THE PSYCHOMETRIC PROPERTIES OF THE NONRESTORATIVE SLEEP SCALE AND A PROSPECTIVE OBSERVATIONAL STUDY OF THE PHYSIOLOGICAL CORRELATES OF NONRESTORATIVE SLEEP

BY

KATE WILKINSON

A THESIS SUBMITTED IN CONFORMITY WITH THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE
GRADUATE DEPARTMENT OF CELL AND SYSTEMS BIOLOGY
UNIVERSITY OF TORONTO

© KATE WILKINSON 2012
NONRESTORATIVE SLEEP SCALE AND A PROSPECTIVE OBSERVATIONAL STUDY OF THE PHYSIOLOGICAL CORRELATES OF NONRESTORATIVE SLEEP

Kate Wilkinson, Master of Science, 2012
Graduate Department of Cell and Systems Biology, University of Toronto
Toronto, Ontario, Canada

ABSTRACT

Nonrestorative sleep refers to the experience of sleep as insufficiently refreshing, often despite the appearance of normal sleep according to objective parameters. As a result, a valid and reliable measure of the subjective experience of NRS is required in order to allow for standardization and comparability in its assessment. This thesis reports the results of a study involving the development and validation of a scale to assess NRS, the Nonrestorative Sleep Scale (NRSS). The psychometric properties of the NRSS were assessed in a group of 256 participants recruited from a sleep clinic population. Principal component analysis revealed four domains. The scale demonstrated good internal and test-retest reliability and reasonable validity compared to other measures. Overnight polysomnographic variables were also compared to scores on the scale and a few were found to be weakly correlated with scale scores. These included alpha EEG, sleep efficiency, and REM latency.
ACKNOWLEDGMENTS

I would like to thank my supervisor Dr. Colin Shapiro as well as my committee members Dr. Keith Gordon and Dr. John Flanagan for their support. I would also like to thank the staffs of both the Toronto Sleep and Alertness Clinic in Toronto and the International Sleep Clinic in Parry Sound for their support and contributions to the project. Finally, this thesis was made possible by a Graduate Award from CIHR.
# TABLE OF CONTENTS

**MANUSCRIPTS APPEARING IN THESIS** ................................................................. vi
**LIST OF FIGURES** .............................................................................................. vii
**LIST OF TABLES** ............................................................................................... viii
**ABBREVIATIONS** ............................................................................................ ix

## 1 LITERATURE REVIEW ..................................................................................... 1

1.1 OVERVIEW ........................................................................................................ 1

1.2 DEFINING NONRESTORATIVE SLEEP .............................................................. 2

1.2.1 *Clinical Diagnosis* ...................................................................................... 2

1.2.2 *Defining NRS in Research* .......................................................................... 5

1.3 PREVALENCE ..................................................................................................... 8

1.4 CORRELATES .................................................................................................... 9

1.4.1 *Gender* ......................................................................................................... 9

1.4.2 *Age* ............................................................................................................... 10

1.4.3 *Insomnia* ..................................................................................................... 11

1.4.4 *Daytime Symptoms* .................................................................................... 13

1.4.5 *Psychiatric Correlates* ................................................................................ 16

1.4.6 *Fibromyalgia Syndrome and Chronic Fatigue Syndrome* ......................... 17

1.5 PHYSIOLOGICAL MARKERS OF NRS ............................................................... 19

1.5.1 *Alpha Wave Intrusion in Non-REM Sleep* .................................................. 20

1.5.2 *Cyclic Alternating Pattern* ........................................................................ 22

1.5.3 *Immunological Markers of NRS* ................................................................. 23

1.5.4 *Slow Wave Sleep* ....................................................................................... 25

1.5.5 *Growth Hormone* ....................................................................................... 26

1.5.6 *Summary of Review* .................................................................................. 26

1.6 STUDY RATIONALE ........................................................................................ 28

1.7 THESIS HYPOTHESIS AND OBJECTIVES ....................................................... 29

## 2 METHODS ........................................................................................................ 30

2.1 SAMPLE SIZE DETERMINATION ..................................................................... 30
MANUSCRIPTS APPEARING IN THESIS

- Literature review adapted from an original manuscript titled “Potential physiological correlates of nonrestorative sleep: a review” submitted to the journal Sleep Medicine.
LIST OF FIGURES

**Figure 1** – Scree plot showing the amount of variance accounted for by each factor. Arrow indicates point chosen for cutoff. ...........................................................................................................46

**Figure 2** – ROC curve plotting sensitivity vs. specificity when using NRS screening question to predict NRSS global scores ...........................................................................................................51
LIST OF TABLES

Table 1 – Demographic characteristics of patient and control groups…………………………..41

Table 2 – Descriptive statistics for individual NRSS items…………………………………………45

Table 3 – Standardized regression coefficients of the rotated factor pattern…………………..47

Table 4 - Corrected item-total correlations for each item within its factor and in global scale…49

Table 5 – Sensitivity and specificity for NRS based on specific cutoff scores…………………..50

Table 6 – Correlations between global scores on the NRSS and global scores on other subjective measures………………………………………………………………………………….53

Table 7 – Mean and standard deviations for PSG variables reported for the NRS and non-NRS groups………………………………………………………………………………….54
ABBREVIATIONS

AI  Arousal Index
AIS  Athens Insomnia Scale
CAP  Cyclic Alternating Pattern
CES-D  Centre for Epidemiological Studies Depression Scale
CFS  Chronic Fatigue Syndrome
EEG  Electroencephalography
FS  Fibromyalgia Syndrome
FSS  Fatigue Severity Scale
KMO  Kaiser-Meyer-Olkin
NRS  Nonrestorative Sleep
NRSS  Nonrestorative Sleep Scale
PASW  Predictive Analytics Software
PLMS  Periodic Leg Movement Syndrome
PSG  Polysomnography
PSQI  Pittsburgh Sleep Quality Index
PTSD  Post-Traumatic Stress Disorder
REM  Rapid Eye Movement
ROC  Receiver Operating Characteristic
SAQ  Sleep Assessment Questionnaire
SD  Sleep Disturbance
SDB  Sleep-Disordered Breathing
SE  Sleep Efficiency
SOL  Sleep Onset Latency
THAT  Toronto Hospital Alertness Test
TST  Total Sleep Time
WASO  Wake After Sleep Onset
WT  Wake Time
1 LITERATURE REVIEW

1.1 OVERVIEW

Conditions like sleep apnea, narcolepsy, and insomnia are increasingly recognized for their detrimental effects on both physical and mental health, and this recognition has led to the development of gold-standard approaches to diagnosis and treatment. However, issues with sleep often defy simple categorization and, increasingly, researchers approaching the diagnosis and treatment of sleep disorders have begun to recognize that objective sleep data obtained through polysomnography (PSG), actigraphy, or otherwise may not be sufficient to explain poor subjective quality of sleep. Nonrestorative sleep (NRS)—or the subjective feeling that sleep has been insufficiently refreshing—is a complex sleep problem that remains contested in the literature and difficult to define using objective physiological data. In order to better understand and treat experiences of NRS, a valid and reliable self-report tool is necessary.

The research presented in this thesis will discuss the current literature available on NRS and will outline the development and validation of a scale to assess NRS. Results on this NRS questionnaire will also be compared with physiological sleep data collected via overnight PSG in a sleep clinic population in order to determine if any relationship between this subjective self-report measure and objective sleep data exists.
1.2 DEFINING NONRESTORATIVE SLEEP

A general definition of NRS is sleep that is subjectively experienced as unrefreshing, light, restless, or of poor quality. However, such a definition lacks specificity, making it applicable to almost any individual in the general population. Thus, it becomes difficult to distinguish occasional poor sleep quality from clinically significant NRS. The definitions of NRS used in clinical and research contexts have tended to diverge due to differences in perspective. Any attempt to understand NRS should take both the clinical and research spheres into account.

1.2.1 Clinical Diagnosis

From a clinical point-of-view, NRS is most typically treated as a symptom of another primary sleep or medical disorder. In particular, it is often regarded as a symptom of insomnia. NRS receives mention in almost every diagnostic classification manual, though the way it is defined and the degree of attention it receives varies.

In the ICSD-2 NRS is referred to as sleep which is “unrefreshing” and is listed as a symptom of obstructive sleep apnea, periodic limb movement disorder, shift work sleep disorder, and fibromyalgia. The ICSD-2 also lists NRS as a common complaint of patients with insomnia due to a medical condition. However, the manual offers no guidelines for determining how frequency of symptoms may factor into diagnosis, though it suggests that symptoms should persist for at least one month in order to receive a diagnosis.

The ICD-10 Classification of Mental and Behavioural Disorders addresses NRS using slightly different wording. The manual suggests that quantity of sleep alone should not be the primary criteria for diagnosis, as even those who experience a sufficient amount of sleep both
objectively and subjectively may still “suffer immensely from the poor quality of their sleep.”

A complaint of “poor quality of sleep” is listed as one of the three primary clinical features of insomnia, along with difficulties falling asleep and maintaining sleep. The manual suggests that the sleep disturbance should occur at least three times per week for a period of one month before the diagnosis of insomnia is given.

Of all three diagnostic manuals, the DSM-IV-TR is the only one to specifically define what is meant by “nonrestorative sleep.” The label is used to describe sleep that is “restless, light, or of poor quality” and is listed one of the predominant complaints associated with primary insomnia. The manual suggests that NRS should be present for at least one month in order to merit a diagnosis of insomnia, though no frequency criteria are given. NRS is also listed as a feature of primary hypersomnia, and “unrefreshing” sleep is offered as one of the less frequent symptoms of breathing-related sleep disorder. Finally, the complaint of NRS associated with fibromyalgia is supplied as an example of a sleep disorder due to a general medical condition.

Currently, NRS is most often addressed as a symptom of insomnia. However, some researchers question the utility of these insomnia criteria in identifying individuals with sleep complaints. A study by Ohayon and Reynolds used operational definitions of insomnia found in both the ICSD and the DSM-IV to determine its prevalence in a general European population. They observed that many individuals who report chronic issues with their sleep remain unidentified according to current diagnostic criteria. As many as 60% of all insomnia sufferers experience a problem with sleep for more than five years, suggesting that these issues go unrecognized. For this reason, researchers have recently begun to recommend the addition of more discriminating criteria to diagnostic manuals to help improve their specificity and
reliability. These would include guidelines on frequency, chronicity, comorbidity, and/or severity, which the majority of manuals currently lack.

With the upcoming release of the DSM-5 in 2013, certain proposed changes to the manual may help to clarify the label of NRS and its use in the diagnosis of a number of sleep disorders. One of the changes suggested by the Sleep-Wake Disorders Work Group is the addition of a frequency criterion for the diagnosis of insomnia, requiring nonrestorative sleep (or other insomnia-related sleep problems) to be present at least three times per week. In conjunction with this, the group has recommended a greater emphasis on the importance of dissatisfaction with sleep in insomnia, moving away from a focus on difficulties initiating sleep and maintaining sleep as the primary criteria.

Another proposed change includes a greater role for NRS in the diagnosis of primary hypersomnia/narcolepsy without cataplexy. Sleep periods that are nonrestorative or unrefreshing (despite being of sufficient or greater-than-average length) could potentially become one of the primary symptoms listed in this category, though hypersomnia and excessive sleepiness would remain the predominant criteria. While these proposed revisions are still under development, many of these suggested changes appear to reflect a greater emphasis on the importance of subjective sleep quality. This would grant greater clinical significance to NRS—a sleep problem that, without many satisfactory objective markers, is currently subjective in nature. Additionally, this greater focus on sleep quality may allow for more sensitive diagnosis and screening, as individuals who report sleep dissatisfaction have higher prevalence rates of mental disorders and present with complaints of poor quality sleep and daytime sleepiness than those with insomnia symptoms alone.
1.2.2   Defining NRS in Research

Similar to the context of clinical diagnosis, standard definitions of NRS in research have shifted over time. Thus, efforts to draw parallels between the results of these investigations into the epidemiology and presentation of NRS have been difficult. In a recent review of the literature on NRS, Stone and colleagues found that, of 10 studies attempting to determine the prevalence rate of NRS, each used a different method for screening.\(^1\) Some of the earliest assessments of NRS were performed by Ohayon and colleagues using the Sleep-EVAL system, a computer program that could be implemented for large-scale phone surveys.\(^8\) The Sleep-EVAL program queries sleep habits, schedule, and hygiene, while also collecting information about psychiatric issues to be used in differential diagnosis. The system has been implemented in a number of large epidemiological studies. However, it is unavailable to clinicians and researchers and, from summaries performed for literature reviews, appears to use different criteria for NRS with every study.\(^1\)

Approaches to measuring NRS taken by other research groups vary, but typically they consist of a short screening tool ranging from one to four items in length. For example, in a study of insomnia in adolescents, Roberts and colleagues asked participants to report the frequency with which they were “Not feeling really rested” over the past month, with responses of “often” or “almost every day” indicating NRS.\(^9\) Other researchers have chosen pre-existing questionnaires for the purpose of assessment, focusing on items that query feelings of refreshment or fatigue following sleep in order to screen for NRS.\(^10\)

Though knowledge to the contents of the Sleep-EVAL system is limited, much of the work conducted by Ohayon and colleagues has helped to improve the specificity of the definition
of NRS, allowing for increased standardization between research groups. As part of a large-scale telephone survey conducted across several European countries, Ohayon and Roth developed a more discriminating method for evaluating the presence of NRS (and insomnia symptoms in general) than the one offered in the DSM-IV.\textsuperscript{11} Using their new definition of insomnia, they found that the number of individuals reporting one insomnia symptom who also had a sleep or mental disorder diagnosis increased by approximately 20%. In this case, NRS was queried using three distinct criteria: the report that sleep feels unrefreshing or nonrestorative at least three nights per week despite sufficient sleep duration, and difficulty getting started in the morning almost every day.

