SCREENING FOR OBSTRUCTIVE SLEEP

APNEA: A VALIDATION

STUDY

by

CAITLIN E. MORAN

JAMES HAMILTON, COMMITTEE CHAIR
JAMES GEYER
THEODORE TOMENY

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts
in the Department of Psychology
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2018
ABSTRACT

Obstructive sleep apnea (OSA) is a debilitating sleep disorder that, when left untreated, causes a myriad of health, safety, and financial problems. Effective treatments for this disorder exist; the largest barrier to treatment is currently diagnosis. The DSM-5 requires polysomnography, an expensive and time-consuming measure, for diagnosis. Research indicates that up to 93% of women and 82% of men with moderate to severe sleep apnea remain undiagnosed (Young, Evans, Finn, & Palta, 1997). Questionnaires are frequently used as screening tools to identify individuals at risk for OSA, but few have both high sensitivity and specificity. Spoormaker and colleagues (2005) developed the SLEEP-50 questionnaire to assess symptoms of various sleep disorders; they found the sleep apnea subscale to have both high specificity (.88) and sensitivity (.85). The current research tested the psychometric properties of the SLEEP-50 sleep apnea subscale in a sample of 300 sleep center patients and found lower specificity (.64) and sensitivity (.56) than Spoormaker (2005).
# LIST OF ABBREVIATIONS AND SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>CI95%</td>
<td>Confidence interval of 95%</td>
</tr>
<tr>
<td>M</td>
<td>Mean: the sum of a set of measurements divided by the number of measurements in the set</td>
</tr>
<tr>
<td>N</td>
<td>Number</td>
</tr>
<tr>
<td>p</td>
<td>Probability associated with the occurrence under the null hypothesis of a value as extreme as or more extreme than the observed value</td>
</tr>
<tr>
<td>r</td>
<td>Pearson product-moment correlation</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiving Operator Characteristics</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation:</td>
</tr>
<tr>
<td>$X^2$</td>
<td>Chi squared:</td>
</tr>
<tr>
<td>%</td>
<td>Percent</td>
</tr>
<tr>
<td>=</td>
<td>Equal to</td>
</tr>
<tr>
<td>$$</td>
<td>Dollar value</td>
</tr>
<tr>
<td>&lt;</td>
<td>Less than</td>
</tr>
<tr>
<td>=</td>
<td>Equal to</td>
</tr>
<tr>
<td>/</td>
<td>Divided by</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I am pleased to have this opportunity to thank the colleagues, faculty members, and community partners who have helped me with this research project. I would like to thank Kenny Lichstein for developing this project with me and sharing his advice, research expertise, and excellent feedback; and my committee chair, Jim Hamilton, for finishing this project with me. I would like to thank my committee members Jim Geyer and Ted Tomeny for their time and effort. I would also like to thank Jim Geyer, Monica Henderson, and the office administrators at DCH Northport Sleep Disorders Center for hosting and facilitating my data collection, as well as Caitie Tighe and Ian Sherwood for their assistance in data collection.

This research would not have been possible without the support of my friends and family, who have encouraged and supported me every step of the way. Finally, I would like to thank the patients at DCH Northport Sleep Disorders Center who participated in this project.
CONTENTS

ABSTRACT...........................................................................................................................................ii
LIST OF ABBREVIATIONS...........................................................................................................iii
ACKNOWLEDGMENTS.......................................................................................................................iv
LIST OF TABLES..........................................................................................................................vi
LIST OF FIGURES..........................................................................................................................vii
INTRODUCTION..............................................................................................................................1
METHOD............................................................................................................................................6
Participants.......................................................................................................................................6
Measures.........................................................................................................................................6
Procedure.........................................................................................................................................7
Data Analysis Plan..........................................................................................................................7
RESULTS..........................................................................................................................................10
DISCUSSION.....................................................................................................................................14
REFERENCES.....................................................................................................................................17
APPENDIX..........................................................................................................................................19
LIST OF TABLES

1. Correlations Among SLEEP-50 Sleep Apnea Subscale Items and Physician Diagnosis.........21

2. Accuracy of the SLEEP-50 Sleep Apnea Subscale and Alternatives…………………………22
LIST OF FIGURES

1. The revised sleep apnea subscale and impact subscale........................................19

2. Format for sensitivity and specificity analyses.........................................................20
INTRODUCTION

Obstructive sleep apnea (OSA) is a debilitating sleep disorder characterized by obstructive apneas, a complete lack of airflow caused by a blockage of the upper airway, and hypopneas, a partial blockage, throughout sleep. This disorder affects 2% to 26% of the population in the US, and estimates of undiagnosed individuals are high (Abrishami, Khajehdehi, & Chung, 2010; Young et al., 1997). The DSM-5 and most insurance companies require assessment via polysomnography for a diagnosis of OSA, which is an expensive and time-consuming procedure (American Psychiatric Association, 2013). To address the waiting lists for and financial burdens of polysomnography, researchers have developed screening questionnaires to identify individuals at risk for OSA. These questionnaires, based on risk factors for this disorder, have not proved adequate; none boasts both high specificity and high sensitivity. However, Spoormaker, Verbeek, van den Bout, and Klip (2005) recently developed the SLEEP-50 questionnaire, a tool designed to identify a variety of sleep disorders including OSA. Spoormaker et al. claim that the sleep apnea subscale of the SLEEP-50 has both high sensitivity and high specificity. However, to date, no other researchers have attempted to replicate those results; the current study will attempt to do so.

