorative, unrefreshing, reliability, validity, sleep scale

**Citation:** Wilkinson K; Shapiro C. Development and validation of the Nonrestorative Sleep Scale (NRSS). *J Clin Sleep Med* 2013;9(9):929-937.

Nonrestorative sleep (NRS) is defined as the subjective experience that sleep has not been sufficiently refreshing or restorative.<sup>1,2</sup> NRS is conventionally recognized as a peripheral symptom of insomnia or as a feature of medical conditions like fibromyalgia and chronic fatigue syndrome.<sup>3,4</sup> However, it has increasingly gained attention as a diagnostic entity in its own right, with recent studies investigating both its prevalence and presentation in the absence of other disorders, as well as its interactions with comorbid conditions.<sup>5-8</sup>

Preliminary findings from these investigations suggest that reports of frequent NRS are related to deficits in cognitive and physical functioning, as well as affective symptoms, sleepiness, and fatigue.<sup>7,9</sup> Other studies have indicated a high degree of psychiatric comorbidity, with NRS occurring more frequently in individuals with mood disorder, anxiety, and substance abuse disorder.<sup>10,11</sup> Unfortunately, attempts to identify objective physiological correlates have not been successful. One recent polysomnographic study found overnight sleep results for individuals with NRS to be similar to those collected for healthy controls.<sup>5</sup> Given its continued diagnostic and definitional elusiveness, some have suggested a new paradigm for the evaluation of NRS—one that characterizes it both as a symptom of various medical and psychiatric conditions and as a disorder in its own right.<sup>2</sup> This would be akin to current approaches to insomnia, which treat the complaint both as a symptom of a number of illnesses and as a diagnostic category in its own right.

Despite growing evidence of its clinical relevance, the assessment of NRS has yet to be standardized. Though Stone and colleagues have suggested several criteria for the appli-

## BRIEF SUMMARY

**Current Knowledge/Study Rationale:** With the imminent publication of the DSM-V poised to grant more importance than ever to subjective sleep quality, appropriate tools for evaluating and tracking the associated symptoms of nonrestorative sleep have become vital. This study was initiated to address the current paucity of public-domain self-assessment instruments validated for the measurement of nonrestorative sleep.

**Study Impact:** As a reliable and valid tool for the assessment of nonrestorative sleep, the NRSS will enable clinicians to identify and track experiences of nonrestorative in their patients. Further, given a burgeoning debate regarding the nature of nonrestorative sleep as both a symptom and a unique diagnostic entity, the NRSS will provide researchers with a standardized evaluative tool for investigating the relevance of nonrestorative sleep and its position within the clinical picture.

cation of the label “NRS,” these criteria have yet to be universally adopted and methods chosen for the evaluation of NRS vary.<sup>1</sup> In a recent review of 26 questionnaires selected for their relevancy to NRS as a symptom of insomnia, Vernon and colleagues found that almost all of the scales evaluated contained only one or two items relating to unrefreshing sleep, and none were considered adequate for assessing symptom severity or treatment response, leading researchers to conclude that there is currently no satisfactory public domain questionnaire available for the assessment of NRS.<sup>12</sup> Further, reviewers focused exclusively on questionnaires relating to NRS as a symptom of insomnia. If, as some have suggested, NRS should be investigated for its potential as a unique diagnostic entity, a valid questionnaire specifically designed for this purpose is necessary.

This study was undertaken to meet the demand for such a standardized instrument. This article describes the development and validation of the Nonrestorative Sleep Scale (NRSS) and offers some preliminary findings regarding its relationship to several PSG variables.

## METHODS

### Questionnaire Development

The original pool of 34 items (see **Appendix 1**) was developed by a team of sleep and psychiatric experts. Items were derived from clinical experience, a comprehensive review of the literature regarding NRS, and an evaluation of instruments previously identified as relevant to the construct. The resulting 34 questions were then evaluated by two focus groups as part of a qualitative analysis of their face validity. Focus groups were recruited from two sources: from a support group for patients with narcolepsy and from an obesity therapy group. A total of 17 participants were recruited to respond to two issues: (1) whether the contents of the questionnaire were considered representative of their experiences with sleep, and (2) whether the scale adequately addressed the issues they identified as most relevant to the experience of NRS. As a result of the item-screening and focus group feedback phase, 22 items were marked for removal due to word choice, lack of relevance, and other issues (in the original item pool given in **Appendix 1**, those marked with an asterisk were used in the final scale). One of these 22 items was retained as a potential screening item (discussed below).

### Subjects

Over the course of a 12-month study period, the NRSS was administered to a consecutive sample of 226 patients (118 males, 108 females; mean age 46.7 years, range 18-85) who attended two Ontario sleep clinics for overnight assessment of their sleep problems. The study was approved by an institutional review board, and informed consent was obtained from all participants. Patients received the questionnaire at the time of their first consultation with the sleep specialist. Participants were also asked to complete an additional battery of questionnaires that included the Pittsburgh Sleep Quality Index (PSQI), Toronto Hospital Alertness Test (THAT), Centre for Epidemiological Studies Depression Scale (CES-D), and Athens Insomnia Scale (AIS).