These new guidelines have been used repeatedly for subsequent studies, including those performed by other researchers. They offer a more focused definition of NRS, creating parameters regarding the frequency of the sleep event, offering a description of potential daytime symptoms, and stipulating that the complaint of NRS be present despite sufficient sleep duration. This new definitions suggested that NRS should be a more specific construct relating to sleep quality, rather than an issue relating to quantity.

Following the publication of this study, Roth and colleagues developed a four-item scale that encompasses most of the criteria contained in this proposed definition.\textsuperscript{12} Items queried the frequency with which respondents experienced the following issues over the previous year: having trouble getting up in the morning, waking up not feeling rested, feeling as though they had not slept long enough despite having enough time in bed, and not feeling refreshed after sleep. Respondents could select “often,” “sometimes,” “rarely,” and “never” in response. While the scale does not specifically address the frequency of these problems during the week, it does
incorporate some criteria for frequency that could be used to satisfy this requirement. The scale has been subjected to factor analysis and found to have a single-factor structure. However, it has not been systematically analyzed for reliability and validity and no questionnaire to date has been designed and subjected to psychometric analysis for NRS. Thus, measurement of NRS has been left to the discretion of individual researchers and has yet to be standardized, making it difficult to draw comparisons between studies.

Despite these issues, several generally accepted principals have emerged from extended research on NRS, as suggested initially in a review by Stone and colleagues. These are: 1) the term NRS is applied to individuals who feel unrested or unrefreshed upon awakening despite having received a sufficient duration of sleep, 2) the problem must be present for a minimum of one month, and 3) the complaint must be experienced at least three times per week. Stone and colleagues have also suggested that, in order for the label NRS to be applied, individuals should not present with difficulties initiating or maintaining sleep and should not be diagnosed with an organic sleep disorder. However, these guidelines call into question whether NRS should be considered its own primary condition or whether it occurs predominantly as a symptom of other sleep and medical disorders. This is similar to the historical perspective of insomnia as a feature of many other conditions (e.g. depression, pain) rather than a co-occurring condition that may in fact represent a primary disorder.

Researchers continue to treat NRS as a symptom of insomnia and to address its connection to disorders like insomnia, fibromyalgia syndrome (FS), and chronic fatigue syndrome (CFS), meaning that much of the current literature regarding NRS has emerged from its relationship to these other primary conditions. Similarly, some of the only potential biological markers for NRS that have been located to date (alpha EEG sleep anomaly, cyclic
alternating pattern) have been identified through studies of patients with FS and CFS.\textsuperscript{14-15} Only a combination of approaches—those that investigate both the presentation of NRS in a variety of sleep and medical disorders and its prevalence as a primary complaint—will clarify the fundamental nature of NRS.

1.3 PREVALENCE

Without a clear consensus on the definition of NRS, estimates regarding its prevalence differ widely. In their 2008 review, Stone and colleagues observed rates of NRS ranging from 1.4\% to 35\% in the general population.\textsuperscript{1} The highest estimates emerged from studies using single, nonspecific items to detect NRS such as “do you ever wake up with a feeling of exhaustion or fatigue?” and how frequently do you awaken “not feeling really rested?”\textsuperscript{9, 10} Lower estimates were found in studies conducted by groups led by Ohayon and Roth, most of which employed at least a minimum criterion for frequency, if not more stringent stipulations about duration, daytime symptoms, and the presence of a normal sleep duration.\textsuperscript{11-12}

Since the publication of their review, several new studies following most of the guidelines for NRS specified above (sufficient sleep duration, frequency of three nights per week, occurring for one month) have been published for public use. Researchers have found rates as high as 18\% in Sweden and as low as 3\% in Spain,\textsuperscript{16-17} with rates between 8.9\% and 11.1\% found across seven European countries.\textsuperscript{5, 11}
1.4 CORRELATES

NRS is strongly related to other sleep and medical disorders, particularly psychiatric conditions. In one survey of the general population, NRS was found to occur on its own in the absence of other insomnia symptoms (difficulties initiating and maintaining sleep) in only 1.4% of participants.\textsuperscript{11} Similarly, in a survey of primary care patients who had been diagnosed with insomnia, only 1.4% reported having NRS alone.\textsuperscript{18} NRS appears to be related to a number of disorders, symptoms, and demographic categories.

1.4.1 Gender

The interaction between NRS and gender has been investigated in several ways. Among the sufferers of sleep and medical disorders most frequently associated with NRS—insomnia, FS, CFS—women tend to be overrepresented. Prevalence estimates suggest that more than 90% of all cases of FS occur in women, while studies of CFS suggest a gender ratio that is less skewed but still indicative of greater rates in females.\textsuperscript{19-20} Similarly, a large meta-analysis of insomnia epidemiology found that women are at 41% greater risk for developing the sleep disorder than men.\textsuperscript{21} Thus, there is added support for the few studies investigating the prevalence of NRS that have found greater rates in women than men. In the case of two of the large epidemiological studies conducted by Ohayon, reported odds ratios have been as high as 1.5.\textsuperscript{8,16}

It is possible that the higher rates of NRS in women seen in these epidemiological studies are simply reflective of a female propensity towards sleep and medical disorders associated with NRS symptoms. However, even studies evaluating individuals experiencing NRS in absence of other symptoms and diagnoses seem to suggest a gender difference. Roth and colleagues
evaluated a cohort of insomnia patients reporting only NRS symptoms using their more stringent criteria for NRS. They found that 61% of those patients with NRS were women.\textsuperscript{22} One study of 541 insomnia patients conducted by Sarsour and colleagues assessed NRS by asking participants how often in the past month they had awoken feeling unrested despite a sufficient amount of sleep. The researchers found that 75% of those who reported NRS were women.\textsuperscript{23} Similarly, the presence of children under the age of six in the home has been found to be a non-significant factor in the determination of NRS, suggesting that sleep disruption due to childcare cannot explain these differences.\textsuperscript{16}

Though a gender disparity has been replicated in quite a few studies, it is important to note that in some cases an almost 1:1 relationship has also been observed.\textsuperscript{18, 24} Whether the gender difference seen in NRS reflects some innate quality or is simply an artefact of the higher rates of depression and anxiety seen in women remains unclear. Additional studies on the subject will help to clarify how the disparity should be interpreted.

\subsection*{1.4.2 Age}

The interaction between age and NRS has been well established through a number of prevalence studies. Ohayon and colleagues have demonstrated repeatedly that NRS prevalence rates seem to peak in the period between ages 25 to 44 years and that rates tend to be lowest among those 55 years and up.\textsuperscript{8, 16, 24} Similarly, in one population survey Roth and colleagues found an inverse relationship between age and NRS, with those in the 18 to 29 years cohort demonstrating significantly elevated odds for NRS.\textsuperscript{12}

The trend may appear counterintuitive as PSG studies have confirmed that older age is significantly associated with shorter sleep time, reduced sleep efficiency, increased arousals, and
increased reports of feeling unrested after sleep, suggesting that the identification of NRS should also increase in this population. However, it may be these age-related declines in sleep quality and quantity that actually cause older adults to be ineligible for inclusion in NRS classifications. Many elderly individuals may no longer meet the criterion for normal sleep duration that typically accompanies NRS diagnosis. This could leave a significant population of interest unrepresented in efforts to define NRS.

The trend of younger adults experiencing higher rates of NRS has not been replicated in all cases—one study of insomnia prevalence in Spain found nearly equivalent rates of NRS in all age groups, with the lowest rates occurring in the 25-34 year age bracket. This same study also found comparable rates of NRS between women and men, suggesting a possible cultural component reflected in the presentation of NRS.

1.4.3 Insomnia

As mentioned, research on NRS indicates that it frequently occurs in the presence of other symptoms of insomnia. A recent study examining a sample of insomnia patients in a primary care setting used logistic regression to demonstrate that all symptoms of insomnia are significantly related to one another. The researchers found that the most frequently reported type of insomnia was a combination of all symptoms (difficulties initiating and maintaining sleep, early morning awakenings, and NRS), with 38.6% of the population reporting all four. In the same study, only 1.4% of the population were categorized as having NRS alone. Similarly, a study of insomnia patients conducted by Sarsour and colleague demonstrated that, of those reporting NRS symptoms, more than 80% also reported at least one other symptom of
insomnia.$^{23}$ The frequency of reported NRS symptoms was also found to be significantly correlated with difficulties initiating and maintaining sleep and early morning awakenings.

This high degree of interconnectedness between symptoms may present further complications to the definition of NRS. Many issues associated with insomnia—difficulties initiating and maintaining sleep, early morning awakenings—have the potential to disrupt sleep continuity and duration and to interfere with individuals’ attempts to sleep at times that are optimal in terms of their circadian cycles. In this way, and specifically in this context, it becomes unclear whether NRS is its own symptom/diagnosis or simply a product of the disturbances in sleep caused by other symptoms of insomnia.

In an attempt to clarify the issue and to isolate NRS as a distinct entity, Roth and colleagues conducted an analysis of 579 subjects reporting unrefreshing or unrestful sleep.$^{22}$ After ruling out any sleep disorders or medical conditions that could be causing the subjective experience of unrefreshing sleep, they identified a cohort of individuals experiencing “pure NRS.” Using PSG, researchers confirmed that individuals in the NRS-only group had normal mean sleep latency, a normal amount of wakefulness following sleep onset, and a normal sleep duration when compared to healthy controls. Despite having received an adequate duration of sleep, the levels of daytime impairment, depression, and anxiety in this NRS-only group were comparable to levels found in individuals with other insomnia symptoms (difficulties initiating and/or maintaining sleep). These results suggest not only that NRS exists as a distinct entity, but that it causes equivalent issues with functioning as other better recognized sleep disorder symptoms.
Similarly, in their study of 1258 insomnia patients with and without NRS, Sarsour and colleagues found that almost 15% of those belonging to the NRS group were classified as having no insomnia using the Insomnia Severity Index. Further, 12% were found to have no insomnia and no depression, allowing for the elimination of at least one psychiatric diagnosis as an explanation for poor subjective sleep quality. Finally, when they controlled for insomnia severity and the presence of depression, researchers found that NRS was associated with an increase in total sleep duration. Though all sleep variables were subjectively reported and therefore not verified using objective data, these findings suggest that NRS does indeed exist independently of other insomnia symptoms and diagnoses, despite the presence of a reasonable duration of sleep. One might consider a classification of NRS-P (pure), NRS-I (with other insomnia symptoms), NRS-SD (with other sleep disorders), or NRS-C (comorbid with other disorders, e.g. fibromyalgia)

1.4.4 Daytime Symptoms

Very few studies have examined impairment of daytime function specifically within NRS. In studies that have investigated NRS along with comorbid issues, NRS appears to be strongly linked to issues with daytime functioning. For example, Ohayon found that individuals with NRS rated all spheres of daytime impairment as worse than those who had other symptoms of insomnia only. However, the presentation of NRS in the absence of other symptoms or diagnoses has not been well characterized.

A few recent studies have investigated this issue, evaluating the presentation and outcomes of NRS in the absence of other insomnia symptoms. In a survey of primary care insomnia patients, Léger and colleagues asked participants to rate the intensity of the negative
impact that their sleep problems had on several spheres of daytime functioning, including: daily activities, work, relationships, leisure, memory, concentration, irritability, tension, and depressed mood. When compared to individuals with other combinations of insomnia symptoms, patients with NRS alone reported less impairment to daytime functioning. For each category of impairment, between 40 and 60% of those belonging to the NRS-only group endorsed moderate to severe impairment, suggesting that NRS is still an issue in its own right. However, though researchers did screen for psychiatric and sleep disorder comorbidities, participants were not excluded on this basis. Thus, any conclusions drawn regarding their NRS group must take into account the possibility of confounding influences.

In contrast, Roth and colleagues examined a “pure NRS” group for daytime correlates, using both questionnaire data and overnight PSG to screen for individuals experiencing only NRS symptoms. They employed several questionnaires to assess daytime functioning including scales that evaluate impact on productivity and efficiency in work activities, daytime sleepiness, fatigue symptoms, psychiatric outcomes, and physical and mental quality of life. Additionally, they administered two questionnaires still under development—the Restorative Sleep Questionnaire and the Daytime Consequences of Sleep Questionnaire. Using these outcome measures, researchers found equivalent impairments between the NRS-only group and the cohort with difficulties initiating and/or maintaining sleep in all areas of daytime functioning, including work productivity, mood, daytime sleepiness, and fatigue. Similarly, the NRS group was found to have significantly greater daytime impairment than healthy controls, as well as lower scores on both mental and physical quality of life.
In their recent study of insomnia patients with and without NRS, Sarsour and colleagues also excluded those with other sleep diagnoses in order to better isolate NRS. Daytime function was assessed by asking respondents how frequently they experienced the following three issues due to their sleep pattern: 1) feeling sleepy, tired, or low energy during the day, 2) having poor concentration or memory, or needing more effort to get things done, and 3) feeling irritable, stressed, or being in a bad mood. Researchers found that the NRS group had more individuals with decreased daytime physical function (73% vs. 33%), cognitive function (51% vs. 20%), and emotional function (53% vs. 22%). This association remained even after controlling for insomnia severity, sleep duration, and depression.

One of the biggest issues confronting research on NRS is that of patient perception. In their recent study, Léger and colleagues observed that the strongest factor associated with severe daytime impairment was the perception that sleep quality was poor. As with any medical condition, patient perception is highly significant in determining the course and outcome of illness. Daytime functioning is strongly linked with perceived quality of sleep, regardless of its objective quality as determined by biological measures. Thus, the impression that sleep has not been restorative may, in itself, contribute to outcomes resembling nonrestorative sleep, complicating the task of addressing NRS even further.