Early diagnosis and treatment of OSA are critical to patient outcome. Individuals with OSA experience many co-occurring health problems, many of which stem from the lack of adequate oxygen saturation in the blood caused by the apneas and hypopneas. Systemic hypertension, diabetes, heart failure, stroke, coronary artery disease, and increased rates of mortality have all been consistently correlated with OSA (Gottlieb et al., 2010; Punjabi et al.,
There are also increased rates of depression in individuals with OSA; rates in clinical populations range from 21% to 41% (Harris, Glozier, Ratnavadivel, & Grunstein, 2009). Additionally, individuals who suffer from OSA for extended periods of time can exhibit personality changes, problems with memory, increased irritability, and a diminished libido (Cartwright, 2003).

Not only does OSA take a toll on the physical and mental health of individuals with this disorder, but it also puts their physical safety at risk. Individuals often experience cognitive deficits resulting from the poor sleep and daytime sleepiness caused by OSA, which can lead to increased rates of motor vehicle and workplace accidents. When those individuals adhere to treatment for their OSA, however, their driving performance returns to normal levels (Leger et al., 2011). OSA also has economic implications. Kapur (2010) identifies polysomnography, various treatments, co-occurring medical problems, car and motorcycle accidents, decreased quality of life, and work-related losses as sources of economic costs incurred by individuals with this disorder; one estimate of the cost of OSA-related car and motorcycle accidents alone in 2000 was $15.9 billion (Kapur, 2010).

OSA also impacts quality of life. Multiple studies have shown that subjects with untreated OSA tend to score lower on all dimensions of the Short Form-36 Health Survey (SF-36) than the general population; treatment for OSA improved those individuals’ quality of life ratings (see Smith & Shneerson, 1995). More recently, Sin and colleagues (2002) conducted a prospective longitudinal cohort study that found that individuals with OSA who received three months of CPAP therapy scored significantly higher on the vitality domain of the SF-36 than those who received no treatment.
Risk factors for developing OSA include being male, overweight, and having a short and thick neck, a large tongue, narrow airway, and a receding chin (Cartwright, 2003). The male-to-female ratio of incidence of OSA has been estimated to be 8:1, and other estimates are even higher (Kapur, 2010). One study found a 3% change in the severity of OSA over four years for every 1% change in weight (Kapur, 2010). As rates of obesity in the United States continue to increase, rates of OSA will also continue to increase (Cartwright, 2003).

Estimates of undiagnosed OSA are high; Young and colleagues estimated that 93% of women and 82% of men with OSA have not yet been diagnosed (Young et al., 1997). Unfortunately, resources are limited for individuals seeking assessment—most sleep centers have waiting lists, and there are few sleep centers and sleep specialists in rural areas. In response to the limited supply and high demand, researchers have developed questionnaires to identify individuals who may be at risk for developing OSA (Leger, Bayon, Laaban, & Philip, 2011). As comorbid mental and physical health problems increase in number and severity as OSA goes untreated and increases in severity, using valid and reliable assessment measures to identify at-risk individuals becomes critical.

Unlike many other sleep apnea questionnaires, the SLEEP-50 does not focus on the individual’s physical characteristics. Others, like the STOP-Bang questionnaire or the American Society of Anesthesiologists’ questionnaire, rely on questions about the physical characteristics that put individuals at risk for developing OSA (Abrishami et al. 2010). Abrishami, Khajehdehi, and Chung (2010) conducted a systematic review of research on screening questionnaires used for OSA. Eight questionnaires were developed and/or validated in the studies selected; the most common measures were the Berlin questionnaire and the Wisconsin questionnaire. Judging by their sensitivity, the rate of detecting individuals positive
for OSA, and specificity, the detection rate for individuals negative for OSA, most of the questionnaires were inadequate. Sensitivities ranged from 59-81% and specificities ranged from 46-80%. Most questionnaires had either high specificity or high sensitivity, but few had both. When tested in sleep disorder patients, Haraldsson’s questionnaire had a sensitivity level of .81 and specificity level of .80 (Abrishami et al., 2010). When tested in patients without a history of sleep disorders, the Berlin questionnaire had a sensitivity level of .85 and specificity level of .95 in a study by Sharma (2006), but a sensitivity level of .69 and specificity level of .56 in a study by Chung (2008).

