Sleep Disturbances and Sleep Disorders in Older Adults: Epidemiology, Identification and Associations with Inpatient Healthcare Utilization

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Misti Lyn Paudel

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Kristine E. Ensrud, M.D., MPH
Co-Adviser: Russell Luepker, M.D., MS

September 2014

Acknowledgments

I am extremely appreciative for all the support, mentorship and contributions from the members of my dissertation committee. I especially thank my advisor, Dr. Kristine Ensrud for being such an incredible mentor and source of support during this process. As a result of her guidance, I have grown and transformed into an independent, passionate researcher. I also extend a very special thanks to Dr. Brent Taylor, who has been a constant source of support, guidance and encouragement throughout the program. I would also like to thank my co-advisor, Dr. Russell Luepker for his thoughtful critiques and support. I am also very appreciative of other committee members, Dr. Alvaro Alonso and Dr. John Schousboe for all the thoughtful and valuable insights, advice and support throughout the dissertation.

There are many others who also contributed to this work. I thank Dr. Beth Virnig for teaching me about Medicare, and I thank Dr. Katie Stone, Dr. Sonia Ancoli-Israel and Dr. Susan Redline for teaching me about sleep research. I thank Dr. Chuck McCulloch and Li-Yung (Lily) Lui, for helping me to learn about two-part modeling methods.

I also thank the participants of the Study of Osteoporotic Fractures (SOF) and the Osteoporotic Fractures in Men (MrOS) studies for their willingness to participate in this research. I also thank the study Coordinating Centers, and staff for their devotion to high quality research and their passion for the work they do.

The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute on Aging (NIA), the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Center for Advancing Translational Sciences (NCATS), and NIH Roadmap for Medical Research under the following grant numbers: U01 AG027810, U01 AG042124, U01 AG042139,

U01 AG042140, U01 AG042143, U01 AG042145, U01 AG042168, U01 AR066160, and UL1 TR000128. The National Heart, Lung, and Blood Institute (NHLBI) provides funding for the MrOS Sleep ancillary study "Outcomes of Sleep Disorders in Older Men" under the following grant numbers: R01 HL071194, R01 HL070848, R01 HL070847, R01 HL070842, R01 HL070841, R01 HL070837, R01 HL070838, and R01 HL070839.

The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, R01 AG027576, R01 AG026720 and R01 AG05394.

Finally, I would like to thank my family, and especially my husband, Roshan, for their unwavering love, support and encouragement to keep pushing ahead and persevere.

Table of Contents

Acknowledgments iii
List of Tables vii
List of Figures ix
Chapter 1: Background and Significance 1
Introduction 1
Burden of Sleep Disturbances in Older Adults: 2
Burden of Sleep Apnea in Older Adults: 2
Risk Factors for Sleep Apnea: 2
Diagnosis and Screening of Sleep Apnea: 3
Health Care Utilization in Older Adults 3
Chapter 2: Study Design and Methods 5
The Outcomes of Sleep Disorders in Older Men (MrOS Sleep) Study 5
The Osteoporotic Fractures in Women (SOF) Study 6
Chapter 3: Performance of the STOP and STOP-BANG questionnaires in detecting moderate-severe sleep disordered breathing in a cohort of older men 9
Introduction: 11
Methods: 15
Results 22
Discussion 27
Tables and Figures 31
Chapter 4: The epidemiology of variability in sleep/wake patterns in older adults: definitions, prevalence and correlates. 44
Introduction: 46
Methods: 48
Results 52
Discussion 59
Tables and Figures 65
Chapter 5: The Impact of Sleep Disturbances on Inpatient Health Care Utilization in Older Women 83
Introduction 85

Methods 86
Results 93
Discussion 96
Tables and Figures 101
Chapter 6: Conclusions 112
Bibliography 116
Appendices 125

List of Tables

Tables	**Page**
Table 1. Characteristics of MrOS Sleep cohort by severity of OSA	32
Table 2. Frequency of STOP and STOPBANG Components	34
Table 3. Predictive parameters of Different STOP-BANG score cut-offs	35
Table 4. Predictive parameters of STOP-BANG Components	38
Table 5. Examination of alternative cut-points for STOP-BANG components for the detection of severe OSA	40
Table 6. Associations between STOP-BANG, AHI and Excessive Daytime Sleepiness	43
Table 7. Baseline characteristics of MrOS Participants	65
Table 8. Distributions and Variability in Actigraphic Sleep/Wake Parameters in Older Men	68
Table 9. Age and site adjusted odds ratios (95% CI) of predictors with measures of variability in sleep/wake patterns	69
Table 10. Multivariable adjusted associations between predictors and variability in sleep/wake parameters	72
Table 11. Multivariable adjusted models with adjustment for mean total sleep time adjusted odds ratios (95% CI) of predictors with measures of variability in sleep/wake patterns	75
Table 12. Associations between predictors and continuous variability measures of WASO, sleep latency and total sleep time	78
Table 13. Associations between predictors and continuous variability measures of in and out of bed timing	81
Table 14. Baseline Characteristics of 2,103 by Hospitalization Status	102
Table 15. Association between sleep disturbances and odds of being hospitalized	105
Table 16. Association between sleep disturbances and rate ratio of inpatient hospital days among those with $\geq$1 hospital admissions	106
Table 17. Association between sleep disturbances and cause-specific	108

hospitalizations
Table 18. Association between sleep disturbances and odds of being hospitalized among FFS enrollees 110
Table 19. Association between sleep disturbances and rate ratio of inpatient hospital days among those with ≥1 hospital admissions among FFS enrollees 111

List of Figures

Figures	Page
Figure 1. MrOS & Medicare Data Collection Timeline	5
Figure 2. SOF & Claims Data (Medicare/KPNW) Collection Timeline	8
Figure 3. Roadmap of participants included in analyses	31
Figure 4. ROC Curve Results for Continuous STOP-BANG Scores and Severe OSA.	37
Figure 5. Roadmap of Analytical Cohort	101

Chapter 1: Background and Significance

Introduction

Complaints of poor or insufficient sleep are common in older adults, with up 50% reporting symptoms such as fragmented sleep, difficulty falling asleep, early morning awakening and daytime sleepiness[1, 2]. Although the biological and physiological functions of sleep are not well understood, sleep does appear to have a role in several processes, including memory consolidation, energy conservation, restoration, endocrine function and brain development. The consequences of insufficient sleep are especially of concern in older adult populations, who are often burdened by multiple comorbidities. Prior research in middle-aged and older adult populations has suggested associations between poor sleep/sleep-related disorders and depression[3, 4], cardiovascular disease[5, 5, 6], frailty[7, 8], impaired cognitive functioning[9, 10] and mortality[8].

Therefore, adequate sleep is an important aspect of healthy aging in older adults, and the field of sleep research is striving to better understand the pathways, correlates and consequences of insufficient sleep. The field also faces many challenges, including lack of a consistent and concrete measure of sleep disturbances and disorders across studies, inconsistent results and use of screening tools that have not been studied in older populations. This dissertation will focus on improving our understanding of the epidemiology of sleep disturbances in older adults, by exploring variability in sleep/wake parameters as a novel measure of sleep disturbance, assessing the implications of a widely used sleep-apnea screening questionnaire in an older adult population, and evaluating associations between measures of self-reported and actigraphy assessed sleep disturbances and health-related outcomes using a cohort-linked to Medicare dataset. Data from two large cohort studies of older adults will be used to address these questions.

Burden of Self-Reported Sleep Disturbances in Older Adults:

Self-reported sleep disturbances include difficulties getting to sleep, staying asleep (sleep fragmentation), early morning awakening and reduced total sleep time. Over 50% of older adults report one or more sleep disturbances[1, 11], and >22% report waking up feeling unrefreshed[1]. Surveys have also shown that despite a high prevalence of complaints, older adults typically report getting an average of 7 hours of sleep per night[1, 11].

Research has also suggested that that the amount of sleep needed (and not necessarily the ability to sleep) naturally declines with older age, and that the inability to obtain enough sleep in older adults may be attributed to a higher burden of comorbidities, psychiatric conditions (such as depression or dementia), medications and life changes (bereavement)[11]. Additional research has suggested that the sleep patterns of older adults (night-to-night variability) are more stable[11] than in younger populations, but this has not been directly assessed in a community-based population of older adults.

Burden of Sleep Apnea in Older Adults:

Sleep disordered breathing (SDB), or sleep apnea, is characterized by repetitive partial or complete airway obstruction during sleep. The prevalence of SDB in middle aged adults ranges from 2-4%[12], and in older adults from 6-70% depending on the definitions being used and characteristics of the population studied (referral or community-based)[13-18]. In the Outcomes of Sleep Disorders in Older Men (MrOS Sleep) cohort study, we found a prevalence of moderate-severe sleep apnea of 26.4% (Apnea Hypopnea Index$\geq$15)[18].

Risk Factors for Sleep Apnea:

There are several risk factors and correlates of SDB[18, 19], including older age, non-Caucasian race, obesity, sleepiness, hypertension, snoring, cardiovascular disease, nasal congestion, smoking, breathing pauses, male

gender, neck circumference>40cm, and alcohol use. SDB has also been linked with several health-related outcomes[19] including an increased risk of hypertension[20-23], stroke[6, 24], cardiovascular disease[6, 25], lower quality of life[26, 27], excessive sleepiness[19, 28] and mortality[13].

Diagnosis and Screening of Sleep Apnea:

A diagnosis of obstructive sleep apnea is made by overnight polysomnography (PSG), typically performed in a hospital or referral center-based setting. PSG studies are expensive, require highly trained personnel, equipment and considerable resources for staff, hospitals (or centers performing the study) and time since they involve an overnight stay. There can also be a lengthy waiting list to get a PSG depending on the availability and capacity of the center. Therefore, there is a great need for an effective screening tool for use in the primary care clinical practice setting to determine which patients most likely have SDB and warrant evaluation with polysomnography.

Several simple questionnaires have been developed to screen and identify patients who have sleep apnea[29]. The STOP-Bang is a widely used questionnaire, because of its simplicity and high sensitivity, but was developed in a surgical population as a way to assess risk of sleep disordered breathing prior to giving anesthesia. In general, screening tools for sleep apnea have performed moderately well[29], and to our knowledge, only the Berlin has been previously studied in a population of older adults[30]. Due to the cost and limited resources of polysomnography, the need for a reliable and accurate screening tool for sleep apnea in older adult populations is warranted.

Health Care Utilization in Older Adults

According to the 2007 National Hospital Discharge Survey, older adults comprise about 13% of the US population, but account for 37% of all hospital discharges[31]. Healthcare utilization, and especially inpatient admissions, may be

a good indicator of major disease events, and signal unsuccessful aging in older adults. Examining how factors not available in the medical record or administrative claims data such as sleep disturbances, are associated with hospitalizations in older adults will increase our understanding of determinants of inpatient health care use in the elderly and better quantify the impact that sleep disturbances have on health in older adults.

Therefore, to improve our understanding of the epidemiology of sleep disturbances in older adults, this dissertation will explore variability in sleep/wake measures as a novel measure of sleep disturbance in older men, assess the implications of a widely used sleep-apnea screening questionnaire in a cohort of community-dwelling older men, and evaluate associations between measures of self-reported and actigraphy assessed sleep disturbances and inpatient admissions in a cohort of older women concurrently enrolled in fee-for-service Medicare.

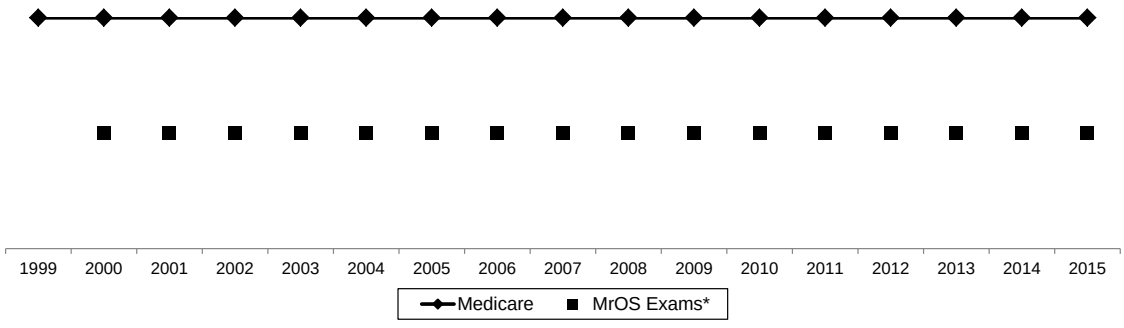
Chapter 2: Study Design and Methods

The Outcomes of Sleep Disorders in Older Men (MrOS Sleep) Study

Cohort Recruitment and Composition

The prospective Osteoporotic Fractures in Men (MrOS) study enrolled 5,994 men aged 65 years and older from March 2000 through April 2002. The study was primarily initiated to better understand the extent to which fracture risk is related to bone mass, bone geometry, lifestyle, anthropometric and neuromuscular measures and fall propensity, as well as to determine how fractures affect quality of life in older men. Men were recruited from population-based listings in six areas of the United States: Birmingham, Alabama; the Monongahela Valley near Pittsburgh, Pennsylvania; Minneapolis, Minnesota; Palo Alto, California; San Diego, California; and Portland, Oregon. Men with a history of bilateral hip replacement ad men who were unable to walk without assistance of another person were excluded from study enrollment. Several additional follow up visits and sub studies were conducted following the initial baseline exam, with visits occurring approximately every 1-2 years (Figure 1).

Figure 1. MrOS & Medicare Data Collection Timeline



Medicare claims data currently through 12/31/2007; the renewal grant has funding to purchase additional years of claims data (1/1/2008 through 12/31/2015)

*5,994 men aged 65 and older were enrolled at the baseline exam (2000 to 2002) at the Minneapolis, Pittsburgh, Portland, Palo Alto, San Diego, and Birmingham clinical centers. MrOS participants are followed-up every 4 months via mail or phone to track endpoints of falls, fractures, prostate cancer, and death.

From December 2003 through March 2005, a subset of MrOS participants were invited to participate in an ancillary study Outcomes of Sleep Disorders in Older Men (MrOS Sleep), to better understand how sleep disorders impact health related outcomes in older adult men. To be eligible to participate in the sleep ancillary study, participants had to report not using oxygen therapy in the past three months, no history of an open tracheotomy, not sleeping with a mouthpiece for snoring or sleep apnea in the past three months, or not sleeping with a CPAP or BiPAP mask in the last three months. Some exceptions were made for participants who intermittently used CPAP, or who were willing to forgo wearing the CPAP or BiPAP mask during the sleep study. Of the 5,994 men enrolled in the MrOS Parent study, 3,135 (>100% of recruitment goal) participated in the MrOS Sleep exam.

The Study of Osteoporotic Fractures in Women (SOF)

The Study of Osteoporotic Fractures (SOF) is a landmark longitudinal epidemiologic study designed to examine risk factors for osteoporotic fractures. Women were recruited from four U.S. clinical centers (Baltimore, Maryland; Minneapolis, Minnesota; the Monongahela Valley nears Pittsburgh, Pennsylvania; and Portland, Oregon)[32]. The SOF study enrolled 9,704 community-dwelling white women aged 65 years and older from 1986-1988. Women were excluded if they were unable to walk without assistance, or if they had undergone a previous bilateral hip replacement. Initially African American women were excluded from the study due to their low incidence of hip fractures, but from 1997-1998, 662 African American women aged 65 years and older were recruited[33].

After completion of the baseline clinic visit, additional follow-up visits were conducted approximately every 1-4 years (Figure 2). This dissertation will utilize

data from the Year 16 (or Visit 8) SOF exam that was conducted between 2002 and 2004. At this visit, 3,137 women attended a clinic visit and 539 attended a home visit.

Linkage to Medicare Claims and Kaiser Permanente Encounter Data

Linkage of the SOF cohort to Medicare claims data was completed in 2008, by submitting social security and/or Medicare (HIC) numbers for SOF participants who were alive as of 1/1/1991 (first date that Medicare claims were available), to *Centers for Medicare and Medicaid Services (CMS)*. A linkage was determined to be valid if there was an exact match on SSN/HIC, and sufficient agreement on DOB, gender, last known residence (ZIP code), and date of death (when available). Medicare data was purchased from January 1991-December 2010. Of the 10,366 women enrolled in the SOF study, 9,986 were alive and actively enrolled in SOF as of January 1, 1991, and of those 9,228 (92.4%) were determined to be valid linkages to Medicare claims data.

Women at the SOF Portland site were originally recruited into the SOF study through Kaiser Permanente, and thus we observed a high rate of Medicare Advantage enrollment at this site (96%). Linkage of SOF Portland participants to Kaiser Permanente inpatient encounter records was completed in 2014 by submitting social security numbers to Kaiser Permanente. Of the 2,464 women enrolled at the Portland SOF site who were alive as of 1/1/1991, 2,180 (88.5%) were enrolled in a Kaiser Permanente plan. Kaiser Permanente inpatient encounter data was obtained from January 1991-December 2010. In combining Medicare and Kaiser Permanente encounter records we were able to successfully link 9,381 (93.9% of 9,986) SOF participants to Medicare and/or Kaiser encounter records.

We required that during the month of the SOF V8 exam, participants be observable in claims/encounters data, meaning that they were either enrolled in Kaiser, or enrolled in a Part A Medicare plan for which CMS processes all of the inpatient claims. Of the 3,123 women who attended Visit 8 and had technically

adequate actigraphy data, 2,103 (67.3%) linked successfully to Medicare and/or Kaiser and were enrolled in a Part A plan, or Kaiser plan for at least one month after their sleep visit, until death, enrollment in a Medicare Advantage plan, or the end of follow up, whichever came first (Figure 4).

Figure 2. SOF & Claims Data (Medicare/KPNW) Collection Timeline

*Claims data currently available from 1/1/1991 through 12/31/2007; the renewal grant has funding to purchase additional years of claims data (1/1/2008 through 12/31/2014)

†7,280 Caucasian women were enrolled at the baseline exam at the Minneapolis, Pittsburgh, and Portland clinical centers between 1986-1988; an additional 480 African American women were enrolled at the Year 10 exam between 1997-1998

NOTE: Until Fall 2009, SOF participants were followed-up every 4 months via mail or phone to track endpoints of falls, fractures, breast cancer, and deaths. Since Fall 2009, SOF participants have been followed-up every 6 months via questionnaire administered over the telephone to collect additional data focused on living situation, health status, and functional status.

Chapter 3: Performance of the STOP and STOP-BANG questionnaires in detecting moderate-severe sleep disordered breathing in a cohort of older men

Objectives: To evaluate the ability of the STOP-BANG questionnaire to accurately identify older men with Obstructive Sleep Apnea (OSA), and to examine the association between STOP-BANG scores and excessive daytime sleepiness.

Design: Cross-sectional study

Participants: Two thousand nine hundred fifty three men aged 67 years and older enrolled in the Outcomes of Sleep Disorders in Older Men (MrOS Sleep) cohort study.

Measurements: OSA was assessed via overnight, in-home polysomnograpy, and the STOP-BANG questionnaire was recreated using data collected during the clinic exam. Severe OSA was defined as an Apnea Hypopnea Index (AHI)$\geq$30, and primary analyses examined STOP-BANG scores using a cut point of $\geq$3. Secondary analyses examined alternative cut points in OSA severity and STOP-BANG scores. Excessive daytime sleepiness (ESS) was defined as an Epworth Sleepiness Scale score >10, and analyses examining associations between STOP-BANG and ESS used logistic regression.

Results: Severe OSA was prevalent (17.5%) in this population, as were most STOP-BANG components of Snoring loudly (41.8%); Tired (45.7%); Observed apneas (20.9%); Pressure (50.0%); BMI>35 kg/m^2 (3.6%); Age>50 y (100%); Neck circumference>40 cm (37.9%); and male gender (100%). At a cut point of $\geq$3, the STOP-BANG identified 88.4% of the men as having a high likelihood of OSA. Furthermore, this cut point resulted in a large number of false positives, as evidenced by a sensitivity of 94.0% and a specificity of 12.7%, and little impact on probability revision as evidenced by a positive predictive value (PPV) that approximated the prevalence of OSA (18.6% vs. 17.5% for PPV vs. prevalence respectively). Secondary analyses did not suggest improved functionality with

higher STOP-BANG cut points due to the resultant high rate of missed cases and indicated that BMI>35 kg/m^2 was more accurate in identifying OSA than STOP-BANG.

Conclusion: The STOP-BANG questionnaire has poor discriminatory ability in detecting OSA in community-dwelling older men. Additional research into risk factors and characteristics of OSA in older populations is warranted.

Introduction:

Sleep apnea is a chronic age-related medical condition characterized by repeated episodes of pauses in breathing, or shallow/infrequent breathing that occur during sleep. These episodes can last from seconds to minutes, and can occur up to 30 or more times per hour[34]. Patients with sleep apnea may report symptoms of excessive daytime sleepiness, fatigue, snoring, and/or disturbed sleep[19, 34]. Emerging evidence from primarily middle-aged populations also suggests that sleep apnea is associated with and increased risk of mortality[35-37], hypertension[38-41], stroke[42, 43], cardiovascular disease[43, 44], traffic accidents[19, 45], cognitive impairment[46, 47] and diminished quality of life[48, 49]. Studies have also suggested that patients with sleep apnea have a higher risk of post-surgical complications[50].

The prevalence of obstructive sleep apnea (OSA) that has been reported in the literature varies depending on definition and population. The results of observational studies have estimated that 17-26% of adult men and 7-14% of adult women have at least mild sleep apnea[12, 19]. The prevalence of OSA rises with increasing age (up to about age 70 years), and as many as 50% or more of older adults may have at least mild sleep apnea[18]. The prevalence of severe OSA is often not the focus of many epidemiologic studies, but was observed to be 7.2% in a population of mostly middle-aged adults enrolled in the Sleep Heart Health Study[51].

While sleep apnea is associated with an increased risk of adverse health outcomes in middle-aged adults, associations between OSA and outcomes such as mortality may be weaker in older populations[35, 36, 52, 53]. Findings regarding the association in older adults may reflect an underlying difference in the pathophysiology of the condition, comorbidities, survival bias or competing risks[18]. For example, in the Outcomes of Sleep Disorders in Older Men (MrOS Sleep) cohort, significant associations were observed between measures of severe OSA and greater nocturnal hypoxemia and increased risk of mortality

after 3.4 years of follow-up, but there was no association of at least mild OSA and moderate/severe OSA with risk of mortality in this cohort[52].

To date, the gold standard methodology to diagnose sleep apnea requires an overnight polysomnogram, which typically, but not always involves an overnight stay in an sleep laboratory, and the attachment of several monitors to measure breathing, eye movements, brain waves, chest and abdominal respiratory effort, oxygen saturation and heart rate. Overnight polysomnography is also expensive, and there can often be lengthy wait lists, which make it especially not feasible for pre-operative sleep apnea testing. Therefore, there is a need for a simple initial screening tool that could be used in the busy clinical practice setting to identify patients who are at high likelihood of having sleep apnea.

In an effort to address this need, several check-list style screening risk assessment tools have been developed. The focus of our study was on the STOP-BANG questionnaire, due to its simplicity for use, comprehensive criteria and existing clinical practice guidelines that recommend its use for sleep apnea screening in pre-anesthesia patients [54 55]. Furthermore, results of two meta-analyses comparing multiple questionnaires both reported that the STOP-BANG performs the best. The first meta-analyses, conducted by Dr. Abrishami and colleagues[56] evaluated the STOP-BANG, STOP, Berlin, Wisconsin Sleep Questionnaire, and American Society of Anesthesiologists (ASA) checklist and suggested that the STOP-BANG had higher quality evidence than the other tools assessed, was simple to use and had the greatest sensitivity. Another meta-analysis published in 2009 evaluated several tools (including many of the same studies that were included in the initial meta-analysis) and suggested that the STOP-BANG had a high sensitivity, and was an excellent screening test for severe OSA, but had an unacceptable false negative rate for the diagnosis of any or moderate/severe OSA[57]. Both meta-analyses observed significant heterogeneity across studies and thus were unable to present pooled results.

The STOP-BANG questionnaire was developed in 2008 as a tool to identify pre-anesthesia patients who had undiagnosed sleep apnea[50]. It consists of eight yes or no questions (scored as 1/0) and summary scores ranging from 0-8. The eight components of the STOP-BANG assess presence of Snoring, Tiredness, Observed apneas, Pressure (hypertension), BMI, Age, Neck circumference and Gender. The STOP-BANG (as well as the shorter STOP questionnaire) was initially validated in a population of 177 preoperative adult patients without a prior diagnosis of OSA undergoing elective procedures. Participants had a mean age of 55 $\pm$ 13 years, and about 50% were male. Using a cut off of AHI>30 to define severe sleep apnea (prevalence=22.0%) and a cut point of $\geq$3 on the STOP-BANG (prevalence=71.2%), results indicated that in this population, STOP-BANG had a sensitivity of 100%, specificity of 37.0%, positive predictive value (PPV) of 31.0% and negative predictive value (NPV) of 100%. Thus, the STOP-BANG cutpoint of $\geq$3 identified all participants who had severe OSA, but also 63% of the participants who didn't have severe OSA as false positive[50].

In a follow-up study of 746 preoperative patients (mean age 60 years, 49% male, 18.0% prevalence of severe OSA), researchers observed that at a cut-point of $\geq$3, the sensitivity, specificity, PPV and NPV was observed to be 94.8%, 27.6%, 22.3% and 96.0% respectively for the identification of severe OSA. They concluded that a cut point of 5 or more might be optimal to identify patients with OSA the sensitivity, specificity, PPV and NPV at a cut point of $\geq$5 for detection of severe OSA was observed to be 56.0%, 74.2%, 32.2% and 88.5%[58]. It is important to note that at this cut point, use of the STOP-BANG resulted in a false positive rate of 25.8%, and a false negative rate of 44.0%.