In this way, patient reports are crucial to understanding NRS. Roth and colleagues have observed that clinicians tend rate the severity of symptoms as lower for NRS-only patients compared to patients with difficulties initiating and/or maintaining sleep. In contrast, Sarsour and colleagues found that close to 50% of patients with NRS rated their NRS symptoms as severe or very severe. It is possible that lower clinician ratings may result from the physicians'
perception of NRS as a vague complaint—one without reliable biological markers or validated treatment approaches that is better treated in conjunction with other issues than addressed separately.

1.4.5 

Psychiatric Correlates

According to an analysis of eight epidemiological studies, the presence of insomnia symptoms is significantly predictive of an increased risk for developing depression up to three years after initial assessment. In one survey of primary care patients with insomnia, 22% of participants were being treated for depression and 32% reported being anxious. Similarly, NRS also appears to be linked to a number of psychiatric issues. In one prevalence study, individuals with depressive and bipolar disorders were approximately four times as likely to present with NRS as other individuals in the general population. Those with anxiety were also found to report higher rates of NRS, though the rates were not as significantly elevated as those found within the mood disorder groups.

Similarly, in a study of insomnia patients, Sarsour and colleagues found depression to be significantly correlated with NRS, even after controlling for the severity of insomnia symptoms and patients’ total sleep duration. However, researchers in this study used only two items to screen for depression: 1) “During the past month have you often been bothered by feeling down, depressed, or hopeless?” and 2) “During the past month, have you been bothered by little interest or pleasure in doing things?” Though these screening items have been validated for their utility in detecting the presence or absence of depression, they do not allow for an in-depth exploration of the connection between specific depressive symptoms and NRS.
In their study, Roth and colleagues used data collected as part of the National Comorbidity Survey to evaluate the connection between DSM-IV psychiatric diagnosis and individual symptoms of insomnia. Participants received diagnoses following a structured clinical interview and insomnia was assessed using the more stringent criteria for duration and frequency specified above. Of those participants found to have NRS, 54% also had at least one DSM-IV diagnosis (e.g. mood disorder, anxiety disorder, substance use disorder) with an odds ratio of 4.4. Additionally, 17% of those individuals with NRS were found to have three or more diagnoses, indicating a high degree of psychiatric comorbidity in the NRS population.

These proportions are greater than those observed in a study that examined the relationship between DSM-IV diagnoses and NRS alone (in the absence of other insomnia symptoms). In this case, researchers observed that 34% of NRS-only subjects met the criteria for a mental disorder, suggesting that some of the psychiatric comorbidity observed in other studies may be related to other symptoms of insomnia and not just NRS. Similarly, other studies have found no relationship at all between NRS and depressive or anxious mood. In order to satisfactorily characterize the connection between psychiatric illness and NRS, future studies will need to focus on isolating and examining “pure NRS” (in the absence of other sleep disorders and medical conditions).

1.4.6 Fibromyalgia Syndrome and Chronic Fatigue Syndrome

Fibromyalgia syndrome (FS) is a heterogeneous illness characterized by complaints of widespread pain, unrefreshing sleep, fatigue that significantly impairs daytime function, and increased sensitivity to stimuli. Similarly, chronic fatigue syndrome (CFS) is defined by persistent fatigue unexplained by other medical disorders and a variety of other symptoms
including joint and muscle pain, depression, and unrefreshing sleep. For many, these supposedly separate conditions represent a single entity due to their high degree of symptom overlap (see reference 29 for a discussion of this). Regardless of their etiologies, with NRS and persistent fatigue as two of the defining features of these disorders, FS and CFS have been investigated as potential keys to the understanding of NRS.

In a recent survey study, researchers administered the Medical Outcomes Study Sleep Scale to more than 3000 FS patients and found that only about 30% reported “optimal sleep.” On a scale of 0-100, participants reported a mean score of 32.2 to an item evaluating “sleep adequacy.” In another study of 43 women with FS, 72% of those surveyed reported “nonrestorative” sleep—a significantly greater proportion than those found in both a rheumatoid arthritis cohort and a group of healthy controls. And in one internet survey study of FS, respondents labelled “nonrestorative sleep” as one of their most severe symptoms (ranked only slightly below morning stiffness and fatigue). In terms of CFS, approximately 85-95% of patients report unrefreshing or “nonrestorative” sleep. Indeed, one study found that individuals with CFS were 28 times as likely to score above threshold on a factor evaluating NRS than individuals with no fatigue.

However, attempts to draw comparisons between NRS in insomnia patients and chronic pain or fatigue patients is complicated by the very different methods used to define and measure NRS within these two separate disciplines. In the majority of research studies examining FS, the criteria used to evaluate “nonrestorative” or unrefreshing sleep are less stringent than those currently employed in insomnia research. In the study performed by Drewes and colleagues, “nonrestorative sleep” was assessed with only two items: “feeling unrefreshed or tired after
getting up” and “difficulties staying awake after getting up,” with no guidelines regarding the frequency or duration of problems and no attempts made to take into account sleep quantity.\textsuperscript{31} Other studies make no mention at all of how they have defined “nonrestorative” or unrefreshing sleep, likely using only a single item.\textsuperscript{33}

The Sleep Assessment Questionnaire (SAQ), which features one factor devoted to measuring NRS, represents a somewhat more focused tool for assessing sleep in those with FS. Its NRS factor threshold was selected based on its sensitivity in evaluating unrefreshing sleep in patients with FS and CFS.\textsuperscript{35} Scores on the SAQ were also found to be higher for patients diagnosed with sleep apnea, periodic leg movements, and snoring,\textsuperscript{36} suggesting it may be useful in evaluating NRS across several sleep and medical conditions. Unfortunately, the SAQ has been employed in very few studies thus far, and some have criticized it for using only one item that seems to specifically address NRS.\textsuperscript{13}

\subsection*{1.5 PHYSIOLOGICAL MARKERS OF NRS}

Unlike conditions such as obstructive sleep apnea or restless legs syndrome, there are no definitive guidelines regarding the biological mechanisms of NRS. Further complicating this issue is the fact that the criteria of NRS proposed by some researchers suggest that many of the parameters recorded by objective measures should appear very similar for both NRS patients and those with normal, healthy sleep.

Only one study to date has analyzed polysomnographic data in patients with “pure NRS” (meeting the criteria suggested by Roth and Ohayon and occurring in the absence of other sleep or medical disorders). They confirmed that many factors evaluated by PSG are similar between
these NRS-only patients and normal controls, including total sleep time, sleep latency, and wakefulness after sleep onset. In terms of the differences found, the NRS-only group spent somewhat less time in stages 3 and 4 sleep and less time in rapid eye movement (REM) sleep, though the distinctions were minimal.

Due to the lack of studies examining “pure NRS” in the literature, the majority of evidence for physiological markers of NRS has emerged from studies of patients with fibromyalgia and chronic fatigue. Using this patient population, researchers have identified several PSG markers that appear to be tentatively associated with NRS, including arousal disturbances and disruptions to non-REM or slow wave sleep. More recent theories have also attempted to explain FS and CFS using immunological and endocrinological models, though it remains unclear if this approach could be generalized to a population with NRS from other causes. Characteristic markers of FS have been observed repeatedly in PSG studies, including reduced sleep efficiency, prolonged sleep latency, increased stage 1 sleep, reduced REM and slow wave sleep percentages, and increased arousals. Similar findings are supported in studies of patients with CFS as well, with participants demonstrating reduced total sleep time and sleep efficiency, and decreased REM and slow wave sleep. However, these findings have not been replicated in all cases, and some researchers have observed no differences in objective PSG data.

1.5.1  
Alpha Wave Intrusion in Non-REM Sleep

In further attempts to pinpoint the mechanisms of unrefreshing sleep, investigators have moved beyond these macroarchitectural parameters to evaluate the microarchitecture of sleep. In 1973, Hauri and Hawkins examined the overnight electroencephalographic (EEG) recordings of
nine psychiatric patients suffering from fatigue or “chronic, somatic malaise.” Researchers observed a novel EEG pattern that did not match the patterns previously described for different stages of sleep, referring to it as “alpha-delta sleep.” Following the initial recognition of this anomaly, a number of sleep microarchitectural parameters have been investigated for their contribution to the feeling of unrefreshing sleep.

Since it was first described almost 40 years ago, elevated levels of “alpha-delta sleep” have been found both in patients with FS and, to a lesser extent, in those with CFS. Also referred to as alpha EEG non-REM sleep, alpha-delta sleep is an arousal disturbance characterized by intrusions of alpha waves—waves with a frequency of 7 to 11 Hz and minimum peak-to-peak amplitude of 5uV—into stages 2, 3, and/or 4 sleep. Roizenblatt and colleagues have observed three distinct patterns of alpha sleep—a phasic pattern which occurs episodically in conjunction with delta waves, a tonic pattern that presents continuously throughout non-REM sleep, and low alpha activity which predominates for the majority of “normal” sleepers.

However, the evidence linking alpha-delta EEG patterns and unrefreshing sleep remains controversial. Roizenblatt et al. found that the phasic pattern of activity occurred in 50% of FS patients and was present in only 7% of controls. In the phasic group, 100% of patients reported nonrefreshing sleep. They gave significantly lower scores to their sleep quality and were found to have reduced total sleep time, sleep efficiency and slow wave sleep (SWS) compared to patients in the other alpha activity groups. Thus, these phasic alpha intrusions appear to disrupt slow wave sleep and lead to unrefreshing or nonrestorative sleep. In their validation study of the SAQ, Cesta and colleagues observed that all patients in the CFS and FS groups demonstrated this alpha sleep anomaly.
A number of other studies have failed to replicate these findings.\textsuperscript{45} Rather, they indicate links between alpha activity and conditions such as pain and anxiety, but not unrefreshing sleep.\textsuperscript{43,46} One of the biggest complications inherent in interpreting these studies lies in the distinction between different types of alpha activity. While phasic alpha intrusions may be linked to fatigue, myalgia, and unrefreshing sleep, some studies suggest that continuous tonic alpha patterns may have sleep maintaining effects.\textsuperscript{47} In many cases, the distinction is not made while scoring. Still, alpha non-REM intrusions represent a tentative clue to the nature of NRS and additional studies will help to elucidate this relationship further.

1.5.2  \textit{Cyclic Alternating Pattern}

Cyclic alternating pattern (CAP) refers to periodic, stereotyped EEG patterns of higher amplitude separated by intervening periods of lower amplitude background activity.\textsuperscript{46} Each cycle (the repetitive activity phase plus the background activity phase) lasts between 2 and 60 seconds and can be characterized by the EEG elements present in the activity phase (e.g. K-complexes, spindles, vertex waves). A few different subtypes of CAP have been identified. Much like alpha EEG anomaly, one of these subtypes appears to have sleep maintaining effects.\textsuperscript{48} Other subtypes, however, represent arousal disturbances to non-REM sleep that could potentially hold the key to the experience of unrefreshing sleep.

Periodic K-alpha EEG sleep has received the most attention as a candidate for the relationship between NRS and CAP. With this arousal disturbance, a non-REM stage 2 K-complex is not followed by a sleep spindle as expected but is instead associated with a burst of alpha activity.\textsuperscript{44} A high frequency of CAP has been found in both FS and CFS patients.\textsuperscript{49-50} Similarly, several studies have found CAP to be associated with complaints of unrefreshing sleep.
and sleep disorders, including primary insomnia and periodic limb movements during sleep.\textsuperscript{51-52} Still, research on CAP is only beginning to elucidate its relationship to sleep quality, and investigations into the microstructure of sleep are complicated by the lack of standardized recording and scoring techniques as well as the differences in lab equipment used by researchers.\textsuperscript{53}

\textbf{1.5.3 Immunological Markers of NRS}

While the analysis of sleep microarchitecture offers some tentative evidence for the mechanisms of NRS, any attempts to understand sleep function should also take neuroendocrine and immune systems into account. The relationship between sleep and immune function has been established for decades, but evidence in support of an immunological explanation of NRS has recently begun to build. Both CFS and FS have been linked to immune system function. The onset of CFS is often preceded by some form of acute inflammatory insult, and individuals with CFS appear to have abnormalities in their immune cell function.\textsuperscript{53-54} Similarly, autoimmune disease and acute infections have been correlated with the onset of fibromyalgia.\textsuperscript{55} Even in healthy individuals, experimentally induced sleep loss has been shown to interfere with the proper functioning of the immune system.\textsuperscript{56}

A growing body of experimental evidence suggests a bi-directional relationship between changes in sleep regulation and pro-inflammatory cytokines such as interleukin-1 (IL-1), interleukin-6 (IL-6), and tumor necrotic factor. Studies have found that increasing circulating levels of IL-6 correlate with decreases in SWS, and increases in REM percentage and density.\textsuperscript{57-59} Subjective sleep quality also appears to be affected, with studies showing a negative correlation between circulating levels of IL-6 and subjective ratings of sleep (assessed using
questions regarding “feeling refreshed in the morning,” among other symptoms). Perhaps most strikingly, the stimulation of evening IL-6 production in healthy controls is also associated with decreased SWS and increased REM sleep, as well as increased fatigue reported during the following day. Though a causal relationship cannot be assumed from these results, they are highly suggestive of an influence on sleep architecture by circulating levels of cytokines and they provide preliminary support for the role of immune dysregulation in NRS.