Polysomnography has both high sensitivity and high specificity in diagnosing OSA, which is why it is considered the gold standard of diagnosis. Because of the nature of OSA, a diagnosis cannot currently be given solely from questionnaire results; however, there are numerous benefits to using questionnaires with high sensitivity and specificity. By more accurately identifying individuals at risk for OSA, fewer patients will have to undergo unnecessary polysomnography, reducing both the economic and personal burden.

Spoormaker et al. (2005) developed a new questionnaire, the SLEEP-50, designed to identify multiple sleep disorders as described in the DSM-IV-TR, including OSA. The sleep apnea subscale includes eight statements about snoring, sweating, gasping, and coughing during sleep, as well as asking about waking up with a headache, sour taste, and with a dry mouth. Other questionnaires tend to focus on factors such as age, body mass index, blood pressure and snoring, and include approximately 4-12 items each (see Abrishami et al., 2009).

The SLEEP-50 also includes a subscale that addresses the impact of sleep complaints on daily functioning; it includes statements that address daytime sleepiness, irritability, and concentration. The specific disorder subscales address sleep complaints only; the impact subscale
was included to address the DSM-IV requirement of impairment of daily functioning for the diagnosis of a sleep disorder. In order to qualify for being at-risk for OSA, both the sleep apnea and impact subscale scores must be 15 or higher, and at least one item on each subscale must be rated 3 or 4.

To determine its construct validity and internal consistency, the SLEEP-50 was completed by 336 college students. To determine the test-retest reliability, 41 master’s level students completed the SLEEP-50 twice, with a period of three weeks between each testing. 300 sleep disorder center patients were asked during intake to complete the SLEEP-50, and 252 patients completed and returned the questionnaire. 32 nightmare-sufferers and 44 healthy volunteers also completed the SLEEP-50, and their results were included with the sleep disorder center patients when evaluating the questionnaire. One criticism of the SLEEP-50 sleep apnea subscale is that it relies upon the responder having a bedmate; three of the eight items begin with the phrase, “I am told that I…” (Spoormaker et al., 2005). To adjust for this shortcoming, the results of the proposed study will be analyzed separately for patients with and without bedmates.

The SLEEP-50 sleep apnea subscale had a sensitivity level of .85 and a specificity level of .88, both of which are unusually high. The current research replicated the Spoormaker et al. (2005) study by using a sample of 300 sleep clinic patients from a local sleep center. Patients were given only the sleep apnea and impact subscales of the SLEEP-50, and the results of the SLEEP-50 subscales were compared to a physician’s diagnosis based on PSG in order to calculate the sensitivity and specificity of the questionnaire. With this sample, it is hypothesized that the OSA subscale will again have both high sensitivity and high specificity in identifying individuals with OSA. Spoormaker et al.’s findings will be considered replicated if both the sensitivity and specificity values exceed .80.
METHOD

Participants

Spoormaker et al. (2005) recruited 300 sleep clinic patients to complete the SLEEP-50 questionnaire and received usable data from 246 individuals. For this study, 300 new patients over the age of 19 (\(M_{\text{age}} = 46.3\), range: 19-84) were recruited from a local sleep center that sees approximately 30 new patients each week. 136 participants were female (45.3%) and 164 male (54.75%); 225 were White American (75%), 65 African American (21.7%), 4 Hispanic American (1.3%), and 1 Asian American (.3%). Patients were excluded from the study if they were currently receiving treatment for OSA, if they did not respond to all items on the questionnaire, if they completed the questionnaire but did not attend the subsequent physician appointment, if the physician recommended PSG but the patient did not complete the PSG, and if PSG results were inconclusive.

Measures

SLEEP-50 Questionnaire. Sleep center patients received the sleep apnea subscale and impact subscale from the SLEEP-50 questionnaire (Spoormaker et al., 2005). Patients responded to each statement on the subscales on a four-point scale based on the applicability of each statement to themselves over the last four weeks: 1 (not at all), 2 (a little), 3 (rather much), or 4 (very much). For example, a patient could respond to the item, “I wake up coughing” with a response of 3, indicating that they wake up coughing rather much. The sleep apnea subscale has eight items, and the impact subscale has seven items. Three items on the sleep apnea subscale begin with “I am told that...”, an item asking whether the participant sleeps with a bedmate was
included after the subscales. Numerical responses to the subscales items were tallied separately for each scale. In accordance with Spoormaker (2005), individuals were considered to meet SLEEP-50 criteria for OSA if they scored 15 or higher on both subscales, and if they rated at least one item on each subscale 3 or 4. See Figure 1 for a list of items.