Of those in the patient group, 43 consecutive participants were approached again on the night of their sleep study to complete the NRSS for a second time. This second administration was used to assess the test-retest reliability of the scale. An additional 30 questionnaires were also administered to normal control participants (14 male, 16 female; mean age 36.9 years) who were recruited with a posted flyer. Control participants were screened on the basis of their response to two questions: (1) Have you ever been diagnosed with a sleep disorder? (2) Have you ever thought you might have a sleep disorder? Those who responded “no” to both questions were included in the study.

At the completion of their sleep study analysis, patient charts were consulted and overnight polysomnographic (PSG) data were collected for 215 individuals. Eleven patients were not included in this sample due to the unavailability of charts or

because overnight studies were markedly unusual (e.g. patients who did not sleep during the overnight assessment).

### Factor Analysis

All statistical analysis was conducted using SPSS Statistics Version 19. During the item-screening phase, a correlation matrix was created to identify items that correlated too highly with one another ( $r > 0.9$ ) or items that correlated with very few other variables. Several questions were also targeted for elimination based on factors that included redundancy, lack of clarity, and frequent skipping during response. In combination with the focus group feedback stage, this step resulted in the removal of 22 items (in **Appendix 1**, those that are not marked with an asterisk). One item was retained in the final scale as a potential screening question, but was not included in the analysis. Following the data-screening phase, principal component analysis was conducted on the remaining 12 items (see **Appendix 2**). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was calculated and the Bartlett test of sphericity was completed to ensure that a factor model was appropriate. A KMO value  $> 0.5$  and a significance of  $p < 0.005$  on the Bartlett test were considered adequate. Using a scree test, eigenvalues for each factor were plotted and a visual break between factors was identified, as per Kaiser’s recommendation cited in Field and an approach previously described by Zwick and Velicer.<sup>13,14</sup> It was also specified that a minimum of 70% of the cumulative variance should be explained by the retained factors.

Though the precise connection between factors remains unclear (e.g., the possible bi-directional relationship between affective symptoms and NRS), a promax (oblique) rotation was selected for principal component analysis based on the assumption that different scale factors were likely to be related. Both structure and pattern matrices were consulted when interpreting rotated results, with pattern loadings  $> 0.40$  considered during analysis of the conceptual meaning of different factors.

### Analysis of Reliability and Validity

This phase of analysis was conducted on the final 12-item version of the NRSS developed during data screening. Internal reliability of the overall scale was evaluated using Cronbach  $\alpha$  statistic, and corrected inter-item correlations were calculated for each factor to ensure that items within these groups related well. To assess test-retest reliability, Spearman correlations were used to compare factor and global scores on the first and second administrations of the scale. Sample size for test-retest reliability was chosen based on a method described by Walter, Eliasziw, and Donner.<sup>15</sup> Minimal acceptable reliability was set at 0.6, with  $n = 2$  repeats of testing, and power values were set such that  $p$  values that differ by a value of 0.2 would be detected. Sensitivity and specificity of the scale were assessed using a receiver operating characteristic (ROC) curve as in the approach described by Hanley and McNeil.<sup>16</sup> As the questionnaire was designed to distinguish between individuals with and without unrefreshing sleep—a complaint that is not necessarily present in all the sleep referral participants—the presence or absence of NRS was defined instead by response to a screening question. Individuals who reported experiencing unrefreshing or nonrestorative sleep three times per week or more were placed in the NRS group. This cutoff was chosen based on recommendations made in a previous review of the literature on NRS and on

**Table 1**—Rotated structure matrix for the NRSS

Item	F1 - Refreshment from Sleep	F2 - Physical/Medical Symptoms	F3 - Daytime Functioning	F4 - Affective Symptoms
	Factor Loadings			
How would you rate the quality of your sleep? 2	0.86	0.04	0.07	0.05
Usually, do you think your sleep is restoring or refreshing? 5	0.83	0.04	0.07	-0.23
Have you felt rested if you've slept for your usual amount of time? 19	0.78	0.04	-0.04	0.14
Have you had physical sensations or unusual feelings in your body that you couldn't identify? 27	-0.38	0.87	0.35	-0.03
In the past month, how often have you had one or more of the following: headaches, body pain, numbness or tingling in parts of your body, nausea, racing heart/palpitations, sore throat, frequent cough? 28	0.35	0.74	-0.27	-0.09
Do you feel that physical or medical problems are dragging you down? 29	0.25	0.64	0.35	0.01
Do you ever have a sense of panic, or physical symptoms of panic such as heart racing, for no apparent reason? 33	0.05	0.46	-0.06	0.43
How is your memory and concentration during the daytime? 15	0.01	0.11	0.87	0.08
What is your usual level of daytime energy? 17	0.39	-0.06	0.72	-0.12
Do you usually feel alert during the daytime? 4	0.50	-0.12	0.59	0.04
Do you feel depressed or down if you didn't sleep well the night before? 6	-0.19	-0.03	-0.01	0.94
How often have you felt irritable or gotten the "blahs" if you didn't sleep well the night before? 10	0.32	-0.07	0.11	0.64
Eigenvalue	5.47	1.41	1.16	0.75
Variance	45.58	11.74	9.67	6.27

criteria employed repeatedly by Ohayon and Roth in a number of recent studies regarding the complaint of unrefreshing sleep.<sup>7,10,17</sup>