Efforts to validate the performance of the STOP-BANG in a more general population of middle-aged adults enrolled in the Sleep Heart Health Study (SHHS) suggested an even poorer ability to detect individuals with severe OSA. In a cohort of 4,770 SHHS participants (mean age 62.4 years, 51.5% male), results suggested that at a cut point of $\geq$3, the STOP-BANG identified 72.4% of

the cohort as having severe OSA, with a sensitivity of 70.4%, specificity of 59.5%, PPV of 11.9% and NPV of 96.3%. The prevalence of severe OSA in this population was observed to be 7.2%, which was much lower than the number identified as having a high likelihood of OSA on the STOP-BANG, which resulted in a high false positive rate (40.5%), and missed about one-third of individuals with severe OSA[51].

To our knowledge, the performance of the STOP-BANG in a population of older adults has not been directly assessed. Given the higher prevalence of sleep apnea and STOP-BANG components in aged adults, it is plausible that the ability of this tool to identify patients with OSA may be further diminished beyond its performance in younger populations. On the other hand, the simplicity of this tool makes it particularly attractive for use in primary care practice settings to evaluate older patients presenting with sleep complaints suggestive of sleep apnea to determine need for further testing, such as referring the patient for overnight polysomnography. Hence, a better understanding of the performance of this tool in an aged population is essential.

In addition, a detailed examination of the association between STOP-BANG scores and self-reported sleep complaints in an older population may provide insight into how well the STOP-BANG questionnaire predicts sleep-disordered breathing phenotypes that impair daytime functioning. For example, excessive daytime sleepiness (ESS) is a measure of propensity for sleep onset, and is a key symptom of sleep apnea that often drives patients to seek medical care. The mechanism by which sleep apnea is thought to cause ESS is through arousals that result in fragmented sleep, although studies attempting to elucidate this pathway have not confirmed that frequency of arousals explains the variation between sleep apnea and ESS[59]. The association between STOP-BANG and excessive daytime sleepiness in older adults is unknown, as well as how this association compares to that between OSA defined using AHI and excessive daytime sleepiness.

Therefore, the aim of this study was to comprehensively evaluate the ability of the STOP-BANG questionnaire to identify older men with and without OSA. A second aim was to examine if higher STOP-BANG scores are associated with a measure of excessive daytime sleepiness, a key symptom of OSA. To assess these aims, we measured OSA using overnight, in-home polysomnography and recreated the STOP-BANG questionnaire using data collected in 2,853 men aged 67 years and older who were enrolled in the Outcomes of Sleep Disorders in Older Men (MrOS Sleep) cohort study.

Methods:

Study Population

Participants in the Osteoporotic Fractures in Men (MrOS) study were recruited at six U.S. clinical centers (Birmingham, AL; Minneapolis, MN; Palo Alto, CA; Monongahela Valley near Pittsburgh, PA; Portland, OR; and San Diego, CA) to complete a baseline examination between March 2000 and April 2002. The primary aim of the MrOS study was to identify risk factors for osteoporosis and fractures in older community-dwelling men. To be eligible, men had to be able to provide consent, walk without assistance from another person or aid, be aged 65 years and older, and not have had bilateral hip replacements. A total of 5,994 men were enrolled in MrOS, and details regarding recruitment and study design have been published elsewhere[60, 61].

From December 2003 through March 2005, a subset of MrOS participants were invited to participate in an ancillary study to identify outcomes of sleep disorders in older men (MrOS Sleep). To be eligible, participants had to report not using oxygen therapy in the past three months, have no history of an open tracheotomy, report not sleeping with a mouthpiece for snoring or sleep apnea in the past 3 months, and not sleeping with a CPAP or BiPAP mask in the last 3 months. Approval was obtained from the institutional review board at each site. Written informed consent was obtained for all individuals.

Of the 5,994 men enrolled in MrOS, 3,135 (>100% recruitment goal) participated in the MrOS Sleep ancillary study. A total of 1,997 men declined participation in the sleep study, 150 were not eligible, 344 died prior to the sleep study, 36 terminated from the MrOS study before being contacted and 332 were not contacted because enrollment goals had been met (Figure 1). Of the 3,135 participants who participated in the sleep visit, 2,911 (92.9%) had useable PSG data, 179 did not participate in PSG, and 45 had PSG data gathered, but it was not technically adequate. Furthermore, of the 2,911 participants with PSG data, 2,853 had complete measures of STOP-BANG components and comprise the analytical cohort for this paper. Compared to the 2,853 participants included in the analytic cohort, the 1,997 MrOS Study participants who refused participation in MrOS Sleep were slightly older at enrollment in the MrOS study (74.0 $\pm$ 5.9 vs. 73.0 $\pm$ 5.5 years, p<.001), more often reported having poor, very poor or fair health (14.0% vs. 11.5%, p=0.01), were slightly less likely to be Caucasian race (90.1% vs. 91.9%, p=0.025), and did not differ with respect to body mass index (27.2 $\pm$ 3.7 vs. 27.4 $\pm$ 3.7 kg/m^2, p=0.116).

Measurement of Obstructive Sleep Apnea

Single night, in-home sleep studies using unattended polysomnography (Safiro, Compumedics Inc., Melbourne, Australia) were performed successfully on 2,911 MrOS Sleep participants. The PSG recordings were to be gathered within 1 month of the clinic visit (mean 6.9 $\pm$ 15.8 days from visit), with 78% of recordings gathered within 1 week of the clinic visit. Data was gathered in 30-second epochs. The recording montage consisted of C_3/A_2 and C_4/A_1 electroencephalograms, bilateral electrooculograms, a bipolar submental electromyogram, thoracic and abdominal respiratory inductance plethysmography, airflow (using nasal-oral thermocouple and nasal pressure cannula), finger pulse oximetry, electrocardiogram, body position (mercury switch sensor), and bilateral leg movements (piezoelectric sensors). Trained certified staff members performed home visits for setup of the sleep study units. After

sensors were placed and calibrated, signal quality and impedance were checked, and sensors were repositioned as needed to improve signal quality, replacing electrodes if impedances were greater than 5kΩ, using approaches similar to those in the Sleep Heart Health Study[62]. Staff returned the next morning to collect the equipment and download the data to the Central Sleep Reading Center (Cleveland, OH) for centralized scoring by a trained technician. Polysomnography data quality had a failure rate of less than 4% and more than 70% of studies were graded as being of excellent or outstanding quality.

The diagnosis of Sleep apnea was determined using the apnea hypopnea index (AHI), which represents the average number of apneas and hypopneas per hour of sleep. Apneas were defined as a complete or almost complete cessation of airflow for more than 10 seconds. Obstructive apneas were scored if persistence of effort on abdominal or thoracic inductance plethysmography was noted, and central apneas were scored if there was no evident effort on either the abdominal and thoracic plethysmography bands. We did not differentiate between obstructive and central apnea in primary analyses, since central apneas occurred very infrequently in this cohort (<5%). Hypopneas were scored using SHHS criteria[63], requiring a >30% reduction in amplitude of either respiratory effort or airflow for more than 10 seconds, and a $\geq$3% oxygen desaturation. Secondary analyses examined AHI events associated with a $\geq$4% oxygen desaturation. Severity of AHI was classified as: 0-4 (none); $\geq$5 (Any: mild/moderate/severe); $\geq$15 (moderate/severe); and $\geq$30 (severe).

Self-Reported Sleep Measures

Data from the Pittsburgh Sleep Quality Index (PSQI) was used to construct STOP-BANG questions for **S**nore loudly and **O**bserved apneas, as well as information on bed partners. The PSQI is a validated measure of subjective sleep quality and sleep disturbances over a 1-month period. The questionnaire is divided into sections that assess subjective sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbances, use of sleeping medications and

daytime dysfunction. Global PSQI scores range from 0 to 21[64, 65]. Data from the Functional Outcomes of Sleep Questionnaire (FOSQ) were used to construct STOP-BANG questions for **T**ired. The FOSQ is a disease specific quality of life questionnaire that is used to assess functional status of daily behaviors that are impacted as a result of excessive sleepiness. The questionnaire consists of 26 items and 4 factor subscales that assess difficulties with activity level, vigilance, general productivity and social outcomes[66].

The Epworth Sleepiness Scale (ESS) is a self-administered questionnaire that classifies subjective daytime sleepiness in people with sleep disorders, and is a measure of propensity for sleep onset. Participants are asked to rate how likely (from 1 to 3, with 1 being unlikely and 3 being highly likely) they are to doze off in eight typical daily situations. Scores range from 0 to 24, with a standard cutoff of greater than 10 indicative of excessive daytime sleepiness[67, 68].

Additional Measures

At the time of their clinic visit, participants completed several clinical measurements and self-administered questionnaires. Body weight was measured using a standard balance beam, or digital scale and height using a wall mounted Harpenden stadiometer. Body mass index (**B**MI) was calculated as kg/m^2. **N**eck circumference was also measured at the clinic visit and expressed in cm. Participants completed a self-administered questionnaire which ascertained information on medical history (including history of hypertension/**P**ressure), current health status, smoking and alcohol use, and physical activity. Information from the MrOS baseline visit was used to assess age, race and educational level.

Construction of the STOP-BANG Scores

To create the STOP-BANG questionnaire for the purposes of this study, we utilized data collected in the MrOS Sleep Visit that was similar but not always identical to questions in the STOP-BANG instrument. A comparison of the

questions in STOP-BANG and similar questions collected in MrOS are provided in Appendix A. **S**nore loudly was considered positive if the participant reported having trouble sleeping during the past month because they cough or snore loudly (*less than once a week, once or twice a week, or three or more times a week=Yes, not during the past month=No)*. **T**ired was considered positive if the participant reported difficult being as active as they want to be in the morning (or afternoon) because they are sleepy or tired (*extreme difficulty or moderate difficulty or a little difficulty=Yes, No difficulty= No*). **O**bserved apneas were considered positive if the participant reported that their bed partner observed that during the past month they had long pauses between breaths while asleep (*less than once a week, or once or twice a week or three or more times a week =Yes, not during the past month=no).* **P**ressure was considered positive if the participant self-reported having a physician diagnosis of hypertension. **B**MI was considered positive if the participants' BMI was greater than 35 kg/m^2. **A**ge was considered positive for all participants since the minimum age in the cohort was 67 years. **N**eck circumference was considered positive if greater than 40 cm. Male **G**ender was considered positive for all participants since only males were enrolled in the cohort.

Sensitivity analyses explored alternative variables available in the MrOS Sleep cohort, such as a positive response to **T**ired if the participant reports having difficulty concentration on things because they are sleepy or tired (*a little difficulty, or moderate difficulty or extreme difficulty=Yes, no difficulty*=No), or a positive response to **P**ressure if the participant reports use of antihypertensive medications or has a systolic blood pressure>140 mm Hg.

Global STOP-BANG scores were calculated by summing the number of affirmative answers to each of the questions.

Statistical Analysis

The ability of STOPBANG scores to identify men with and without OSA was evaluated using epidemiologic parameters of sensitivity, specificity, positive

predictive value (PPV), negative predictive value (NPV) and likelihood ratios. PPV and NPV estimates provide clinically useful information on the performance of a screening tool, given a positive or negative finding. PPV is defined as the probability that participants with a positive test truly have OSA, whereas the NPV is the probability that participants with a negative test don't have OSA. If the PPV roughly equals the prevalence of disease in the cohort (prevalence of OSA in this case), and NPV equals 1-prevalence of disease then the screening tool yields little or no added value for clinical decision making, beyond the prior information that is already known (OSA prevalence in the population).

The likelihood ratio uses sensitivity and specificity to determine whether a test results in a meaningful change in the probability of the disease. The likelihood ratio positive (LR+) is defined as the ratio of the probability that a person who has the disease tests positive divided by the probability that a person who does not have the disease tests positive. Similarly, the likelihood ratio negative (LR-) is the ratio of the probability that person who has the disease tests negative divided by the probability that a person who does not have the disease tests negative. Likelihood ratios greater than 1 indicate that the test is associated with the presence of disease. A likelihood ratio less than 1 indicates that the test is associated with absence of the disease, and likelihood ratios close to 1 have little diagnostic value since the prevalence of the disease is equivalent to the post-test probability (positive predictive value).

For primary analyses we assumed a STOP-BANG cut-point $\geq$3 was indicative of a positive test, and OSA was defined as an AHI$\geq$30. Secondary analyses examined alternative STOP-BANG cut points as well as lower AHI cut points (AHI$\geq$5 and $\geq$15).

Logistic regression was used to examine receiver operating characteristic (ROC) curves and to calculate the area under the curve (AUC). The area under a ROC curve is used as a measure of how well a test, such as the STOP-BANG, can distinguish between two groups (OSA/normal). For models expressing STOP-BANG scores as continuous variables, the AUC represents a global

summary statistic of the diagnostic accuracy of the test, irrespective of any particular cut point. For models examining dichotomous cut points in STOP-BANG scores, the AUC represents the diagnostic accuracy of the test at that particular cut point. Models with AUC statistics of 0.50 do no better than chance alone, whereas models with higher AUC statistics do much better than chance, and an AUC of 1.0 indicates a perfect test.

Secondary analyses assessed the predictive ability of individual STOP-BANG components, expressed as dichotomous variables based on STOP-BANG criteria and in their original form (continuous or categorical) to determine if alternative cut points improved the performance of the STOPBANG tool. Secondary analyses also examined the association of specific STOP-BANG thresholds and severe OSA with odds of excessive daytime sleepiness, as defined by an Epworth Sleepiness Scale score of 11 or higher. Analyses examining the odds of excessive daytime sleepiness were adjusted for known potential confounders, and included age, site, Caucasian race, smoking status, diabetes, history of myocardial infarction, BMI, self-reported sleep quality (PSQI), and physical activity (PASE).

We performed several sensitivity analyses. Since results of sensitivity analyses were similar to those of the primary analyses, they are not included in the results section but can be found in Appendices. These sensitivity analyses included examination of OSA using alternate definitions of AHI, such as a more stringent desaturation criteria (4% oxygen desaturation criteria) (Appendix B), expanding the definition to include all apneic events (Appendix G), excluding central sleep apnea events (Appendix H), excluding men with unusual occurrences of periodic breathing (i.e. Cheynes-Stroke) (Appendix I), and restricting to men who reported having a bed-partner (Appendix J).

We also performed several sensitivity analyses redefining and exploring alternative cut points in the STOP-BANG components, such as evaluating varying combinations of components (Appendix C), redefining the Tired and Pressure components (Appendices E and F), using alternative cohort questions,

and examining ROC curve and AUC statistics for continuous components of Age, BMI and Neck circumference (Appendix D).

Furthermore, we examined analyses stratified by excessive daytime sleepiness (ESS>10 vs. ≤10) (Appendix K), and age (≥80 vs. <80 years) (Appendix L). Finally, we performed analyses using STOP (Appendix M), instead of STOP-BANG, and examined ROC curves and AUC statistics for STOP scores (Appendix N).

Results

Subject Characteristics and Prevalence of OSA

The mean (SD) age of the cohort was 76.4 (5.6) years; 91.2% of the men identified themselves as Caucasian, and 86.6% rated their health status as excellent/good (Table 1). Men in the cohort were generally well educated, with 78.4% having completed at least some years of college education.

OSA was common among men in the cohort, with 17.5% of men meeting the criteria for severe OSA (AHI≥30); 43.4% had at least moderate OSA (AHI≥15); and 78.8% had at least mild OSA [AHI≥5]. With the exception of **B**MI>35 kg/m^2, STOP-BANG components were also highly prevalent in this population: **S**noring loudly (41.8%), **T**ired (45.7%), **O**bserved apneas (20.9%), hypertension/**P**ressure (50.0%), **N**eck circumference>40 cm (37.9%), **A**ge>50 (100%), male **g**ender (100%) and **B**MI>35 kg/m^2 (3.6%).

Several baseline characteristics were significantly associated with greater AHI severity, including older age, lower education, poorer health status, lower physical activity, higher body mass index, larger neck circumference, greater excessive daytime sleepiness, and greater prevalence of diabetes (Table 1). With the exception of **T**ired, individual STOP-BANG components were significantly associated with greater AHI severity, and there was evidence of linear trend (p-trend<.001) across AHI categories for **S**nore loudly, **O**bserved apneas, **P**ressure, **B**MI>35 kg/m^2, and **N**eck circumference>40 cm. Higher total

STOP-BANG scores were also significantly associated with greater AHI severity (p-value<.001).

Prevalence of STOP-BANG Scores and Components

The frequency distribution of total STOP-BANG and STOP scores among men in the cohort are presented in Table 2. STOP-BANG scores ranged from 2-8 (all participants met the age and male gender components), and 88.4% of men had 3 or more components. STOP scores ranged from 0-4, and 50.1% of men endorsed 2 or more STOP components.

Discriminative ability of STOP-BANG to detect OSA

The predictive parameters for STOP-BANG scores are presented in Table 3. Using the conventional cut point of $\geq$3 on the STOP-BANG identified nearly all men in the cohort (88.4%) as 'high risk' for severe OSA, and had a sensitivity of 94.0%, a specificity of 12.7%, a PPV of 18.6% and NPV of 90.0% for the identification of severe sleep apnea (AHI$\geq$30 vs. <30). Using this cut-point, the PPV was very similar to the prevalence of severe OSA in the cohort (18.6% for PPV vs. 17.5% for prevalence of OSA), and the NPV (90.9%) was similar to the prevalence of not having OSA in the cohort (82.5%). Furthermore, the area under the curve [AUC (95% CI)] using a STOP-BANG score of $\geq$3 was 0.53 (0.52-0.55).

Raising the cut point to 7 decreased the sensitivity to 6.4% and specificity increased to 97.4%. Furthermore, the PPV increased from 18.6% to 34.4% (an increase of 15%), and the NPV decreased from 90.9% to 83.1% (7.6% decrease). Lastly, the AUC (95% CI) at a cut point of 7 was 0.52 (0.51-0.53) (Table 3).

In general, results were similar when the definition of OSA was made less stringent (AHI$\geq$5: Any sleep apnea; and AHI$\geq$15: Moderate/severe sleep apnea) (Table 3).

The area under the ROC expressing STOP-BANG scores as a continuous variable was estimated to be 0.610 (95% CI 0.584-0.637) for severe OSA (AHI≥30). A plot of the ROC curve for the prediction of severe OSA is presented in Figure 4.

Discriminative ability of STOP-BANG components to detect OSA

Secondary analyses examined the predictive parameters of the individual STOP-BANG components, and those results are presented in Table 4. Overall, **S**nore loudly, **T**ired, **O**bserved apneas, **P**ressure (hypertension), **B**MI and **N**eck circumference had sensitivities of 49.7%, 49.5%, 26.3%, 57.1%, 8.0% and 52.5% respectively, and specificities of 59.9%, 55.1%, 80.3%, 51.5%, 97.3% and 65.2% respectively. With the exception of **B**MI, PPV's did not differ dramatically from the prevalence of OSA in the cohort (PPV's ranged from 18.9% to 24.2% vs. prevalence=17.5%) and NPV's did not differ from the prevalence of not having OSA in the cohort (range 83.3-86.6 vs. prevalence=82.5%). There was some evidence that BMI>35 kg/m^2 had a moderate ability to detect severe OSA, with a PPV of 38.8% (21% increased predictive probability), but BMI≤35 had limited ability to identify men without sleep apnea (NPV=83.3% and prevalence of not having sleep apnea=82.5%).

In general, results were similar when the definition of OSA was made less stringent (AHI≥5: Any OSA; and AHI≥15: Moderate/severe OSA) (Table 4).

Examination of Alternative Cut Points for STOP-BANG Components

For each component, we examined alternative cut points and observed that results were generally similar, with some cut points demonstrating small, but insignificant gains in performance. Results are presented in Table 5. For example, increasing the cut point for **S**nore loudly from any loud snoring or coughing per week (>0 per week vs. none) to frequent snoring (≥3 times per week vs. <3 times per week) increased the specificity (59.9% to 86.8%) and also resulted in a small increase in PPV (20.8% to 26.0%). A similar pattern was

observed for **O**bserved apneas (specificity increased by 14.6% and PPV increased by 6.3%).

Since the **T**ired component in the cohort consisted of two questions assessing tiredness in the morning and afternoon, we specifically examined each question, and resultant levels separately. Although the prevalence of extreme difficulty in the afternoon due to being tired was only 1.1%, it had the highest specificity (99.0%) and PPV (25.0%), yet missed most cases of severe OSA (sensitivity=1.6%).

In primary analyses, we defined **P**ressure as a self-report of physician diagnosis of hypertension, and in sensitivity analyses we examined varying definitions of hypertension based on systolic blood pressure cut points and medication use, and results were similar to those of the primary analyses.

Despite having a low prevalence (3.6%), results suggested that **B**MI$\geq$35 kg/m^2 had the greatest ability to identify men with OSA as compared with alternative cut points (PPV=38.8%), and **B**MI$\geq$25 kg/m^2 had the greatest ability to identify men who did not have OSA (NPV=89.9%). More specifically, lowering the **B**MI cut point to $\geq$30 kg/m^2 and $\geq$25 kg/m^2 resulted in increased sensitivity (5.2%, 27.4% and 79.4% for cut points of $\geq$35 kg/m^2, $\geq$30 kg/m^2 and $\geq$25 kg/m^2 respectively) and reduced specificity (97.6%, 84.9% and 37.3% respectively). We also examined AUC and 95% confidence intervals using continuous BMI, and results are presented in appendix D. The area under the ROC curve for continuous BMI was observed to be 0.634 (95% CI 0.607-0.661), which was higher than results for continuous STOP-BANG scores (AUC for continuous STOP-BANG score was 0.610 (95% CI 0.584-0.637).

Neck circumference$\geq$40 cm had a sensitivity of 52.5%, specificity of 65.2%, PPV of 24.2% and NPV of 86.6% in primary analyses (Table 5). Results examining lower cut points based on quartiles resulted in increased sensitivity (65.7% and 83.6% for cut points of $\geq$39.2 cm and $\geq$37.5 cm respectively), and decreased specificity (53.1% and 27.4% for cut points of $\geq$39.2 cm and $\geq$37.5 cm respectively). Furthermore, a cut point of $\geq$40 cm was associated with the

greatest PPV (24.2%, vs. 22.9% and 19.6% for cut points of ≥40.0 cm, ≥39.2 cm and ≥37.5 cm respectively) and the lowest NPV (86.6%, 88.0% and 88.7%). Results of ROC curve analyses for the prediction of severe OSA using continuous neck circumference are presented in appendix D. The area under the ROC curve for continuous neck circumference was 0.617 (95% CI 0.591-0.644).

Although **a**ge ≥76 vs. <76 years and ≥80 vs. 80 years did not result in meaningful improvements in identification of men with and without OSA (Table 6), we examined AUC analyses using age expressed as a continuous variable. Results suggested that the AUC (95% CI) for the prediction of severe OSA for continuous age was 0.563 (95% CI 0.535-0.590) (Appendix D).

Evaluation of the STOP Questionnaire for the Detection of Severe OSA

We also examined the discriminative ability of the STOP questionnaire to identify older men in the cohort with severe OSA, and these results are presented in Appendix M. The standard cut-point of ≥2 on the STOP questionnaire identified half (50.1%) of the cohort as having a high risk for OSA; this cut-point had a sensitivity of 60.1%, a specificity of 52.0% a PPV of 21.0% and NPV of 86.0% for the identification of severe OSA (AHI≥30 vs. <30). Using this cut-point, the PPV was very similar to the prevalence of severe OSA in the cohort (21.0% vs. 17.5% for PPV vs. prevalence of OSA respectively), and the NPV was similar to the prevalence of not having OSA in the cohort (86.0% vs. 82.5% for NPV vs. prevalence of not having OSA respectively). Furthermore, the area under the curve [AUC (95% CI)] using a STOP score of ≥2 was 0.56 (0.54-0.59) (Appendix M).

Raising the cut point to 4 decreased the sensitivity to 6.6% and specificity increased to 95.9%. Furthermore, the PPV increased from 21.0% to 25.6%, and the NPV decreased from 86.0% to 82.9%. Lastly, the AUC (95% CI) at a cut point of 4 was 0.56 (0.54-0.58).

In general, results were similar when the definition of OSA was made less stringent (AHI≥5: Any sleep apnea; and AHI≥15: Moderate/severe sleep apnea) (Appendix M).

The area under the ROC curves expressing STOP as a continuous variable were estimated to be 0.575 (95% CI 0.549-0.602) for severe OSA (AHI≥30) (results not shown). A plot of the ROC curve for continuous STOP score and the identification of severe OSA is presented in Appendix N.

Association between OSA (defined using STOP-BANG) vs. AHI, and Excessive Daytime Sleepiness

Results examining the association between OSA and excessive daytime sleepiness are presented in Table 6. Despite varied prevalence of OSA, in general results were similar, albeit stronger, when OSA was defined using STOP-BANG criteria rather than AHI. In multivariable-adjusted models, men with a STOP-BANG score ≥3 had a 71% greater odds of having excessive daytime sleepiness than men with STOPBANG scores <3 (OR=1.71, 95% CI 1.09-2.69). Similarly, men with a STOPBANG score ≥7 had a nearly 2.2-fold greater odds of excessive daytime sleepiness than men with STOPBANG scores <7 (OR=2.18, 95% CI 1.31-3.61). Finally, men with AHI ≥30 had a 41% increased odds of excessive daytime sleepiness than men with AHI<30 (OR=1.41, 95% CI 1.07-1.88).

Discussion

The objective of this study was to examine how well the STOP-BANG questionnaire identified community-dwelling older men with severe OSA, and results suggest that while the STOP-BANG had a high sensitivity at a cut point of ≥3 (94%), it also had an unacceptably high false positive rate (87.3%) in this population. Furthermore, use of higher STOP-BANG cut points resulted in high specificities (97.4% at a cut point of ≥5) but unacceptably high false negative rates (93.6%). Furthermore, the STOP-BANG had little impact on probability

revision beyond the prevalence of OSA in the cohort. Therefore, we conclude that the STOP-BANG is insufficient for the screening of OSA in older adults.