**Physician Diagnosis.** All patients were examined by a neurologist specializing in sleep medicine; those identified as possibly having sleep apnea received a PSG and subsequent diagnosis if criteria were met. There is no way to know the number of patients who were incorrectly excluded from PSG testing. However, following standards of care, the physician set a liberal threshold for those referred for PSG testing.

**Demographics.** Researchers collected demographic and health information from patient records, including age, gender, race, neck circumference, body mass index (BMI), and the presence of other sleep disorder diagnoses.

**Procedure**

All eligible incoming patients at the DCH Northport Sleep Disorders Center (AASM accredited) received a consent form, the SLEEP-50 sleep apnea subscale and impact subscale, and the bedmate item (see Figure 1) along with the center’s regular intake paperwork. Participants submitted the completed questionnaires and other paperwork to office administrators; the physician did not see questionnaire responses prior to assessment and diagnosis. Demographic and other patient information was later collected from electronic medical records and matched to questionnaire responses.

**Data Analysis Plan**

To determine the validity of the SLEEP-50 subscales, I calculated the sensitivity and specificity of the measure by comparing the SLEEP-50 subscale results to the physician
diagnosis (see Figure 2). Participants were classified as “true positives” if they met SLEEP-50 criteria for OSA and received a physician diagnosis of OSA (counted in cell 1 of Figure 2); “false negatives” if they did not meet SLEEP-50 criteria for OSA and received a physician diagnosis of OSA (counted in cell 2 of Figure 2); “false positives” if they met SLEEP-50 criteria for OSA and did not receive a physician diagnosis of OSA (counted in cell 3 of Figure 2); and “true negatives” if they did not meet SLEEP-50 criteria for OSA and did not receive a physician diagnosis of OSA (counted in cell 4 of Figure 2).

Sensitivity was calculated to assess the rate of correctly identified individuals with OSA, by the ratio of cells 1/(1+2). Specificity was calculated to assess the rate of correctly identified individuals without OSA, by the ratio of cells 4/(4+3). The overall error rate, the proportion of misclassifications, was calculated using the ratio of cells (2+3)/(1+2+3+4). The positive predictive value (PPV) was calculated to assess the rate of positive results that were true positives using the ratio of cells 1/(1+2). The negative predictive value (NPV) was calculated to assess the rate of negative results that were true negatives using the ratio of cells 4/(4+2).

Receiver operating characteristics (ROC) graphs were used to visualize and analyze the utility of the SLEEP-50 sleep apnea subscale as compared to physician diagnosis. ROC graphs plot the rate of true positives (sensitivity) on the Y axis and the rate of false positives (specificity) on the X axis in order to illustrate the tradeoffs between sensitivity and specificity. The point (0,0) represents a scenario in which neither true nor false positive classifications are made, and the point (1,1) represents a scenario in which only positive classifications are made. The point (0,1) represents perfect true positive classification, and the point (1, 0) represents entirely false positive classification. The area under the ROC curve (AUC) represents the probability that the
classifying variable will rank a random positive instance higher than one that is negative (Fawcett, 2006).
RESULTS

**Preliminary analyses.** Total scores on the sleep apnea subscale ($M = 16.77$, $SD = 5.26$) ranged from 8 to 32, and total scores on the impact subscale ($M = 18.28$, $SD = 4.83$) ranged from 7 to 28. Most participants received a physician diagnosis of OSA ($N = 234$, 67.4%), 131 of whom (56%) also received a SLEEP-50 diagnosis of OSA. Other physician diagnoses, frequently comorbid with OSA, included hypsomolence (72.3%), snoring (57.1%), insomnia (56.8%), and restless legs syndrome (24.2%). 205 participants reported sleeping with a bedmate (59%) and 94 (27%) without.

**Sensitivity and specificity.** I next conducted analyses to determine the validity of the SLEEP-50 sleep apnea diagnosis with respect to the correspondence of SLEEP-50 sleep apnea results and physician diagnosis of sleep apnea. Of the 58% of participants accurately classified, 131 were true positives and 42 true negatives. The sensitivity of the SLEEP-50 was .56, specificity was .64, and the error rate was .42. Of the falsely classified participants (42%), 103 participants were false negatives and 24 participants were false positives.

The sensitivity and specificity of the measure were equally poor for males and females, and for those with and without a bedmate. For women, the sensitivity was .57 and specificity .64; for men, sensitivity was .55 and specificity .63. For participants who reported having a bedmate, the sensitivity was .57 and specificity .59. For those who reported no bedmate, the sensitivity was .54 and specificity .73.
Additional analyses were conducted to determine the validity of the SLEEP-50 sleep apnea measure. There was a significant relation between physician diagnosis of OSA and SLEEP-50 diagnosis of OSA, $X^2 (1, N = 300) = 7.94, p = .005$. When the impact subscale was excluded from analysis, there was a significant relation between physician diagnosis and meeting SLEEP-50 criteria on the sleep apnea subscale, $X^2 (1, N = 300) = 19.01, p < .01$.