Cytokine production is not only stimulated by the immune system, but also has a strong relationship with the stress response. In this way, the interaction between cytokines and sleep may be mediated by a “hyperarousal” effect. Chronic stress leads to increased activity of the sympathetic nervous system, which in turn stimulates production of cytokines like IL-6. A large body of evidence has shown that stress leads to PSG changes that are similar to those associated with increased cytokine production—specifically, reduced SWS and sleep efficiency. Studies have confirmed alterations in brain function among patients with primary insomnia that suggest hyperarousal of the central nervous system.

The high degree of interconnectedness between immune function, stress, and sleep makes it very difficult to determine the direction of any potential cause and effect relationships. Longitudinal evaluations of the interplay between these factors are necessary before any conclusions can be drawn. Still, these kinds of neurobiological explanations of NRS offer new perspectives on issues of sleep quality and may even help to explain some of the sex differences inherent in the presentation of disorders associated with NRS. Since women experience greater immune cell activation in response to sleep loss, it is possible that their higher rates of reported insomnia, CFS, and FS could be related.
1.5.4 Slow Wave Sleep

What many theoretical models of NRS share in common is their effect on sleep architecture. Alpha EEG non-REM sleep, CAP, and increased cytokine production all seem to share a similar polysomnographic characteristic—namely, the reduction or fragmentation of SWS. This has led many researchers to identify SWS as a potential candidate involved in the restorative effect of sleep. Slow wave sleep activity occurs primarily during non-REM sleep and is characterized by the presence of high amplitude delta waves with a frequency ranging from ~0.5-5 Hz. This activity is greatest following periods of sleep deprivation or disruption, showing significant rebound effects and indicating its powerful homeostatic necessity.\(^{65}\)

The implications of this are unclear. Studies in which SWS is experimentally disrupted or reduced have been inconclusive thus far. Some have demonstrated a number of subsequent daytime impairments following SWS disturbance including greater daytime sleep propensity (as assessed by MSLT) and poorer performance on learning, attention, and memory tasks.\(^{18,66}\) In their study of the links between IL-6 production and fatigue, Thomas and colleagues found that, of those PSG features assessed, SWS alone mediated the relationship between reported daytime fatigue and cytokine production.\(^{67}\) However, other studies have found no link with daytime functioning, and none appear to have specifically queried feelings of being unrefreshed or unrested following sleep.\(^{68}\) Similarly, of two pharmacological interventions found to increase SWS, one (tiagbine) appears to relieve cognitive impairments but not sleepiness while another (gaboxodol) improves sleepiness but not cognitive performance.\(^{65}\) Thus, there is likely to be a third factor mediating the relationships between restorative sleep and SWS activity that remains unknown.
1.5.5 *Growth Hormone*

Directly related to the interruption of SWS as a potential marker of NRS, the secretion of growth hormone (GH) also seems to have significant bearing on feelings of fatigue. Individuals with GH deficiency report poorer sleep quality and obtain lower scores on measures of quality of life than healthy controls.\(^6^9\) Some researchers have also observed a higher prevalence of GH dysfunction among individuals with FS, suggesting a possible explanation for the high degree of symptom overlap between the two conditions.\(^7^0\)

In many ways, this is likely due to the close association between GH release and SWS. Daily secretion of GH peaks at the first appearance of delta waves during sleep, and the quantity of GH secreted is directly related to length of the SWS episode.\(^7^1\) Similarly, night time injections of GH-releasing hormone (GHRH) increases time spent in slow wave sleep, while substances that suppress endogenous GHRH cause decreases in non-REM sleep.\(^7^2\) Issues with GH secretion represent an interesting avenue for potential pharmacological intervention. However, many preliminary studies of GH replacement therapy have failed to find significant changes to sleep architecture,\(^7^3\) meaning that much work is yet to be done before the interaction between sleep and GH is adequately understood.

1.5.6 *Summary of Review*

Sleep is a multi-factorial process that is not easily captured by the diagnostic tools sleep specialists currently have at their disposal. Polysomnographic cannot currently capture the extent of this process. Electroencephalography, for example, primarily records surface electrical potential, while deep brain structures contribute only minimally to the signal. Similarly, the impracticality of monitoring endocrine variables during routine PSG dictates that an integral
element of the sleep quality picture must be neglected during diagnosis. A number of promising physiological markers of NRS have been identified. Alpha EEG sleep, CAP, and various endocrine and immunological variables offer researchers new perspectives on sleep and its functions. However, many additional studies are required before these factors are sufficiently confirmed as biomarkers that can be employed in a clinical setting. Thus, with no satisfactory confirmatory markers currently established to detect NRS, an instrument for the assessment of subjective NRS is necessary.

Recent studies have brought sleep researchers to relative consensus regarding the features of NRS. That is, patients should awaken feeling unrefreshed or unrested three times per week and the problem should persist for the duration of at least one month. However, these guidelines have not yet been adopted by researchers in other fields, making interdisciplinary communication more difficult. Additionally, these criteria remain arbitrarily chosen for the purpose of facilitating research and it is unclear if they will change how NRS is treated in the practical setting of the clinic. Such symptoms may result from a highly heterogeneous group of conditions and disorders, making it difficult to determine cause-and-effect relationships. In order to improve the outcomes of patients experiencing NRS and to enhance understanding of the underlying function of sleep itself, certain research directions should be made high priority for the future. These include: a) the impact of NRS on illness intrusiveness, b) the impact of NRS on daytime sleepiness, fatigue, and alertness, and c) the relationship between NRS and psychiatric disorders.
1.6 STUDY RATIONALE

Without a gold-standard biological marker for the subjective experience of NRS, poor subjective sleep quality remains a self-report symptom. However, there is currently no definitively accepted tool for the measurement of NRS across medical, psychiatric, and sleep disorders, and many of the tools used to assess its presence in specific disorders are unavailable to researchers and clinicians. A recent review evaluated 59 different sleep-related questionnaires for content relating to NRS. Of the 26 questionnaires found to query unrefreshing sleep and its daytime consequences, only one featured a domain score specifically devoted to NRS—the Sleep Assessment Questionnaire (SAQ). The SAQ has been found to be sensitive to FMS and CFS compared to healthy controls. However, some have argued that the scale includes only one item that technically queries NRS—a question asking respondents to indicate if they awaken feeling unrefreshed or unrested.

A number of other questionnaires have also been identified for their potential to assess NRS. The Pittsburgh Insomnia Rating Scale (PIRS) contains four NRS-related questions relating to quality of sleep and feelings of being refreshed from sleep, asking about both daytime and night time consequences of NRS. The PIRS does ask respondents to specify the frequency with which they awaken feeling unrefreshed, but no duration criteria are specified. Similarly, the Insomnia Symptom Questionnaire includes only one item relating to unrefreshing sleep, though it asks for both the number of times per week that the symptom is experienced and for the duration of the problem, making it potentially useful in identifying chronic NRS.

Beyond these limited options, no readily available standardized tool has been designed and implemented specifically for the purpose of evaluating NRS. A scale validated for use at the
clinical level—across all manner of sleep and medical conditions—would provide an invaluable resource for physicians as they screen for NRS and attempt to evaluate its progress over the course of treatment. It would also permit more in-depth evaluations of potential objective markers of NRS—investigations that could then be standardized across different studies.

1.7 THESIS HYPOTHESIS AND OBJECTIVES

Two research questions will be addressed as part of this thesis:

1. Can a valid and reliable measure of NRS be developed from the original item pool of the NRSS?

2. Are specific overnight polysomnographic (PSG) variables correlated with the subjective experience of nonrestorative sleep?

Thus, two hypotheses accompany this research:

1. Developed through principal component analysis, the reduced version of the NRSS will prove to be a statistically valid and reliable instrument for measuring nonrestorative sleep in a sleep clinic population.

2. Arousal index and level of non-REM alpha EEG will be correlated with the subjective experience of nonrestorative sleep.

The objective is to create a reliable, valid, and manageable instrument for assessing NRS at a clinical level so that such a tool can be made available to researchers and clinicians alike in their efforts to diagnose, to treat, and to further investigate the presentation of NRS in a number of disorders and conditions.
2 METHODS

2.1 SAMPLE SIZE DETERMINATION

Prior to commencing the current project, the literature contained a number of studies investigating the prevalence of NRS in both the general population and in a primary care setting.\textsuperscript{8,11,16,18} However, no studies to date have examined the prevalence of NRS in a sleep clinic population, making traditional sample size calculation unfeasible. Therefore, in lieu of this calculation, I consulted the methods used by Buysse and colleagues during development of the Pittsburgh Sleep Quality Index (PSQI) as well as recommendations made by Field that suggest between 200 and 300 participants should be recruited for the purposes of scale validation.\textsuperscript{76-77} The questionnaire was administered to 225 patients and 30 control participants for a total of 255 participants.

2.2 SAMPLE POPULATION

Patient participants were recruited to the study based on their attendance at one of two different sleep clinics for diagnosis of a sleep complaint. Patients were approached consecutively at the time of their initial consultation with the sleep physician and a total of 226 individuals (108 female) completed the questionnaire battery at this time. Of those in the patient group, 43 consecutive patients were approached again on the night of their sleep study to complete the NRSS for a second time. The scales collected during this second administration
were used to assess the test-retest reliability of the scale. A group of 30 control individuals (16 female) were also asked to complete the NRSS. This group was recruited on the basis of their oral 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. Demographic information for participants is reported in the results section.

2.3 STUDY DESIGN

The study was divided into four phases: 1) questionnaire design, 2) factor analysis and assessment of reliability, sensitivity, and specificity, 3) evaluation of questionnaire validity, and 4) assessment of relationship between subjective measures of NRS and objective sleep data collected through PSG.

2.3.1 Questionnaire Design

The initial version of the NRSS was developed by Dr. Colin Shapiro, a second sleep specialist, and a medical student. Beginning with a review of the literature, several issues associated NRS were identified: subjective reports of sleep quality, physical/medical outcomes associated with NRS, affective symptoms of NRS, and impairment of daytime functioning. In consideration of these factors, a pool of items was generated based on both clinical experience and recommendations made in previous reviews evaluating approaches to assessing NRS.\textsuperscript{1,13} Following item development, the scale was introduced to a small panel of six experts in the field and items were revised, added, or cut based on suggestions. A larger group of approximately 30 individuals working as health care professionals in both sleep and psychiatric departments and
associated with Dr. Shapiro’s laboratory were also asked to examine the scale and consult their clinical experience to help revise items and improve its face validity.

Finally, two patient focus groups were consulted in an effort to determine 1) if the contents of the questionnaire were considered representative of their experiences with sleep, and 2) if the scale adequately addressed the issues they identified as most relevant to the experience of NRS. The first focus group included 7 individuals who had previously been diagnosed with narcolepsy, while the second group included approximately 10 individuals who were taking part in a support group for obesity-related issues organized through the sleep clinic.

Based on the initial development phase, a pool of 34 items was generated (see Appendix, page 86). Of those 34 items, 28 employ Likert scales that range from one to ten, while an additional six items ask respondents to select one option (of five or six) that best describes their experiences. These differently-formatted items were interspersed throughout the questionnaire in order to break up the uniformity of the instrument and to encourage respondents to attend to each item more carefully. A number of items were worded positively (with ten indicating very good sleep or very high alertness), while others were worded negatively (where ten refers to very poor sleep or refreshment) to ensure that respondents paid closer attention to their answers and not develop a reference set.

The scale’s scoring system has been designed so that, when evaluating responses, all items are given a weighted score from one to five. Therefore, when scoring those questions with Likert scales ranging from one to ten, responses of one or two are given a score of one, responses of three or four are given a score of two, etc. Though this system slightly complicates the process of scoring, it ensures that all questions are weighted equally, while scales ranging from one to ten were selected in order to provide a finer degree of choice to respondents and hopefully
discourage the selection of multiple responses for each question. Negatively worded items are reversed before scoring, meaning that higher scores on the scale are indicative of less NRS.

In terms of the content of the scale, questions were designed to fit into one of roughly four different subscales. The first subscale directly addresses feelings of refreshment and restoration from sleep, asking respondents to indicate how frequently they feel refreshed, at what times during the day, and what factors tend to affect how refreshed they feel. The final three subscales assess constructs that—while distinct from restorative sleep—are found to be influenced by feelings of subjective sleep quality. The first of these subscales refers daytime functioning and its impairment. These items relate to reported energy levels, feelings of alertness, and ability to concentrate and remember things during the day. The final two subscales concern the physical and affective symptoms of NRS. These symptoms include unexplained illness, medical issues, panic, depression, irritability, and negative mood.

2.3.2 Factor Analysis and Evaluation of Scale Reliability, Sensitivity, and Specificity

During the second phase, a total of 255 completed questionnaires were obtained. Of these, 226 were collected from patients consecutively attending two sleep clinics—the Toronto Sleep and Alertness Clinic in Toronto and the International Sleep Clinic in Parry Sound. An additional 30 questionnaires were obtained from healthy control participants. Using these collected scales, statistical analysis was performed to determine the best factor solution of the scale. All analysis was conducted using Predictive Analytics Software (PASW) Statistics Version 18.

Initially, a correlation matrix was created in order to screen for items that correlated too highly with other items (r > 0.9) or for items that correlated with very few other variables. At
this point, several items were also identified and eliminated based on factors like redundancy, lack of clarity, frequently skipped during response, etc. Following data screening, principal component analysis was conducted. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and the Bartlett’s test were calculated. A KMO value greater than 0.5 and a significance of p < .005 on the Bartlett’s test were considered adequate.

In order to determine which factors should be retained, Kaiser’s recommendation as cited in Field was tested. This led to the selection of those factors with eigenvalues greater than 1. Following this, a scree plot was consulted to confirm the validity of this approach and to make adjustments in order to improve the total variance explained by the scale (to a minimum of 70%). Though the precise connection between factors remains unclear (e.g. the bi-directional interaction between affective symptoms and NRS), an oblique rotation was selected for principal component analysis as opposed to an orthogonal rotation. This was based on the assumption that different scale factors were likely to be related, which was later confirmed during analysis.