The results of our study also suggested that in this population individual STOP-BANG components including **S**noring, **T**ired, **O**bserved apneas, **P**ressure and **A**ge were not strongly predictive of OSA. Furthermore, the area under the ROC curve demonstrate that **b**ody mass index and **n**eck circumference had fair discriminatory power for detecting severe OSA in older men, with an AUC (95% CI) of 0.634 (95% CI 0.607-0.661) for **B**MI, and 0.617 (95% CI 0.591-0.644) for **n**eck circumference. These results suggest that **B**MI and **n**eck circumference individually are statistically slightly better, but not clinically meaningfully different at identifying men with sleep apnea than the STOP-BANG questionnaire in this population (AUC, 95% CI for the detection of severe OSA was 0.610, 0.584-0.637 for STOP-BANG).

We also observed significant cross-sectional associations between OSA and excessive daytime sleepiness. Results were stronger when OSA was defined using the STOP-BANG, as opposed to AHI. These results suggest that while the STOP-BANG may not be particularly useful for detecting OSA, it is strongly associated with excessively sleepiness, and may be picking up on other sleep-related disorders in addition to OSA.

Our results were weaker, albeit overall similar to those of prior studies in selected and middle-aged populations that evaluated the performance of the STOP-BANG questionnaire, including those from the Sleep Heart Health study (SHHS)[51]. Focusing only on the conventional cut point of $\geq$3, the SHHS observed a sensitivity and specificity of 70.4%, 59.5% respectively for the detection of severe OSA, which varied from our results of 94.0% and 12.7%. The prevalence of severe OSA was also lower in the SHHS compared to MrOS Sleep (7.2% vs. 17.5%). While both cohorts consist of community-dwelling adults, participants in the SHHS are on average over 10 years younger than MrOS Sleep participants (mean age=62.4 y in SHHS vs. 76.4 y in MrOS Sleep). Prior studies have suggested that risk factors for OSA are weaker in older, as

opposed to middle-aged adults[18]. While the SHHS used more stringent criteria to define OSA (4% desaturations), we performed sensitivity analyses using this criteria and results were unchanged. Furthermore, the SHHS did not assess the performance of the STOP-BANG at alternative cut points.

The performance of the STOP-BANG has been previously assessed in various selected populations. In a population of 177 preoperative patients, Chung et. al.[50] observed a very high sensitivity (100%), and specificity of 37% for the identification of patients with severe sleep apnea (AHI $\geq$30). While the cohort used for Chung's analyses was generally younger than MrOS Sleep (mean age=55 y) the prevalence of severe OSA was approximately similar (22.0% vs. 17.5% in MrOS Sleep). It is difficult to know how participant characteristics compare between the MrOS Sleep and the Chung 2008 study, as participants included in the Chung study were a convenience sample of patients scheduled to undergo elective surgical procedures at two Toronto hospitals.

The high sensitivity of the STOP-BANG questionnaire (at a cut point of $\geq$3) might seem to make it useful for ruling out OSA for negative test results. However, its ability as a rule-out tool is greatly hampered by its high false positive rate, indicating that the test identifies nearly everyone as having OSA. Conversely, at higher cut points, the high specificity of the STOP-BANG appears to make it a useful tool for ruling in OSA, but again, this is hampered by its high false negative rate that results in many missed cases. Overall, the questionnaire appears to be little better than chance at discriminating OSA in an aged male population.

Currently, the AASM guidelines[69] do not suggest using any screening tool for OSA, beyond simpler polysomnograms, and instead state that anyone whom OSA is suspected should undergo a comprehensive sleep evaluation. While other screening tools exist, they contain the same general information and there is little evidence to suggest that any of them work any better than STOP-BANG[56, 57]. Therefore, more research is needed to better understand the development of

OSA and risk factors: especially in older adults, as well as cost effective screening tools that can more accurately identify those with sleep apnea.

This study has several strengths, including a large sample size, use of validated measures of sleep apnea (in-home polysomnography) and self-reported sleep (ESS). Furthermore, participants were not selected on the basis of a sleep, or any other disorder or medical condition.

There are also several important limitations to consider. The STOP-BANG questionnaire was created based on similar questions collected in the MrOS Sleep study and it is possible that participants completing the questionnaire itself might respond differently. However, most questions used in MrOS were also based on self-report from the participant, and so discrepancies are unlikely. Also, results are limited to older, Caucasian, generally healthy and highly educated males. Due to the types of information collected in MrOS, we were unable to assess additional sleep-disordered breathing tools including the Berlin or ARES Questionnaires. However, given the results of the two meta analyses we previously discussed[56, 57], we would not expect these tools to perform differently or better in this population.

In summary, the STOP-BANG questionnaire has poor discriminatory ability in detecting severe OSA in this cohort of community-dwelling older men. Additional research into the risk factors and characteristics of OSA in older populations is warranted.

Tables and Figures

Figure 3. Roadmap of participants included in analyses

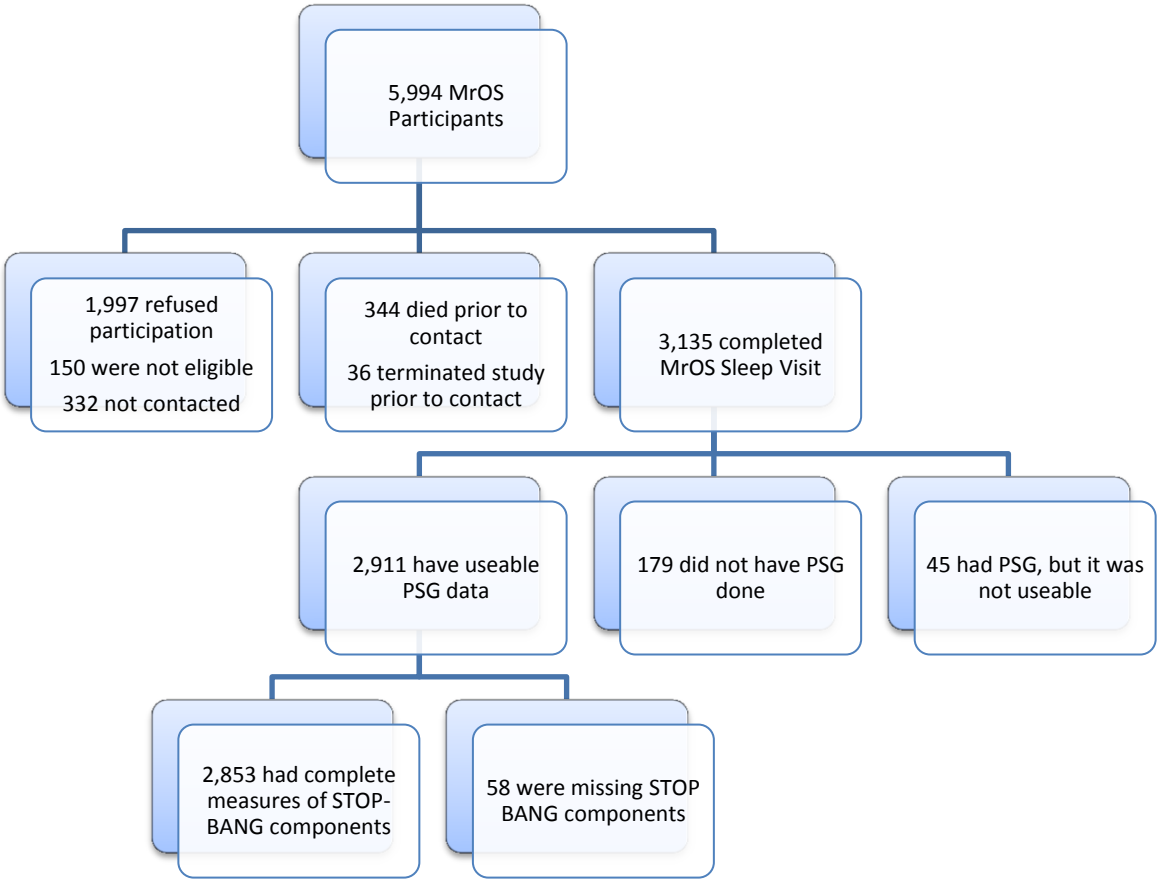


Table 1. Characteristics of MrOS Sleep cohort by severity of OSA

Characteristic	Entire Cohort (n=2853)	None AHI 0-4 (n=604)	Mild AHI 5-14 (n=1012)	Moderate AHI 15-29 (n=738)	Severe AHI≥30 (n=499)	P-value*
Age, y, mean (SD)	76.4 (5.5)	76.2 (5.6)	76.1 (5.5)	76.3 (5.4)	77.3 (5.5)	<.001
Caucasian, %	91.9	92.1	91.7	91.9	92.4	0.973
Education, %						<.001
Less than HS	5.3	4.6	4.3	6.0	7.4	
HS diploma/GED	16.3	11.4	15.0	19.5	20.0	
Some college or beyond	78.4	83.9	80.7	74.5	72.6	
Excellent/good health status, %	86.6	89.2	87.1	86.6	82.6	0.013
Smoking status, %						0.435
Current	2.0	2.5	2.3	1.8	1.2	
Former	58.6	56.1	60.3	58.6	57.9	
Never	39.5	41.4	37.5	39.6	40.9	
Physical Activity score, mean (SD)	145.6 (71.2)	144.3 (71.0)	148.1 (72.5)	148.9 (68.7)	137.4 (71.9)	0.023
Body Mass Index, mean (SD)	27.2 (3.8)	25.7 (3.2)	27.0 (3.7)	27.6 (3.7)	28.8 (4.1)	<.001
Neck circumference, cm, mean (SD)	39.4 (2.8)	38.5 (2.6)	39.3 (2.7)	39.8 (2.7)	40.4 (2.9)	<.001
Excessive daytime sleepiness (Epworth >10), %	13.0	11.4	12.3	13.0	16.4	0.071

Comorbid conditions						
Diabetes, %	13.2	9.4	12.9	13.8	17.2	0.002
Myocardial infarction, %	17.5	15.2	18.5	17.3	18.6	0.351
STOP-BANG Components						
Snore (loudly), %	41.8	33.1	40.8	45.0	49.7	<.001
Tired, %	45.7	45.4)	45.1	44.3	49.5	0.299
Witnessed stop breathing, %	20.9	15.7)	20.6	21.8	26.3	<.001
Hypertension, %	50.0	43.4)	48.5	52.7	57.1	<.001
BMI>30, %	3.6	1.0	3.3	3.3	8.0	<.001
Neck circumference >40 cm, %	37.9	24.5	35.3	42.7	52.5	<.001
STOP total score, mean (SD)	1.6 (1.1)	1.4 (1.1)	1.5 (1.0)	1.6 (1.1)	1.8 (1.1)	<.001
STOP-BANG total score, mean (SD)	4.0 (1.3)	3.6 (1.2)	3.9 (1.2)	4.1 (1.3)	4.4 (1.3)	<.001

SD=Standard Deviation; HS=High School, PASE= Physical Activity Scale for the Elderly; SDB= Sleep Disordered Breathing; AHI= Apnea Hypopnea Index.

*P-values were computed using ANOVA, and have 3 d.f.

Table 2. Frequency of STOP and STOPBANG Components

Number of components	STOP-BANG N (%)		STOP N (%)	
	N (%)	Cumulative N (%)	N (%)	Cumulative N (%)
0	-	0 (0)	482 (16.9)	482 (16.9)
1	-	0 (0)	942 (33.0)	1424 (49.9)
2	330 (11.6)	330 (11.6)	839 (29.4)	2263 (79.3)
3	765 (26.8)	1095 (38.4)	461 (16.2)	2724 (95.5)
4	804 (28.2)	1899 (66.6)	129 (4.5)	2853 (100)
5	586 (20.5)	2485 (87.1)	-	-
6	275 (9.6)	2760 (96.7)	-	-
7	85 (3.0)	2845 (99.7)	-	-
8	8 (0.3)	2853 (100)	-	-

Table 3. Predictive parameters of Different STOP-BANG score cut-offs

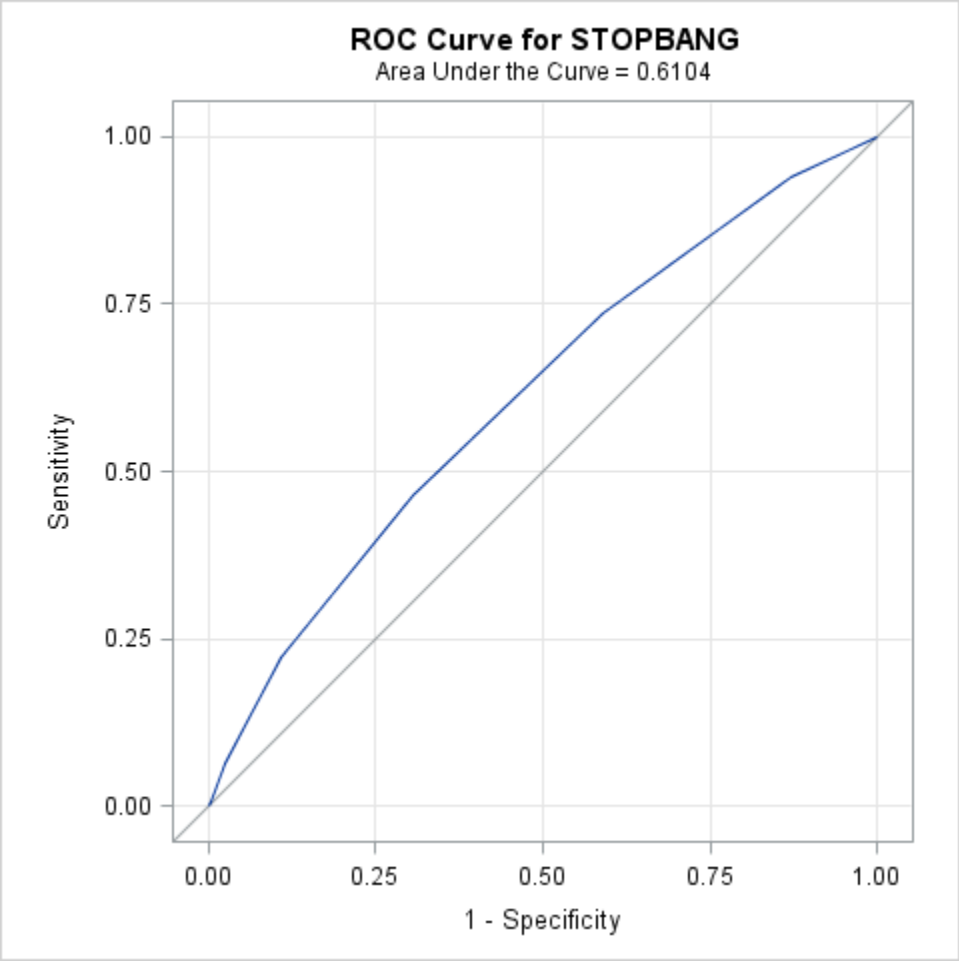
STOP-BANG score cut-off	Prevalence of OSA (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR-	AUC (95% CI)
Severe OSA (AHI≥30) vs. Moderate/Mild/None (AHI<30)*								
≥3 vs. 2	17.5	94.0	12.7	18.6	90.9	1.08	0.47	0.53 (0.52-0.55)
≥4 vs. 0-3		73.6	40.9	20.9	88.0	1.24	0.65	0.57 (0.55-0.59)
≥5 vs. 0-4		46.5	69.3	24.3	85.9	1.52	0.77	0.58 (0.56-0.60)
≥6 vs. 0-5		22.0	89.0	29.9	84.4	2.01	0.88	0.56 (0.54-0.57)
≥7 vs. 0-6		6.4	97.4	34.4	83.1	2.47	0.96	0.52 (0.51-0.53)
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)*								
≥3 vs. 2	43.4	91.8	14.1	45.0	69.1	1.07	0.58	0.53 (0.52-0.54)
≥4 vs. 0-3		67.6	43.0	47.6	63.4	1.18	0.75	0.55 (0.53-0.57)
≥5 vs. 0-4		40.3	71.8	52.3	61.1	1.43	0.83	0.56 (0.54-0.58)
≥6 vs. 0-5		18.0	91.0	60.6	59.2	2.01	0.90	0.55 (0.53-0.56)
≥7 vs. 0-6		5.1	98.1	67.7	57.5	2.74	0.97	0.52 (0.51-0.52)
Any OSA (AHI≥5) vs. None (AHI<5)*								
≥3 vs. 2	78.8	90.2	18.2	80.4	33.3	1.10	0.54	0.54 (0.53-0.56)
≥4 vs. 0-3		64.8	50.2	82.9	27.7	1.30	0.70	0.57 (0.55-0.60)
≥5 vs. 0-4		35.8	75.3	84.4	24.0	1.45	0.85	0.56 (0.54-0.58)

≥6 vs. 0-5		14.7	93.9	90.0	22.8	2.40	0.91	0.54 (0.53-0.56)
≥7 vs. 0-6		4.0	99.3	95.7	21.7	5.98	0.97	0.52 (0.51-0.52)

*AHI defined using ≥3% oxygen desaturation criteria

PPV=Positive predictive value; NPV=Negative predictive value; LR+=Likelihood Ratio for positive test; LR-=Likelihood ratio for a negative test; AUC=Area under the curve; CI= Confidence Intervals; OSA=Obstructive Sleep Apnea; AHI=Apnea Hyponea Index

Figure 4. ROC Curve Results for Continuous STOP-BANG Scores and Severe OSA.



Straight line depicts no discriminative ability (i.e. Area under the ROC Curve= 0.5). This figure models the area under the ROC curve for continuous STOP-BANG scores and severe OSA. The AUC (95% CI) across all cut points was 0.610 (0.584-0.637).

Table 4. Predictive parameters of STOP-BANG Components

STOP-BANG score cut-off	Prevalence (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR-	AUC (95% CI)
Severe OSA (AHI≥30) vs. Moderate/Mild/None (AHI<30)								
Snore loudly	41.8	49.7	59.9	20.8	84.9	1.24	0.84	0.55 (0.52-0.57)
Tired	45.7	49.5	55.1	18.9	83.7	1.10	0.92	0.52 (0.50-0.55)
Observed apneas	20.9	26.3	80.3	22.0	83.7	1.33	0.92	0.53 (0.51-0.55)
Hypertensive	50.0	57.1	51.5	20.0	85.0	1.18	0.83	0.54 (0.52-0.57)
BMI>35 kg/m^2	3.6	8.0	97.3	38.8	83.3	3.00	0.95	0.53 (0.51-0.54)
Neck circum>40 cm	37.9	52.5	65.2	24.2	86.6	1.51	0.73	0.59 (0.56-0.61)
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)								
Snore loudly	41.8	46.9	62.1	48.6	60.4	1.24	0.86	0.54 (0.53-0.56)
Tired	45.7	46.4	54.8	44.0	57.2	1.03	0.98	0.51 (0.49-0.52)
Observed apneas	20.9	23.6	81.3	49.1	58.2	1.26	0.94	0.52 (0.51-0.54)
Hypertensive	50.0	54.5	53.4	47.2	60.5	1.17	0.85	0.54 (0.52-1.56)
BMI>35 kg/m^2	3.6	5.2	97.6	62.1	57.4	2.14	0.97	0.51 (0.51-0.52)
Neck circum>40 cm	37.9	46.7	68.8	53.3	62.7	1.49	0.78	0.58 (0.56-0.59)
All OSA (AHI≥5) vs. None (AHI<5)								
Snore loudly	41.8	44.2	66.9	83.2	24.3	1.33	0.83	0.56 (0.53-0.58)

Tired	45.7	45.8	54.6	79.0	21.3	1.01	0.99	0.50 (0.48-0.52)
Observed apneas	20.9	22.2	84.3	84.0	22.5	1.41	0.92	0.53 (0.52-0.55)
Hypertensive	50.0	51.8	56.6	81.6	24.0	1.19	0.85	0.54 (0.52-0.56)
BMI>35 kg/m^2	3.6	4.3	99.0	94.2	21.8	4.34	0.97	0.52 (0.51-0.52)
Neck circum>40 cm	37.9	41.5	75.5	86.6	25.3	1.69	0.77	0.59 (0.57-0.61)

*AHI defined using 3% desaturation criteria

PPV=Positive predictive value; NPV=Negative predictive value; LR+=Likelihood Ratio for a positive test; LR-=Likelihood Ratio for a negative test; AUC= Area under the curve; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index; BMI=Body Mass Index

Table 5. Examination of alternative cut-points for STOP-BANG components for the detection of severe OSA

STOP-BANG Components	Prevalence (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+ (%)	LR- (%)
Snore loudly							
>0 per wk vs. none*	41.8	49.7	59.9	20.8	84.9	1.24	0.84
≥1x per wk vs. <1x per wk	27.1	34.5	74.5	22.3	84.3	1.35	0.88
≥3x per wk vs. <3x per wk	14.7	21.8	86.8	26.0	84.0	1.66	0.90
Tired							
Tired in the morning:							
Extreme/moderate/mild difficulty vs. none	17.7	21.8	83.2	21.6	83.4	1.30	0.94
Moderate/extreme difficulty vs. mild/none	3.7	5.0	96.7	23.6	82.7	1.46	0.98
Extreme difficulty vs. moderate/mild/none	1.1	1.4	99.0	23.3	82.6	1.44	1.0
Tired in the afternoon:							
Extreme/moderate/mild difficulty vs. none	41.8	44.9	58.9	18.8	83.5	1.09	0.94
Moderate/extreme difficulty vs. mild/none	6.8	8.8	93.6	22.6	82.9	1.37	0.97
Extreme difficulty vs. moderate/mild/none	1.1	1.6	99.0	25.0	82.6	1.57	0.99
*Tired: Any difficulty in morning Or afternoon**	45.7	49.5	55.1	18.9	83.7	1.10	0.92
Observed apneas							
>0 per wk vs. none*	20.9	26.3	80.3	22.0	83.7	1.33	0.92

≥1x per wk vs. <1x per wk	12.2	16.8	88.7	24.1	83.4	1.50	0.94
≥3x per wk vs. <3x per wk	5.8	9.4	94.9	28.3	83.2	1.86	0.95
Pressure							
SysBP>120 vs. ≤120	67.2	67.7	32.9	17.6	82.8	1.01	0.98
SysBP>140 vs. ≤140	19.5	19.0	80.4	17.1	82.4	0.97	1.01
SysBP>160 vs. ≤160	3.9	5.6	96.5	25.2	82.8	1.59	0.58
Any HTN Medication use	50.1	61.3	52.3	21.4	86.5	1.29	0.74
ARB user vs. nonuser	9.6	13.6	91.2	24.7	83.3	1.55	0.95
Beta blocker user vs. nonuser	28.0	34.1	73.3	21.3	84.0	1.27	0.90
Calcium channel blocker user vs. nonuser	15.2	19.2	85.7	22.2	83.4	1.34	0.94
Diuretic user vs. nonuser	18.6	23.7	82.5	22.3	83.6	1.35	0.93
Self-reported HTN vs. none*	50.0	57.1	51.5	20.0	85.0	1.18	0.83
BMI, kg/m^2							
≥25 vs. <25	70.0	82.6	32.7	20.6	89.9	1.23	0.53
≥30 vs. <30	20.4	35.1	82.7	30.0	85.7	2.02	0.79
≥35 vs. <35*	3.6	8.0	97.3	38.8	83.3	3.00	0.95
Age, years							
≥76 vs. <76	28.2	36.1	73.5	22.4	84.4	1.36	0.87
≥80 vs. <80	50.9	59.7	51.0	20.5	85.7	1.22	0.79
Neck Circumference, cm							

≥37.5 vs. <37.5	74.5	83.6	27.4	19.6	88.7	1.15	0.60
≥39.2 vs. <39.2	50.2	65.7	53.1	22.9	88.0	1.40	0.65
≥40 vs. <40 *	37.9	52.5	65.2	24.2	86.6	1.51	0.73

*****Indicates the level used in the original calculation of STOP-BANG primary analyses

PPV=Positive predictive value; NPV=Negative predictive value; LR+=Likelihood Ratio for a positive test; LR-=Likelihood Ratio for a negative test; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index; BMI=Body Mass Index; wk=week; HTN=Hypertension.

Table 6. Associations between STOP-BANG, AHI and Excessive Daytime Sleepiness

	Odds Ratio of Excessive Daytime Sleepiness OR (95% CI)*		
	STOPBANG		**AHI**
	≥3 vs. <3 (n=2,523 with OSA)	**≥5 vs. <5 (n=93 with OSA)**	**≥30 vs. <30 (n=499 with OSA)**
OSA, yes vs. no	1.71 (1.09-2.69)	2.18 (1.31-3.61)	1.41 (1.07-1.88)
Age, per 5 year increase	1.01 (0.91-1.13)	1.02 (0.91-1.13)	1.00 (0.89-1.11)
Non-White vs. White	1.43 (0.98-2.08)	1.41 (0.97-2.06)	1.39 (0.95-2.02)
Current vs. never smoker	1.28 (0.61-2.70)	1.31 (0.62-2.76)	1.31 (0.63-2.76)
Former vs. never smoker	1.03 (0.82-1.30)	1.02 (0.81-1.28)	1.03 (0.82-1.29)
Diabetic	1.10 (0.80-1.51)	1.11 (0.81-1.52)	1.11 (0.80-1.52)
Myocardial Infarction	1.02 (0.76-1.36)	0.99 (0.74-1.33)	1.03 (0.77-1.37)
BMI, per 3.8 kg/m^2	1.11 (0.99-1.24)	1.08 (0.96-1.22)	1.11 (0.99-1.24)
PSQI, per unit increase	1.07 (0.99-1.24)	1.07 (1.04-1.11)	1.08 (1.05-1.11)
PASE, per 71 unit decrease	1.09 (0.97-1.23)	1.10 (0.97-1.23)	1.09 (0.97-1.23)

*Models additionally adjusted for clinic site.

Chapter 4: The epidemiology of variability in sleep/wake patterns in older adults: definitions, prevalence and correlates.