**Alternative scoring algorithms.** In light of the poor performance of the SLEEP-50 sleep apnea and impact subscales in predicting physician diagnosis, I performed additional exploratory analyses to identify alternative scoring approaches that more accurately predicted physician diagnosis. I will describe these briefly below; however, none of these approaches yielded acceptable levels of specificity and sensitivity (see Table 2).

When the SLEEP-50 diagnosis using Spoormaker’s criteria was plotted against physician diagnosis, $AUC = .60 (CI_{95\%}: .52 - .68), p = .015$. Using a cutoff score of 16 or greater on the sleep apnea subscale, requiring at least one sleep apnea scale item be scored 3 or 4, and including the impact subscale resulted in slightly greater, but still unacceptably low, specificity (.59) and sensitivity (.65) than Spoormaker’s (2005) cutoff score of 15.

Second, I explored scoring algorithms that used the sleep apnea subscale without the impact factor. When the sleep apnea subscale scores were plotted against physician diagnosis, $AUC = .69 (CI_{95\%}: .55 - .82), p = .009$. When the sleep apnea subscale criterion of 15 or more was used to predict physician diagnosis, the sensitivity increased to .68 and specificity decreased to .62. When the sleep apnea subscale criterion of 16 or more was used to predict physician diagnosis, sensitivity was .59 and sensitivity .65. Similar results were found when ROC analyses were performed using participants with bedmates and participants without bedmates.
Next I conducted analyses to determine whether alternative scoring algorithms could be developed using subsets of the original SLEEP-50 sleep apnea subscale items. I started by conducting point biserial correlations testing the relations of individual items to physician diagnosis (see Table 1). Four sleep apnea subscale items were significantly correlated with physician diagnosis of sleep apnea at \( p < .01 \), including item 1 \( (r = .34) \), item 3 \( (r = .32) \), item 4 \( (r = .21) \), and item 5 \( (r = .15) \). Two items were significantly but weakly correlated with a physician diagnosis at \( p < .05 \), including item 6 \( (r = .12) \) and item 8 \( (r = -.11) \).

Based on these correlations, I created alternate scales using items highly correlated with physician diagnosis to determine whether sensitivity and specificity could be improved. When a continuous scale consisting of items 1, 3, 4, 5, 6, and 8 was plotted against physician diagnosis on a ROC curve, \( AUC = .69 \) \( (CI_{95\%}: .62 - .77) \), \( p < .01 \), the ideal cutoff score was 12 with a sensitivity of .67 and a specificity of .64. When a continuous scale consisting of items 1, 3, 4, and 5 was plotted against physician diagnosis on a ROC curve, \( AUC = .74 \) \( (CI_{95\%}: .67 - .81) \), \( p < .01 \), the ideal cutoff score was 9 with a sensitivity of .63 and a specificity of .71.

I created another alternate scale using the total impact subscale score, divided by the highest possible total score, to weight the sleep apnea subscale score. When this scale was plotted against physician diagnosis using ROC analysis, \( AUC = .59 \) \( (CI_{95\%}: .51 - .67) \), \( p = .02 \). When using a cutoff score of 9, the sensitivity was .65 and specificity .55.

Lastly, I added physical indicators of sleep apnea as criteria to the sleep apnea subscale. First, I identified ideal cutoff scores for BMI and neck circumference for each gender by plotting each measure against physician diagnosis on an ROC curve. When BMI was plotted against physician diagnosis on a ROC curve for women, \( AUC = .74 \) \( (CI_{95\%}: .65 - .83) \), \( p = .05 \), the ideal cutoff score was 31. When BMI was plotted against physician diagnosis on a ROC curve for
men, $AUC = 0.78$ ($CI_{95\%}: 0.64 - 0.89$), $p = 0.06$, the ideal cutoff score was 29. When neck circumference was plotted against physician diagnosis on a ROC curve for women, $AUC = 0.78$ ($CI_{95\%}: 0.70 - 0.86$), $p = 0.04$, the ideal cutoff score was 14. When neck circumference was plotted against physician diagnosis on a ROC curve for men, $AUC = 0.81$ ($CI_{95\%}: 0.71 - 0.91$), $p = 0.05$, the ideal cutoff score was 16. When I added participant BMI as an independent criterion to the sleep apnea subscale using cutoff scores by gender, sensitivity was 0.50 and specificity 0.86. When I added neck circumference as an independent criterion to the sleep apnea subscale using cutoff scores by gender, sensitivity was 0.58 and specificity 0.78.
DISCUSSION

Spoormaker’s (2005) findings of high sensitivity (.88) and specificity (.85) for the sleep apnea and impact subscales of the SLEEP-50 were not replicated in my sample. When the SLEEP-50 criteria for sleep apnea were used to determine whether participants met criteria for sleep apnea, the measure’s sensitivity was .56 and specificity .64. Analyzing sensitivity and specificity separately by participant gender and whether the participant slept with a bedmate did not improve either measure. However, removing the impact subscale and including only the sleep apnea subscale results in analyses resulted in slightly improved sensitivity (.68) and similar specificity (.62).