The construct validity of the scale was evaluated by calculating correlations between several previously validated instruments and scores on the different factors of the NRSS to ensure that it assessed a construct that is unique from those evaluated by the various sleep- and mood-related questionnaires administered during the initial data collection stage. However, the selection of specific subscales of instruments for validity analysis was performed post hoc following the identification of the scale's factors. All validity analysis was performed on the basis of guidelines published previously by McIntire and Miller.<sup>18</sup>

Finally, overnight sleep study data was collected and a preliminary comparison between the NRSS and objective sleep data was conducted. PSG variables included sleep onset latency, total sleep time, sleep efficiency, percentages of stages 1 through 4 sleep, percent wake, and arousal index. As part of the sleep report, level of alpha-EEG sleep was scored by 1 of 3 scoring technicians on a scale ranging from 1 through 5 (with 5 representing the highest levels of alpha intrusion). This scaled score was also collected for analysis.

## RESULTS

### Factor Analysis

Factor analysis was conducted using the entire sample, including both the control and the patient group. A KMO value

of 0.88 was observed and the Bartlett test of sphericity was found to be significant ( $p < 0.001$ ), suggesting that a factor model should be applied. To determine the number of factors best supported by the scale, an approach evaluating eigenvalues  $> 1$  was initially consulted, resulting in the extraction of three factors explaining 67.0% of the total variance. This did not meet the previously specified minimum of 70%; therefore, a visual inspection of the scree plot was also conducted. This approach led to the selection of a four-factor solution, which provided the best comprehensibility, allowed for at least two items within each factor, and explained 73.2% of the total variance. The rotated pattern matrix is reported in **Table 1**.

All items were shown to load above 0.40 on their extracted factor. Factor one consisted of three items relating to quality of sleep and feelings of being restored or refreshed after sleep, so this factor was labeled "refreshment from sleep." The second factor included four items querying physical symptoms of poor sleep such as body pain and frequent illness and it also incorporated questions relating to medical problems and anxiety. This factor was labeled "physical/medical symptoms of NRS." The third factor queried three issues related to daytime function: cognitive abilities, energy, and alertness. This factor was labeled "daytime functioning." The final factor included two items relating to feelings of depression or irritability that follow from unrefreshing sleep and so it was labeled "affective symptoms of NRS." Correlation coefficients between the different factors ranged from 0.38 to 0.83 (**Table 2**).

**Table 2**—Correlations between factor scores on the NRSS

	Global score	Refreshment from sleep	Physical/medical symptoms of NRS	Daytime functioning
Refreshment from sleep	0.83 ( $p < 0.001$ )			
Physical/medical symptoms of NRS	0.83 ( $p < 0.001$ )	0.51 ( $p < 0.001$ )		
Daytime functioning	0.83 ( $p < 0.001$ )	0.71 ( $p < 0.001$ )	0.52 ( $p < 0.001$ )	
Affective symptoms of NRS	0.61 ( $p < 0.001$ )	0.38 ( $p < 0.001$ )	0.40 ( $p < 0.001$ )	0.41 ( $p < 0.001$ )

**Table 3**—Correlations between NRSS scores and global scores on other subjective scales

	Pittsburgh Sleep Quality Index	Athens Insomnia Scale	Fatigue Severity Scale	Centre for Epidemiological Studies-Depression Scale	Toronto Hospital Alertness Test
Global score	0.57 ( $p < 0.001$ )	0.68 ( $p < 0.001$ )	0.70 ( $p < 0.001$ )	0.74 ( $p < 0.001$ )	0.62 ( $p < 0.001$ )
Factor 1: refreshment from sleep	0.52 ( $p < 0.001$ )	0.61 ( $p < 0.001$ )	0.53 ( $p < 0.001$ )	0.51 ( $p < 0.001$ )	0.55 ( $p < 0.001$ )
Factor 2: physical/medical symptoms of NRS	0.46 ( $p < 0.001$ )	0.50 ( $p < 0.001$ )	0.54 ( $p < 0.001$ )	0.54 ( $p < 0.001$ )	0.33 ( $p < 0.001$ )
Factor 3: daytime functioning	0.33 ( $p < 0.001$ )	0.46 ( $p < 0.001$ )	0.60 ( $p < 0.001$ )	0.59 ( $p < 0.001$ )	0.68 ( $p < 0.001$ )
Factor 4: affective symptoms of NRS	0.42 ( $p < 0.001$ )	0.48 ( $p < 0.001$ )	0.43 ( $p < 0.001$ )	0.66 ( $p < 0.001$ )	0.46 ( $p < 0.001$ )

In cases where scales are scored in an opposite direction from the NRSS (i.e., higher scores indicate poorer sleep quality), directions of correlations have been reported in reverse for clarity.