Objectives: To evaluate the within-person variability of sleep/wake parameters in older men, and to examine potential correlates of greater variability in sleep measures.

Design: Cross-sectional study

Participants: Two thousand eight hundred four men aged 67 years and older enrolled in the Outcomes of Sleep Disorders in Older Men (MrOS Sleep) cohort study with at least five nights of actigraphy data.

Measurements: Objectively measured sleep parameters from wrist actigraphy were total sleep time, sleep latency, nighttime wakefulness (wake after sleep onset), in bed timing and out of bed timing. Variability was defined as the intra-individual standard deviation in each of these measures, and greater variability was defined as being in the highest quintile of the distribution. Associations between participant characteristics and odds of being in the highest quintile of variability were assessed using logistic regression.

Results: Substantial within-person variability in sleep parameters was found among the cohort of older men. The sleep parameters with the greatest amount of within-person variability were sleep latency and nighttime wakefulness; mean (SD) variability in sleep latency, nighttime wakefulness, total sleep time, in bed timing and out of bed timing were 24 (26); 27(20); 47(24); 37(25) and 39(26) minutes respectively. Several characteristics were associated with greater variability, with the strongest factors being African American race, living alone, smoking, antidepressant use, benzodiazepine use, depression, greater body mass index and greater comorbidity burden. Lower education, heavy alcohol use, not getting up at night to use the bathroom, cognitive impairment and greater impairments in Instrumental Activities of Daily living were also associated with at least one measure of variability.

Conclusion: Significant within-person variability exists in older, community dwelling men, and is associated with several potentially modifiable demographic and health-related factors. Future prospective research studies should examine whether variability in sleep-wake patterns is associated with risk of health outcomes among older adults, and if so, whether or not interventions aimed at reducing variability in sleep improve health outcomes.

Introduction:

Little is known about the intra-individual variability of sleep, either among the general population, or more specifically in older adults who often complain of sleep disturbances. However, establishing regular stable sleep patterns is a critical component of behavioral therapy for insomnia[11, 70], and night-to-night variability has been used as a measure of adherence to a treatment regimen[71].

There are many reasons why it is important to advance understanding of individual variability in measures of sleep-wake patterns. First, greater variability in night-to-night sleep may promote the development of insomnia. If variable sleep patterns result in compensatory behaviors (i.e. catching up on sleep by staying in bed longer), then over time these behaviors may lead to conditioned arousal, in which the individual has difficulty falling asleep or staying asleep in a normal sleep environment, and eventually may result in insomnia[72-75].

Second, variability in sleep may mask the association of sleep measures such as sleep duration with health-related outcomes, including the risk of mortality. Several studies have examined the association between sleep duration and mortality, and have reported inconsistent findings[52, 76-80]. If variability in sleep duration is independently associated with mortality, and this hypothesis has not been studied, then variable sleep/wake patterns may be one explanation for the inconsistent findings across studies.

Self-reported sleep data and objective parameters of sleep-wake patterns measured using actigraphy are typically collected over several consecutive nights in research studies of sleep. The traditional analytical approach is to characterize sleep/wake parameters as aggregate means which omits information on the night-to-night variability of sleep, However, measures of variability may provide additional information on disturbed sleep, especially in those individuals whose aggregate means appear to be normal, but who otherwise complain of sleep-related disturbances.

Only a few studies have directly assessed variability of sleep measures. These studies have been limited to younger populations[81], by small sample sizes[82], or by selection on the basis of insomnia[83]. None of the aforementioned

studies have examined potential correlates of variability. In a study of 669 participants aged 38-50 years enrolled in the CARDIA study[81], investigators examined the within-subject and between-subject variability of actigraphic measures of sleep duration, sleep latency, sleep efficiency and time in bed. The authors found substantial variation in day-to-day sleep measures within individuals in the cohort. For example, the standard deviation (SD) of sleep duration was 1.26 hrs. within individuals and 0.70 hrs. between individuals. Likewise, the SD of sleep efficiency was 8.4% in individuals and 8.1% between individuals, and the SD of sleep latency was 30.7 minutes within individuals and 22.2 minutes between individuals.

Buysse et. al examined the variability of diary and actigraphic measures of sleep in 61 older adults with chronic insomnia and 31 controls[83]. Results from this case-control study suggested that insomniacs exhibited greater variability on most self-reported diary measures of sleep, and on actigraphy measures of awakening after sleep onset and sleep efficiency, but did not differ from controls with respect to measures of variability in objective measures of sleep duration, and sleep latency. On average, the SD of wake after sleep onset was 22.6 vs. 19.3 mins ($p<.001$); sleep efficiency was 7.0% vs. 5.6% ($p=0.003$); sleep duration was 55.6 vs. 53.9 mins ($p=0.5$); and sleep latency was 26.2 vs. 16.8 mins ($p=0.13$) in insomniacs vs. controls respectively.

In summary, the aforementioned studies have observed variability in the sleep/wake patterns of adults, but whether or not such variability exists in an unselected sample of older community-dwelling adults is uncertain. Furthermore, the associations of demographic and health-related factors with variability in sleep/wake parameters and in older adults have not been assessed.

Therefore, we conducted a cross-sectional analysis of the Outcomes of Sleep Disorders in Older Men (MrOS Sleep) study to examine the characteristics and correlates of variability in sleep/wake parameters among community-dwelling older men.

Methods:

Study Population

The Osteoporotic Fractures in Men (MrOS) study recruited 5,994 community-dwelling men between March 2000 and April 2002 at six U.S. clinical centers (Birmingham, Alabama; Minneapolis, Minnesota; Palo Alto, California; Monongahela Valley near Pittsburgh, Pennsylvania; Portland, Oregon and San Diego, California). To be eligible to participate in the MrOS Study, men had to be aged 65 years and older, able to walk without assistance and not have had a bilateral hip replacement. Recruitment details and study design have been published elsewhere[60, 61].

From December 2003 through March 2005, MrOS enrollees were invited to participate in an ancillary study to identify outcomes of sleep disorders in older men. To be eligible to participate in the Sleep study, participants had to report not using oxygen therapy in the past three months, no history of an open tracheotomy, not sleeping with a mouthpiece for snoring or sleep apnea in the past 3 months, or not sleeping with a CPAP or BiPAP mask in the last 3 months. Some exceptions were made for participants who intermittently use CPAP, or who were willing to not wear the CPAP mask during the study. A total of 3,135 (>100% recruitment goal) were enrolled in the MrOS Sleep ancillary study. Of these, 3,058 men wore a wrist actigraph and had technically adequate actigraphy data, and of these 2,804 had at least five nights of measurements and comprised the cohort for this analysis. The institutional review board at each center approved the study protocol and written informed consent was obtained from all participants.

Measurement of Sleep/Wake Parameters

Activity patterns were measured using an octagonal wrist actigraph (Ambulatory Monitoring, Inc., Ardsley, NY), which is a small device resembling a wrist watch that is worn on the wrist of the non-dominant hand. Participants were instructed to wear the actigraph continuously for a minimum of five nights (mean $\pm$ SD = 5.2 $\pm$ 0.9 nights) and to remove it only for bathing or situations in which it might get submerged in water. Actigraphs contain accelerometers that measure

and record movement in 1-minute epochs, and have been shown to provide reliable estimates of sleep-wake activity in comparison to polysomnography, which is currently the gold standard[84]. Data were collected in three modes but are reported here based on digital integration mode (also known as proportional integration mode)[85]. In the MrOS Sleep study, data collected in PIM was most correlated with PSG[86]. Actigraphy data were transferred to the San Francisco Coordinating Center for centralized processing. Centralized training and certification were also required for clinic staff gathering actigraphy data. Activity data from the actigraph was analyzed using Action W-2 software (Ambulatory Monitoring, Inc.).

In addition, participants were also asked to complete a sleep diary for the time period in which they wore the actigraphs. The diaries included information on time into and out of bed, as well as times in which the actigraph was removed. This information was used in editing the actigraphy data files. Time periods in which participants removed the actigraphs for >10% of the time during the day or for over 2 hours during the night are not included in the analyses. Interscorer reliability for editing the actigraphy data has been excellent in our group and actigraphic sleep duration has been shown to have good agreement with PSG (gold standard)[87, 88].

Several sleep/wake parameters were computed from the actigraphy data and defined as follows: In bed timing was defined as the time in which the participant reported getting into bed and trying to sleep. In general, this was based entirely off from the sleep diary, although adjustments were made in cases where the diary data clearly did not match the actigraphy (i.e. reported in bed time occurred after sleep onset). Sleep onset was defined as the time when the first 20 minute block containing >19 minutes of sleep began. Sleep latency was the number of minutes from the time the participant reported getting in bed until sleep onset. Wake after sleep onset (WASO), a measure of nighttime wakefulness, was the number of minutes scored as wake from sleep onset until the end of the last sleep episode while in-bed. Out of bed timing was the time when the participant reported getting out of bed. In cases where the out of bed

time did not match the actigraphy data (i.e. occurred during a period scored as sleep), out of bed timing was adjusted to be the last minute scored as sleep. Total sleep time was the number of minutes scored as sleep during the in-bed interval. Data for these variables was averaged over the total number of nights the actigraph was worn.

Self-Reported Sleep Measures

Participants completed the Pittsburgh Sleep Quality Index (PSQI) and the Epworth Sleepiness Scale (ESS) questionnaires at their sleep study clinic visit. The PSQI is a validated measure of subjective sleep quality and sleep disturbances over a 1-month period. The questionnaire is divided into sections that assess subjective sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbances, use of sleeping medications and daytime dysfunction. Global PSQI scores range from 0 to 21, and a standard cut-off of greater than 5 is indicative of poor sleep quality. This cutoff has a sensitivity of 89.6% and a specificity of 86.5% in distinguishing good sleepers from poor sleepers[64, 65].

The ESS is a self-administered questionnaire that classifies subjective daytime sleepiness in people with sleep disorders. Participants are asked to rate how likely (from 1 to 3, with 1 being unlikely and 3 being highly likely) they are to doze off in eight typical daily situations. Scores range from 0 to 24, with a standard cutoff of greater than 10 indicative of excessive daytime sleepiness[67, 68].

Other Measures

Additional measures were collected on the day of the sleep study clinic visit. All participants completed questionnaire data, which included questions about their medical history, current health status, smoking and alcohol use and social support. Comorbidity burden was computed as a sum of the following medical conditions: Parkinson's disease, chronic kidney disease, chronic obstructive pulmonary disease, stroke, diabetes, congestive heart failure, myocardial infarction, hypertension and hypercholesterolemia, and expressed as

a three-level variable (0,1, 2+ conditions). Information from the MrOS baseline visit (mean (SD) between MrOS baseline and sleep study visits 3.4 (0.5) years) was used to assess age, race and education. At the sleep study visit, body weight was measured using a standard balance beam, or digital scale and height using a wall mounted Harpenden stadiometer. BMI was calculated as kg/m^2 and was expressed as a three-level variable (<25, 25-29, 30+). Participants were asked to bring in all medications used within the past two weeks, and a computerized medication coding dictionary was used to categorize the medications[89]. The Geriatric Depression Scale was used to assess depressive symptoms, with scores $\geq$6 indicative of depression[90]. Functional status was measured using information collected on six independent activities of daily living (IADL)[91, 92] and was expressed as a three-level variable (0,1-2, 3+ impairments). Cognitive function was using the Modified Mini-Mental State examination (3MS). The 3MS is a global measurement of cognitive function, with components for orientation, concentration, praxis and immediate and delayed memory. Scores range from 0 to 100, with higher scores representing better cognitive functioning. A cut point of >80 is indicative of cognitive impairment[93]. Finally, men were asked to record the number of times they typically get up during the night to use the bathroom (0, 1,2,3,4, 5 or more), and was expressed as a 4-level categorical variable (0, 1-3, 4 or 5 or more) based on its distribution.

Statistical Analysis

Our analytic cohort consists of 2,804 men who had at least 5 nights of actigraphy data (91.7% of participants with useable actigraphy data). Differences between the analytic cohort and the remaining surviving MrOS population (N=1,360) were examined using t-tests for normally distributed continuous variables, Wilcoxon rank sum tests for skewed continuous variables and chi-square tests for categorical variables.

The within-person standard deviation (SD) was calculated for each sleep/wake parameter using data from each night that the actigraph was worn. Higher SD values indicate greater variability in sleep/wake patterns over the

measured time period. The coefficient of variation (CV) for each sleep/wake parameter was also computed as an alternative measure of variability, and it was calculated as the between-person standard deviation divided by the overall between-person mean.

Logistic regression models were used to examine associations between potential predictor variables and odds of greater variability (being in the highest quintile vs. quintiles 1-4). We first analyzed the data using models adjusted for age and clinic site, and then a multivariable adjusted model that included all predictors. Additionally, to examine whether the association between predictors and variable sleep/wake parameters was explained by sleep duration, we added mean total sleep time to multivariable adjusted models. Finally, multivariable adjusted analyses were repeated expressing variability measures as continuous variables and linear regression was used to examine these associations.

For models examining associations with sleep latency variability and in-bed timing variability, additional sensitivity analyses were performed adjusting for the quality of the self-reported time to bed. Since quality of self-reported in bed timing did not alter findings, these analyses were not reported.

All analyses were conducted using SAS version 9.3 (SAS Institute Inc., Cary, North Carolina).

Results

The mean (SD) age of the cohort was 76.3 (5.5) years, and 3.7% were African American. Characteristics of the cohort are provided in Table 7. Of the 5,994 men originally recruited for the MrOS Study, 1,830 men died prior to the MrOS Sleep visit. Compared to the 1,360 MrOS men who were alive, but were not included in the analytic cohort, men in the analytic cohort tended to be younger (72.9 vs. 76.0 years for analytic vs. other cohort, $p<.001$), slightly more educated (78.1% vs. 72.1% attended college, $p<.001$), slightly more likely to live with others (87.3% vs. 82.9%, $p<.001$), and reported fewer impairments in Instrumental Activities of Daily Living (0.4 vs. 0.6 mean IADL impairments, $p<.001$), but did not differ on Caucasian race (91.3% vs. 92.2%, $p=0.322$).

The distribution of actigraphic sleep/wake variables including mean (SD), median, interquartile range, coefficient of variation and within-subject standard deviation are provided in Table 8. In examining the sleep/wake variables expressed as averages across the nights in which the participant wore the actigraph (mean $\pm$ SD = 5.3 $\pm$ 0.7 nights), the greatest amount of variability was observed for sleep latency (CV=104.7%), and the least amount for time to bed (CV=5.0%).

Greater variability was observed in between-persons sleep/wake measures rather than within-person, suggesting that nightly sleep measures are more similar within individuals (Table 8). For example, the between-persons standard deviation in total sleep time was 1.2 hours (or 72 minutes), and the mean within-persons standard deviation was 47 minutes. Similar patterns were observed across all sleep/wake measures.

Independent predictors of greater within person variability in sleep/wake parameters (Tables 9 & 10)

Demographics – *Age:* In age and site adjusted models (Table 9), older age was associated with increased odds of highly variable nighttime wakefulness, and lower odds of highly variable in and out of bed timing. With the exception of out of bed timing, these associations persisted in multivariable adjusted models (Table 10). Compared to men aged 67-69 (referent group), men aged 80+ had a 58% increased odds of highly variable nighttime wakefulness (being in the highest quintile of variability vs. the lower four quintiles), and a 56% lower odds of highly variable in bed timing in multivariable adjusted models. Furthermore, compared to the referent group, men aged 75-79 had 38% lower odds of highly variable out of bed timing. Age was not associated with variability in total sleep time or sleep latency in either age and site, nor multivariable adjusted models.

Race: Compared to Caucasians (referent group), African Americans had an increased odds of greater variability in all sleep/wake parameters, and men

who identified themselves as 'other' race had an increased odds of greater variability in in-bed timing, in age and site adjusted models. These associations persisted in multivariable-adjusted models. African Americans had 1.6 to 3.3-fold increased odds of being in the highest quintile of variability across all sleep/wake measures than Caucasian men in multivariable adjusted models. Furthermore, men who were 'other' race had 1.7-fold increased odds of greater variability in in-bed timing than Caucasian men.

Education: Compared to men who attended college, men who did not complete high school had increased odds of greater variability in total sleep time, nighttime wakefulness and out of bed timing, whereas men who finished high school vs. those who attend college did not have an increased odds of greater variability in any sleep/wake parameter in age and site adjusted models. In multivariable adjusted models, men who had less than a high school education had a 1.5-fold increased odds of highly variable total sleep time than men who attended college, but associations with nighttime wakefulness and out of bed timing no longer reached statistical significance.

Lifestyle- *Social Support:* Living with others, as opposed to living alone, was associated with increased odds of greater variability across all sleep/wake parameters in age and site adjusted models. In multivariable adjusted models however, living with others was associated with a 43-84% increased odds of greater variability in total sleep time, nighttime wakefulness, sleep latency and in bed timing, but was not independently associated with variable out of bed timing.

Smoking: Past and current smoking was also associated with increased odds of greater variability in nearly all sleep/wake parameters in age and site adjusted models. In multivariable adjusted models, current smokers (vs. never smokers) had a 97% increased odds of greater variability in in-bed timing. However, despite not reaching statistical significance, the magnitude of other point estimates suggested that current smoking might be associated with all variability measures (range of OR=1.5 to 1.8). In multivariable adjusted models, former smoking was not associated with variability measures.

Alcohol consumption: Greater alcohol consumption (14+ drinks per week vs. 0-1) was associated with 73% increased odds of greater variability in nighttime wakefulness, but was not associated with other measures of variability in age and site adjusted models. In multivariable adjusted models, this association was slightly attenuated, but otherwise remained statistically significant (OR=1.62).

Medications- *Antidepressant use:* Men who reported current use of antidepressants had about a 2-fold greater odds of highly variable total sleep time, sleep latency and out of bed timing in age and site adjusted models. In multivariable adjusted models, these associations were slightly attenuated (1.7 to 1.8-fold) but remained statistically significant. Antidepressant use was not independently associated with variability in nighttime wakefulness and in-bed timing in any models.

Benzodiazepine use: Men who reported using benzodiazepines had a 1.9 to 2-fold increased odds of greater variability in total sleep time, sleep latency, in bed timing and out of bed timing in age and site adjusted models. Benzodiazepine use was not associated with variability in nighttime wakefulness in age and site adjusted models. Multivariable adjusted models associations were slightly attenuated, but benzodiazepine use remained associated with a 1.6 to 1.8-fold increased odds of greater variability in total sleep time, in bed timing and out of bed timing. Benzodiazepine use was not independently associated with night time wakefulness or sleep latency variability in multivariable adjusted models.

Non-benzodiazepine anxiolytic hypnotic use: Although non-benzodiazepine anxiolytic hypnotic use was not significantly associated with variability in sleep/wake parameters, the magnitude of point estimates suggest that non-benzodiazepine anxiolytic hypnotic use may be associated with an increased odds of greater variability in sleep latency, total sleep time and in-bed timing (34-53% increased odds in age and site adjusted models and 32-64% increased odds in multivariable adjusted models).

Anthropometric and physical functioning- *Body Mass Index:* Higher BMI was associated with increased odds of greater variability in all sleep/wake parameters in age and site adjusted models. In multivariable adjusted models compared to men with a body mass index (BMI) less than 25 kg/m^2 (referent group), men with a BMI of 25-29 kg/m^2, had a 1.3 to 1.4-fold increased odds of greater variability in total sleep time, nighttime wakefulness and out of bed timing, whereas men with a BMI 30+ kg/m^2 had a 1.3 to 2.6-fold increased odds of greater variability in total sleep time, nighttime wakefulness, sleep latency and in bed timing.

Impairments in Instrumental Activities of Daily Living (IADLs): Greater burden of IADLs were associated with a 1.4 to 2.5-fold increased odds of greater variability in all sleep/wake parameters in age and site adjusted models, although associations with out of bed timing did not reach statistical significance. In multivariable adjusted models, associations were attenuated, and having 2 or more IADL impairments (vs. none) was associated with a 74-77% increased odds of greater variability in sleep latency and in-bed timing. IADL impairments were not independently associated with variability in total sleep time, night time wakefulness or out of bed timing in multivariable adjusted models.

Health related factors- *Self-reported health status:* Compared to men with excellent or good health, having a fair, poor or very poor health status was associated with a 1.3 to 1.7 fold increased odds of greater variability in total sleep time, nighttime wakefulness, sleep latency and out of bed timing, but was not associated with variability in in bed timing. In multivariable adjusted models, these associations were attenuated and no longer reached statistical significance.

Number of times up to use bathroom: In age and site adjusted models, men who reported getting up 5 or more times per night to use the bathroom had a 55% increased odds of greater out of bed variability than men who reported getting up 1-3 times at night to use the bathroom. Furthermore, men who

reported not getting up to use the bathroom had a 50% lower odds of greater variability in nighttime wakefulness than men who reported getting up 1-3 times per night. In multivariable adjusted models, this latter association remained statistically significant, but all other associations failed to reach statistical significance.

Depression: Being depressed as defined by a GDS score >6 was associated with 2-fold increased odds of greater variability across all sleep/wake parameters in age and site adjusted models. In multivariable adjusted models, these associations were slightly attenuated, but remained statistically significant (1.5 to 1.7-fold increased odds).

Cognitive impairment: Cognitive impairment (MMSE<80) was associated with a 1.3 to 1.8-fold increased odds of greater variability in all sleep/wake parameters, but statistical significance was only reached for associations with variability in sleep latency in age and site adjusted models. In multivariable adjusted models, none of the associations were statistically significant, but the magnitude of the point estimates suggested that cognitive impairment may be associated with a 58% increased odds of greater variability in sleep latency.

Comorbidity burden: Greater burden of comorbidity was associated with increased odds of greater variability in several measures, in age and site adjusted models. In multivariable adjusted models the strongest associations were observed for multimorbidity. Multimorbidity (2 or more comorbid medical conditions) was associated with a 1.4 to 1.7-fold increased odds of greater variability in total sleep time, nighttime wakefulness, sleep latency and out of bed timing, compared to men with no comorbid medical conditions.

Impact of Mean Total Sleep Time on Multivariable Associations

In multivariable adjusted models, mean total sleep time was an independent predictor of variability in nighttime wakefulness and sleep latency, but not for variability in total sleep time, in-bed timing and out of bed timing. Each standard deviation increase in total sleep time (SD=1.2 hrs) was associated

with a 60% reduced odds of greater variability in nighttime wakefulness and a 57% reduced odds of greater variability in sleep latency (Table 11).

The addition of mean total sleep time to multivariable adjusted models had some impact on predictors associated with greater variability in nighttime wakefulness and sleep latency. After adjusting for mean total sleep time, African American race, living alone, current smoking, BMI 25-29 kg/m^2 were no longer independently associated with variability in nighttime wakefulness. Furthermore, after additional adjustment for mean total sleep time, African American race, living alone, current smoking, BMI>30 kg/m^2, greater IADL impairments and greater burden of comorbidities were no longer independently associated with variability in sleep latency. Although mean total sleep time was not independently associated with variability in in and out of bed timing, the addition of this measure to those models resulted in associations with greater BMI no longer reaching statistical significance.

Associations with continuous variability measures

Multivariable adjusted associations between predictors and continuous sleep/wake variability measures are presented in Tables 12 and 13. In general, the statistical significance of most of the primary associations were unchanged in models expressing variability outcomes as continuous measures. However, some evidence became stronger due to increased power of expressing outcomes as continuous, rather than dichotomous variables. For example, in primary models, older age was not associated with increased odds of being in the highest quintile of variability in total sleep time. However, in models expressing total sleep time variability as a continuous measure, men aged 75-79 years had about an average of 4 minutes greater total sleep time variability than men aged 67-69 years, in multivariable adjusted models. Furthermore, associations between older age and variability in in-bed timing and out of bed timing were strengthened in analyses expressing variability as a continuous, rather than categorical variable. The magnitude of a few other associations were

slightly strengthened in these analyses, and included education, alcohol use and IADL impairments.

Associations between depression and variability in in-bed timing were slightly attenuated in models expressing in-bed timing variability as a continuous measure, and were no longer statistically significant (p=0.093).

Discussion

In this population of older community-dwelling men, these results suggest that while the majority of variability in sleep/wake parameters is derived from between-persons differences, a substantial amount of within-person variability exists. Furthermore, several demographic, anthropometric, lifestyle, medical, medication and physical functioning factors were independently associated with measures of greater sleep/wake variability. Overall, African American race, living alone, antidepressant use, depression, greater body mass index and multimorbidity were the strongest independent predictors of parameters of greater sleep/wake variability.

Compared to results of prior studies, our findings suggest that the sleep/wake patterns of older men exhibit slightly less variability than that of a younger (middle aged) cohort[14]. In comparison to the 669 CARDIA participants (mean age 42.9 y) included in this prior study, we observed slightly less variability (Coefficient of Variation (CV) for total sleep time was 21% vs. 19%, and sleep latency 136% vs. 105% for CARDIA vs. MrOS Sleep participants, respectively). Additionally, the within-person SD of sleep measures were also greater for CARDIA participants (SD total sleep time 86 vs. 47 mins; sleep latency 31 vs. 24 mins, for CARDIA vs. MrOS Sleep, respectively). The study in CARDIA participants did not assess variation in bed time, wake time or nighttime wakefulness and did not examine associations between potential predictors and variability measures.

While we are unable to assess directionality and causality in this analysis our results also suggest that there may be potentially modifiable determinants of sleep variability in older adults including smoking, depression, greater BMI,

functional impairments and antidepressant use. These findings might help direct the design of future intervention studies aimed at improving sleep among older adults.