Following principal component analysis, both the internal reliability and the test-retest reliability of the scale were assessed. Internal reliability was evaluated using Cronbach’s $a$ statistic and corrected inter-item correlations were calculated both within factors and for the scale overall. To evaluate test-retest reliability, comparisons of factor and global scores were made between first and second administrations of the scale using Spearman correlations.

Sensitivity and specificity were assessed using a Receiver Operating Characteristic (ROC) curve to evaluate the capacity for different cutoff scores to distinguish between patient and control groups. Sensitivity is defined as the proportion of individuals belonging to the patient group who are correctly identified as having difficulties with NRS, while specificity
represents the proportion of individuals in the control group who are correctly identified as not having problems with NRS.

However, as NRS should not necessarily be considered a symptom that all individuals within a sleep clinic population experience, an additional assessment of sensitivity and specificity was also performed. In this case, patients were separated into groups based on their response to the question: “How often have you felt really refreshed upon awakening in the morning?” According to criteria suggested in a recent review by Stone and colleagues, in order to be considered to have NRS, an individual should report feeling unrefreshed from sleep at least three times per week, despite receiving a sufficient duration of sleep. Therefore, individuals who endorsed experiencing refreshing sleep two to three days per week or less were placed in the NRS group. A second ROC curve was created to evaluate the scale’s capacity to distinguish between the two groups. Though a single question regarding NRS does not represent the kind of gold standard diagnostic measure usually required for the creation of an ROC curve, this approach represents a tentative attempt to create a distinction between individuals with a sleep disorder and those actually experiencing NRS.

2.3.3 Evaluation of Questionnaire Validity

This phase involved collection of supplementary questionnaire data for participants belonging to the patient group in order to determine the validity of the NRSS. As no scale to date has been validated specifically for the measurement of NRS in a sleep clinic population, the questionnaires selected for validation purposes were chosen based on a review performed by Vernon and colleagues suggesting their relevancy to restorative sleep, as well as for their relationship to additional factors evaluated as part of the NRSS. These supplementary
questionnaires were completed either simultaneously with the NRSS or within approximately the same two-week period.

**Pittsburgh Sleep Quality Index (PSQI):** The PSQI consists of 19 items chosen to evaluate seven different domains of sleep, including one item that specifically queries subjective sleep quality. Additional scale domains are: sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction. The scale includes both Likert-type items and open-ended questions, which are later converted to scaled scores using questionnaire guidelines. A global score is calculated by summing each of the seven component scores. Respondents are required to refer to their experiences over the past month when answering. Scores on the restorative sleep domain of the NRSS will be compared using Spearman correlations to scores on one item of the PSQI (“During the past month, how would you rate your sleep quality overall?”). This item was identified in a recent review of almost 100 scales as specifically addressing the subjective experience of restorative sleep. Additionally, the daytime dysfunction factor of the PSQI will be compared to the relevant factor on the NRSS through Spearman correlation (factor queries difficulties staying awake and enthusiasm for getting things done). Finally, global scores on the PSQI will be compared using Spearman correlations to global scores on the NRSS to ensure that scales evaluate distinct constructs.

**Fatigue Severity Scale (FSS):** Respondents use a Likert scale ranging from 1 to 7 to indicate how strongly they agree with nine statements regarding fatigue (higher scores demonstrate a greater degree of agreement). Items relate to a variety of issues surrounding fatigue, focusing particularly on its interference with daily functioning. A total score is calculated by summing scores on each item. Scores on the FSS will be compared to the NRSS factor most closely related to daytime functioning using Spearman correlations. Additionally,
global scores on the FSS will be compared to global scores on the NRSS using Spearman correlations in order to confirm that they assess different constructs. Finally, one of the items of the FSS has been selected for its relevancy to the physical/medical factor of the NRSS: “My fatigue prevents sustained physical functioning.” Spearman correlations will be used to compare scores on this item to scores on the physical/medical factor of the NRSS. While this item is possibly not satisfactorily relevant to the domain being assessed, it has been chosen for this capacity because very few questions deemed relevant were initially administered to participants. This issue is discussed further in the limitations section of this thesis.

**Center for Epidemiological Studies Depression Scale (CES-D):** The CES-D was designed as a screening tool for depression and it consists of 20 items that are divided into four domains: depressed affect, positive affect, somatic and retarded activity, and interpersonal factors. Respondents use a Likert scale ranging from 0 to 3 to indicate how often they have experienced certain cognitions or physical states during the previous month with higher scores indicating greater depressive symptoms. Summing scores on each item will give a total score, with a cutoff of 16 considered suggestive of high levels of depression. Global scores on the CES-D will be compared using Spearman correlations to scores on the affective factor of the NRSS. Additionally, Spearman correlations will be used to compare global scores on both scales to ensure that they measure separate constructs.

**Athens Insomnia Scale (AIS):** Designed to identify potential cases of insomnia, the AIS is composed of 8 questions that evaluate both night time symptoms of insomnia as well as its daytime consequences. Each item includes a Likert scale ranging from 0 to 3, with higher scores indicating a greater severity of symptoms. Respondents are asked to refer to their experiences over the previous month when answering and are to indicate a response only if it
occurred at least three times per week. In their review of scales to assess NRS, Vernon and colleagues identified a single item of the AIS as particularly relevant to the construct: “Overall quality of sleep (no matter how long you slept).” Scores on this item will be compared using Spearman correlations to global scores on the NRSS. Similarly, Spearman correlations will be used to compare global scores on both scales to ensure that they assess independent domains.

Toronto Hospital Alertness Test (THAT): A 10-item measure to assess the subjective experience of alertness, the THAT evaluates factors that include: ability to concentrate, cognitive capabilities, energy levels, and focus. The questionnaire uses a Likert scale ranging from 0 to 5, with 0 indicating that the symptom or cognitive state is never experienced and 5 meaning that it occurs all the time. A total score is calculated by summing scores on each item. As the entire scale assesses issues such as energy level, alertness, ability to concentrate, and memory, global scores on the THAT will be compared using Spearman correlations to the daytime functioning factor of the NRSS. Global scores will also be compared for the two scales to evaluate divergent validity.

2.3.4 Relationship between Subjective Measures of NRS and Objective Sleep Quality

In this final phase, clinical charts were consulted for all 226 individuals belonging to the patient group. Of those, 215 full sleep reports were collected. For each individual seen at the clinic, overnight PSG consisted of continuous sleep data collection using a standard montage. Recording measures included electroencephalogram (EEG), electrooculogram, submental electromyogram and electrocardiogram. Ancillary channels were used to document respiratory effort (by thoracoabdominal excursions, respiratory inductive plethysmography), oronasal airflow
(by nasal airflow pressure transducer) and oxygen saturation (by pulse oxymetry). The PSG record was analyzed according to standardized scoring techniques.

Outcome measures collected from each PSG study included the following:

- Sleep Onset Latency (SOL), in minutes
- Total Sleep Time (TST), in minutes
- Sleep Efficiency (SE), percentage calculated as TST/time in bed
- Rapid Eye Movement (REM) latency, in minutes
- Percentage of stages 1, 2, 3, 4, and REM sleep
- Percentage of time in bed spent awake (WT)
- Minutes spent awake after sleep onset (WASO)
- Arousal Index (AI)
- Level of alpha EEG sleep (assigned on a scale of 1-5 by scoring technician, with 5 indicating the most severe level of alpha intrusion).

Statistical analysis of objective sleep data involved determining Spearman correlations between global scores on the NRSS and each of the preceding PSG variables (all correlations reported are given as $r$ values). Though PSG represents only a single night of sleep quality, this analysis allowed for the identification of any significant relationships between subjective sleep quality as measured by the NRSS and objective measures of sleep quality.
3 RESULTS

3.1 PARTICIPANT DEMOGRAPHICS

Ages of all participants ranged from 18-85 years. Participants in the patient group had a mean age of 46.7 ± 14.9, while those in the control group had a mean age of 36.9 ± 12.5. A Mann Whitney U test confirmed a significant difference in age between the two groups (p = .001). Of the 226 individuals recruited in the patient group, the data from overnight polysomnography (PSG) were collected for 215. The missing 11 were not collected due to unavailability of charts or because overnight studies were markedly unusual (for example, one patient did not sleep during their study). By referring to information collected during the initial clinical interview, across the overnight PSG, and at the time of the follow-up consultation, patients were placed in groups based on their diagnosis. Where more than one diagnosis was given, group placement was decided based on the primary issue listed following the sleep study as well as the concern chosen to be addressed in treatment at the follow-up appointment.

The largest diagnostic category was the Sleep-Disordered Breathing (SDB) group, which included patients diagnosed with obstructive sleep apnea, central sleep apnea, and primary snoring. Of the patient sample, 112 were placed in the SDB diagnostic category (52.1%), while the three next largest groups included those with primary insomnia (17.7%), Periodic Leg Movement Syndrome (PLMS) (7.9%), and sleep disturbance due to psychiatric illness (11.6%). 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, Post-Traumatic Stress Disorder (PTSD), and other psychiatric illness.
The remaining groups accounted for only 10.7% of the patient sample. These diagnostic categories were circadian rhythm disturbances, parasomnia, narcolepsy, fibromyalgia, and non-specific 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.” Demographic data for all groups are summarized in Table 1.

Table 1 – Demographic characteristics of patient and control groups.

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

SD – Sleep Disturbance

In terms of the patient group, exclusion criteria were minimal. Only those individuals who were unable to complete questionnaires due to language, vision, or literacy conflicts were excluded from the study. Though some may consider this a limitation, the aim was to garner a
sample that would accurately represent patients typically seen in a sleep clinic environment—essentially, those individuals who experience sufficient issues with sleep quality to seek help for their concerns. Due to these limited exclusion criteria, a number of participants in the patient group were receiving medications that could have impacted the outcome of their overnight sleep studies.

In order to describe the pharmaceutical tendencies of the sample, each medication was given one of five designations based on its ability to influence sleep study outcomes: stimulant, sedative, melatonin, non-specific sleep architectural effects (this group included both first- and second-generation antidepressants as well as other medications that list alterations to sleep as a primary side effect), and benign (medications like aspirin for which there is no evidence suggesting sleep-related side effects). Of the patient sample, 25.5% were taking one or more sedative, 4.6% were taking stimulants, 3.2% were taking melatonin, and 28.7% were taking medications that could potentially alter sleep quantity, quality, or architecture. Overlap between groups did occur. A total of 99 sleep study participants (45.8%) were taking one or more of these sleep-altering medications, while 117 (54.2%) were taking either a benign medication or no medication at all.

Along with medication considerations, many participants in the patient group also demonstrated comorbid psychiatric conditions identified both prior to their attendance at the sleep clinic and during. Of the 215 who had sleep studies performed, 26.9% were found to have sleep architectural markers of psychiatric illness. Within this psychiatric group, all had markers of depression while only a few patients were also identified as having markers of post-traumatic stress disorder, panic disorder, and/or obsessive compulsive disorder.
3.2 FACTOR ANALYSIS

A number of questions were eliminated even prior to the data screening phase due to their unsuitability for the Likert format. These included items 22, 26 and 23 from the original pool. After creating a correlation matrix, additional items were excluded based on issues with redundancy (items 17-19, and 32), confusing language (items 20 and 25), and a communality of less than 0.40 with other items (7-9, 30, 31, and 34). One item (21) was skipped by 11 individuals, possibly due to its stigmatizing phrasing (“Do you feel that having unrefreshing sleep has led you to perform poorly at work or school?”). It was also eliminated.

After this phase of analysis, 13 questions were retained for their relevance to the construct of NRS. However, one of these questions was also targeted for removal from the scale proper. In their review of the literature regarding NRS, Stone and colleagues suggested that in order for the label of NRS to be applied, individuals should experience unrefreshing sleep at least three times per week despite receiving sufficient sleep. For this reason, a question that addresses this factor (“How often have you felt really refreshed upon awakening in the morning?”) was removed for use as a screening item. In administrations of the scale, this question could be treated as a preliminary screening tool to determine the presence or absence of traditionally-defined NRS, while the remaining questions provide for a finer degree of assessment of NRS itself. Scores on this item do not contribute to global scores on the NRSS. Thus, the scale proper was reduced to a total of 12 items. Table 2 provides descriptive statistics for the entire population of participants for these 12 items and for the Global Score.

The remaining 12 items were subjected to principal component analysis. Testing revealed a KMO Measure of Sampling Adequacy of 0.88 and a significance of $p < 0.001$ on the Bartlett’s Test of Sphericity, suggesting that partial correlations between variables are large
enough to support a relationship between them and also indicating that use of a factor model is appropriate. According to Kaiser’s recommendations, initially eigenvalues greater than 1 were consulted in order to determine the appropriate number of factors. This approach led to a three-factor solution. However, three factors explained only 66.4% of the total variance—less than the originally determined 70%. In order to manage this discrepancy, the scree plot was consulted, leading to the selection of four factors which explained 72.7% of the total variance (Figure 1). This particular cutoff was deemed appropriate since factors below this point began to explain less than 6% of total variance and corresponded to eigenvalues of less than 0.7.

Once a four-factor solution had been chosen, an oblique rotation of factors was consulted. The rotated factor pattern obtained through the promax rotation is given in Table 3. As observed in the table, the first factor includes items related to quality of sleep and feelings of being restored or refreshed after sleep, so this factor was labeled “refreshment from sleep.” The second factor involves items that query physical symptoms of fatigue or poor sleep such as body pain and frequent illness and it also incorporates questions relating to medical problems and anxiety. Therefore, this factor was labeled “physical/medical symptoms of NRS.” The third factor queries three issues related to daytime function: cognitive abilities like memory and concentration, energy, and alertness. This factor was labeled “daytime functioning.” The final factor relates to feelings of depression or irritability that follow from unrefreshing sleep and so it was labeled “affective symptoms of NRS.”
Table 2 – Descriptive statistics for individual NRSS items.