Modifying the SLEEP-50 sleep apnea subscale to include only the items highly correlated with physician diagnosis also resulted in slightly improved sensitivity and specificity. When the subscale consisted of items 1, 3, 4, and 5 and the cutoff score was set at 9, the sensitivity was .63 and specificity .71; neither measure met psychometric standards nor approached levels of sensitivity and specificity reported by Spoormaker (2005). Modifying the measure by weighting the sleep apnea subscale with the impact subscale item score as a ratio; this scale yielded similarly unimpressive results.

When the BMIs of participants were compared based on whether the SLEEP-50 diagnosis matched the physician diagnosis, BMIs were significantly higher for those who received a physician diagnosis than participants who did not receive a physician diagnosis, regardless of SLEEP-50 results. Similarly, participant neck circumferences were significantly larger for participants who received a physician diagnosis than those who did not, regardless of
SLEEP-50 results. Using BMI, which is easily attained, as the sole predictor of a sleep apnea diagnosis was found to be as predictive as Spoormaker’s subscales. These physical indicators of sleep apnea were not addressed by the SLEEP-50; including them improved the specificity of the measure but not its sensitivity.

The incongruences between these results and Spoormaker’s (2005) could have resulted in part from differences in sample population. Spoormaker’s sample included healthy volunteers and nightmare sufferers in addition to sleep clinic patients, resulting in a lower base rate of individuals with sleep apnea in his sample. The majority of patients at the sleep center that was used to recruit participants for this study are diagnosed with sleep apnea. The higher base rates of sleep apnea may have increased the error rate. Both studies used physician diagnosis via PSG to confirm the SLEEP-50 sleep apnea measure results, and neither used a secondary physician’s opinion to validate the diagnosis. The poor performance of the measure in the current study may have been due to poorer reliability of the criterion diagnosis, but there is no particular reason to question the clinical diagnoses made for this study.

Because maximizing a measure’s sensitivity leads to decreased specificity and vice-versa, the purpose and impact of the measure must be considered to determine which kind of error to minimize. If the goal is to minimize unnecessary PSG, the specificity of the measure should be prioritized. However, if it is more important to prevent individuals with sleep apnea from remaining undiagnosed, sensitivity should be prioritized. Maximizing specificity would result in fewer individuals undergoing PSG, and thereby fewer individuals unnecessarily spending their resources on this test. However, the negative health impacts of remaining undiagnosed increase substantially over time, and the long-term cost of those health problems could outweigh the cost of increased PSG screenings. Maximizing the sensitivity of a measure leads to more unnecessary
PSGs, but prevents individuals with sleep apnea from slipping through the cracks and experiencing the physical and psychological health problems that result from untreated sleep apnea.

For example, the cutoff score could be adjusted to 16. In a population with a high base rate of sleep apnea the sensitivity could be increased, sacrificing specificity, which would result in a larger number of unnecessary PSGs. On the other hand, the cutoff score could be adjusted to 14, allowing more individuals with sleep apnea to slip through the cracks. In the current sample with 78% base rate of sleep apnea, simply concluding that every patient had sleep apnea would produce a higher correct prediction rate of sleep apnea than Spoormaker’s criteria.

Using a more reliable and valid screening questionnaire would prevent some individuals presenting with sleep problems not due to OSA from unnecessarily undergoing polysomnography, saving them both time and money. Arguable more importantly, it would also prevent individuals with sleep apnea from slipping through the cracks and experiencing deteriorating health over time. By improving the diagnostic process, the rates of undiagnosed individuals should decrease and the rates of individuals receiving treatment for this disorder should increase, which would have positive effects on the various comorbid health problems associated with OSA.
REFERENCES


APPENDIX

Please read every statement below and indicate to what extent it applied to you during the last four weeks:

<table>
<thead>
<tr>
<th></th>
<th>(not at all)</th>
<th>(a little)</th>
<th>(rather much)</th>
<th>(very much)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My bed partner informed me that I regularly snore</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. I sweat during the night and wake up with a wet skin.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. My bed partner informed me that I hold my breath when sleeping.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. I wake up coughing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. I have a dry mouth when I wake up.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. I’m sometimes short of breath when I wake up during the night.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. I have a sour taste in my mouth when I wake up in the morning.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8. I have headaches when I wake up in the morning.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(not at all)</th>
<th>(a little)</th>
<th>(rather much)</th>
<th>(very much)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I feel tired at getting up.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. I feel sleepy during the day.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. I feel I have too little energy during the day.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. I am quickly aggravated.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. I have difficulty in concentrating at work.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. I worry about sleeping enough.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. Generally, I sleep badly.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Do you sleep with a bedmate? (yes/no) _________________