African American race was a strong, independent predictor of greater sleep/wake variability in this cohort. These findings were similarly noted in the CARDIA study, although results were not presented. In CARDIA, investigators observed a significant race-sex group effect in models, with white males and white females having lower daily variability in sleep/wake parameters than black males and black females, though the absolute magnitude of differences across race and sex groups was not presented[94]. We are not able to assess race-sex differences in this study of older men, but we did observe that African American men have significantly more variability in sleep than Caucasian men. These findings were independent of several other measures of health, education and physical functioning. The associations of race with variability in nighttime wakefulness and sleep latency were explained in part, by the average amount of total sleep the individual attained during the study period. In our study, African American men slept an average of 28.1 minutes less per night than Caucasian men (data not shown), after adjusting for age, clinic site and education. This observation that African Americans have worse sleep than Caucasians is supported by findings of other studies[95]. These results suggest that African American men have more variability in night to night nighttime wakefulness, which may lead to an overall reduction in the average amount of sleep attained over a period of time. Future analyses are needed to confirm these findings and to evaluate the effects of specific socio-economic and lifestyle stressors that may have a greater impact on these associations in African American populations.

Living alone was also a strong, independent predictor of greater sleep/wake variability in this population, and there are several reasons that may explain this observation. While living alone might be an indication of the ability to live independently, it may also be indicative of recent bereavement or a lack of social support. Questions regarding lifestyle stressors such as bereavement in the past year, were not assessed at the sleep exam. Given the amount of time

between this assessment and the sleep exam, it is uncertain if these measures would have any meaningful impact on the associations we observed. Therefore, future research should explore these in greater detail on the extent to which lifestyle stressors may impact sleep.

Smoking may be associated with greater sleep/wake variability in older adults though associations in this study did not reach significance.. In this cohort, only 1.9% of men reported being current smokers. This low prevalence contributes to the lack of statistical power, and also makes associations a little more difficult to interpret. Currently there is no literature examining the effects of smoking on sleep/wake variability, however smoking is sleep disrupting due to the stimulating effects of nicotine, which could promote insomnia and/or variable sleep/wake patterns[96].

Antidepressant use remained a strong, independent predictor of variability in total sleep time, sleep latency and out of bed timing in multivariable adjusted models. While some antidepressants are sleep promoting, others may be sleep disturbing, and thus these associations consist of a mixture of the two possibilities. Due to a low prevalence of antidepressant use in the cohort (<8%), we are unable to examine and compare associations across different classes of antidepressants. Associations with antidepressant use were fairly similar in magnitude to those observed for depression. Given the association between depression and insomnia symptoms[97] these results raise the question of whether insomnia may be underlying this association. In our study we did not assess insomnia, and are unable to examine this possibility further.

Benzodiazepine use was also associated with greater variability in total sleep time, in bed timing and out of bed timing. We also had a low prevalence of benzodiazepine use in this cohort (4.3%) which prohibits us from exploring associations with specific subclasses of medications such as whether the association between benzodiazepine use depends on duration of action (i.e. long vs short acting agent). However, benzodiazepines have sedating and sleep-promoting effects, and their impact on sleep variability may be a direct result of their mechanistic action[98]. It is also a possibility that intermittent benzodiazepine

use may contribute to greater sleep/wake variability, but this has not been confirmed in studies.

Higher body mass index was associated with greater variability in nearly all sleep/wake measures, although most associations appeared to be due, in part to reduced total sleep time. After additional adjustment for mean total sleep time in multivariable adjusted models, associations between BMI and variability in sleep latency, in bed timing and out of bed timing were attenuated and no longer reached statistical significance. Associations between being overweight (BMI 25-29 kg/m^2) and nighttime wakefulness were also no longer statistically significant with further adjustment for total sleep time. However, associations between being obese (BMI 30+ kg/m^2) and greater variability in nighttime wakefulness, as well as being overweight or obese and greater variability in total sleep time, remained significant. A prior study in the MrOS cohort observed a strong cross sectional association between actigraphy measured short total sleep time and greater adiposity, which was explained, in part by an increased prevalence of sleep apnea among older men with shortened sleep duration[99]. Given these results, sleep apnea may be one pathway that may mediate the association between higher BMI and greater variability in sleep/wake measures. Future projects should explore this potential mechanism for the association.

In our cohort, associations between cognitive impairment and variability in sleep/wake parameters appear to be due, in large part, to other factors such as older age, lower education and poorer health given the attenuation of the associations in age and site as compared to multivariable adjusted models. Despite this, however, the magnitude of the point estimate for variability in sleep latency (OR=1.58) in multivariable adjusted models suggests that an association may exist. In secondary analyses expressing variability measures as continuous variables, this association was statistically significant. Men who were cognitively impaired had on average a standard deviation for sleep latency variability that was 7.3 minutes greater than men who were not cognitively impaired (mean(SD) = 7.3 (2.8), p=0.008). Since the algorithm to calculate sleep latency involves utilizing the daily sleep diaries it is possible that errors in self-reporting,

influenced by cognitive impairment, have contributed to increased variability in this measure, despite the fact that actigraphy data were edited whenever large discrepancies were observed. This seems unlikely however, since adjusting for the quality of the time reported in-bed data (i.e. the amount of disagreement between self-report and actigraphy data) did not impact these results. A prior study using data from the Study of Osteoporotic Fractures cohort observed that older women who took longer to fall asleep at night had worse cognitive scores (0.8% worse MMSE score for every ½ hour increase in sleep latency) that was not explained by health and other related factors. Our findings suggest that variability in sleep latency may also be important to consider in this population. Future research should explore this further as well as the potential mechanisms underlying this association.

Greater comorbidity burden, and more importantly multimorbidity, was independently associated with an increased odds for greater variability in total sleep time, nighttime wakefulness, sleep latency and out of bed timing. Further sensitivity analyses suggested that associations with sleep latency variability were largely explained by reduced total sleep time, However other associations were not impacted by additional adjustment for mean total sleep time. There are several potential mechanistic pathways that may explain these associations, including medication side effects, anxiety, life stressors and the impact of specific medical conditions, and more work is needed to better understand these associations.

In addition to the results highlighted above, several other associations were observed, such as that between heavy alcohol use and greater variability in nighttime wakefulness and sleep latency, older age and greater WASO variability and less variability in bed timing, and lower education with greater total sleep time variability. Each of these should be further examined in future research, although not all may be modifiable.

There are several strengths worth noting in this study, including a large sample size of community-dwelling (non-institutionalized) older men, objective estimates of sleep and assessment of multiple potential predictors. However,

there are also some limitations. Results are not necessarily generalizable to other population groups. This study had a cross-sectional design, and thus we are unable to assess causality in any of these associations. In addition, actigraphy is a measure of activity and inactivity, and is not a definitive characterization of sleep/wake status. Actigraphs were worn for at least 5 consecutive nights, which may be a relatively short time frame to characterize variability in sleep/wake patterns, and could be influenced by one extreme value. Future studies should examine these associations in studies that collect measures of sleep-wake patterns over longer time frames (>2 weeks). We performed multiple comparisons and some of the observed associations may be due to chance.

In conclusion, we observed significant within-person variability in sleep latency and nighttime wakefulness in a cohort of older, community dwelling men. We also observed that several demographic and health-related predictors, some of which may be modifiable, were significantly associated with greater variability in sleep/wake parameters. Future prospective research studies should examine whether variability in sleep-wake patterns is associated with risk of adverse health outcomes among older adults, and if so, whether or not interventions aimed at reducing variability in sleep improve health outcomes.

Tables and Figures

Table 7. Baseline characteristics of MrOS Participants

Characteristic	Value
Age groups, y, %	
67-69	8.9
70-74	35.0
75-79	28.8
80+	27.3
Race, %	
Caucasian	89.9
African American	3.7
Other	6.4
Education, %	
Less than HS	5.4
High School	16.4
College	78.1
Lives alone, %	13.2
Smoking status, %	
Current	1.9
Former	58.9
Never	39.1
Alcohol use, drinks/wk, %	
0-1	47.2
2-13	47.0
14+	5.8
Number of times up at night to use bathroom, %	
0	4.4
1-3	83.1
4	6.2
5 or more	6.4

Self-reported health status, %	
Good/excellent	86.6
Fair/Poor/Very Poor	13.4
Antidepressant user, %	7.9
Benzodiazepine user, %	4.3
Non-benzodiazepine anxiolytic/hypnotic user, %	1.9
Depressed (GDS≥6), %	6.6
BMI categories, %	
<25 (underweight to normal weight)	29.6
25-29 (overweight)	49.6
30+ (obese)	20.8
Cognitive impairment (Modified Mini Mental State Examination Score <80%), %	3.5
Number of impairments in Instrumental Activities of Daily Living, %	
0	78.9
1-2	16.7
3 or more	4.4
Comorbidity burden, %	
0	26.2
1	31.7
2 or more	42.1
Parkinson's disease, %	1.2
Chronic kidney disease, %	1.0
Chronic obstructive pulmonary disease, %	5.4
Stroke, %	3.5
Diabetes, %	13.4
Congestive heart failure, %	6.0
Myocardial infarction, %	17.4
Hypertension, %	50.7

Hypercholesterolemia	42.1

Table 8. Distributions and Variability in Actigraphic Sleep/Wake Parameters in Older Men

	Between-subject			Within-subject
Sleep/wake parameter	**Mean (SD)**	**Median (IQR)**	**CV**	**Variability Mean (SD)**
Time to bed*	10:47 PM (1.3 hrs)	10:44 (10:05-11:28) PM	5.0	37 (25) mins
Time out of bed*	6:59 AM (1.3 hrs)	6:58 (6:19-7:37) AM	15.5	39 (26) mins
Total sleep time	6.4 (1.2) hours	6.5 (5.8-7.2) hours	19.2	47 (24) mins
Nighttime wakefulness	78.2 (43.7) mins	68.7 (46-101) mins	55.8	27 (20) mins
Sleep latency	30.7 (32.1) mins	21.0 (12-37) mins	104.7	24 (26) mins

SD=Standard deviation; IQR=Interquartile Range; hrs=hours; mins=minutes.

*Time parameters are recorded as clock time (HH:MM).

Table 9. Age and site adjusted odds ratios (95% CI) of potential predictors with measures of variability in sleep/wake patterns

	Odds Ratio (95% CI)				
	Measures of high sleep/wake variability (Quintile 5 vs. Quintiles 1-4)*				
	Total sleep time >63.3 vs. ≤ 63.3 mins	**Nighttime Wakefulness >39.4 vs. ≤39.4 mins**	**Sleep Latency >34.7 vs. ≤34.7 mins**	**In BED >53.5 vs. ≤53.5 mins**	**Out Bed >55 vs. ≤55 mins**
Age groups, y					
70-74 vs. 67-69	1.06 (0.75-1.50)	0.95 (0.66-1.37)	0.98 (0.69-1.39)	0.75 (0.54-1.03)	0.87 (0.63-1.21)
75-79 vs. 67-69	0.94 (0.66-1.34)	1.04 (0.72-1.51)	0.90 (0.63-1.29)	0.65 (0.46-0.90)	0.74 (0.53-1.04)
80+ vs. 67-69	1.06 (0.74-1.52)	1.54 (1.07-2.22)	1.08 (0.76-1.55)	1.50 (0.35-0.70)	0.67 (0.47-0.95)
Race					
African American vs. Caucasian	3.63 (2.41-5.47)	2.43 (1.57-3.76)	1.79 (1.14-2.81)	3.13 (2.07-4.73)	3.30 (2.19-4.97)
Other vs. Caucasian	1.32 (0.89-1.96)	1.48 (1.00-2.18)	1.06 (0.71-1.60)	1.73 (1.20-2.48)	1.43 (0.96-2.11)
Education					
Less than high school vs. College	2.06(1.43-2.97)	1.90 (1.32-2.74)	1.34 (0.91-1.98)	1.44 (0.97-2.14)	1.83 (1.26-2.65)
High School vs. College	1.17 (0.90-1.51)	1.12 (0.86-1.45)	1.16 (0.90-1.51)	0.98 (0.75-1.29)	0.99 (0.75-1.29)
Social Support					
Lives with others vs. alone	1.75 (1.36-2.25)	1.49 (1.15-1.92)	1.55 (1.20-2.00)	1.98 (1.54-2.55)	1.40 (1.08-1.82)
Smoking status					
Former vs. Never	1.10 (0.91-1.34)	1.32 (1.08-1.61)	1.29 (1.06-1.58)	0.97 (0.80-1.17)	1.23 (1.01-1.50)

Current vs. Never	2.06 (1.14-3.71)	1.95 (1.04-3.65)	1.93 (1.05-3.57)	2.41 (1.36-4.27)	1.94 (1.06-3.54)
Health status					
Fair/poor/very poor vs. good/excellent	1.60 (1.24-2.05)	1.70 (1.32-2.18)	1.33 (1.03-1.72)	1.18 (0.91-1.54)	1.54 (1.20-1.99)
Alcohol use, drinks/wk					
2-13 vs. 0-1	0.84 (0.69-1.02)	0.90 (0.81-1.21)	1.14 (0.94-1.40)	0.89 (0.73-1.09)	1.02 (0.84-1.25)
14+ vs. 0-1	0.90 (0.60-1.37)	1.62 (1.11-2.36)	1.29 (0.86-1.93)	1.08 (0.72-1.61)	0.89 (0.58-1.37)
Number of times up to use bathroom					
0 vs. 1-3	0.86 (0.53-1.39)	0.50 (0.28-0.89)	0.66 (0.39-1.12)	0.96 (0.61-1.53)	1.03 (0.65-1.62)
4 vs. 1-3	1.10 (0.75-1.60)	1.14 (0.78-1.65)	1.09 (0.75-1.59)	1.09 (0.74-1.60)	0.80 (0.53-1.22)
5 or more vs. 1-3	1.41 (0.99-2.01)	1.33 (0.93-1.91)	1.23 (0.86-1.77)	1.09 (0.75-1.59)	1.55 (1.09-2.20)
Antidepressant use					
User vs. non-user	2.03 (1.50-2.74)	1.17 (0.84-1.64)	1.88 (1.38-2.55)	1.34 (0.97-1.85)	2.14 (1.59-2.89)
Benzodiazepine use					
User vs. non-user	2.28 (1.54-3.38)	1.09 (0.69-1.71)	1.81 (1.21-2.73)	1.84 (1.23-2.76)	1.95 (1.30-2.91)
Non-Benzodiazepine Anxiolytic/Hypnotics					
User vs. non-user	1.34 (0.71-2.54)	0.80 (0.39-1.67)	1.53 (0.82-2.85)	1.36 (0.72-2.58)	1.11 (0.57-2.18)
Depressed					
GDS≥6 vs. <6	2.14 (1.55-2.96)	2.09 (1.51-2.89)	2.09 (1.51-2.89)	2.00 (1.44-2.79)	2.23 (1.61-3.09)
Body Mass Index, kg/m^2					
25-29 vs. <25	1.30 (1.03-1.64)	1.40 (1.10-1.79)	1.21 (0.96-1.53)	1.13 (0.90-1.41)	1.26 (1.01-1.58)
30+ vs. <25	1.98 (1.51-2.59)	2.85 (2.17-3.74)	2.03 (1.55-2.65)	1.39 (1.06-1.82)	1.31 (1.00-1.73)
Cognitive impairment					

mMMSE<80 vs. ≥80	1.41 (0.88-2.25)	1.56 (0.99-2.45)	1.79 (1.15-2.81)	1.30 (0.79-2.11)	1.54 (0.96-2.47)
IADL impairments					
1 vs. none	1.38 (1.08-1.75)	1.35 (1.06-1.72)	1.23 (0.96-1.58)	1.39 (1.09-1.77)	1.15 (0.89-1.47)
2 or more vs. none	2.16 (1.45-3.22)	1.94 (1.30-2.98)	2.51 (1.70-3.71)	2.38 (1.59-3.58)	1.52 (0.99-2.32)
Comorbidity burden, %					
1 vs. 0	1.06 (0.85-1.32)	1.32 (1.07-1.64)	1.28 (1.04-1.59)	1.13 (0.92-1.41)	1.05 (0.84-1.30)
2 or more vs. 0	1.84 (1.38-2.45)	2.06 (1.54-2.75)	1.68 (1.25-2.26)	1.32 (0.96-1.79)	1.67 (1.24-2.24)

All analyses were adjusted for age and clinic site.

*Each variability outcome is expressed as highest quintile vs. remaining four quintiles

Table 10. Multivariable adjusted associations between potential predictors and variability in sleep/wake parameters

	Odds Ratio (95% CI)**				
	Measures of greatest sleep/wake variability (Quintile 5 vs. Quintiles 1-4)*				
	Total sleep time >63.3 vs. ≤ 63.3 mins	**WASO >39.4 vs. ≤39.4 mins**	**Sleep Latency >34.7 vs. ≤34.7 mins**	**In Bed Timing >53.5 vs. ≤53.5 mins**	**Out of Bed Timing >55 vs. ≤55 mins**
Age groups, y					
70-74 vs. 67-69	1.01 (0.70-1.44)	0.88 (0.60-1.27)	0.96 (0.67-1.38)	0.72 (0.52-1.01)	0.83 (0.59-1.16)
75-79 vs. 67-69	0.92 (0.63-1.33)	1.02 (0.70-1.50)	0.88 (0.61-1.28)	0.62 (0.44-0.88)	0.73 (0.51-1.04)
80+ vs. 67-69	1.07 (0.73-1.56)	1.58 (1.07-2.34)	1.06 (0.72-1.56)	0.44 (0.31-0.64)	0.69 (0.48-1.01)
Race					
African American vs. Caucasian	3.26 (2.09-5.06)	1.98 (1.23-3.19)	1.63 (1.01-2.65)	2.88 (1.86-4.47)	3.13 (2.02-4.85)
Other vs. Caucasian	1.27 (0.84-1.93)	1.50 (0.99-2.25)	1.12 (0.73-1.71)	1.71 (1.17-2.51)	1.44 (0.95-2.17)
Education					
Less than high school vs. College	1.51 (1.01-2.25)	1.29 (0.86-1.93)	0.96 (0.62-1.47)	1.09 (0.71-1.68)	1.34 (0.89-2.01)
High School vs. College	1.02 (0.78-1.34)	0.98 (0.75-1.29)	1.06 (0.81-1.38)	0.90 (0.68-1.19)	0.90 (0.68-1.18)
Social Support					
Lives with others vs. alone	1.57 (1.20-2.04)	1.43 (1.09-1.87)	1.45 (1.11-1.89)	1.84 (1.42-2.39)	1.29 (0.98-1.69)
Smoking status					
Former vs. Never	1.04 (0.84-1.27)	1.17 (0.95-1.44)	1.19 (0.97-1.46)	0.93 (0.76-1.14)	1.18 (0.96-1.46)

Current vs. Never	1.79 (0.94-3.40)	1.53 (0.78-3.00)	1.72 (0.90-3.26)	1.97 (1.07-3.63)	1.55 (0.81-2.98)
Health status					
Fair/poor/very poor vs. good/excellent	1.10 (0.83-1.47)	1.25 (0.94-1.66)	0.90 (0.67-1.21)	0.84 (0.62-1.14)	1.20 (0.90-1.60)
Alcohol use, drinks/week					
2-13 vs. 0-1	0.88 (0.72-1.09)	1.08 (0.87-1.33)	1.21 (0.98-1.49)	0.98 (0.80-1.21)	1.10 (0.89-1.35)
14+ vs. 0-1	0.98 (0.64-1.51)	1.73 (1.16-2.58)	1.34 (0.88-2.03)	1.20 (0.79-1.81)	0.97 (0.62-1.51)
Number of times up to use bathroom					
0 vs. 1-3	0.83 (0.50-1.37)	0.51 (0.28-0.93)	0.67 (0.39-1.15)	0.85 (0.52-1.39)	0.97 (0.60-1.57)
4 vs. 1-3	0.99 (0.66-1.48)	1.09 (0.74-1.61)	1.04 (0.70-1.54)	1.00 (0.67-1.50)	0.73 (0.47-1.13)
5 or more vs. 1-3	1.34 (0.92-1.95)	1.24 (0.85-1.82)	1.19 (0.82-1.75)	0.98 (0.66-1.46)	1.40 (0.97-2.03)
Antidepressant use					
User vs. non-user	1.79 (1.29-2.48)	1.04 (0.72-1.49)	1.67 (1.20-2.32)	1.12 (0.79-1.59)	1.86 (1.35-2.56)
Benzodiazepine use					
User vs. non-user	1.76 (1.15-2.69)	0.85 (0.52-1.39)	1.31 (0.84-2.03)	1.58 (1.03-2.43)	1.55 (1.01-2.38)
Non-Benzodiazepine Anxiolytic/Hyp use					
User vs. non-user	1.45 (0.76-2.80)	0.90 (0.43-1.91)	1.64 (0.86-3.10)	1.32 (0.68-2.54)	1.18 (0.59-2.35)
Depressed					
GDS≥6 vs. <6	1.48 (1.03-2.13)	1.61 (1.12-2.32)	1.61 (1.12-2.31)	1.54 (1.06-2.23)	1.65 (1.15-2.38)
Body Mass Index, kg/m^2					
25-29 vs. <25	1.36 (1.07-1.73)	1.38 (1.08-1.77)	1.22 (0.96-1.54)	1.17 (0.92-1.47)	1.29 (1.02-1.63)
30+ vs. <25	1.91 (1.44-2.53)	2.61 (1.97-3.46)	1.95 (1.48-2.57)	1.34 (1.01-1.77)	1.27 (0.95-1.69)
Cognitive impairment					

mMMSE<80 vs. ≥80	0.86 (0.51-1.47)	1.11 (0.67-1.84)	1.58 (0.97-2.57)	0.93 (0.54-1.60)	1.05 (0.62-1.78)
IADL impairments					
1 vs. none	1.09 (0.84-1.42)	1.01 (0.78-1.31)	1.01 (0.77-1.31)	1.26 (0.96-1.64)	0.91 (0.70-1.20)
2 or more vs. none	1.27 (0.81-2.00)	1.13 (0.72-1.78)	1.74 (1.13-2.69)	1.77 (1.13-2.78)	0.85 (0.52-1.38)
Comorbidity burden, %					
1 vs. 0	0.98 (0.78-1.23)	1.24 (0.99-1.55)	1.25 (1.00-1.56)	1.10 (0.88-1.37)	0.99 (0.79-1.25)
2 or more vs. 0	1.39 (1.01-1.90)	1.68 (1.23-2.30)	1.38 (1.00-1.91)	1.11 (0.80-1.56)	1.40 (1.02-1.93)

*Models adjusted for all covariates presented in above table plus clinic site.