<table>
<thead>
<tr>
<th>Retained Item</th>
<th>Mean ± S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>* How often have you felt really refreshed upon awakening in the morning?</td>
<td>2.32 ± 1.23</td>
</tr>
<tr>
<td>1. How would you rate the quality of your sleep?</td>
<td>2.50 ± 1.16</td>
</tr>
<tr>
<td>2. Usually, do you think your sleep is restoring or refreshing?</td>
<td>2.49 ± 1.26</td>
</tr>
<tr>
<td>3. Have you felt rested if you’ve slept for your usual amount of time?</td>
<td>3.00 ± 1.32</td>
</tr>
<tr>
<td>4. Have you had physical sensations or unusual feelings in your body that you couldn’t identify?</td>
<td>3.94 ± 1.41</td>
</tr>
<tr>
<td>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?</td>
<td>2.72 ± 1.46</td>
</tr>
<tr>
<td>6. Do you feel that physical or medical problems are dragging you down?</td>
<td>2.61 ± 1.46</td>
</tr>
<tr>
<td>7. Do you ever have a sense of panic, or physical symptoms of panic such as heart racing, for no apparent reason?</td>
<td>3.95 ± 1.31</td>
</tr>
<tr>
<td>8. How is your memory and concentration during the daytime?</td>
<td>3.29 ± 1.13</td>
</tr>
<tr>
<td>9. What is your usual level of daytime energy?</td>
<td>2.95 ± 1.04</td>
</tr>
<tr>
<td>10. Do you usually feel alert during the daytime?</td>
<td>3.08 ± 1.21</td>
</tr>
<tr>
<td>11. Do you feel depressed or down if you didn’t sleep well the night before?</td>
<td>3.17 ± 1.30</td>
</tr>
<tr>
<td>12. How often have you felt irritable or gotten the “blahs” if you didn’t sleep well the night before?</td>
<td>3.07 ± 1.19</td>
</tr>
<tr>
<td>Global Score</td>
<td>36.61 ± 9.54</td>
</tr>
</tbody>
</table>

Mean ± S.D. for patients and controls. Item marked with * is retained as a screening question only and is not included in calculation of a Global Score. Questions 1, 2, 3, 8, 9, 10 are scored positively, with higher scores corresponding to greater alertness, better sleep quality, etc. Items 4, 5, 6, 7, 11, and 12 are reversed when scored.
Figure 1 – Scree plot showing the amount of variance accounted for by each factor. Arrow indicates point chosen for cutoff.
### Table 3 – Standardized regression coefficients of the rotated factor pattern.

<table>
<thead>
<tr>
<th>Retained Items</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How would you rate the quality of your sleep?</td>
<td>86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Usually, do you think your sleep is restoring or refreshing?</td>
<td>83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Have you felt rested if you’ve slept for your usual amount of time?</td>
<td>78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Have you had physical sensations or unusual feelings in your body that you couldn’t identify?</td>
<td>87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>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?</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Do you feel that physical or medical problems are dragging you down?</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Do you ever have a sense of panic, or physical symptoms of panic such as heart racing, for no apparent reason?</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. How is your memory and concentration during the daytime?</td>
<td></td>
<td>87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. What is your usual level of daytime energy?</td>
<td></td>
<td></td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>10. Do you usually feel alert during the daytime?</td>
<td></td>
<td></td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>11. Do you feel depressed or down if you didn’t sleep well the night before?</td>
<td></td>
<td></td>
<td></td>
<td>94</td>
</tr>
<tr>
<td>12. How often have you felt irritable or gotten the “blahs” if you didn’t sleep well the night before?</td>
<td></td>
<td></td>
<td></td>
<td>64</td>
</tr>
</tbody>
</table>

Standardized coefficients have been rounded and multiplied by 100. Values less than 40 have been suppressed for clarity.
3.3 RELIABILITY

3.3.1 Internal Reliability

Internal consistency for the scale was found to be satisfactory, demonstrating a Cronbach’s $\alpha$ statistic of 0.88 for the total scale. Corrected item-total correlations ranged from 0.65 to 0.81 for Factor 1, from 0.44 to 0.60 for Factor 2, from 0.66 to 0.80 for Factor 3, and were 0.48 for Factor 4. With the exception of item 11, item-total correlations for the global scale ranged from 0.40 to 0.77 (summarized in Table 4). While item 11 did demonstrate a lower correlation of 0.28, its item-total correlation within the “affective symptoms” factor was 0.47, supporting its retention. These results demonstrate reasonably strong support for the suggestion that items on the same factors—and on the global scale itself—were measuring the same construct.

3.3.2 Test-Retest Reliability

For 43 participants in the patient group, the NRSS was administered a second time (within at least two weeks of the first administration). Factor and global scores for initial and second administrations were compared using Spearman correlations to determine the test-retest reliability of the scale. “Test” and “retest” scores were significantly associated from Factors 1, 2, and 3 with correlations of 0.76, 0.77, and 0.69 respectively ($p < 0.01$). Scores on Factor 4 were only moderately associated with a correlation of 0.27. In terms of the overall scale, global scores between administrations were also significantly associated, showing a correlation of 0.72 ($p < 0.01$). In general, these results suggest that scores on the NRSS remain relatively stable over time.
Table 4 - Corrected item-total correlations for each item within its factor and in global scale.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Correlation within Factor</th>
<th>Correlation with Global Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>0.71</td>
<td>0.67</td>
</tr>
<tr>
<td>Item 2</td>
<td>0.81</td>
<td>0.77</td>
</tr>
<tr>
<td>Item 3</td>
<td>0.65</td>
<td>0.57</td>
</tr>
<tr>
<td>Factor 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 4</td>
<td>0.53</td>
<td>0.41</td>
</tr>
<tr>
<td>Item 5</td>
<td>0.56</td>
<td>0.51</td>
</tr>
<tr>
<td>Item 6</td>
<td>0.60</td>
<td>0.66</td>
</tr>
<tr>
<td>Item 7</td>
<td>0.44</td>
<td>0.49</td>
</tr>
<tr>
<td>Factor 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 8</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>Item 9</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>Item 10</td>
<td>0.80</td>
<td>0.67</td>
</tr>
<tr>
<td>Factor 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 11</td>
<td>0.48</td>
<td>0.28</td>
</tr>
<tr>
<td>Item 12</td>
<td>0.48</td>
<td>0.61</td>
</tr>
</tbody>
</table>

3.4 SENSITIVITY AND SPECIFICITY

In this phase, the goal was to determine if Global Scores on the NRSS could predict whether participants belonged to the patient or control group. The area under the ROC curve was 0.89 (p < 0.01), suggesting that Global Score is a good indicator of patient status. It is possible that the significant difference in age between the patient and control groups could have some bearing on these findings. However, a Spearman correlation revealed no significant association between Global Score and age (r = 0.04, p = 0.56), suggesting that the difference likely has little influence on the outcome of the ROC curve.
A similar ROC curve was also calculated based on the selected screening item: “How often have you felt really refreshed upon awakening in the morning?” Participants who responded with “2-3 days per week” (a score of three on a five-point scale) or less were placed in the NRS groups. This was deemed to be a more accurate gauge of NRS, given that it does not depend solely on patient status. Though a Chi-square test revealed that sleep clinic patients were more often placed in the NRS group (Pearson Chi-square statistic = 59.4, p < 0.01), referral to a sleep clinic should be distinguished from the construct of NRS. In this case, the area under the ROC curve was 0.90 (p < 0.01) (Figure 2). For this second ROC curve, potential values for sensitivity and specificity were generated based on several potential cutoff scores. These are summarized in Table 5. In terms of maximizing both sensitivity and specificity, a score of 43 or less should be considered indicative of NRS, giving a sensitivity of 84% and a specificity of 85%.

Table 5 – Sensitivity and specificity for NRS based on specific cutoff scores.

<table>
<thead>
<tr>
<th>Cutoff Score</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>0.81</td>
<td>0.89</td>
</tr>
<tr>
<td>43</td>
<td>0.84</td>
<td>0.85</td>
</tr>
<tr>
<td>44</td>
<td>0.85</td>
<td>0.80</td>
</tr>
<tr>
<td>45</td>
<td>0.87</td>
<td>0.79</td>
</tr>
<tr>
<td>46</td>
<td>0.91</td>
<td>0.75</td>
</tr>
<tr>
<td>47</td>
<td>0.94</td>
<td>0.69</td>
</tr>
<tr>
<td>48</td>
<td>0.95</td>
<td>0.67</td>
</tr>
</tbody>
</table>

A score that is less than or equal to the listed cutoff score indicates NRS.
Figure 2 – ROC curve plotting sensitivity vs. specificity when using NRS screening question to predict NRSS global scores. Diagonal line included for reference. Arrow shows the values for a cutoff point of 43.

3.5 QUESTIONNAIRE VALIDITY

In order to determine the construct validity of the NRSS, each factor was examined independently for its relationship to factors that evaluate similar constructs. In terms of the questionnaire’s discriminant validity, global NRSS scores were compared to global scores on other scales.
3.5.1  

*Factor 1: Refreshment from Sleep*

Scores on Factor 1 were compared to two questionnaire items that had previously been identified for their relevance to NRS, item 9 on the PSQI (the sleep quality factor) and item 5 on the AIS. Factor 1 scores were significantly associated with scores on both items, showing a correlation of 0.69 for the PSQI item and a correlation of 0.66 for the AIS item. Similarly, scores on the PSQI and AIS items showed a significant association with each other and had a correlation of 0.72—a value very close to those observed for the NRSS. Finally, scores on this factor were also significantly correlated with the singly screening item, with a value of 0.76.

3.5.2  

*Factor 2: Physical/Medical Symptoms Associated with NRS*

Factor 2 scores were assessed based on their relationship to item 6 of the FSS, which refers to the tendency of fatigue to interfere with sustained physical functioning. A significant association was found, with a correlation of 0.46 between the two items.

3.5.3  

*Factor 3: Daytime Functioning*

THAT global scores and factor scores on the PSQI daytime functioning domain were used to assess the construct validity of Factor 3. Significant associations were found in both cases. THAT global scores showed a correlation of 0.68 with Factor 3. The correlation with the PSQI daytime functioning domain was significant at 0.48.

3.5.4  

*Factor 4: Affective Symptoms of NRS*

Scores on the affective symptoms factor were compared to global scores on the CES-D. A significant association was found with a correlation of 0.66, suggesting a fairly strong relationship between the NRSS affective factor and established measures of depression.
3.5.5 *Divergent Validity of the NRSS*

To determine if the NRSS measures a construct that is discrete from those assessed by already existing questionnaires, global scores on the scale were compared to global scores on the PSQI, AIS, FSS, THAT, and CES-D. Correlations were all significant and ranged from 0.57 for the PSQI to 0.74 for the CES-D. Values are summarized in Table 6.

**Table 6** – Correlations between global scores on the NRSS and global scores on other subjective measures.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Spearman Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSQI</td>
<td>0.57</td>
</tr>
<tr>
<td>AIS</td>
<td>0.68</td>
</tr>
<tr>
<td>FSS</td>
<td>0.70</td>
</tr>
<tr>
<td>CES-D</td>
<td>0.74</td>
</tr>
<tr>
<td>THAT</td>
<td>0.62</td>
</tr>
</tbody>
</table>

All correlations significant at p < .001

3.5.6 *Relationship between Subjective Measures of NRS and PSG Results*

In order to determine if any gender or age differences existed between individuals with and without NRS, a cut off of 43 on the NRSS was chosen to divide questionnaire respondents into two groups for the purposes of evaluation. A chi-square test indicated that individuals in the NRS group were more frequently female ($\chi^2 = 4.27, p = 0.04$), despite there being no significant difference in distribution of the two genders between the control and patient groups.

A Wilcoxon Signed Rank Test found no significant difference in mean age between the NRS and non-NRS group. However, when it was considered that the difference in age between the patient and control group may have some influence on these findings, mean ages were
compared just for the patient group. In this case, a significant difference was found ($p = 0.009$). Individuals placed in the non-NRS group based on the cutoff of 43 had a mean age of $50.5 \pm 13.5$ years, while individuals placed in the NRS group appeared to be only slightly younger with a mean age of $49.9 \pm 14.4$. PSG variables for the two groups are reported in Table 7.

Table 7 – Mean and standard deviations for PSG variables reported for the NRS and non-NRS groups.

<table>
<thead>
<tr>
<th></th>
<th>NRS Group</th>
<th>Non-NRS Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep Onset Latency</td>
<td>$34.5 \pm 51.6$</td>
<td>$31.3 \pm 32.2$</td>
</tr>
<tr>
<td>Total Sleep Time</td>
<td>$369.1 \pm 82.2$</td>
<td>$363.2 \pm 75.1$</td>
</tr>
<tr>
<td>Sleep Efficiency</td>
<td>$80.6 \pm 15.4$</td>
<td>$78.8 \pm 14.7$</td>
</tr>
<tr>
<td>Rapid Eye Movement Latency</td>
<td>$135.6 \pm 87.3$</td>
<td>$112.8 \pm 64.8$</td>
</tr>
<tr>
<td>Stage 1, %</td>
<td>$6.9 \pm 3.9$</td>
<td>$7.3 \pm 4.9$</td>
</tr>
<tr>
<td>Stage 2, %</td>
<td>$47.8 \pm 12.7$</td>
<td>$46.9 \pm 14.1$</td>
</tr>
<tr>
<td>Stage 3, %</td>
<td>$5.8 \pm 3.8$</td>
<td>$5.1 \pm 3.2$</td>
</tr>
<tr>
<td>Stage 4, %</td>
<td>$8.9 \pm 6.3$</td>
<td>$8.8 \pm 7.1$</td>
</tr>
<tr>
<td>Wake Time, %</td>
<td>$13.7 \pm 13.6$</td>
<td>$15.4 \pm 11.9$</td>
</tr>
<tr>
<td>Arousal Index</td>
<td>$18.8 \pm 13.8$</td>
<td>$22.4 \pm 17.9$</td>
</tr>
<tr>
<td>Level of Alpha EEG Sleep</td>
<td>$2.0 \pm 0.8$</td>
<td>$1.8 \pm 0.7$</td>
</tr>
</tbody>
</table>

PSG variables reported for NRS and non-NRS groups. Group placement determined based on cut-off score of 43 on the NRSS.