## Reliability and Validity

### Internal Reliability

The NRSS demonstrated a high internal consistency, with a Cronbach  $\alpha$  of 0.88. In terms of the reliability of specific items within their particular subscales, corrected item-total correlations ranged from 0.65 to 0.81 for the refreshment from sleep factor with a Cronbach  $\alpha$  of 0.85; from 0.44 to 0.59 for the physical/medical symptoms factor with an  $\alpha$  of 0.74; and from 0.66 to 0.80 for the daytime functioning factor with an  $\alpha$  of 0.85. The two affective symptoms items demonstrated a coefficient of 0.48 and an  $\alpha$  of 0.64.

### Test-Retest Reliability

The average interval between first and second administration of the NRSS was 6.8 days (range 2 to 31 days; average 5.1 days). A Spearman correlation coefficient of 0.72 was observed for global scores on the scale. Correlation coefficients for the subscales were 0.76 for the refreshment from sleep factor, 0.77 for the medical/physical symptoms factor, 0.69 for the daytime dysfunction factor, and 0.27 for the affective symptoms factor. Paired samples *t*-tests performed for each item confirmed that there were no significant differences in means between the first and second administration.

### Sensitivity and Specificity

ROC analysis revealed an area under the curve of 0.90, suggesting that the NRSS has excellent diagnostic accuracy in terms of its ability to distinguish between individuals who report NRS frequently ( $\geq 3$  times per week) from those who do not. A cutoff point of 46 was selected, as it appeared to maximize both sensitivity and specificity, giving values of 0.91 and 0.75, respectively.

### Construct Validity

To assess the value of individual factors of the NRSS, scores on those subscales were compared to existing subjective mea-

asures designed to evaluate similar constructs. See **Table 3** for a summary of these relationships. Factor 3, the daytime functioning subscale, showed its highest correlation with the THAT ( $r = 0.68$ ,  $p < 0.001$ ), while the affective symptoms factor demonstrated its strongest relationship with the CES-D ( $r = 0.66$ ,  $p < 0.001$ ). The physical/medical symptoms factor showed moderate correlations with three of the scales—the AIS, FSS, and CES-D. The refreshment from sleep factor demonstrated similarly moderate correlations with all scales, with a slightly stronger relationship seen for the AIS ( $r = 0.61$ ,  $p < 0.001$ ).

### PSG Variables and the NRSS

Due to the relatively small sample size and lack of a priori selection of participants based on existing diagnoses or confirmed medical status, detailed analysis relating to comorbid conditions, medication use, and other clinical features was not possible. However, as participant demographic information is likely to be relevant to NRSS validation, they are summarized in **Table 4**. The largest diagnostic category was the sleep disordered breathing (SDB) group, with 112 participants, while the three next largest groups included those with primary insomnia, periodic leg movement syndrome (PLMS), and sleep disturbance due to psychiatric illness. Though a number of individuals were found to have comorbid psychiatric conditions, this label encompasses those patients whose overnight sleep studies were characterized primarily by sleep markers of depression, posttraumatic stress disorder (PTSD), and other psychiatric illness. The remaining groups represented diagnostic categories such as circadian rhythm disturbances, parasomnia, narcolepsy, fibromyalgia, and nonspecific disturbances to sleep architecture that do not fall into a diagnostic category. These were factors identified in the sleep report as “fragmented sleep” or “poor sleep quality.”

In terms of PSG analysis, a few very small but significant correlations emerged when controlling for participant age and

**Table 4**—Demographic characteristics of patient and control groups

	N (%)	Male (%)	Female (%)	Age	
				Mean ± SD	Range
Patient Group	226 (88.3)	118 (52.2)	108 (47.8)	46.7 ± 14.9	18-85
Sleep disordered breathing	112 (52.1)	71 (63.4)	41 (36.6)	51.0 ± 12.6	25-85
Insomnia	38 (17.7)	15 (39.5)	23 (60.5)	46.0 ± 16.5	19-84
Sleep disturbance due to psychiatric illness	25 (11.6)	7 (28.0)	18 (72.0)	39.3 ± 12.4	19-58
Periodic leg movement syndrome	17 (7.9)	7 (41.2)	10 (58.8)	46.8 ± 17.6	18-82
Circadian rhythm disorder	7 (3.3)	7 (100.0)	0	29.7 ± 9.3	20-43
Nonspecific sleep disturbance	7 (3.3)	3 (42.9)	4 (57.1)	39.7 ± 17.0	18-66
Parasomnia	2 (1.0)	2 (100.0)	0	24.0 ± 2.8	22-26
Narcolepsy	4 (1.9)	1 (25.0)	3 (75.0)	27.5 ± 9.8	21-42
Fibromyalgia	3 (1.4)	0	3 (100.0)	45.3 ± 20.4	20-60
No Sleep Study	11 (4.9)	5 (45.5)	6 (54.5)	47.7 ± 13.3	23-66
Control Group	30 (11.7)	14 (46.7)	16 (53.3)	36.9 ± 12.5	21-72