**Each variability outcome is expressed as highest quintile vs. remaining four quintiles

Cohort N is 2,776 in above analyses due to some missing covariates

mMMSE=modified Mini Mental State Examination; IADL=Instrumental Activities of Daily Living

Table 11. Multivariable adjusted models with adjustment for mean total sleep time adjusted odds ratios (95% CI) of predictors with measures of variability in sleep/wake patterns

	Odds Ratio (95% CI)				
	Measures of high sleep/wake variability (Quintile 5 vs. Quintiles 1-4)**				
	Total sleep time >63.3 vs. ≤ 63.3 mins	**WASO >39.4 vs. ≤39.4 mins**	**Sleep Latency >34.7 vs. ≤34.7 mins**	**In Bed Timing >53.5 vs. ≤53.5 mins**	**Out Bed Timing >55 vs. ≤55 mins**
Age groups, y					
70-74 vs. 67-69	1.01 (0.70-1.45)	0.85 (0.57-1.27)	0.95 (0.65-1.40)	0.72 (0.52-1.01)	0.83 (0.59-1.16)
75-79 vs. 67-69	0.92 (0.63-1.33)	1.02 (0.68-1.54)	0.86 (0.58-1.27)	0.62 (0.44-0.88)	0.73 (0.51-1.04)
80+ vs. 67-69	1.07 (0.73-1.57)	1.64 (1.08-2.48)	1.03 (0.69-1.55)	0.44 (0.31-0.64)	0.69 (0.48-1.01)
Race					
African American vs. Caucasian	3.18 (2.04-4.95)	1.60 (0.96-2.67)	1.29 (0.77-2.17)	2.82 (1.81-4.37)	3.04 (1.96-4.72)
Other vs. Caucasian	1.25 (0.82-1.90)	1.29 (0.84-2.00)	0.95 (0.61-1.48)	1.69 (1.15-2.48)	1.41 (0.93-2.12)
Education					
Less than high school vs. College	1.51 (1.01-2.26)	1.32 (0.85-2.03)	0.94 (0.60-1.49)	1.09 (0.71-1.68)	1.34 (0.89-2.02)
High School vs. College	1.02 (0.78-1.34)	0.95 (0.70-1.28)	1.03 (0.77-1.38)	0.89 (067-1.19)	0.89 (0.68-1.18)
Social Support					
Lives with others vs. alone	1.54 (1.18-2.01)	1.26 (0.94-1.69)	1.29 (0.96-1.71)	1.82 (1.40-2.36)	1.27 (0.96-1.67)
Smoking status					

Former vs. Never	1.03 (0.84-1.27)	1.15 (0.92-1.44)	1.18 (0.95-1.47)	0.93 (0.76-1.13)	1.18 (0.96-1.45)
Current vs. Never	1.74 (0.92-3.31)	1.06 (0.50-2.24)	1.25 (0.61-1.26)	1.92 (1.04-3.55)	1.50 (0.78-2.89)
Health status					
Fair/poor/very poor vs. good/excellent	1.10 (0.83-1.48)	1.33 (0.97-1.80)	0.90 (0.66-1.23)	0.84 (0.62-1.14)	1.20 (0.90-1.60)
Alcohol use, drinks/wk					
2-13 vs. 0-1	0.88 (0.72-1.09)	1.03 (0.83-1.30)	1.18 (0.95-1.47)	0.98 (0.80-1.20)	1.09 (0.89-1.34)
14+ vs. 0-1	0.99 (0.64-1.52)	1.87 (1.21-2.89)	1.36 (0.86-2.14)	1.20 (0.79-1.81)	0.97 (0.62-1.52)
Number of times up to use bathroom					
0 vs. 1-3	0.82 (0.50-1.36)	0.41 (0.21-0.78)	0.58 (0.32-1.02)	0.85 (0.52-1.38)	0.96 (0.60-1.55)
4 vs. 1-3	0.99 (0.66-1.48)	1.13 (0.74-1.72)	1.07 (0.70-1.63)	1.01 (0.67-1.50)	0.73 (0.47-1.13)
5 or more vs. 1-3	1.35 (0.93-1.96)	1.35 (0.90-2.02)	1.26 (0.84-1.89)	0.98 (0.66-1.46)	1.41 (0.97-2.04)
Antidepressant use					
User vs. non-user	1.80 (1.30-2.50)	1.12 (0.75-1.68)	1.92 (1.34-2.75)	1.13 (0.79-1.60)	1.88 (1.36-2.59)
Benzodiazepine use					
User vs. non-user	1.78 (1.16-2.72)	0.89 (0.52-1.51)	1.49 (0.93-2.39)	1.60 (1.04-2.45)	1.57 (1.02-2.42)
Non-Benzodiazepine Anxiolytic/Hypnotics					
User vs. non-user	1.43 (0.75-2.76)	0.75 (0.34-1.63)	1.49 (0.77-2.89)	1.30 (0.68-2.51)	1.16 (0.58-2.31)
Depressed					
GDS≥6 vs. <6	1.49 (1.03-2.15)	1.82 (1.23-2.69)	1.82 (1.24-2.69)	1.55 (1.06-2.24)	1.67 (1.16-2.40)
Body Mass Index, kg/m^2					
25-29 vs. <25	1.34 (1.05-1.71)	1.22 (0.94-1.59)	1.06 (0.82-1.36)	1.15 (0.91-1.46)	1.27 (1.01-1.61)
30+ vs. <25	1.83 (1.38-2.44)	1.81 (1.33-2.46)	1.31 (0.97-1.78)	1.29 (0.97-1.72)	1.21 (0.90-1.62)

Cognitive impairment					
mMMSE<80 vs. ≥80	0.87 (0.51-1.48)	1.16 (0.68-1.99)	1.69 (1.01-2.81)	0.93 (0.54-1.60)	1.06 (0.62-1.79)
IADL impairments					
1 vs. none	1.09 (0.84-1.41)	0.99 (0.75-1.31)	0.98 (0.74-1.30)	1.25 (0.96-1.63)	0.91 (0.69-1.19)
2 or more vs. none	1.26 (0.80-1.97)	0.95 (0.58-1.55)	1.58 (0.99-2.52)	1.75 (1.11-2.75)	0.84 (0.51-1.36)
Comorbidity burden					
1 vs. 0	0.98 (0.78-1.23)	1.16 (0.91-1.47)	1.18 (0.93-1.50)	1.09 (0.87-1.37)	0.99 (0.79-1.24)
2 or more vs. 0	1.38 (1.00-1.89)	1.63 (1.16-2.29)	1.30 (0.92-1.83)	1.11 (0.79-1.55)	1.39 (1.01-1.92)
Total sleep time, mins, per SD increase*	0.92 (0.84-1.02)	0.40 (0.36-0.45)	0.43 (0.39-0.48)	0.93 (0.84-1.02)	0.91 (0.82-1.00)

All analyses were additionally adjusted for clinic site. Associations significant at P<0.05 level are italicized

*Standard Deviation for Total Sleep Time = 1.2 hrs, or 72 minutes

**Each variability outcome is expressed as highest quintile vs. remaining four quintiles

Table 12. Associations between predictors and continuous variability measures of WASO, sleep latency and total sleep time

	WASO, mins		Sleep Latency, mins		Total sleep time, mins	
	Estimate (SE)	P-value	Estimate (SE)	P-value	Estimate (SE)	P-value
Intercept	18.54		15.61		42.93	
Age groups, y						
70-74 vs. 67-69	0.65 (1.36)	0.636	-0.58 (1.82)	0.749	-2.42 (1.66)	0.145
75-79 vs. 67-69	1.64 (1.40)	0.239	-1.27 (1.87)	0.497	-3.99 (1.70)	0.019
80+ vs. 67-69	4.29 (1.45)	0.003	0.89 (1.94)	0.647	-1.34 (1.77)	0.447
Race						
African American vs. Caucasian	9.06 (1.99)	<.001	6.44 (2.67)	0.016	13.53 (2.43)	<.001
Other vs. Caucasian	2.86 (1.58)	0.070	0.37 (2.11)	0.861	2.79 (1.92)	0.147
Education						
Less than high school vs. College	2.29 (1.12)	0.041	1.88 (1.50)	0.211	3.58 (1.36)	0.009
High School vs. College	-0.80 (0.82)	0.327	-0.99 (1.09)	0.367	-2.73 (1.00)	0.006
Social Support						
Lives with others vs. alone	4.40 (1.09)	<.001	5.11 (1.45)	<.001	4.21 (1.32)	0.002
Smoking status						
Former vs. Never	1.40 (0.76)	0.068	1.66 (1.02)	0.104	1.75 (0.93)	0.060

Current vs. Never	1.83 (2.73)	0.503	5.36 (3.66)	0.143	8.16 (3.33)	0.014
Health status						
Fair/p/vp vs. good/excellent	1.45 (1.16)	0.210	0.82 (1.55)	0.599	0.12 (1.41)	0.932
Alcohol use, drinks/week						
2-13 vs. 0-1	0.99 (0.78)	0.206	2.79 (1.05)	0.008	-0.38 (0.96)	0.693
14+ vs. 0-1	4.61 (1.62)	0.005	3.05 (2.17)	0.161	-0.23 (1.98)	0.907
# of times up to use bathroom						
0 vs. 1-3	-4.45 (1.78)	0.012	-3.08 (2.38)	0.195	-0.65 (2.17)	0.763
4 vs. 1-3	1.65 (1.52)	0.278	-0.003 (2.04)	0.999	-1.09 (1.86)	0.557
5 or more vs. 1-3	2.85 (1.50)	0.058	1.14 (2.01)	0.571	3.29 (1.83)	0.073
Antidepressant use						
User vs. non-user	0.47 (1.40)	0.736	6.73 (1.87)	<.001	5.88 (1.70)	<.001
Benzodiazepine use						
User vs. non-user	2.20 (1.84)	0.231	2.79 (2.46)	0.257	5.53 (2.24)	0.014
Non-Benzo Anxiolytic/Hyp use						
User vs. non-user	1.95 (2.65)	0.462	4.81 (3.54)	0.174	4.01 (3.23)	0.214
Depressed						
GDS≥6 vs. <6	3.31 (1.55)	0.033	3.19 (2.08)	0.125	5.79 (1.89)	0.002
Body Mass Index, kg/m^2						
25-29 vs. <25	2.82 (0.85)	<.001	3.24 (1.14)	0.005	3.13 (1.04)	0.003
30+ vs. <25	8.16 (1.08)	<.001	9.43 (1.44)	<.001	6.61 (1.31)	<.001
Cognitive impairment						

mMMSE<80 vs. ≥80	4.03 (2.06)	0.050	7.33 (2.75)	0.008	0.23 (2.51)	0.928
IADL impairments						
1 vs. none	0.70 (1.03)	0.498	0.15 (1.38)	0.911	1.31 (1.25)	0.295
2 or more vs. none	4.97 (1.90)	0.009	8.40 (2.54)	<.001	7.74 (2.31)	<.001
Comorbidity burden, %						
1 vs. 0	2.37 (0.85)	0.005	1.03 (1.14)	0.365	0.28 (1.03)	0.790
2 or more vs. 0	4.58 (1.31)	<.001	3.10 (1.75)	0.076	5.88 (1.59)	<.001
Scale	19.0		25.43		23.15	

Table 13. Associations between predictors and continuous variability measures of in and out of bed timing

	In bed timing, mins		Out of bed timing, mins	
	Estimate (SE)	**P-value**	**Estimate (SE)**	**P-value**
Intercept	37.73		37.88	
Age groups, y				
70-74 vs. 67-69	-3.58 (1.73)	0.039	-1.43 (1.79)	0.426
75-79 vs. 67-69	-5.61 (1.77)	0.002	-3.38 (1.83)	0.065
80+ vs. 67-69	-6.97 (1.84)	<.001	-5.09 (1.90)	0.008
Race				
African American vs. Caucasian	18.72 (2.53)	<.001	12.79 (2.61)	<.001
Other vs. Caucasian	5.75 (2.00)	0.004	3.39 (2.07)	0.102
Education				
Less than high school vs. College	1.89 (1.42)	0.185	5.82 (1.47)	<.001
High School vs. College	-0.45 (1.04)	0.071	-3.64 (1.07)	<.001
Social Support				
Lives with others vs. alone	6.85 (1.38)	<.001	2.65 (1.43)	0.063
Smoking status				
Former vs. Never	-0.20 (0.97)	0.839	1.18 (1.00)	0.240
Current vs. Never	6.75 (3.47)	0.052	4.77 (3.59)	0.184
Health status				
Fair/p/vp vs. good/excellent	1.49 (1.47)	0.313	1.01 (1.52)	0.508
Alcohol use, drinks/week				
2-13 vs. 0-1	0.57 (1.00)	0.565	-0.17 (1.03)	0.869
14+ vs. 0-1	0.89 (2.06)	0.666	-0.49 (2.13)	0.819
# of times up to use bathroom				
0 vs. 1-3	-2.28 (2.26)	0.314	-0.40 (2.34)	0.866
4 vs. 1-3	1.20 (1.93)	0.536	-1.86 (2.00)	0.352
5 or more vs. 1-3	2.71 (1.91)	0.156	1.38 (1.97)	0.483
Antidepressant use				
User vs. non-user	0.39 (1.77)	0.828	5.49 (1.83)	0.003

Benzodiazepine use				
User vs. non-user	5.54 (2.33)	0.018	4.55 (2.41)	0.059
Non-Benzo Anxiolytic/Hyp use				
User vs. non-user	-2.93 (3.36)	0.383	1.79 (3.47)	0.606
Depressed				
GDS≥6 vs. <6	3.31 (1.97)	0.093	7.66 (2.04)	<.001
Body Mass Index, kg/m^2				
25-29 vs. <25	1.54 (1.08)	0.155	3.44 (1.12)	0.002
30+ vs. <25	3.22 (1.37)	0.018	4.24 (1.41)	0.003
Cognitive impairment				
mMMSE<80 vs. ≥80	1.36 (2.61)	0.603	0.06 (2.70)	0.845
IADL impairments				
1 vs. none	2.86 (1.31)	0.028	-0.66 (1.35)	0.625
2 or more vs. none	9.26 (2.41)	<.001	3.60 (2.49)	0.149
Comorbidity burden, %				
1 vs. 0	-0.21 (1.08)	0.847	0.31 (1.11)	0.778
2 or more vs. 0	1.60 (1.66)	0.335	6.93 (1.71)	<.001
Scale	24.12		24.93	

Chapter 5: The Impact of Sleep Disturbances on Inpatient Health Care Utilization in Older Women

Background: Sleep disturbances are common in aged populations and often associated with comorbid medical conditions. However, little is known about the extent to which sleep disturbances impact inpatient health care utilization (HCU), especially among an unselected population of community-dwelling older adults.
Methods: This analysis included 2,103 women (mean age 84.2±3.9 years) enrolled in the Study of Osteoporotic Fractures (SOF). Sleep parameters were assessed at the Year 16 exam (2002-2004) using wrist actigraphy (mean 4.2±0.7 nights) and by self-report (Pittsburgh Sleep Quality Index (PSQI) and Epworth Sleepiness Scale (ESS)). Sleep disturbances were defined as being in the worst quartile of a given sleep measure. Inpatient HCU was obtained from Medicare and/or Kaiser Permanente data. Risk of being hospitalized and the rate ratio of inpatient days during the three years after the clinic exam were estimated using logit-Poisson Hurdle Models, and bootstrapping was used to obtain 95% confidence intervals for rate ratio outcomes.
Results: Significant sleep disruption was observed in this population, with 25.9% sleeping less than 6 hours per night, and 50.2% reported poor sleep quality (PSQI>5). 1,157 (55%) of the sample was hospitalized during an average of 2.8 ± 0.6 years of follow-up. In age and site adjusted models, women in the worst quartiles of sleep efficiency, sleep latency, wake after sleep onset (WASO) and short total sleep time had a 31-72% greater odds of being hospitalized. Associations were largely explained by health status, comorbidities and depression. In analyses restricted to women who were hospitalized, being in the worst quartile of sleep efficiency, sleep latency, WASO and Pittsburgh Sleep Quality Index (PSQI) were each associated with a 14-24% increased rate of inpatient days. These results were not statistically significant in multivariable adjusted models. Sleep efficiency and WASO were each associated with greater odds of hospitalization related to CHF (OR=1.98, 95% CI 1.29-3.05 for reduced

sleep efficiency and OR=2.48, 95% CI 1.61-3.80 for increased WASO) and COPD (OR=2.15, 95% CI 1.19-3.91 for reduced sleep efficiency and OR=2.38, 95% CI 1.31-4.31 for increased WASO) in multivariable adjusted models. Conclusion: Associations between sleep disturbances and all-cause hospitalizations are explained, in a large part by health-related factors such as comorbidities, physical functioning, depression and health status, but sleep disturbances are independently associated with a greater odds of hospitalizations due to COPD and CHF in older women. Future studies are needed to determine whether sleep disturbance is a marker of more severe chronic diseases, or an exacerbating factor that results in increased hospitalization risk. Future studies should also examine associations between sleep disturbances and other measures of utilization including nursing home, hospice and home health care.

Introduction

Given rising health care costs, high prevalence and chronic nature of sleep-related complaints, and association of sleep disorders with comorbid medical conditions, the impact of sleep-disturbances on health care utilization has been the focus of several studies. In a 3-month follow-up study of 373 young and middle-aged adults enrolled in a Health Maintenance Organization (HMO), researchers found that patients with self-reported insomnia complaints had greater disability, greater functional impairment, greater number of self-reported days in bed, and greater total health care costs, compared to patients without sleep complaints [100]. A similarly designed longitudinal study in the UK observed that participants with sleep complaints had a 1.7-fold greater odds of healthcare use (consult or prescription for insomnia or mood in 12 months following survey), and this association was higher among those who also self-reported symptoms of anxiety/depression[101].

Similar patterns of higher healthcare utilization among poor sleepers were observed in a cross-sectional study of 12,643 participants conducted in Hungary. Researchers observed that participants with self-reported sleep complaints also self-reported greater hospitalization days in the past year (11.1 vs. 3.7), and greater number of sick leave days (16.8 vs. 10.0) than participants without sleep complaints[102].

These studies have not focused specifically on older adult populations, where the burden of sleep disorders and comorbid medical conditions are highest. Most studies have examined associations with a single sleep disorder, such as insomnia[100-106]. Studies examining more global measures of self-reported sleep disturbance (such as the Pittsburgh Sleep Quality Index) have been performed[104], but were conducted in a primarily younger population. Furthermore, several prior studies have relied on self-reported measures of healthcare utilization[102-105], which may be subject to reporting bias.

One study of 14,355 older adults enrolled in the Health and Retirement Study (HRS) examined associations between self-reported measures of insomnia symptoms and self-reported measures of healthcare utilization[107]. The

HRS study observed that participants with one (vs. no) insomnia symptoms had a 1.3-fold greater odds of being hospitalized, and that participants with two or more insomnia symptoms (vs. none) had a 1.7-fold greater odds of being hospitalized. These associations were independent of age, gender, education, race, comorbidities and depression. To our knowledge, no study has examined associations between objective measures of sleep and health-care utilization, Furthermore, having the ability to comprehensively adjust for potential mediators and/or confounders such as physical functioning, cognitive impairment, health status, medications and dementia would also further our understanding of associations between sleep and HCU. Additionally, analyses examining associations between specific sleep disturbances and cause-specific inpatient admissions may provide further insight into the potential mechanisms that may underlie the connections between sleep disturbances and chronic diseases.

Therefore, to examine whether sleep disturbances are associated with overall, as well as cause-specific inpatient health care utilization, we used data from the longitudinal Study of Osteoporotic Fractures (SOF), linked with Medicare claims and Kaiser Permanente encounters.

Methods

Participants

The Study of Osteoporotic Fractures (SOF) is a landmark longitudinal epidemiologic study designed to examine risk factors for osteoporotic fractures. Women were recruited from four U.S. clinical centers (Baltimore, Maryland; Minneapolis, Minnesota; the Monongahela Valley nears Pittsburgh, Pennsylvania; and Portland, Oregon)[32]. The SOF study enrolled 9,704 community-dwelling white women aged 65 years and older from 1986-1988. Women were excluded if they were unable to walk without assistance, or if they had undergone a previous bilateral hip replacement. Initially African American women were excluded from the study due to their low incidence of hip fractures, but from 1997-1998, 662 African American women aged 65 years and older were recruited[33].

The focus of this analysis will be on data collected at visit 8 (year 16), which was conducted from 2002-2004 and invited surviving SOF participants to attend. Of the 10,366 women recruited in SOF, 3,676 attended the visit 8 exam (Figure 4).

Measurement of Sleep/Wake Parameters

Activity patterns were measured using an octagonal wrist actigraph SleepWatch-O (Ambulatory Monitoring, Inc., Ardsley, NY), which is a small device resembling a wrist watch that is worn on the wrist of the non-dominant hand. Actigraphs contain accelerometers that measure and record movement in 1-minute epochs, and have been shown to provide reliable estimates of sleep-wake activity in comparison to polysomnography, which is currently the gold standard[84]. Actigraphy data were transferred to the San Francisco Coordinating Center for centralized processing. Centralized training and certification were also required for clinic staff gathering actigraphy data. Activity data from the actigraph was analyzed using Action W-2 software (Ambulatory Monitoring, Inc.). Actigraphs collect data in 3 modes, with different methodologies and sensitivities and thresholds to determine movement. The University of California at San Diego sleep scoring algorithm was used for data collected in the digital integration mode (also known as the proportional integration mode, or PIM), and time-above-threshold (TAT), and the Cole-Kripke algorithm was used for data collected in the zero-crossings mode (ZCM)[85].

Women were asked to wear the actigraph continuously for at least 72 hours, and to remove it only for bathing or situations in which it might get submerged in water. Time periods in which participants removed the actigraphs are not included in the analyses, and if the actigraph was removed for greater than 10% of the time during the day or for over 2 hours during the night, the data from that night is not included in the analyses.

Several sleep/wake parameters were computed from the actigraphy data and defined as follows: total sleep time (the hours per night spent sleeping while in bed), sleep efficiency (the percentage of time in bed spent sleeping), sleep

latency (the number of minutes from the time when the participant reported getting into bed (and attempting to sleep) until sleep onset, wake after sleep onset (minutes of wake after sleep onset occurring during the time in bed). Sleep onset was defined as the first 20 continuous minutes of sleep after getting into bed.

In a subset of SOF participants who had concurrent actigraphy and polysomnography (gold standard) data during the same night, the intraclass correlation was highest for total sleep time (r=0.76) and more moderate in magnitude for sleep efficiency and wake after sleep onset (r=0.61 and 0.58, respectively).[87] Of the 3,676 women who attended the visit 8 exam, 85% (n=3,123) had technically adequate wrist actigraphy data.

Self-Reported Sleep Measures

Women enrolled in SOF also completed the Pittsburgh Sleep Quality Index (PSQI) and the Epworth Sleepiness Scale (ESS) questionnaires at their sleep study visit. The PSQI is a validated measure of subjective sleep quality and sleep disturbances over a 1-month period. The questionnaire is divided into sections that assess subjective sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbances, use of sleeping medications and daytime dysfunction. Global PSQI scores range from 0 to 21, and a standard cut-off of greater than 5 is indicative of poor sleep quality. This cutoff has a sensitivity of 89.6% and a specificity of 86.5% in distinguishing good sleepers from poor sleepers[64, 65].

The ESS is a self-administered questionnaire that assesses propensity for sleep onset. Participants are asked to rate how likely (from 1 to 3, with 1 being unlikely and 3 being highly likely) they are to doze off in eight typical daily situations. Scores range from 0 to 24, with a standard cutoff of greater than 10 indicative of excessive daytime sleepiness[67, 68].

Linkage of SOF Cohort Data to Medicare Claims Data and Kaiser Permanente Encounter Data

Linkage of the SOF cohort to Medicare claims data was completed in 2008, by submitting social security and/or Medicare (HIC) numbers for SOF participants who were alive as of 1/1/1991, to *Centers for Medicare and Medicaid Services (CMS)*. A linkage was determined to be valid if there was an exact match on SSN/HIC, and sufficient agreement on DOB, gender, last known residence (ZIP code), and date of death (when available). Medicare data was purchased from January 1991-December 2010. Of the 10,366 women enrolled in the SOF study, 9,986 were alive and actively enrolled in SOF as of January 1, 1991, and of those 9,228 (92.4%) were determined to be valid linkages to Medicare claims data.

Women at the SOF Portland site were originally recruited into the SOF study through Kaiser Permanente, and thus we observed a high rate of Medicare Advantage enrollment at this site (96%). Linkage of SOF Portland participants to Kaiser Permanente inpatient encounter records was completed in 2014 by submitting social security numbers to Kaiser Permanente. Of the 2,464 women enrolled at the Portland SOF site who were alive as of 1/1/1991, 2,180 (88.5%) were enrolled in a Kaiser Permanente plan. Kaiser Permanente inpatient encounter data was obtained from January 1991-December 2010. In combining Medicare claims and Kaiser Permanente encounter records we were able to successfully link 9,381 (93.9% of 9,986) SOF participants to Medicare and/or Kaiser encounter records.

We required that during the month of the SOF V8 exam that participants be observable in claims/encounters data, meaning that they were either enrolled in Kaiser, or enrolled in a Part A Medicare plan for which CMS processes all of the inpatient claims. Of the 3,123 women who attended Visit 8 and had technically adequate actigraphy data, 2,103 (67.3%) linked successfully to Medicare and/or Kaiser and were enrolled in a Part A plan, or Kaiser plan for at least one month from their sleep visit, until death, disenrollment or the end of follow up (Figure 4).

Inpatient Health Care Utilization

Inpatient health care utilization was obtained for the three years following the SOF visit 8 exam. Data on hospitalizations and cumulative inpatient days were assessed using the MedPAR file for participants enrolled in a Part A Medicare plan, and from Kaiser Permanente inpatient encounters data for participants enrolled in a Kaiser plan. We computed the cumulative inpatient days observed during the three years following the clinic visit.

Other Measures

Additional measures were collected at the time of the visit 8 exam. All participants completed questionnaire data, which included questions about their current health status, smoking, alcohol use and medical history. Medical information included a self-reported history of a physician diagnosis of cardiovascular disease (including myocardial infarction, angina, congestive heart failure, other heart disease), stroke, diabetes, Parkinson's disease, Alzheimer's disease, chronic obstructive pulmonary disease and cancer. A variable was created to indicate presence of 0, 1 or 2 or more selected medical conditions. Caffeine intake was estimated based on self-report of the average daily number of cups of caffeinated coffee, tea or cans of caffeinated soda[32]. The Geriatric Depression Scale (GDS) was used to assessed depressive symptoms, with scores $\geq$6 indicative of depression[90]. Functional status was measured using information collected on six independent activities of daily living (IADL)[91, 92], and $\geq$1 IADL impairments were indicated if a woman reported having any difficult with performing any of the 6 IADL abilities (walking 2 to 3 blocks on level ground, climb up 10 steps, walk down 10 steps, prepare meals, do heavy housework, and shop for groceries or clothes). Tests of physical function included walk speed, which is the time in seconds to walk 6 meters at usual pace. The Mini-Mental State Examination[108] was administered. This is a brief, global cognitive function test with concentration, language, and memory components designed to screen

for cognitive impairment. The Mini-Mental State Examination scale ranges from 0 to 30, with higher numbers indicating better performance. Participants were asked to bring in all medications used within the past two weeks, and a computerized medication coding dictionary was used to categorize the medications[89]. Possible dementia was defined as a Mini-Mental State Examination score lower than 26, self-reported history of dementia or use of medications commonly prescribed for dementia. Participants were also asked to indicate the type of residence they live in (private home/apartment, retirement home/senior complex, nursing home, personal care home, other), and an indicator was created for independent living in a private home/apartment status. Body weight was measured using a standard balance beam, or digital scale and height using a wall mounted Harpenden stadiometer. Body Mass Index (BMI) was calculated as kg/m^2. In a subset of women enrolled in fee-for-service Medicare, we calculated an Elixhauser Comorbidity Summary Score[109] (range 0-30), and expressed categories of Elixhauser Comorbidities as 0, 1, 2+. An indicator for prior hospitalization indicated that the woman was enrolled in fee-for-service or Kaiser in the full 12 months prior to Visit 8, and was hospitalized at least once during that time. Information from the SOF baseline visit was used to assess age, self-reported race/ethnicity and highest level of education attained.

Statistical Analysis

In primary analyses, measures of sleep disturbances were expressed as dichotomous predictors based on worst quartile. The following cut points pertained to the worst quartile for each sleep/wake parameter: sleep efficiency <79.4% vs. $\geq$79.4%, sleep latency >53.1 vs. $\leq$53.1 minutes, wake after sleep onset $\geq$100 vs. <100 minutes, short sleep duration <6 hours vs. 6-7.5 hours, long sleep duration >7.5 hours vs. 6.0-7.5 hours; PSQI>8 vs. $\leq$8; and ESS >7 vs. $\leq$7. Secondary analyses examined associations between sleep/wake parameters and inpatient health care utilization using sleep/wake cut points that have been previously published in the Study of Osteoporotic Fractures. The following cut points were specified: sleep efficiency <70% vs. $\geq$70%, sleep latency $\geq$1 hours

vs. <1 hours, awakening after sleep onset ≥90 vs. <90 mins, and sleep duration was expressed as <6 hours (short) vs. 6-8 hours (normal: referent group) vs. >8 hours (long). The following cut points were used for self-reported sleep disturbances: PSQI>5 vs. ≤5, ESS >10 vs. ≤10.