*Figure 1.* The revised sleep apnea subscale and impact subscale from V. Spoormaker’s SLEEP-50 questionnaire (personal communication, April 24, 2014).
<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>SLEEP-50 Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
</tr>
<tr>
<td>+</td>
<td>1</td>
</tr>
<tr>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>

*Figure 2.* Format for sensitivity and specificity analyses, modeled after Figure 1 from Lichstein et al. (2003)
Table 1

**Correlations Among SLEEP-50 Sleep Apnea Subscale Items and Physician Diagnosis**

<table>
<thead>
<tr>
<th></th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Item 7</th>
<th>Item 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDdx</td>
<td>.335**</td>
<td>.051</td>
<td>.322**</td>
<td>.209**</td>
<td>.153**</td>
<td>.119*</td>
<td>-.022</td>
<td>-.114*</td>
</tr>
<tr>
<td>Item 1</td>
<td>.270**</td>
<td>.482**</td>
<td>.367**</td>
<td>.303**</td>
<td>.296**</td>
<td>.210**</td>
<td>.034</td>
<td></td>
</tr>
<tr>
<td>Item 2</td>
<td>.244**</td>
<td>.238**</td>
<td>.314**</td>
<td>.315**</td>
<td>.291**</td>
<td>.237**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 3</td>
<td>.662**</td>
<td>.322**</td>
<td>.416**</td>
<td>.229**</td>
<td>.005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 4</td>
<td>.306**</td>
<td>.518**</td>
<td>.338**</td>
<td>.126*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 5</td>
<td>.450**</td>
<td>.387**</td>
<td>.253**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 6</td>
<td>.429**</td>
<td>.247**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 7</td>
<td></td>
<td></td>
<td>.236**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.
<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Cutoff score (&gt;=)</th>
<th>TP</th>
<th>FN</th>
<th>FP</th>
<th>TN</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLEEP-50-SA-IM-34</td>
<td>15</td>
<td>13</td>
<td>10</td>
<td>24</td>
<td>42</td>
<td>.56</td>
<td>.64</td>
<td>.85</td>
<td>.29</td>
<td>.42</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>13</td>
<td>96</td>
<td>23</td>
<td>43</td>
<td>.59</td>
<td>.65</td>
<td>.86</td>
<td>.31</td>
<td>.40</td>
</tr>
<tr>
<td>SLEEP-50-SA</td>
<td>14</td>
<td>17</td>
<td>59</td>
<td>29</td>
<td>37</td>
<td>.75</td>
<td>.56</td>
<td>.86</td>
<td>.39</td>
<td>.29</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>15</td>
<td>76</td>
<td>25</td>
<td>41</td>
<td>.68</td>
<td>.62</td>
<td>.86</td>
<td>.35</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>13</td>
<td>96</td>
<td>23</td>
<td>43</td>
<td>.59</td>
<td>.65</td>
<td>.86</td>
<td>.31</td>
<td>.40</td>
</tr>
<tr>
<td>SLEEP-50-SA-BMI</td>
<td>15</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>57</td>
<td>.50</td>
<td>.86</td>
<td>.93</td>
<td>.33</td>
<td>.42</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>10</td>
<td>12</td>
<td>8</td>
<td>58</td>
<td>.46</td>
<td>.88</td>
<td>.93</td>
<td>.32</td>
<td>.45</td>
</tr>
<tr>
<td>SLEEP-50-SA-NECK</td>
<td>15</td>
<td>13</td>
<td>94</td>
<td>14</td>
<td>50</td>
<td>.58</td>
<td>.78</td>
<td>.90</td>
<td>.35</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>11</td>
<td>10</td>
<td>13</td>
<td>51</td>
<td>.52</td>
<td>.80</td>
<td>.90</td>
<td>.32</td>
<td>.42</td>
</tr>
<tr>
<td>SLEEP-50-SA-4</td>
<td>8</td>
<td>17</td>
<td>56</td>
<td>29</td>
<td>37</td>
<td>.76</td>
<td>.56</td>
<td>.86</td>
<td>.40</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>14</td>
<td>86</td>
<td>19</td>
<td>47</td>
<td>.63</td>
<td>.71</td>
<td>.89</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td>SLEEP-50-SA-6</td>
<td>11</td>
<td>17</td>
<td>62</td>
<td>29</td>
<td>37</td>
<td>.73</td>
<td>.56</td>
<td>.86</td>
<td>.37</td>
<td>.30</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>15</td>
<td>78</td>
<td>24</td>
<td>42</td>
<td>.67</td>
<td>.64</td>
<td>.87</td>
<td>.35</td>
<td>.34</td>
</tr>
<tr>
<td>SLEEP-50-SA*IM</td>
<td>8</td>
<td>16</td>
<td>70</td>
<td>34</td>
<td>32</td>
<td>.70</td>
<td>.48</td>
<td>.82</td>
<td>.31</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>15</td>
<td>83</td>
<td>30</td>
<td>36</td>
<td>.65</td>
<td>.55</td>
<td>.83</td>
<td>.30</td>
<td>.38</td>
</tr>
</tbody>
</table>