total sleep time. Scores on the refreshment factor of the scale possessed a significant negative association with levels of alpha EEG sleep ( $r = -0.16$ ,  $p = 0.027$ ). Scores on the daytime functioning factor ( $r = -0.15$ ,  $p = 0.042$ ) and the affective symptoms factor were also significantly, negatively associated with alpha EEG level ( $r = -0.16$ ,  $p = 0.033$ ). Finally, global scores demonstrated a significant negative association with alpha EEG level ( $r = -0.16$ ,  $p = 0.026$ ).

Associations also emerged between several other variables and the affective symptoms domain. This factor was significantly, negatively correlated with sleep efficiency ( $r = -0.15$ ,  $p = 0.042$ ), REM latency ( $r = -0.15$ ,  $p = 0.047$ ), and percentage of stage 2 sleep ( $r = -0.19$ ,  $p = 0.010$ ), and positively correlated with wake after sleep onset ( $r = 0.17$ ,  $p = 0.025$ ). In contrast, global scores on the CES-D were positively correlated only with level of alpha EEG ( $r = 0.18$ ,  $p = 0.026$ ), and global scores on the PSQI shared no significant correlations with any of the overnight PSG variables measured. THAT global scores, however, were significantly, negatively correlated with level of alpha EEG sleep ( $r = -0.17$ ,  $p = 0.042$ ) and positively correlated with percentage of stage 4 sleep ( $r = 0.16$ ,  $p = 0.045$ ). No other associations were found between subjective sleep measures and objective PSG data.

## DISCUSSION

The current study provides strong evidence for the validity and reliability of the NRSS. The scale consists of 12 items evaluating four factors: refreshment from sleep, physical/medical symptoms, daytime functioning, and affective symptoms. Of the scale's twelve items, ten employ Likert scales with values ranging from one to ten, while an additional two items offer five options. Some items are worded positively (with ten indicating very good sleep quality), while others are worded negatively (where ten refers to very poor sleep quality). The scale's scoring system was designed such that all items are given a weighted score from one to five. Though this complicates scoring, it ensures that all questions are weighted equally. Negatively worded items are reversed before scoring, meaning that higher scores on the scale indicate less NRS. Global scores can range from 12 to 60.

Similar to the diagnostic guidelines employed for insomnia, recent approaches to NRS have suggested that, in order for the construct to be applied, a patient should demonstrate the symptom at least three nights per week.<sup>1,10,17</sup> However, we have previously questioned the arbitrary distinction between these different frequency cohorts.<sup>2</sup> For this reason, we have excluded this particular item from the scale proper until its suitability for the Likert-type format is better understood. Currently, frequency may be assessed using an additional screening item that can be included or altered at the researcher's discretion (see **Appendix 2**).

A rotated structure matrix revealed that each item loaded at a value of 0.46 or higher on its specific factor and below 0.40 on all others. Though previous recommendations by Zwick and Velicer have suggested that individual factors should not contain fewer than three items, the affective symptoms factor consists of only two.<sup>14</sup> It was decided that the factor would remain, given that preliminary studies suggest that NRS is significantly related to affective and mood complaints.<sup>7,9</sup> In future studies, it may be necessary to evaluate new items for addition to this factor.

The overall NRSS possesses a strong internal consistency ( $\alpha = 0.88$ ) and good test-retest reliability ( $r = 0.72$ ,  $p = 0.01$ ). Reliabilities for each of the subscales were also satisfactory, with the lowest value found between the two items of the affective symptoms factor ( $r = 0.48$ ), suggesting that items related well to one another and supporting their inclusion in particular factors. Unfortunately, due to the location of collection in the clinical setting, the time range between administrations (from 2 to 31 days) was somewhat large. While the majority of respondents ( $n = 36$ ) received their second questionnaire within the span of one week, there were 7 individuals who fell outside of that range, skewing the sample. As demonstrated previously by Backhaus and colleagues, an extended period between administrations is likely to result in different findings for test-retest reliability.<sup>19</sup> Future studies may need to replicate this testing in order to confirm these results.

The sensitivity and specificity of the scale remain more difficult to quantify. Given that definitions of NRS itself are still in flux, there is currently no gold standard for assessing its presence or absence. However, the majority of recent studies have employed criteria stating that unrefreshing sleep should

be reported at least three times per week for a minimum of one month—general criteria that have been applied in almost all studies evaluating the presence, prevalence, and associated comorbidities of NRS.<sup>1,5,9,11</sup> The current study was guided by this relative consensus, and those who reported a symptom frequency of three or more times per week were placed in the NRS group. According to this approach, a cutoff score of 46 or less was found to maximize sensitivity (0.91) while still providing satisfactory specificity (0.75).