Among those in the patient group found to have insomnia as their primary diagnosis, 62.5% were placed in the NRS group, though the difference in distributions was not significant ($p = 0.092$). The sole significant difference in distributions was found among individuals who received a primary diagnosis of sleep disturbance due to psychiatric illness. Among these individuals, 91.7% reported NRS ($p = 0.04$). Similarly, among all participants in the patient
group who were found to exhibit PSG markers of depression (regardless of primary diagnosis), 88.2% were placed in the NRS group based on their cutoff scores ($p = 0.008$). Finally, 100% of patients in the fibromyalgia group reported NRS, though with only four patients in the group this observation is made very tentatively.

In terms of the relationships between scores on the NRSS and PSG variables, a few very small but significant correlations emerged when controlling for participant age and total sleep time. Scores on the refreshment factor of the scale were significantly associated with levels of alpha EEG sleep (scored on a five-point scale by the sleep technologist), demonstrating a correlation of $r = -0.16$ ($p = 0.027$). Scores on the daytime functioning factor and the affective symptoms factor were also significantly associated with alpha EEG level, with correlations of $r = -0.15$ ($p = 0.042$) and $r = -0.16$ ($p = 0.033$) respectively. Finally, global scores were significantly associated with alpha EEG level, showing a correlation of $r = -0.16$ ($p = 0.026$).

Associations also emerged between several other variables and the affective symptoms domain. This factor was significant correlated with sleep efficiency ($r = -0.15$, $p = 0.042$), REM latency ($r = -0.15$, $p = 0.047$), wake after sleep onset ($r = 0.17$, $p = 0.025$), and percentage of stage 2 sleep ($r = -0.19$, $p = 0.010$). In contrast, global scores on the CES-D were 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 correlated with both level of alpha EEG sleep ($r = -0.17$, $p = 0.042$) and percentage of stage 4 sleep ($r = 0.16$, $p = 0.045$). No other associations were found between subjective sleep measures and objective PSG.
4 DISCUSSION

4.1 FACTOR ANALYSIS, RELIABILITY, SENSITIVITY AND SPECIFICITY

The finalized scale was reduced to 12 items (with an optional 13th item that could be used for screening but is not included in global score calculation), representing four factors associated with NRS. With a potential maximum score of 60 and a minimum score of 12, the range of potential scores is relatively manageable. The four factors assessed by the scale—refreshment from sleep, physical/medical symptoms of NRS, daytime functioning, and affective symptoms of NRS—adhered closely to predictions made at the outset of the study.

The approach taken to factor solution differed slightly from that recommended by Kaiser, which involves selection of factors with eigenvalues above 1. Using Kaiser’s method led to a three-factor solution that explained less than 70% of the total variance. However, selecting eigenvalues above 1 remains mostly an arbitrary recommendation. In order to improve the total variance explained, the scree plot was consulted and a four-factor solution deemed more appropriate. This factor solution was satisfactory for a number of reasons. First, each retained factor loaded with at least two items (and in most cases, three). Second, the items that loaded on the same factors appeared to share some conceptual meaning, while those that loaded on separate factors differed in the content they assessed. Finally, with the rotated factor structure, most items demonstrated high pattern loadings for a particular factor (> 0.40) and low loadings on all others.

The four factors selected address several of the most salient issues for NRS questionnaire development identified by Vernon and colleagues in their recent review. Namely, they provide not only a mechanism for determining the presence/absence of NRS, but also a tool for
evaluating its concomitant symptoms. Such an instrument will enable researchers and clinicians both to screen for NRS and to assess the severity of symptoms and their response to treatment.

In terms of the reliability or stability of the scale, the NRSS demonstrated a high internal consistency with a Cronbach $\alpha$ of 0.88. This value is high enough to suggest that the overall scale assesses a single related construct, but it is not so high that it raises concerns about question redundancy (e.g. $>0.90$). Similarly, item-total correlations for separate factors were satisfactory. Correlations on factors were as high as 0.81 for an item belonging to the refreshment from sleep factor and as low as 0.44. This lower correlation was found for an item relating to physical feelings of anxiety on the physical/medical symptoms factor, perhaps explaining its lower degree of association with its factor.

Item-total correlations were also satisfactory for the overall scale, in most cases ranging from 0.40 to 0.77. One item—“Do you feel depressed or down if you didn’t sleep well the night before?”—shared a lower correlation of 0.28 with the global scale. However, this item was not targeted for removal for two reasons: 1) it demonstrated an acceptable item-total correlation within its factor, and 2) it represents one of only two questions in the affective symptoms domain, and this factor was deemed too important to be deleted. In many ways, this domain was expected to have the most complicated relationship with the overall scale in that the relationship between mood and sleep quality appears to be bi-directional and has been somewhat difficult to characterize. However, self-reported sleep quality is significantly related to depression and low mood and, while a smaller association with the global scale may not be ideal, some assessment of affective symptoms remains important, particularly considering the strong association observed between NRS and psychiatric disorders.
Test-retest reliability for the scale was also satisfactory, with a global correlation of 0.72 and correlations on Factors 1 through 3 ranging from 0.69 to 0.77. However, scores on the affective symptoms factor demonstrated a test-retest correlation of only 0.27. One potential explanation for this is the timing of questionnaire administration. Participants, for the most part, completed their first questionnaire while waiting to see the physician for their initial consultation appointment. At this point, they would have experienced limited contact with the physician and would not have been advised of typical outcomes for sleep studies or been made aware of potential treatment approaches. The second questionnaire administration typically took place on the evening of their sleep study. Perhaps, along with a potentially reassuring conversation with the physician, the sense of taking action regarding their sleep issues may have altered their perception of affective issues. In the future, testing that takes place within the same timeline in the treatment process may be warranted.

Scale sensitivity and specificity was also found to be fairly good. Global scores were found to be predictive of patient vs. control participant status but, perhaps more importantly, global scores were also highly predictive of response on the NRS screening item. Though not a gold standard, this screening item comes as close as possible in terms of the available literature regarding NRS and was reasonably similar to items used to assess the presence/absence of NRS in previous studies. For example, in their recent study of NRS, Roth and colleagues classified individuals as having NRS when they reported a complaint of unrefreshing or unrestoring sleep three or more times per week coupled with significant distress or disturbance to daytime functioning. As the sample population for this study was drawn from a sleep clinic, significant distress or disturbance to quality of life can essentially be assumed since these are individuals who have sought treatment for their sleep issues.
Using this second ROC curve, a score of 43 or lower was chosen to indicate NRS. This value appeared to maximize both sensitivity and specificity, giving values of 84% and 85% respectively. However, the scale’s optimal cutoff score will necessarily differ depending on the context of its use. For screening purposes, a higher sensitivity may be desirable in order to detect all cases of NRS, even at the expense of a lower specificity/higher rate of false positives. In this case, a score of 46 or lower may be preferable as it provides a sensitivity of greater than 90%.

### 4.2 SCALE VALIDITY

Validation of a scale is an ongoing process, particularly when the scale has been designed for use in such a diverse population of respondents (patients with fibromyalgia, insomnia, etc.). However, preliminary results from this study demonstrate good support for the scale as a valid measure of the construct of NRS. First, it is the only scale of its kind for which quantitative data has been collected in the form of patient focus groups. Second, each factor was significantly correlated with items or scales that were previously determined to be valid and reliable measures of those domains. The refreshment from sleep factor was significantly associated with two items (one from the PSQI and one from the AIS)
that had been identified as part of a review on instruments to measure NRS,\textsuperscript{13} with correlations of 0.69 and 0.66 respectively. While these values appear somewhat low, this may actually be a point in favour of the NRSS. The two items from these other scales involve ranking one’s sleep quality: “During the past month, how would you rate your sleep quality overall?” and “Overall quality of sleep (no matter how long you slept)”. However, they refer only generally to the idea of “sleep quality” and do not use the words “refreshed” or “restored,” calling into question their value in the assessment of a construct like NRS. In contrast, the correlation between scores on this factor and scores on the NRS screening item were higher, with a value of 0.76. Considering that this item is much closer to being a “gold standard” method of determining the presence or absence of NRS, this higher correlation may be a better indicator of the factor’s validity.

Factor 2, the domain relating to physical/medical symptoms of NRS was assessed using a single item from the FSS, “My fatigue prevents sustained physical functioning.” Scores on Factor 2 were significantly associated with this item, with a correlation of 0.46. While this value is somewhat low, it is not unexpected considering the wide range of issues queried within Factor 2 and the limited scope of the item to which it is being compared. An instrument that might capture the construct being assessed in this factor more accurately was not administered, and further evaluations of its construct validity are not possible at this time. In the future, questionnaires that more specifically query issues relating to physical and medical quality of life would be helpful to clarify the validity of this factor.

Validity of the daytime functioning factor was assessed through comparisons to the daytime functioning factor on the PSQI and to global scores on the THAT. Significant correlations were found in both cases, though the correlation in the case of the PSQI was relatively low at 0.48. This was likely due to the different definition of “daytime functioning”
used for the NRSS compared to the PSQI. The PSQI daytime functioning domain incorporates two items: one that queries difficulties staying awake during social situations and when driving, and a second that refers to “enthusiasm to get things done.” These issues are relevant to the daytime functioning questions of the NRSS, but are not as clearly related as those addressed by the THAT which investigates constructs like concentration, memory, energy, and alertness throughout the day.

Finally, regardless of its lower component-total correlations with the global NRSS, the fourth factor relating to affective symptoms of NRSS demonstrated a significant correlation with scores on the CES-D, a previously validated depression scale. However, while the correlation was reasonably strong at 0.66, it was not so high as to suggest that the two measure equivalent constructs. The intention with those NRSS items created to assess affective symptoms of NRS was not simply to distinguish participants who are depressed. Rather, these items were meant to specifically query mood changes that, in the respondent’s experience, tended occur as a result of poor sleep quality. While it may be difficult for individuals completing the questionnaire to represent their experiences with such a complex bi-directional relationship in the form of one or two Likert-type items, the hope is that the wording of the question will bring respondents to reflect a bit more closely on whether the two issues are related for them, rather than assessing the presence/absence of depression.

To conclude, the scale’s divergent validity was also assessed—namely, whether it measures constructs that are separate from other scales commonly used to evaluate sleep, fatigue, alertness, and depression. No correlations above 0.74 were observed between global scores on the NRSS and scores on these other scales, suggesting that the issue being evaluated is distinct from the issues being measured with these other instruments. Future evaluations of both the
reliability and the validity of the NRSS will need to be conducted within patient groups of interest. However, these preliminary results suggest that the scale offers good psychometric properties for the assessment of NRS in a sleep clinic population.

4.3 DEMOGRAPHIC CORRELATES OF NRS

As anticipated based on the NRS literature to date, significantly more females than males were placed in the NRS group based on the chosen cutoff of 43 on the NRSS. However, the difference was not as striking as those observed in some previous investigations of NRS, barely reaching the level of significance with a p-value of 0.04. For example, one study found that 61% of participants who reported NRS were female,\textsuperscript{23} while another found that proportion to be as high as 75%.\textsuperscript{22} For the most part, the difference appears to result from the populations being surveyed, as well as the criteria for NRS. For both of the previously mentioned studies, the sample group was drawn from a clinical population of individuals with complaints of insomnia—a patient group well recognized as being more heavily dominated by females.\textsuperscript{21} The patient groups in both these studies were already predominantly female (approximately 65% of the sample population). In contrast, the sample for the current study was culled from a group of sleep clinic patients with relatively equal proportions of males and females.

Similarly, in order to be considered part of the NRS group for one of these studies, participants were required to demonstrate no other comorbid conditions, meaning they experienced NRS alone.\textsuperscript{23} As mentioned, both studies focused primarily on patients with insomnia. In the current study, the majority of participants had some comorbid sleep or psychiatric condition beyond nonspecific “poor quality sleep.” It is possible that the gender
difference disappears when participants are drawn from a more generalized population, one that is not specific to a single sleep disorder like insomnia. This appears to be supported by the less significant gender differences seen in population-based studies, some of which approach a 1:1 gender ratio.\textsuperscript{18, 24} A slight age difference was also found between the NRS and non-NRS group when patients alone were examined. The NRS patients demonstrated a significantly lower mean age, though the difference was minimal. This may lend some support to previous studies that have found a similar relationship. Ohayon and colleagues have demonstrated repeatedly in large studies examining the general population that individuals aged 25 to 44 years tend to report higher rates of NRS, while rates drop off above the age of 55 years.\textsuperscript{8, 11, 16, 24} The relatively insubstantial difference observed in the current study could be a product of this same observation. However, the sleep clinic population tends to skew older due to the high prevalence rates of disorders like obstructive sleep apnea in aging adults.\textsuperscript{83} Thus, the mean age of the patient population in this study was relatively elevated from the outset and fewer participants within these younger age ranges were available for inclusion. Additional studies that include more participants from younger age groups are necessary to determine if the trend still holds for a sleep clinic patient population.