The cumulative sum of inpatient days during follow up was expressed as a count variable. We examined the association between sleep disturbances and odds of being hospitalized, as well as the rate ratio of inpatient days among those hospitalized using logit-Poisson hurdle models. All outcomes involving rate ratios used bootstrapping in order to obtain more robust 95% confidence intervals, and p-values due to excess heterogeneity. These models also allowed us to compare the mean rates of inpatient days in women with and without sleep disturbances, and rate ratios in the entire cohort (hospitalized and not hospitalized). Models included a base model adjusted for age and clinic site, and a multivariable adjusted model that included covariates that were associated with hospitalizations and/or sleep disturbances using a threshold of p<.10 to determine retention in the model. Analyses utilized data from both Medicare and Kaiser data sources. Sensitivity analyses restricted to individuals enrolled in Medicare fee-for-service.

Secondary analyses will examine the association between sleep disturbances and inpatient health care utilization associated with cardiovascular disease related events. Cardiovascular disease related events were defined as an inpatient admission with a primary diagnosis of myocardial infarction (ICD-9 410), stroke (ICD-9 430, 431, 434 or 436), congestive heart failure (428) and chronic obstructive pulmonary disease (466, 490-496). Algorithms for these outcomes have been previously validated and are provided in the CMS Chronic Condition Warehouse[110]. Logistic regression models were used and base models were adjusted for age and site, and a model that additionally adjusted for health status, comorbidities and walking speed. All analyses were conducted using SAS version 9.3 (SAS Institute Inc., Cary, North Carolina). In sensitivity analyses we also substituted Elixhauser comorbidity score categories for comorbidities, in the subset of participants that were enrolled in Medicare fee-for-

service as Elixhauser comorbidity information was not available in Kaiser enrollees.

Results

Sleep Disturbances and Characteristics

Characteristics of the cohort of 2,103 women are shown in Table 14. The average age was 84.2 years and 11% of the cohort was African American. 24.8% of women rated their health status as fair, poor or very poor, and 52.8% reported having one or impairments in instrumental activities of daily living (IADL). During an average follow-up of 2.8 (0.6) years, 1,157 (55%) women were hospitalized at least once. Among those hospitalized, the mean (SD) for cumulative inpatient days was 10.7 (11.4) during the entire follow-up period. Characteristics of the cohort by hospitalization status are also presented in Table 14. Factors associated with hospitalization included older age, poorer health status, smoking, lower alcohol consumption, more IADL impairments, slower walk speed, depression, use of antidepressants, use of benzodiazepines, poorer cognition, possible dementia, more comorbidities, not living in a private home/apartment, and having a hospitalization in the prior year.

At the SOF visit 8 exam, sleep disturbances were prevalent. In the cohort, the average total sleep duration was 6.7 (1.3) hours, and 25.9% of women had on average less than 6 hours of sleep per night (short sleep duration), and 14.2% had more than 8 hours of sleep per night (long sleep duration). Furthermore, 20.2% took one hour or longer to fall asleep (prolonged sleep latency), and 30.8% spent 90 or minutes awake during the night (increased WASO), and 9.8% had a sleep efficiency of less than 70% (reduced sleep efficiency). Furthermore, with the exception of excessive daytime sleepiness, greater sleep disturbances were associated with increased hospitalization.

Associations between sleep disturbances and inpatient healthcare utilization

In age and site adjusted models we observed that sleep disturbances were associated with an increased odds of being hospitalized in the three years

after the sleep exam (Table 15). Specifically, women in the lowest quartile of sleep efficiency had a 1.7-fold greater odds of being hospitalized (OR=1.66, 95% CI 1.35-2.05) than women in the top three quartiles of sleep efficiency. Similarly, we observed an increased odds of being hospitalized among women in the worst quartiles of sleep latency (>53.1 m, OR=1.31, 1.07-1.61), WASO (≥100 m, OR=1.72, 95% CI 1.40-2.13), sleep duration (<6 h, OR=1.37, 95% CI 1.10-1.71) and Pittsburgh Sleep Quality Index (>8, OR=1.33, 95% CI 1.08-1.64) in models adjusted for continuous age and clinic site. We did not observe an association between sleep duration≥7.6 h (quartile 4) and excessive daytime sleepiness>7 (quartile 4) and odds of being hospitalized in age and site adjusted models.

After further adjustment for covariates such as depression, use of antidepressants, cognitive functioning, walking speed, health status, impairments in instrumental activities of daily living, comorbidities, probable dementia and living independently, the magnitude of the associations between sleep disturbances and odds of being hospitalized were attenuated (OR range=0.89-1.27) and no longer statistically significant (p≥0.056). The covariates that were the strongest factors in multivariable adjusted models were comorbidities, health status and use of antidepressants.

Among the 1,157 women who were hospitalized at least once during follow-up, we observed that women in the worst quartiles of sleep/wake parameters spent, on average a little over one day per year longer in the hospital than women in the remaining quartiles (Table 16). We also observed a greater rate ratio of inpatient days among women in the worst quartiles of sleep efficiency (Rate Ratio=1.16, 95% CI 1.01-1.32), sleep latency (Rate Ratio=1.15, 95% CI 1.01-1.32), WASO (Rate Ratio=1.24, 95% CI 1.07-1.41), and Pittsburgh Sleep Quality Index (Rate Ratio=1.17, 95% CI 1.02-1.34) in age and site adjusted models. We did not observe a significant association of total sleep time and excessive daytime sleepiness with greater inpatient days in age and site adjusted models.

Upon further adjustment for covariates, the magnitude of all associations were attenuated (rate ratio range = 1.01-1.09) and were no longer statistically significant.

Sensitivity analyses restricting to a fee-for-service population yielded similar results for associations with sleep latency, WASO, total sleep time, PSQI and ESS, but not for sleep efficiency (Tables 18 & 19). In multivariable adjusted models, women in the worst quartile of sleep efficiency had a 36% greater odds of being hospitalized during follow up than women in the remaining quartiles (OR=1.36, 95% CI 1.01-1.83).

Associations between sleep disturbances and cause specific hospitalizations

We observed that women in the worst quartile of sleep/wake parameters had greater odds of being hospitalized with a primary diagnosis of COPD or CHF during follow-up than women in the remaining three quartiles, in age and site adjusted models (Table 4). More specifically, women in the worst quartile of sleep efficiency (OR=2.74, 95% CI 1.62-4.63), sleep latency (OR=1.74, 95% CI 1.01-2.99), and WASO (OR=3.24, 95% CI 1.91-5.47) had a greater odds of being hospitalized for COPD. Furthermore, women in the worst quartiles of sleep efficiency (OR=2.49, 95% CI 1.71-3.63), sleep latency (OR=1.51, 1.02-2.24), WASO (OR=2.94, 95% CI 2.02-4.27), total sleep time (<6 h, OR=2.23, 95% CI 1.45-3.44) and daytime sleepiness (OR=1.66, 95% CI 1.12-2.45) had a greater odds of being hospitalized for CHF, in age and site adjusted models. Pittsburgh sleep quality index was not associated with CHF or COPD hospitalizations, and total sleep time and daytime sleepiness were not associated with COPD hospitalizations.

Although some point estimates suggested that some measures of sleep disturbances were associated with a greater odds of hospitalization due to myocardial infarction or stroke, results were not statistically significant.

Further adjustment for depression, use of antidepressants, walking speed, any impairments in instrumental activities of daily living, health status, cognitive

functioning, medical conditions, probable dementia and living independently attenuated results, but did not alter conclusions. In multivariable adjusted models, measures of night-time wakefulness (sleep efficiency and WASO) were associated with greater odds of COPD and greater odds of CHF related hospitalizations. Sleep efficiency<79.4% was associated with a 2-fold greater odds of COPD hospitalizations (OR=2.15, 95% CI 1.19-3.91) and CHF hospitalizations (OR=1.98, 95% CI 1.29-3.05). WASO≥100 minutes was associated with a 2.4-2.5-fold greater odds of COPD hospitalizations (OR=2.38, 95% Ci 1.31-4.31) and CHF hospitalizations (OR=2.48, 95% CI 1.61-3.80). We also observed that total sleep time <6 hrs (vs. 6.0-7.5 hrs) was associated with a 1.8-fold increased odds of CHF hospitalization in multivariable adjusted models (OR=1.84, 95% CI 1.13-3.01). We did not observe any other significant findings between sleep latency, longer total sleep time (i.e. >7.6 hrs), sleep quality or excessive daytime sleepiness and cause-specific hospitalizations in multivariable adjusted models.

Sensitivity Analyses

Results examining clinically relevant or previously published sleep disturbance cut points, adjusting for having one or more hospitalizations in the year prior to the SOF visit 8 exam, and substituting the Elixhauser comorbidity score for comorbidity burden were similar to primary analyses (data not shown).

Discussion

Older women with sleep disturbances, such as reduced sleep efficiency, prolonged sleep latency, increased night time wakefulness, short sleep duration and poor sleep quality have greater odds of hospitalization. However, these associations are largely explained by a greater burden of comorbidities, poorer health status and depression among older women with disturbed sleep. These results suggest that comorbid medical conditions and depression are potentially stronger predictors of risk of hospitalization, and sleep disturbances yield no additional risk information in this population. We did observe some evidence

that although sleep disturbances may not be independently associated with all-cause hospitalizations, they are independently associated with a greater risk of hospitalizations due to heart failure and hospitalizations due to COPD, and may be important predictors of CHF and COPD-related hospitalizations in older women. These findings help to highlight some potential mechanisms involving COPD and CHF, and indicate a need for further research and into these associations.

Prior studies of sleep and inpatient health care utilization have relied on self-report of sleep disturbances and/or health care utilization, which may be prone to reporting bias[107, 111-117]. In general, these studies observed that sleep disturbances, especially insomnia symptoms, are associated with a greater risk of being hospitalized. For example, 14,355 older adults from the Health and Retirement Study were surveyed about insomnia symptoms and risk of hospitalization, use of home health services and health care utilization in the 2 years prior to the survey[107]. The researchers observed significant associations between greater burden of insomnia symptoms and odds of hospitalization, in that participants reporting one insomnia symptom, and 2 or more symptoms had a 1.28-fold and 1.71-fold greater odds of being hospitalized, respectively, than participants with no insomnia symptoms. These associations were greatly attenuated, but persisted after adjustment for participant characteristics, including demographic variables, comorbidities and depression. Similar to these findings, we also observed significant associations between sleep disturbances and hospitalizations in minimally adjusted models (age and clinic site in our study and age, gender, race and education in the Health and Retirement study). However, the magnitude of these associations were diminished and no longer statistically significant after further adjustment for health related factors, physical functioning, cognitive functioning, depression and comorbidities. Kaufman et. al. also adjusted for depression and comorbidities, and observed that their findings were not substantially impacted. In our results, we observed that number of comorbidities (0, 1, 2 or more), self-reported health status and antidepressant use were the strongest predictors of hospitalizations, independent of sleep

disturbances (data not shown). Kaufman et. al. did not adjust for number of comorbidities or health status. To our knowledge, our study is the first to use both objective assessments of sleep and inpatient health care utilization, and to adjust for a comprehensive set of confounders and mediators including health status, physical and cognitive functioning, depression, and comorbidities.

Our results do not provide causal evidence that sleep disturbances independently increase risk for all-cause inpatient hospitalizations,. The association between sleep disturbances and all-cause hospitalization is explained, in a large part by a variety of patient factors, including depression, greater burden of comorbid medical conditions, poorer health status, poorer functional status, poorer cognition and use of medications.

However, the results of our analyses examining associations between sleep disturbances and cause-specific hospitalizations suggest that sleep disturbances may be independent predictors of hospitalization due to congestive heart failure and hospitalization due to chronic obstructive pulmonary disease. Two proposed mechanistic pathways may explain these observed associations. In one, sleep may be a marker for more severe or advanced disease, suggesting that temporality may be an issue in this study. In our study, we do not have measures of disease severity, such as pulmonary function (for COPD), echocardiograms or ejection fraction (for CHF), and are unable to assess this, and to our knowledge, no other study has attempted to correlate sleep disturbances with disease severity. There are several reasons why sleep disturbances may be a marker for more severe disease progression, particularly in patients with CHF and those with COPD. Sleep disturbances are common (~50%) in patients with COPD[118] and those with CHF[119]. Prior studies have suggested that both conditions can reduce sleep quality through medication side effects, nocturia, nocturnal dyspnea, coughing, chest pain and difficulty breathing[118, 119].

On the other hand, sleep apnea, a condition characterized by repeated pauses in breathing during the night, is also highly prevalent in COPD and CHF patients, and may share similar etiologies. Comorbid sleep apnea and or

nocturnal dyspnea and COPD/CHF may result in significantly impaired sleep[118]. Furthermore, there is potential for the two coexisting conditions to exacerbate each other (i.e. CHF exacerbates sleep disordered breathing and vice versa)[119]. In our study, we only had measures of sleep disordered breathing on a very small subset of participants, and thus were unable to assess the mediating effects of sleep apnea on all cause and cause specific hospitalizations. Furthermore, given the potential for reverse causality, and bidirectional associations between CHF/COPD and sleep disturbances, we are unable to further examine the interrelationships between CHF/COPD and sleep disturbances with inpatient health care utilization.

Taken together, these results signify that in older women, sleep disturbances and comorbid medical conditions are closely linked, and this raises questions of whether concurrent treatment of sleep disorders and comorbid medical conditions improves inpatient health care utilization outcomes in this population. A few randomized studies have examined concurrent treatment of sleep disordered breathing and CHF[119], and insomnia and COPD[120], as well as other comorbid conditions, and have generally observed improvement in quality of life, sleep quality and/or comorbidity. However, much more research is needed in this area to identify whether concurrent treatments improve health outcomes and reduce health care costs in this population.

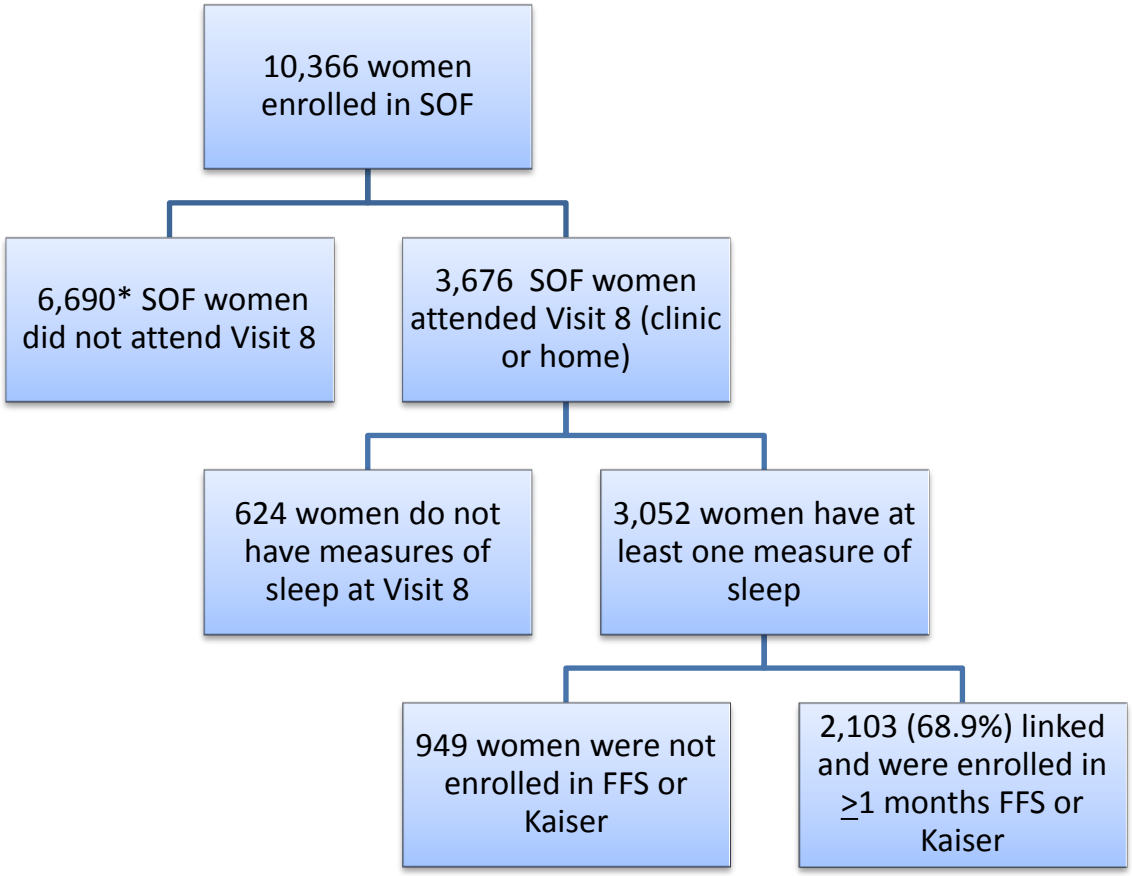
There are several strengths of the current study. The SOF study is a large, well characterized cohort of older women who were not selected on the basis of any sleep-related or other conditions. Additionally, sleep disturbances were assessed using validated measures obtained from wrist actigraphy, or self-reported questionnaires (PSQI and ESS). Furthermore, information from hospitalizations were obtained from linkage with Medicare and/or Kaiser Permanente encounter data, which is a systematic and validated sources of inpatient health care utilization. Finally, our results were adjusted for potentially important confounders and/or mediators of the sleep-hospitalization pathway, including depression, functional impairment and medical conditions. However, a number of limitations should be noted. Our results were limited to inpatient

hospital admissions, and we did not assess associations between sleep disturbances and nursing home, hospice or home health care utilization, which may be particularly important indicators of health care utilization in older populations. Future studies should try to incorporate these additional measures of health care utilization. Furthermore, sleep disordered breathing was assessed in a very small subset of individuals, and we are unable to evaluate the effect of sleep apnea on inpatient health care utilization. We also do not have measures of disease severity, and are unable to assess if sleep disturbances are a marker of more severe disease. We did not examine associations between sleep disturbances and hospitalizations related to other specific causes besides those due to myocardial infarction, stroke, COPD and CHIF, and it is possible that other associations may exist. Finally, results restricted to a fee-for-service yielded slightly increased age and clinic site adjusted mean rates of inpatient days per year, and slightly stronger associations between sleep efficiency and odds of being hospitalized. We did observe lower hospitalization rates and reduced inpatient days among Kaiser enrollees as compared to the Medicare fee-for-service population, which might reflect differences in health systems.

In conclusion, we observed that associations between sleep disturbances and all-cause hospitalizations are mediated by health-related factors such as comorbidities, physical functioning, depression and health status, but sleep disturbances are independently associated with greater odds of hospitalization due to COPD and hospitalization due to CHF in older women. Future studies are needed to determine whether sleep disturbance is a marker of more severe disease, or an exacerbating factor. Future studies should also examine associations between sleep disturbances and nursing home, hospice and home health care utilization.

Tables and Figures

Figure 5. Roadmap of Analytical Cohort



*Of the 6,690 women who did not attend the Visit 8 exam, 4,392 had died, 1,051 were questionnaire only status and 1,247 terminated, refused or were otherwise unable to attend the visit.

Table 14. Baseline Characteristics of 2,103 by Hospitalization Status

Characteristics	Entire Cohort (n=2,103)	Hospitalization		P-Value
		Hospitalized (n=1,157)	Not Hospitalized (n=946)	
Age, y, mean (SD)	84.2 (3.9)	84.5 (4.0)	83.7 (3.7)	<.001
African American, %	10.7	11.6	9.6	0.148
Fair, poor or very poor health status, %	24.8	30.7	17.6	<.001
Current smoker, %	3.1	4.1	1.9	0.004
Alcoholic drinks per day in the last 30 days, mean (SD)	1.2 (2.9)	1.0 (2.7)	1.4 (3.1)	0.004
Daily caffeine intake, mg, mean (SD)	149 (153)	150 (155)	148 (149)	0.859
One or more IADL impairment, %	52.8	61.1	42.6	<.001
Walk speed, m/s, mean (SD)	0.80 (0.27)	0.75 (0.27)	0.86 (0.24)	<.001
Body Mass Index, kg/m^2, mean (SD)	26.9 (5.0)	26.9 (5.2)	26.9 (4.7)	0.947
GDS score (range 0-15), mean (SD)	2.4 (2.6)	2.8 (2.8)	2.0 (2.3)	<.001
Depression, GDS score $\geq$6, %	11.4	14.6	7.6	<.001
Currently taking antidepressants, %	14.3	17.6	10.4	<.001
Currently taking benzodiazepines, %	7.2	8.7	5.5	0.006
MMSE (range 0-30), mean (SD)	27.8 (2.0)	27.6 (2.1)	28.0 (1.8)	<.001

MMSE<26, %	11.0	13.0	8.7	0.002
Possible Dementia, %	11.3	13.1	9.2	0.006
Medical Conditions, %				<.001
0	37.1	29.7	46.1	
1	36.8	36.8	34.9	
2 or more	26.1	26.1	19.0	
Stroke	13.8	16.8	10.0	<.001
Diabetes	10.6	13.4	7.3	<.001
Parkinson's Disease	0.9	1.0	0.9	0.794
Alzheimer's Disease	1.9	2.2	1.5	0.245
COPD	12.3	15.7	8.1	<.001
Cardiovascular Disease	34.4	40.9	26.5	<.001
Cancer	22.8	23.0	22.5	0.783
Elixhauser Comorbidity Summary Score, mean (SD)*	2.13 (2.32)	2.60 (2.45)	1.56 (1.89)	<.001
Elixhauser Comorbidity Categories, %*				<.001
0	34.5	29.0	41.4	
1-2	29.4	26.2	33.5	
3 or more	36.0	44.9	25.2	
Private home/apartment residence, %	75.1	73.5	77.1	0.050

Hospitalized in the year prior, %	23.2	28.7	16.4	<.001
PSQI (range 0-21), mean (SD)	6.2 (3.7)	6.4 (3.8)	5.9 (3.5)	0.003
PSQI>5, %	50.2	52.1	47.9	0.053
ESS (range 0-24), mean (SD)	5.6 (3.8)	5.7 (3.9)	5.6 (3.7)	0.620
ESS>10, %	11.1	11.7	10.3	0.290
Sleep efficiency, %, mean (SD)	83.5 (10.2)	85.0 (9.2)	82.3 (10.8)	<.001
Sleep efficiency<70%, %	9.8	12.7	6.2	<.001
Sleep latency, mins, mean (SD)	42.2 (42.3)	45.4 (46.4)	38.3 (36.4)	<.001
Sleep latency >60 mins, %	20.2	22.5	17.4	0.004
Wake after sleep onset, mins, mean (SD)	78.7 (49.2)	84.8 (52.4)	71.2 (43.9)	<.001
Wake after sleep onset >90 mins, %	30.8	34.8	25.9	<.001
Total sleep time, hrs, mean (SD)	6.7 (1.3)	6.7 (1.4)	6.8 (1.2)	0.018
Total sleep time <6 hrs, %	25.9	27.8	23.5	0.010
Total sleep time >8 hrs, %	14.2	15.1	13.0	

SOF=Study of Osteoporotic Fractures; SD=standard deviation; IADL=instrumental activities of daily living; GDS=Geriatric Depression Scale; COPD=chronic obstructive pulmonary disease; MMSE=Mini-Mental State Examination.

*Available in subset (n=2,065 women) enrolled in FFS for the entire year prior to Visit 8.

Table 15. Association between sleep disturbances and odds of being hospitalized

Sleep/Wake Disturbance	Age and Site adjusted		Multivariable Adjusted	
	OR (95% CI)	p-value	OR (95% CI)	p-value
Sleep efficiency: <79.4 vs. ≥79.4%	1.66 (1.35-2.05)	<.001	1.23 (0.96-1.56)	0.099
Sleep latency: >53.1 vs. ≤53.1 m	1.31 (1.07-1.61)	0.010	1.08 (0.86-1.36)	0.498
WASO: ≥100 vs. <100 m	1.72 (1.40-2.13)	<.001	1.27 (0.99-1.62)	0.056
TST: <6 h vs. 6.0-7.5 h	1.37 (1.10-1.71)	0.008	1.20 (0.94-1.53)	0.136
TST: ≥7.6 h vs. 6.0-7.5 h	1.11 (0.90-1.38)	0.329	1.09 (0.86-1.38)	0.488
PSQI >8 vs. ≤8	1.33 (1.08-1.64)	0.006	1.06 (0.84-1.33)	0.649
ESS >7 vs. ≤7	1.06 (0.86-1.29)	0.598	0.89 (0.71-1.11)	0.285

OR=Odds Ratio; m=minutes; WASO=Wake after sleep onset; TST=total sleep time; h=hours; PSQI=Pittsburgh Sleep Quality Index; ESS=Epworth Sleepiness Scale

Multivariable Models adjusted for age, clinic site, depression, use of antidepressants, cognitive functioning, 6 meter walking speed, self-reported health status, any impairments in instrumental activities of daily living, medical conditions, probable dementia, and living independently.