The questionnaires used to assess construct validity were selected on a post hoc basis. Therefore, only two measures administered as part of the questionnaire battery appeared immediately relevant to specific subscales of the NRSS. These were the THAT and the CES-D. Scores on the THAT were found to be moderately correlated with the daytime functioning subscale ( $r = 0.68$ ,  $p < 0.001$ ), while only mildly correlated with other subscales. Similarly, the CES-D was moderately correlated with the affective symptoms factor ( $r = 0.66$ ,  $p < 0.001$ ), with mild correlations to other factor scores.

No other scale appeared to correlate well with the refreshment from sleep factor—unsurprising since it represents a unique construct. In their review of 26 instruments found to contain content relating to NRS as a symptom of insomnia, Vernon and colleagues concluded that there exists no reliable or valid questionnaire in the public domain to comprehensively assess the construct of NRS.<sup>13</sup> They found that the majority of instruments identified for this purpose contained only one or two items relating to NRS, and none offered satisfactory methods for evaluating symptom severity or treatment response. Furthermore, these questionnaires were developed from a perspective that views NRS as a peripheral symptom of other disorders. The NRSS represents the first such instrument developed with the potential to evaluate the concept of NRS as an entity in its own right. Despite these considerations, in future studies it may be of interest to compare scores on the NRSS to several of those measures previously identified as being relevant to NRS.

In terms of construct validity of the physical/medical symptoms subscale, only mild correlations were found. Additional validation studies should be conducted with the inclusion of a scale to assess physiological complaints in order to confirm that the physical/medical symptoms factor evaluates an appropriate construct. As the constructs of NRS, mood, alertness, insomnia, and physical symptoms of fatigue or sleepiness remain conceptually entangled, the task of separating each into distinct components remains a challenge. Indeed, while the clinical picture of NRS is still relatively undefined, it appears likely that it reflects some element of each of these factors—a suggestion supported by the mild to moderate correlations found here between the NRSS and other previously validated measures.

Finally, a comparison between PSG variables and scores on the NRSS revealed several very small correlations. Of all the objective variables evaluated, level of alpha EEG was negatively associated with the greatest number of subscales, suggesting that lower scores on the scale are related to greater levels of alpha intrusion. This finding may offer some preliminary support of previous studies that have observed connections between alpha EEG NREM sleep and the experience of unrefreshing or nonrestorative sleep.<sup>20,21</sup> However, this relationship has proven to be complex and is still not well understood, with several

studies contradicting the existence of a connection.<sup>22,23</sup> Further, in the current study, none of the correlations found between alpha EEG and the NRSS were larger than 0.19 in magnitude, meaning objective sleep quality values explained less than 4% of the variability in NRSS scores. Additionally, in many cases, significance levels were only slightly below 0.05. As a Bonferroni correction was not conducted in consideration of multiple comparisons, it is possible that these correlations represent false positives. It may also be relevant for future evaluations of the NRSS to examine correlations with stage N3 sleep, the most recent method for assessing slow wave sleep recommended by the American Academy of Sleep Medicine.<sup>24</sup> In this particular study, we felt that the distinction between the two levels of slow wave sleep and the potential interaction with the NRSS could provide a more detailed picture of the influence of varying degrees of deep sleep—a suggestion supported by our observation of significant correlations between scale factors and stage 4 but not stage 3 sleep. However, evaluations of these two stages combined in the future may prove to be clinically relevant and should not be ruled out.

While the presence of very small correlations between the NRSS and various physiological measures is unfortunate from the perspective of those seeking potential biomarkers for poor subjective sleep, this observation is supported by previous research. In one of the only PSG evaluations conducted to date on patients with NRS, researchers found that the majority of objective variables were similar between NRS participants and normal controls.<sup>5</sup> The NRS group spent somewhat less time in stages 3 and 4 sleep, and less time in REM sleep, but distinctions were minimal. This further underlines the necessity of a tool for the assessment of NRS. If the construct is one that can only be observed through subjective report, then there is demand for an instrument that satisfactorily assesses its presence and severity.

One of the greatest limitations of the current study involves its subject pool of consecutive sleep clinical patients. While this participant group allowed for validation of the NRSS in a highly diverse group of individuals representing the range of complaints typically observed in the sleep clinic, this diversity prevented detailed analysis of the relationship between issues such as physical or psychological comorbidities, job status, and medication use, as attempts to partition on these bases resulted in groups that were too small to be analyzed. Future studies will need to evaluate the use of the NRSS in specific populations chosen on an a priori basis.