In terms of the presence of NRS within different sleep disorder diagnosis groups, very few significant relationships were found. While those with insomnia did demonstrate slightly higher rates of NRS than those without, the difference was not significant, suggesting that the construct assessed by the NRSS is distinct from insomnia. In support of this was the relatively low correlation of 0.68 found between scores on the AIS and scores on the NRSS. This was slightly unexpected as previous research has indicated a strong association between difficulties
initiating and maintaining sleep and self-reported NRS.\textsuperscript{18,23} It is possible that a larger sample of insomnia patients may lead to the observation of an association, as the group in this study was relatively small in terms of finding significance. However, one recent study by Roth has confirmed that NRS does occur as an isolated condition distinct from other sleep disorders,\textsuperscript{22} suggesting that studies indicating an independence between the two conditions are not without precedence.

The one patient group that did exhibit significantly greater rates of NRS was the sleep disturbance due to psychiatric illness group. This group consisted of individuals who received no clear primary diagnosis to explain their complaints regarding sleep quality, but who did demonstrate PSG markers of psychiatric disorders including depression and PTSD. Similarly, those individuals who possessed sleep markers of depression—regardless of their primary diagnosis—were significantly more likely to report NRS. According to the 2005 practice parameters laid out by the American Academy of Sleep Medicine, PSG is not routinely indicated in the diagnosis of depression or other psychiatric illness as no characteristics of sleep architecture have been confirmed as specific for the diagnosis of depression.\textsuperscript{84} However, since the publication of these guidelines, the evidence in support of biomarkers of depression has grown and there are several PSG markers of both depression and PTSD that reoccur throughout the literature. These markers are monitored in the laboratory setting in the event that they may provide additional clues in the process of diagnosis.

In the case of depression, potential biomarkers include impaired continuity of sleep, increased stage 1 sleep, decreased slow wave sleep, and increased latency to the first episode of REM sleep.\textsuperscript{85} Abnormalities in the distribution of different sleep stages across the night have also been reported. The markers for PTSD are similar and include reduced sleep efficiency and
decreased slow wave sleep, though these findings have not been confirmed in all cases. In support of the value of these markers, post-hoc testing revealed a significant difference in mean CES-D scores between individuals with and without these biomarkers (p < 0.01). Additionally, the vast majority of patients demonstrating these sleep features of psychiatric illness also report NRS according to the selected cutoff.

While the aforementioned biomarkers were valuable predictors of self-report depression levels, they were also useful in anticipating those patients who are likely to report poor sleep quality on the NRSS. Considering that many of these markers relate to disruptions of sleep continuity and reductions in slow-wave sleep, these findings may lend further support to those theories that suggest a relationship between NRS and decreased amounts of deep sleep. However, they also represent a complicating factor to the labeling of NRS. It is possible that all of these potential physiological markers of NRS are merely associated with depression, which leads subsequently to increasingly negative reports of sleep quality. This issue of patient perception is difficult to separate from any research involving self-report measures, and makes it difficult to infer cause-and-effect relationships from results.

Léger and colleagues have found that the strongest predicting factor for severe daytime impairment from NRS was the perception of sleep as being of poor quality. NRS may itself be self-perpetuating with low affect predicting poor sleep quality rather than the inverse. The relatively strong correlation of 0.74 found between global scores on the CES-D and the NRSS does not help to clarify the issue. Only future research that places emphasis on identifying the sleep biomarkers of individuals in the absence of any psychiatric condition will allow for a better understanding of the bi-directional relationship between negative mood and reports of poor sleep quality.
Finally, the finding that 100% of patients belonging to the fibromyalgia group presented with scores indicating NRS indicates an interesting avenue for further investigation with the NRSS. As there were only four patients total in this group, a much larger sample population would need to be evaluated before any conclusions could be drawn. However, it does provide tentative support for the value of future questionnaire validation in this patient group.

4.4 PSG SLEEP STUDIES

All correlations between PSG variables and subjective measures were calculated by controlling for age and total sleep time as both are likely to have significant effects on sleep architecture findings. The few significant correlations observed tended to be small, but provided some tentative support for at least one existing theory regarding NRS. Namely, scores on three of the NRSS scale factors (refreshment from sleep, daytime functioning, and affective symptoms) as well as global scores were significantly associated with levels of alpha EEG sleep, which was scored on a five-point scale based on severity by the sleep technologist. Alpha EEG was also significantly correlated with scores on the THAT and scores on the CES-D. In all cases, correlations were very small, with the highest being 0.17. Such a small association suggests that any relationship found is likely to be mediated by a number of factors. Further, it provides additional support for the value of an instrument like the NRSS. With such small correlations found between objective and subjective measures of sleep, the utility of a tool that assesses NRS is increased. Still, these results do provide some evidence for a link between self-reported NRS and levels of alpha EEG sleep. One study has indicated previously that individuals with higher levels of phasic alpha EEG intrusions report lower levels of sleep.
quality, but these findings have not been replicated in all cases. This study provides additional support for the value of further investigations into the effects of alpha EEG on reported sleep quality.

While no other significant correlations were found for the global scale, a number of associations were observed for the affective symptoms factor. This factor was significantly associated with reduced sleep efficiency, reduced REM latency, increased WASO and reduced percentage of stage 2 sleep. Many of these changes are similar to EEG biomarkers of depression, which include disruptions to sleep continuity and alterations in the timing of the first REM episode. However, this factor also appeared to be the least stable factor across different administrations, suggesting that the relationships observed may not be quite as relevant in the long term as those seen for other factors.

Though slow wave sleep has potentially been implicated in the subjective experience of NRS, no relationship was observed between the NRSS and percentages of deep sleep. However, global scores on the THAT were significantly associated with percentage of stage 4 sleep, suggesting that subjective alertness may be somewhat linked to amount of slow wave sleep. Similarly, it was interesting to observe that no correlations appeared to exist between scores on the PSQI and objective sleep data. The PSQI has been employed extensively as a measure of subjective sleep quality and, while it functions well in this capacity and provides a good standardized tool for this purpose, no consistent relationship between results on the PSQI and objective sleep variables has been observed. Thus, the validation of a self-report measure that does share some link to potential biomarkers of NRS would be ideal. While the NRSS is far from being used in this capacity, tentative evidence linking it to several physiological variables provides support for further investigations of its potential in this regard.
4.5 LIMITATIONS

As scale design and validation is an ongoing process, many of the issues with the current study are likely to be addressed in future research. However, one of the biggest problems with the current study is the population used. Ideally, a validation of a scale to assess NRS would focus on populations that have previously been determined to experience the symptom at much higher rates. No previous studies have investigated the prevalence of NRS in a general sleep clinic population, meaning that the proportion of identified cases cannot be compared to previous research for confirmation. For this reason, an investigation of the scale in a greater number of fibromyalgia and chronic fatigue patients would be helpful.

Similarly, by evaluating the scale’s validity in a population with disordered sleep, items on the scale did not demonstrate normal distributions. In terms of scale validation, this significantly limits one’s ability to make generalizations about the scale’s use in a general population, meaning that it is currently only valid in a very specific subsection of the population—those seeking treatment for sleep issues. In the future, evaluation in a population that was not experiencing such significant comorbidities might also allow for the identification of a “pure” NRS group, which potentially could have provided better clues to the scale’s sensitivity. However, with only about 1.4% of insomnia sufferers experiencing NRS in the absence of other symptoms,18 such an evaluation would likely need to be done on a very large scale to be successful.

The population examined may also have been problematic due the interaction of medications that were being used over the course of the evaluation. Unfortunately, due to the
wide range of comorbidities and large array of medications employed, an analysis that took into account potential drug interactions was not an option. By dividing the sample into a greater number of groups, achieving statistical significance with the results would become much more difficult. Potentially, these differences could effectively be evaluated in future studies where more participants are recruited. However, in the current study this was not possible.

One of the greatest statistical limitations of the study involves the lack of corrections for multiple statistical tests. For this reason, many of the significant correlations observed could potentially be the result of Type 1 alpha error. This could be particularly problematic for correlations found between PSG variables and scores on the scale, as many of these were low (less than 0.17), with p-values that were very close to 0.05. However, as NRS is a highly multi-faceted issue, low correlations are likely to be anticipated due to a number of other intervening variables. Thus, these results were still considered to be worth discussing, particularly considering the lack of published information regarding PSG variables in individuals with NRS.

An additional limitation involves the determination of the sensitivity and specificity of the scale. ROC curves tend to be applied primarily to gold standard measures. That the presence or absence of NRS was determined using a single screening item that had not previously been used in the literature on NRS could be problematic. It may have been more valuable to have employed a screening item consistent with questions used in previous studies, at the minimum ensuring some degree of standardization. If future validations are performed, applying both the current screening item and the four-item screening measure previously developed by Roth and colleagues would perhaps be more informative.9 Similarly, future validation studies should include measures that have been designed to assess physical and medical symptoms more
effectively. With only a single item relating somewhat tenuously to this construct included in the questionnaire battery, evaluation of this domain’s validity was limited.

The final issue with the current study relates to the assessment of alpha EEG level. All overnight PSG studies were scored by highly skilled sleep technicians with many years’ experience in the field. However, detailed evaluations of sleep microarchitecture are not routinely applied to overnight studies as they can be time-consuming and their diagnostic value remains largely unconfirmed. Thus, the ratings applied to alpha EEG do not reflect some of the fine-grade distinctions between alpha subtypes that have previously been reported. The difference between tonic and phasic alpha activity is highly relevant to the issue of NRS and, without having made that distinction in reporting levels for the current study, it is unclear if the higher levels of severity reported here necessarily correlate with the phasic activity that has been shown to dominate the sleep of NRS reporters in previous research.

4.6 RESEARCH AGENDA

As discussed, additional validation studies with this questionnaire are necessary in order to confirm its potential for use in populations beyond those found at a sleep clinic. Some tentative evidence from this study suggests that the NRSS may be useful in a fibromyalgia patient population. Based on this, studies that investigate its use in chronic fatigue syndrome would also be indicated due to the high degree of symptom overlap between the two disorders. Such an investigation would also likely help with additional validation of the physical/medical symptoms factors as both FS and CFS patients represent populations that often present with
issues such as body pain and fatigue resulting from medical disorders—two problems queried by the domain.

Once the utility of the NRSS has been further evaluated, it will allow for a wide range of new approaches to NRS. First, there are a number of physical and medical conditions that would benefit from an evaluation of their relationship to the construct of NRS as it has been defined in the field of sleep research. While fibromyalgia and chronic fatigue syndrome both appear to be characterized by reports of NRS, at this juncture there have been no studies incorporating any of the specific recommendations regarding the label of NRS that were proposed by Stone and colleagues.\(^1\) Therefore, little is understood regarding how the presentation of NRS in these disorders might relate to NRS as it appears in patients with conditions like insomnia—a population that has received more attention using these recent guidelines. By asking a variety of questions that fit into several different factors, the NRSS allows researchers to collect finer degree of data with which to distinguish the presentation of these different disorders. For example, it is possible that patients with NRS associated with insomnia tend to experience fewer complaints regarding daytime functioning. With the NRSS, such an exploration would be possible.

Similarly, the NRSS allows for the possibility of further investigating the role of psychiatric illness in the experience of NRS. A number of powerful links appear to exist between both self-reported NRS and symptoms of depression. Obviously a cause-and-effect relationship is quite difficult to interpret when it comes to these factors. However, a finer degree of understanding between their connections could be achieved by, for example, comparing results on different subscales of the CES-D or Hamilton Depression Rating Scale to scores on the
different factors of the NRSS. Perhaps scores on one factor are particularly predictive of depressive symptoms.

Similarly, attempts to discover potential biomarkers of NRS have been previously limited due to a lack of standardization in the definition and measurement of the construct. Such a lack of standardization prevents effective communication between different disciplines and makes it difficult to draw conclusions across studies performed by separate research groups. In the future, studies evaluating the construct of NRS could, with the use of the NRSS, evaluate changes in the presentation and symptoms of NRS over the course of different interventions. These changes could also be compared to PSG markers to determine if any objective physiological variables emerge. Use of the questionnaire could be incorporated into finer grade investigations of sleep microarchitecture, allowing for a better understanding of its contributions to subjective, self-report measures.

5 CONCLUSION

The current study provides ample support for the value of the NRSS as a tool for the assessment of nonrestorative sleep in a sleep clinic population. The scale demonstrates good reliability, both in terms of its internal item correlations as well as for repeated administrations. It also correlates well with other previously validated measures to assess similar constructs, though further studies may be necessary in order to confirm the validity of the physical/medical symptoms factor. Finally, evidence from this study provides tentative support for a few
physiological biomarkers of NRS, particularly alpha EEG sleep. However, correlations were very small, suggesting that future analyses of this relationship will need to take into account the interaction of a number of factors.

In summary, the NRSS provides a new standardized tool for the assessment of NRS. While there is currently very little literature examining NRS as its own unique symptom of sleep, medical, and psychiatric disorders, the scale was designed using the recommendations set forward in a recent review of the topic of NRS. In terms of future research, such a tool will allow for comparisons to be made between different approaches to NRS in different clinical populations, allowing researchers to begin to discriminate between these disorders and hopefully to better define the label of NRS.
REFERENCES


APPENDIX

NRSS QUESTIONNAIRE

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?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very poor</td>
<td>Very good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Usually, do you think your sleep is restoring or refreshing?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>Always</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Have you felt rested if you’ve slept for your usual amount of time?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>Absolutely</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Have you had physical sensations or unusual feelings in your body that you couldn’t identify?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>Yes, all the time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>Yes, all the time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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
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?