Note. SA = SLEEP-50 sleep apnea subscale; IM = impact subscale; 34 = criterion of at least one item scored 3 or 4 applied; BMI = body mass index criterion applied; NECK = neck circumference criterion applied; SA-4 = four item sleep apnea subscale; SA-6 = six item sleep apnea subscale; SA*IM = sleep apnea subscale total score, weighted by impact subscale total score divided by 28; TP = true positives; FN = false negatives; FP = false positives; TN = true negatives; PPV = positive predictive value; NPV = negative predictive value.
October 3, 2014

Caitlin Moran
Dept. of Psychology
College of Arts & Sciences
Box 870348

Re: IRB#: 14-OR-342-ME “SLEEP -50: A Validation Study”

Dear Ms. Moran:

The University of Alabama Institutional Review Board has granted approval for your proposed research.

Your application has been given expedited approval according to 45 CFR part 46. Approval has been given under expedited review categories 5 and 7 as outlined below:

(5) Research involving materials (data, documents, records or specimens) that have been collected, or will be collected solely for nonresearch purposes (such as medical treatment or diagnosis)

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies

Your application will expire on October 1, 2015. If your research will continue beyond this date, complete the relevant portions of the IRB Renewal Application. If you wish to modify the application, complete the Modification of an Approved Protocol Form. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, complete the appropriate portions of the IRB Request for Study Closure Form.

Please use reproductions of the IRB approved stamped consent forms to obtain consent from your participants.

Should you need to submit any further correspondence regarding this proposal, please include the above application number.

Good luck with your research.

Sincerely,

Carparinto T. Myles, MSM, CCM, CIP
Director & Research Compliance Officer
Office of Research Compliance
December 16, 2014

Caitlin Moran
Department of Psychology
College of Arts & Sciences
The University of Alabama
Box 870348

Re: IRB # 14-OR-342-ME (Revision) "SLEEP-50: A Validation Study"

Dear Ms. Moran:

The University of Alabama Institutional Review Board has reviewed the revision to your previously approved expedited protocol. The board has approved the change in your protocol.

Please remember that your approval period expires one year from the date of your original approval, October 2, 2014, not the date of this revision approval.

Should you need to submit any further correspondence regarding this proposal, please include the assigned IRB application number. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants.

Good luck with your research.

Sincerely,

Carpenter, T. Myles, MSM, CIM, CIP
Director & Research Compliance Officer
Office for Research Compliance

358 Rose Administration Building
Box 870127
Tuscaloosa, Alabama 35487-0127
(205) 348-8461
fax (205) 348-7189
Toll Free: (877) 320-3066
April 20, 2015

Caitlin Moran  
Department of Psychology  
College of Arts & Sciences  
The University of Alabama  
Box 870348

Re: IRB # 14-OR-342-ME (Revision # 2) “SLEEP-50: A Validation Study”

Dear Ms. Moran:

The University of Alabama Institutional Review Board has reviewed the revision to your previously approved expedited protocol. The board has approved the change in your protocol.

Please remember that your approval period expires one year from the date of your original approval, October 2, 2014, not the date of this revision approval.

Should you need to submit any further correspondence regarding this proposal, please include the assigned IRB application number. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants.

Good luck with your research.

Sincerely,

Carrollato T. Myers, MSM, CIM, CIP  
Director & Research Compliance Officer  
Office for Research Compliance
September 28, 2016

Caitlin Moran, B.S.
Department of Psychology
College of Arts & Sciences
The University of Alabama
Box 870348

Re: IRB # 14-OR-342-ME-R1 “SLEEP-50: A Validation Study”

Dear Ms. Moran:

The University of Alabama Institutional Review Board has granted approval for your renewal application. Your renewal application has been given expedited approval according to 45 CFR part 46. Approval has been given under expedited review category 5 and 7 as outlined below:

(5) Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected, solely for non-research purposes (such as medical treatment or diagnosis).

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Your application will expire on September 27, 2017. If your research will continue beyond this date, complete the relevant portions of the IRB Renewal Application. If you wish to modify the application, complete the Modification of an Approved Protocol Form. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, complete the appropriate portions of the IRB Study Closure Form.

Should you need to submit any further correspondence regarding this proposal, please include the above application number.

Good luck with your research.

Sincerely,

Carapatto T. Myles, MSM, CIM, CIP
Director & Research Compliance Officer
Office of Research Compliance