In conclusion, the NRSS was created using both qualitative and quantitative approaches to the construct of nonrestorative sleep. Through focus group analysis and item screening, an initial item pool was reduced to twelve questions, and factor analysis revealed four factors relating to NRS and its components. Future validation studies are necessary to address the construct validity of all subscales. For example, it may become clear through subsequent studies of NRS that it is more suitably conceived of as a symptom of other disorders—in which case, the physiological and affective symptom subscales may become subsumed under the umbrella of these other diagnostic categories. However, while NRS remains relatively opaque as a construct, the NRSS may serve as a valuable, standardized tool for use in a variety of research and clinical capacities.

## REFERENCES

1. Stone K, Taylor D, McCrae C, Kalsekar A, Lichstein K. Nonrestorative sleep. *Sleep Med* 2008;12:275-88.
2. Wilkinson K, Shapiro C. Nonrestorative sleep: symptom or unique diagnostic entity? *Sleep Med* 2012;13:561-9.
3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders. 4th ed.* Washington: American Psychiatric Association, 2000.
4. Cesta A, Moldofsky H, Sammut C. The sensitivity and specificity of the Sleep Assessment Questionnaire (SAQ) as a measure of nonrestorative sleep. *Sleep* 1999;22:S14.
5. Roth T, Zammit G, Lankford A, et al. Nonrestorative sleep as a distinct component of insomnia. *Sleep* 2010;33:449-58.
6. Ohayon M, Roth T. What are the contributing factors for insomnia in the general population? *J Psychosom Res* 2001;51:745-55.
7. Ohayon M, Riemann D, Morin C, Reynolds C. Hierarchy of insomnia criteria based on daytime consequences. *Sleep Med* 2012;13:52-7.
8. Léger D, Partinen M, Hirshkowitz M, Chokroverty S, Hedner J. Characteristics of insomnia in a primary care setting: EQUINOX survey of 5293 insomniacs from 10 countries. *Sleep Med* 2010;11:987-98.
9. Sarsour K, Van Brunt D, Johnston J, Foley K, Morin C, Walsh J. Associations with nonrestorative sleep with insomnia, depression, and daytime function. *Sleep Med* 2010;11:965-72.
10. Roth T, Jaeger S, Jin R, Kalsekar S, Stang P, Kessler R. Sleep problems, comorbid mental disorders, and role functioning in the National Comorbidity Replication Survey. *Biol Psychiatry* 2006;60:1364-71.
11. Ohayon M. Prevalence and correlates of nonrestorative sleep complaints. *Arch Intern Med* 2005;165:35-41.
12. Vernon M, Dugar A, Revicki D, Treglia M, Buysse D. Measurement of nonrestorative sleep in insomnia: a review of the literature. *Sleep Med Rev* 2010;14:205-12.
13. Field AP. *Discovering statistics using SPSS (2nd edition)*. London: Sage, 2005.
14. Zwick WR, Velicer WF. Comparison of five rules for determining the number of components retain. *Psychol Bull* 1986;99:432-42.
15. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998;17:101-10.
16. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.
17. Ohayon M, Bader G. Prevalence and correlates of insomnia in the Swedish population aged 19-75 years. *Sleep Med* 2010;11:980-6.
18. McIntire SA, Miller LA. *Foundations of psychological testing*. 2nd ed. Sage Publishing Co., 2005.
19. Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *J Psychosomatic Res* 2001;53:737-40.
20. Roizenblatt S, Moldofsky H, Benedito-Sylva A, Tufik S. Alpha sleep characteristics in fibromyalgia. *Arthritis Rheum* 2001;44:222-30.
21. Cesta A, Moldofsky H, Sammut C. The sensitivity and specificity of the Sleep Assessment Questionnaire (SAQ) as a measure of nonrestorative sleep. *Sleep* 1999;22:S14.
22. Van Hoof E, De Becker P, Lapp C, Cluydts R, De Meirleir K. Defining the occurrence and influence of alpha-delta sleep in chronic fatigue syndrome. *Am J Med Sci* 2007;333:78-84.
23. Terzano M, Parrino L, Smerieri A, et al. Atlas, rules and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep. *Sleep Med* 2002;3:187-99.
24. Iber C, Ancoli-Israel S, Chesson A, Quan SF. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. Westchester, IL: American Academy of Sleep Medicine; 2007.

## ACKNOWLEDGMENTS

This research was supported by a graduate student award from CIHR and the Youthdale Foundation Sleep Programme.

## SUBMISSION &amp; CORRESPONDENCE INFORMATION

Submitted for publication October, 2012

Submitted in final revised form April, 2013

Accepted for publication April, 2013

Address correspondence to: Kate Wilkinson, 474 Markham Street, Toronto, ON, M6G 2L3, Canada; Tel: (416) 579-1520; Fax: (416) 703-0507; E-mail: katewilk@gmail.com

## DISCLOSURE STATEMENT

This was not an industry supported study. The authors have indicated no financial conflicts of interest.

## Appendix 1—NRSS Original Item Pool

Those marked with an asterisk were retained.

1. How would you rate the quantity of your sleep?
2. How would you rate the quality of your sleep?\*
3. Usually, what percentage of time in bed do you actually sleep?
4. Do you usually feel alert during the daytime?\*
5. Usually, do you think your sleep is restoring or refreshing?\*
6. Do you feel depressed or down if you didn't sleep well the night before?\*
7. How tired did you feel one hour before going to sleep yesterday?
8. How tired did you feel one hour after you woke up today?
9. How tired did you feel two hours after you woke up today?
10. How often have you felt irritable or gotten the "blahs" if you didn't sleep well the night before?\*
11. How often have you felt really refreshed upon awakening in the morning?\*
12. What is your usual level of daytime energy?\*
13. Does your energy level improve (or would you energy level improve) if you've had a really good night's sleep?
14. Do you feel more refreshed or rested (or would you feel more refreshed or rested) if you've had a really good night's sleep?
15. How is your memory and concentration during the daytime?\*
16. Do you think your memory and concentration improve (or would improve) after you've had a very good night's sleep?
17. What is your energy level on an average day?
18. Do you think your energy level improves (or would improve) after you've had a very good night's sleep?
19. Have you felt rested if you've slept for your usual amount of time?\*
20. Do you think you would feel more rested if you got an extra hour of sleep on most days?
21. Do you feel that having unrefreshing sleep has led you to perform poorly at work or school?
22. Have you ever had an accident (in the car, at work, or at home) because your sleep is not refreshing?
23. How often do you nap during the daytime?
24. Do you feel that napping allows (or would allow) you to catch up on missed sleep?
25. Do you feel that napping allows (or would allow) you to feel more refreshed or restored?
26. Do you think your ability to do well on tasks varies across the day?
27. In the past month, how often do you have one or more of the following: headaches, body pain, numbness or tingling in parts of your body, nausea, racing heart/palpitations, sore throat, frequent cough?\*
28. Do you feel that physical or medical problems are dragging you down?\*
29. Have you had physical sensations or unusual feelings in your body that you couldn't identify?\*
30. Do you ever have problems knowing or identifying exactly which emotions you are feeling?
31. Do you think you are an optimistic person (i.e. you tend to take a hopeful and positive view of what will happen in the future)?
32. Do you ever expect that the worst is going to happen to you no matter what you do?
33. Do you ever have a sense of panic, or physical symptoms of panic such as heart racing, for no apparent reason?\*
34. Do you think that you are easily influenced by others?



## Appendix 2—The Nonrestorative Sleep Scale (NRSS)

Please circle the response that best represents your usual experiences over the past month.

How often have you felt really refreshed upon awakening in the morning?

Never       1 day/week       2-3 days/week       4-5 days/week       6-7 days/week

1. How would you rate the quality of your sleep?  
1      2      3      4      5      6      7      8      9      10  
*Very poor*      *Very good*
2. Usually, do you think your sleep is restoring or refreshing?  
1      2      3      4      5      6      7      8      9      10  
*Never*      *Always*
3. Have you felt rested if you've slept for your usual amount of time?  
1      2      3      4      5      6      7      8      9      10  
*Not at all*      *Absolutely*
- \*4. Have you had physical sensations or unusual feelings in your body that you couldn't identify?  
1      2      3      4      5      6      7      8      9      10  
*Never*      *Yes, all the time*
- \*5. In the past month, how often have you had one or more of the following: headaches, body pain, numbness or tingling in parts of your body, nausea, racing heart/palpitations, sore throat, frequent cough?  
 Never       1 day/week       2-3 days/week       4-5 days/week       6-7 days/week
- \*6. Do you feel that physical or medical problems are dragging you down?  
1      2      3      4      5      6      7      8      9      10  
*Never*      *Yes, all the time*
- \*7. Do you ever have a sense of panic, or physical symptoms of panic such as heart racing, for no apparent reason?  
1      2      3      4      5      6      7      8      9      10  
*Never*      *Yes, all the time*
8. How is your memory and concentration during the daytime?  
1      2      3      4      5      6      7      8      9      10  
*Very poor*      *Very good*
9. What is your usual level of daytime energy?  
1      2      3      4      5      6      7      8      9      10  
*Very low*      *Very high*
10. Do you usually feel alert during the daytime?  
1      2      3      4      5      6      7      8      9      10  
*Not at all*      *Very alert*
- \*11. Do you feel depressed or down if you didn't sleep well the night before?  
1      2      3      4      5      6      7      8      9      10  
*Not at all*      *Very depressed*
- \*12. How often have you felt irritable or gotten the "blahs" if you didn't sleep well the night before?  
 Never       1 day/week       2-3 days/week       4-5 days/week       6-7 days/week

**Scoring:** All items are given a weighted score from 1 to 5 (i.e. for scales ranging from 1-10, responses 1 and 2 are scored as 1, responses 3 and 4 are scored as 2, etc.). Negatively worded items are reversed before scoring (those marked with \*), meaning that higher scores on the scale indicate less severe NRS.