Table 16. Association between sleep disturbances and rate ratio of inpatient hospital days among those with ≥1 hospital admissions

Sleep/Wake Disturbance		Mean (95% CI) rate of inpatient days per year*	Rate Ratio (95% CI)**	
			Age and Site adjusted	Multivariable Adjusted†
Sleep efficiency	<79.4%	4.22 (3.77-4.69)	1.16 (1.01-1.32)	1.04 (0.94-1.14)
	≥79.4%	3.65 (3.38-3.89)	1.0 (ref)	1.0 (ref)
			p=0.138	p=0.500
Sleep latency	>53.1 m	4.22 (3.77-4.70)	1.15 (1.01-1.32)	1.07 (0.97-1.18)
	≤53.1 m	3.66 (3.40-3.89)	1.0 (ref)	1.0 (ref)
			p=0.028	p=0.184
Wake after sleep onset	≥100 m	4.43 (3.93-4.94)	1.24 (1.07-1.41)	1.06 (0.96-1.18)
	<100 m	3.57 (3.32-3.80)	1.0 (ref)	1.0 (ref)
			p=0.004	p=0.296
Total Sleep Time	<6.0 h	3.90 (3.46-4.34)	1.07 (0.91-1.23)	1.02 (0.92-1.14)
	6.0-7.5 hr	3.65 (3.35-3.98)	1.0 (ref)	1.0 (ref)
	≥7.6 hr	4.02 (3.56-4.53)	1.10 (0.95-1.28)	1.07 (0.96-1.20)
			p>0.166	p>0.214
Pittsburgh Sleep Quality Index	>8	4.25 (3.77-4.77)	1.17 (1.02-1.34)	1.09 (0.99-1.20)
	≤8	3.64 (3.38-3.88)	1.0 (ref)	1.0 (ref)

			p=0.016	p=0.100
Epworth Sleepiness Scale	>7	3.91 (3.52-4.00)	1.04 (0.91-1.18)	1.01 (0.92-1.11)
	≤7	3.77 (3.47-4.34)	1.0 (ref)	1.0 (ref)
			p=0.606	p=0.876

*Adjusted for age and clinic site

**Bootstrapping with 1000 samples used to estimate 95% confidence intervals and p-values

†Multivariable Models adjusted for : Age, clinic site, depression, use of antidepressants, cognitive functioning, 6 meter walking speed, self-reported health status, any impairments in instrumental activities of daily living, selected medical conditions, probable dementia, and living independently

Table 17. Association between sleep disturbances and cause-specific hospitalizations

Sleep Disturbance	Myocardial Infarction (n = 58)		Stroke (n = 84)		COPD (n = 60)		CHF (n = 124)	
	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value
Sleep efficiency <79.4 vs. ≥79.4%								
Model 1	1.34 (0.75-2.40)	0.327	1.14 (0.69-1.86)	0.611	2.74 (1.62-4.63)	<.001	2.49 (1.71-3.63)	<.001
Model 2	1.16 (0.61-2.21)	0.653	0.81 (0.45-1.46)	0.481	2.15 (1.19-3.91)	0.012	1.98 (1.29-3.05)	0.002
Sleep latency >53.1 vs. ≤53.1 minutes								
Model 1	1.73 (0.99-3.03)	0.056	0.87 (0.51-1.47)	0.596	1.74 (1.01-2.99)	0.044	1.51 (1.02-2.24)	0.039
Model 2	1.59 (0.86-2.97)	0.142	0.70 (0.38-1.29)	0.255	1.42 (0.77-2.62)	0.265	1.37 (0.88-2.15)	0.167
WASO ≥100 vs. <100 minutes								
Model 1	1.09 (0.59-1.99)	0.789	1.53 (0.95-2.44)	0.078	3.24 (1.91-5.47)	<.001	2.94 (2.02-4.27)	<.001
Model 2	0.90 (0.46-1.77)	0.755	1.19 (0.68-2.06)	0.545	2.38 (1.31-4.31)	0.005	2.48 (1.61-3.80)	<.001
TST <6 vs. 6.0-7.5 hours								
Model 1	1.41 (0.72-2.76)	0.311	0.95 (0.55-1.63)	0.844	1.26 (0.70-2.27)	0.450	2.23 (1.45-3.44)	<.001
Model 2	1.29 (0.63-2.61)	0.487	0.84 (0.46-1.56)	0.584	1.01 (0.52-1.96)	0.973	1.84 (1.13-3.01)	0.015
TST >7.6 vs. 6.0-7.5 hours								
Model 1	1.58 (0.84-2.98)	0.160	0.98 (0.57-1.68)	0.940	0.75 (0.37-1.52)	0.422	1.38 (0.86-2.22)	0.181
Model 2	1.27 (0.63-2.57)	0.509	0.99 (0.55-1.79)	0.984	0.78 (0.37-1.66)	0.518	1.48 (0.88-2.48)	0.141

PSQI >8 vs. ≤8								
Model 1	1.37 (0.77-2.46)	0.287	1.16 (0.71-1.90)	0.559	1.01 (0.55-1.83)	0.983	1.42 (0.95-2.12)	0.091
Model 2	1.04 (0.54-2.01)	0.900	1.26 (0.73-2.17)	0.415	0.68 (0.33-1.40)	0.293	1.19 (0.75-1.88)	0.454
ESS >7 vs. ≤7								
Model 1	1.46 (0.83-2.58)	0.193	0.92 (0.55-1.52)	0.739	1.34 (0.77-2.35)	0.304	1.66 (1.12-2.45)	0.011
Model 2	1.16 (0.62-2.15)	0.646	0.77 (0.43-1.38)	0.384	1.14 (0.62-2.10)	0.679	1.37 (0.88-2.11)	0.161

OR=Odds Ratio; m=minutes; WASO=Wake after sleep onset; TST=total sleep time; h=hours; PSQI=Pittsburgh Sleep Quality Index; ESS=Epworth Sleepiness Scale

Model 1 adjusted for age and clinic site

Model 2 adjusted for age, clinic site, depression, use of antidepressants, cognitive functioning, 6 meter walking speed, self-reported health status, any impairments in instrumental activities of daily living, medical conditions, probable dementia, and living independently.

Table 18. Association between sleep disturbances and odds of being hospitalized among FFS enrollees

Sleep/Wake Disturbance	Age and Site adjusted		Multivariable Adjusted	
	OR (95% CI)	p-value	OR (95% CI)	p-value
Sleep efficiency: <79.4 vs. ≥79.4%	1.74 (1.33-2.26)	<.001	1.36 (1.01-1.83)	0.042
Sleep latency: >53.1 vs. ≤53.1 m	1.25 (0.96-1.61)	0.092	1.00 (0.76-1.33)	0.974
WASO: ≥100 vs. <100 m	1.68 (1.29-2.18)	<.001	1.27 (0.94-1.71)	0.118
TST: <6 h vs. 6.0-7.5 h	1.49 (1.13-1.96)	0.005	1.34 (0.99-1.80)	0.136
TST: ≥7.6 h vs. 6.0-7.5 h	1.17 (0.89-1.53)	0.259	1.11 (0.82-1.49)	0.488
PSQI >8 vs. ≤8	1.32 (1.02-1.72)	0.037	0.96 (0.72-1.29)	0.809
ESS >7 vs. ≤7	1.01 (0.79-1.31)	0.923	0.88 (0.67-1.17)	0.375

OR=Odds Ratio; m=minutes; WASO=Wake after sleep onset; TST=total sleep time; h=hours; PSQI=Pittsburgh Sleep Quality Index; ESS=Epworth Sleepiness Scale

Multivariable Models adjusted for age, clinic site, depression, use of antidepressants, cognitive functioning, 6 meter walking speed, self-reported health status, any impairment in instrumental activities of daily living, medical conditions, probable dementia, and living independently.

Table 19. Association between sleep disturbances and rate ratio of inpatient hospital days among those with ≥1 hospital admissions among FFS enrollees

Sleep/Wake Disturbance		Mean (95% CI) rate of inpatient days per year*	Rate Ratio (95% CI)**	
			Age and Site adjusted	Multivariable Adjusted†
Sleep efficiency	<79.4%	4.81 (3.86-5.62)	1.11 (0.95-1.33)	1.01 (0.89-1.13)
	≥79.4%	4.32 (3.49-4.86)	1.0 (ref)	1.0 (ref)
			0.202	0.882
Sleep latency	>53.1 m	4.78 (3.79-5.62)	1.10 (0.91-1.30)	1.04 (0.91-1.19)
	≤53.1 m	4.36 (3.56-4.89)	1.0 (ref)	1.0 (ref)
			0.324	0.522
Wake after sleep onset	≥100 m	5.25 (4.24-6.07)	1.29 (1.07-1.50)	1.06 (0.93-1.21)
	<100 m	4.08 (3.36-4.58)	1.0 (ref)	1.0 (ref)
			0.020	0.440
Total Sleep Time	<6.0 h	4.14 (3.32-4.83)	0.94 (0.78-1.12)	0.94 (0.84-1.07)
	6.0-7.5 hr	4.42 (3.58-5.03)	1.0 (ref)	1.0 (ref)
	≥7.6 hr	4.86 (3.82-5.72)	1.10 (0.93-1.33)	1.08 (0.95-1.25)
			>0.258	>0.246
Pittsburgh Sleep Quality Index	>8	5.23 (3.94-6.17)	1.24 (1.01-1.45)	1.12 (0.97-1.26)
	≤8	4.22 (3.46-4.77)	1.0 (ref)	1.0 (ref)
			0.038	0.015
Epworth Sleepiness Scale	>7	4.61 (3.62-5.34)	1.05 (0.84-1.27)	1.01 (0.90-1.15)
	≤7	4.40 (3.66-4.94)	1.0 (ref)	1.0 (ref)
			0.712	0.918

*Adjusted for age and clinic site

**Bootstrapping with 1000 samples used to estimate 95% confidence intervals and p-values

†Multivariable Models adjusted for : Age, clinic site, depression, use of antidepressants, cognitive functioning, 6 meter walking speed, self-reported health status, any impairments in instrumental activities of daily living, selected medical conditions, probable dementia, and living independently

Chapter 6: Conclusions

Prior epidemiologic research has shown that sleep disturbances and sleep disorders are common in older adults[1, 2] and obtaining adequate sleep is an important aspect of healthy aging. Although insufficient sleep has been linked to many medical conditions such as depression[3, 4], cardiovascular disease[5, 5, 6], frailty[7, 8], impaired cognitive functioning[9, 10] and mortality[8], additional work is need to better understand the mechanistic pathways, correlates and consequences of insufficient sleep in an older adult population. The field also faces many challenges, including lack of consistent and concrete measures of sleep disturbances and disorders across studies, inconsistent results in similar populations and use of screening tools that have not been validated in older populations.

Utilizing data from the Study of Osteoporotic Fractures in Women (SOF), and the Outcomes of Sleep Disorders in Older Men (MrOS) cohort studies, this dissertation aimed to improve upon our understanding of the epidemiology of sleep disturbances in older adults by focusing on three specific areas that need additional research.

In addressing the first area, this dissertation found that a popular questionnaire for assessing risk of sleep apnea (STOP-BANG) has a high sensitivity at a cut point of $\geq$3 (94%) in an older male population, but also has an unacceptably high false positive rate (87.3%). Furthermore, higher STOP-BANG cut points resulted in high specificities (97.4% at a cut point of $\geq$5) but unacceptably high false negative rates (93.6%) and little impact on probability revision beyond the prevalence of OSA in the cohort. As a result of the work presented in this dissertation, we conclude that the STOP-BANG questionnaire is insufficient for the screening of OSA in older adults.

In addition to answering the question of whether the STOP-BANG questionnaire is an effective screening tool, this dissertation also aimed to examine how well the individual components of the STOP-BANG predict risk of OSA in an older adult male population. Results suggested that in this population

individual STOP-BANG components of Snoring, Tired, Observed apneas, Pressure and Age were not strongly predictive of OSA. Furthermore, body mass index and neck circumference had fair discriminatory power for detecting severe OSA in older men. In comparing the area under the curve (AUC) for each of these measures, with the AUC for the STOP-BANG questionnaire as a whole, these results suggested that BMI and neck circumference individually are only slightly better at identifying older men with sleep apnea than the STOP-BANG questionnaire. Secondary analyses examining associations between the STOP-BANG questionnaire and excessive daytime sleepiness suggest that while the STOP-BANG may not be particularly useful for detecting OSA, it is associated with excessively sleepiness, and may be affected by other sleep-related disorders in this population.

Taken as a whole, the results of this first manuscript confirm that the STOP-BANG questionnaire is insufficient for detecting OSA in an older male population, and that additional research into the risk factors and characteristics of OSA is warranted.

Identifying and exploring novel measures of sleep disturbances in older adults was another area in which this dissertation chose to explore. While the field has chosen to focus on averages, such as the average total sleep time during a specified period of time, very little work has focused on whether greater variability in sleep could also be a measure of sleep disturbance. Using data from the Outcomes of Sleep Disorders in Older Men (MrOS Sleep) study, this dissertation aimed to examine the characteristics and correlates of variability in sleep/wake parameters among community-dwelling older men. Results from this study suggested that qualitatively older men appear to exhibit slightly less variability in sleep than a slightly younger, middle aged cohort (CARDIA)[94], and although the majority of variability in sleep/wake parameters is derived from comparing measures of sleep between individuals, a substantial amount of intra-individual variability exists. Furthermore, several demographic, anthropometric, lifestyle, medical, medication and physical functioning factors were independently associated with greater sleep/wake variability, and African American race, living

alone, antidepressant use, depression, greater body mass index and multimorbidity appear to be the strongest independent predictors of greater sleep/wake variability in this population. Since these are potentially modifiable factors, these findings help direct the design of future intervention studies aimed at improving sleep among older adults. However, it is important that future research explore the mechanisms underlying these associations in order to better understand how these factors impact sleep, or how sleep impacts these factors in older adults.

Finally, the third area this dissertation focused on was improving our understanding of the impact that sleep disturbances have on the risk of all-cause hospitalizations in an older female population. Research into developing a understanding of this association was warranted, based on prior observation of higher rates of health care utilization among adults with insomnia, or self-reported sleep disturbances. Using data from the Study of Osteoporotic Fractures, including objective assessments of sleep disturbances, and taking advantage of the linkage to Medicare Claims and Kaiser encounter data, this dissertation aimed to address the question of whether sleep disturbances were independently associated with all-cause hospitalizations in older women.

The results suggested that sleep disturbances are associated with an increased risk of hospitalization, but that results are explained, in a large part by health-related factors such as comorbidities, physical functioning, depression and health status. In a further analysis, we discovered that sleep disturbances are independently associated with a greater odds of hospitalizations related to COPD and CHF in older women, and although we are unable to assess it, obstructive sleep apnea, as well as nocturnal dyspnea could underlie these associations. Future studies are needed to determine whether improving sleep helps to improve health outcomes, such as reductions in COPD and CHF related hospitalizations. Future studies should also examine associations between sleep disturbances and nursing home, hospice and home health care utilization.

This dissertation is the first to examine these specific topics exclusively in older adult populations. These results provide a foundation from which to

develop additional research into better understanding the role that sleep plays in affecting the health and successful aging of older adult populations.

Bibliography

(1) Foley D, Ancoli-Israel S, Britz P, Walsh J. Sleep disturbances and chronic disease in older adults: results of the 2003 National Sleep Foundation Sleep in America Survey. *J Psychosom Res* 2004;56(5):497-502.

(2) Foley DJ, Monjan AA, Brown SL, Simonsick EM, Wallace RB, Blazer DG. Sleep complaints among elderly persons: an epidemiologic study of three communities. *Sleep* 1995;18(6):425-432.

(3) Paudel ML, Taylor BC, Diem SJ et al. Association between depressive symptoms and sleep disturbances in community-dwelling older men. *J Am Geriatr Soc* 2008;56(7):1228-1235.

(4) Maglione JE, Ancoli-Israel S, Peters KW et al. Depressive symptoms and subjective and objective sleep in community-dwelling older women. *J Am Geriatr Soc* 2012;60(4):635-643.

(5) Malhotra A, Loscalzo J. Sleep and cardiovascular disease: an overview. *Prog Cardiovasc Dis* 2009;51(4):279-284.

(6) Shahar E, Whitney CW, Redline S et al. Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. *Am J Respir Crit Care Med* 2001;163(1):19-25.

(7) Ensrud KE, Blackwell TL, Redline S et al. Sleep disturbances and frailty status in older community-dwelling men. *J Am Geriatr Soc* 2009;57(11):2085-2093.

(8) Ensrud KE, Blackwell TL, Ancoli-Israel S et al. Sleep disturbances and risk of frailty and mortality in older men. *Sleep Med* 2012.

(9) Ferrie JE, Shipley MJ, Akbaraly TN, Marmot MG, Kivimaki M, Singh-Manoux A. Change in sleep duration and cognitive function: findings from the Whitehall II Study. *Sleep* 2011;34(5):565-573.

(10) Tranah GJ, Blackwell T, Stone KL et al. Circadian activity rhythms and risk of incident dementia and mild cognitive impairment in older women. *Ann Neurol* 2011;70(5):722-732.

(11) Neikrug AB, Ancoli-Israel S. Sleep disorders in the older adult - a mini-review. *Gerontology* 2010;56(2):181-189.

(12) Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328(17):1230-1235.

(13) Ancoli-Israel S, Klauber MR, Kripke DF, Parker L, Cobarrubias M. Sleep apnea in female patients in a nursing home. Increased risk of mortality. *Chest* 1989;96(5):1054-1058.

(14) Ancoli-Israel S. Epidemiology of sleep disorders. *Clin Geriatr Med* 1989;5(2):347-362.

(15) Ancoli-Israel S, Kripke DF, Klauber MR, Mason WJ, Fell R, Kaplan O. Sleep-disordered breathing in community-dwelling elderly. *Sleep* 1991;14(6):486-495.

(16) Ancoli-Israel S, Klauber MR, Stepnowsky C, Estline E, Chinn A, Fell R. Sleep-disordered breathing in African-American elderly. *Am J Respir Crit Care Med* 1995;152(6 Pt 1):1946-1949.

(17) Ancoli-Israel S, Kripke DF, Klauber MR et al. Morbidity, mortality and sleep-disordered breathing in community dwelling elderly. *Sleep* 1996;19(4):277-282.

(18) Mehra R, Stone KL, Blackwell T et al. Prevalence and correlates of sleep-disordered breathing in older men: osteoporotic fractures in men sleep study. *J Am Geriatr Soc* 2007;55(9):1356-1364.

(19) Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: a population health perspective. *Am J Respir Crit Care Med* 2002;165(9):1217-1239.

(20) Bixler EO, Vgontzas AN, Lin HM et al. Association of hypertension and sleep-disordered breathing. *Arch Intern Med* 2000;160(15):2289-2295.

(21) Duran J, Esnaola S, Rubio R, Iztueta A. Obstructive sleep apnea-hypopnea and related clinical features in a population-based sample of subjects aged 30 to 70 yr. *Am J Respir Crit Care Med* 2001;163(3 Pt 1):685-689.

(22) Nieto FJ, Young TB, Lind BK et al. Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. Sleep Heart Health Study. *JAMA* 2000;283(14):1829-1836.

(23) Young T, Peppard P, Palta M et al. Population-based study of sleep-disordered breathing as a risk factor for hypertension. *Arch Intern Med* 1997;157(15):1746-1752.

(24) Redline S, Yenokyan G, Gottlieb DJ et al. Obstructive sleep apnea-hypopnea and incident stroke: the sleep heart health study. *Am J Respir Crit Care Med* 2010;182(2):269-277.

(25) Gottlieb DJ, Yenokyan G, Newman AB et al. Prospective study of obstructive sleep apnea and incident coronary heart disease and heart failure: the sleep heart health study. *Circulation* 2010;122(4):352-360.

(26) Baldwin CM, Griffith KA, Nieto FJ, O'Connor GT, Walsleben JA, Redline S. The association of sleep-disordered breathing and sleep symptoms with quality of life in the Sleep Heart Health Study. *Sleep* 2001;24(1):96-105.

(27) Finn L, Young T, Palta M, Fryback DG. Sleep-disordered breathing and self-reported general health status in the Wisconsin Sleep Cohort Study. *Sleep* 1998;21(7):701-706.

(28) Gottlieb DJ, Yao Q, Redline S, Ali T, Mahowald MW. Does snoring predict sleepiness independently of apnea and hypopnea frequency? *Am J Respir Crit Care Med* 2000;162(4 Pt 1):1512-1517.

(29) Abrishami A, Khajehdehi A, Chung F. A systematic review of screening questionnaires for obstructive sleep apnea. *Can J Anaesth* 2010;57(5):423-438.

(30) Sforza E, Chouchou F, Pichot V, Herrmann F, Barthelemy JC, Roche F. Is the Berlin questionnaire a useful tool to diagnose obstructive sleep apnea in the elderly? *Sleep Med* 2011;12(2):142-146.

(31) Hall MJ, DeFrances CJ, Williams SN, Golosinskiy A, Schwartzman A. National Hospital Discharge Survey: 2007 summary. *Natl Health Stat Report* 2010;(29):1-20, 24.

(32) Bauer DC, Browner WS, Cauley JA et al. Factors associated with appendicular bone mass in older women. The Study of Osteoporotic Fractures Research Group. *Ann Intern Med* 1993;118(9):657-665.

(33) Cummings SR, Black DM, Nevitt MC et al. Appendicular bone density and age predict hip fracture in women. The Study of Osteoporotic Fractures Research Group. *JAMA* 1990;263(5):665-668.

(34) NHLBI: Health Information for the Public. Sleep Apnea: What is Sleep Apnea? 2013. U.S. Department of Health and Human Services. 9-6-2013.
Ref Type: Online Source

(35) Punjabi NM, Caffo BS, Goodwin JL et al. Sleep-disordered breathing and mortality: a prospective cohort study. *PLoS Med* 2009;6(8):e1000132.

(36) Young T, Finn L, Peppard PE et al. Sleep disordered breathing and mortality: eighteen-year follow-up of the Wisconsin sleep cohort. *Sleep* 2008;31(8):1071-1078.

(37) Ancoli IS, Klauber MR, Kripke DF. Sleep apnea in female patients in a nursing home. Increased risk of mortality. *Chest* 1989;(96):1054-1058.

(38) Bixler EO, Vgontzas AN, Lin HM et al. Association of hypertension and sleep-disordered breathing. *Arch Intern Med* 2000;160(15):2289-2295.

(39) Duran J, Esnaola S, Rubio R, Iztueta A. Obstructive sleep apnea-hypopnea and related clinical features in a population-based sample of subjects aged 30 to 70 yr. *Am J Respir Crit Care Med* 2001;163(3 Pt 1):685-689.

(40) Nieto FJ, Young TB, Lind BK et al. Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. Sleep Heart Health Study. *JAMA* 2000;283(14):1829-1836.

(41) Young T, Peppard P, Palta M et al. Population-based study of sleep-disordered breathing as a risk factor for hypertension. *Arch Intern Med* 1997;157(15):1746-1752.

(42) Redline S, Yenokyan G, Gottlieb DJ et al. Obstructive sleep apnea-hypopnea and incident stroke: the sleep heart health study. *Am J Respir Crit Care Med* 2010;182(2):269-277.

(43) Shahar E, Whitney CW, Redline S et al. Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. *Am J Respir Crit Care Med* 2001;163(1):19-25.

(44) Gottlieb DJ, Yenokyan G, Newman AB et al. Prospective study of obstructive sleep apnea and incident coronary heart disease and heart failure: the sleep heart health study. *Circulation* 2010;122(4):352-360.

(45) Young T, Blustein J, Finn L, Palta M. Sleep-disordered breathing and motor vehicle accidents in a population-based sample of employed adults. *Sleep* 1997;20(8):608-613.

(46) Fulda S, Schulz H. Cognitive dysfunction in sleep disorders. *Sleep Med Rev* 2001;5(6):423-445.

(47) Redline S, Strauss ME, Adams N et al. Neuropsychological function in mild sleep-disordered breathing. *Sleep* 1997;20(2):160-167.

(48) Baldwin CM, Griffith KA, Nieto FJ, O'Connor GT, Walsleben JA, Redline S. The association of sleep-disordered breathing and sleep symptoms with quality of life in the Sleep Heart Health Study. *Sleep* 2001;24(1):96-105.

(49) Finn L, Young T, Palta M, Fryback DG. Sleep-disordered breathing and self-reported general health status in the Wisconsin Sleep Cohort Study. *Sleep* 1998;21(7):701-706.

(50) Chung F, Yegneswaran B, Liao P et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008;108(5):812-821.

(51) Silva GE, Vana KD, Goodwin JL, Sherrill DL, Quan SF. Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. *J Clin Sleep Med* 2011;7(5):467-472.

(52) Ensrud KE, Blackwell TL, Ancoli-Israel S et al. Sleep disturbances and risk of frailty and mortality in older men. *Sleep Med* 2012;13(10):1217-1225.

(53) Yaggi HK, Concato J, Kernan WN, Lichtman JH, Brass LM, Mohsenin V. Obstructive sleep apnea as a risk factor for stroke and death. *N Engl J Med* 2005;353(19):2034-2041.

(54) Joshi GP, Ankichetty SP, Gan TJ, Chung F. Society for Ambulatory Anesthesia consensus statement on preoperative selection of adult patients with obstructive sleep apnea scheduled for ambulatory surgery. *Anesth Analg* 2012;115(5):1060-1068.

(55) The ASPAN obstructive sleep apnea in the adult patient evidence-based practice recommendation. *J Perianesth Nurs* 2012;27(5):309-315.

(56) Abrishami A, Khajehdehi A, Chung F. A systematic review of screening questionnaires for obstructive sleep apnea. *Can J Anaesth* 2010;57(5):423-438.

(57) Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. *Anesthesiology* 2009;110(4):928-939.

(58) Chung F, Subramanyam R, Liao P, Sasaki E, Shapiro C, Sun Y. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth* 2012;108(5):768-775.

(59) Gottlieb DJ, Whitney CW, Bonekat WH et al. Relation of sleepiness to respiratory disturbance index: the Sleep Heart Health Study. *Am J Respir Crit Care Med* 1999;159(2):502-507.

(60) Blank JB, Cawthon PM, Carrion-Petersen ML et al. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp Clin Trials* 2005;26(5):557-568.

(61) Orwoll E, Blank JB, Barrett-Connor E et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study--a large observational study of the determinants of fracture in older men. *Contemp Clin Trials* 2005;26(5):569-585.

(62) Redline S, Sanders MH, Lind BK et al. Methods for obtaining and analyzing unattended polysomnography data for a multicenter study. Sleep Heart Health Research Group. *Sleep* 1998;21(7):759-767.

(63) Quan SF, Howard BV, Iber C et al. The Sleep Heart Health Study: design, rationale, and methods. *Sleep* 1997;20(12):1077-1085.

(64) Buysse DJ, Reynolds CF, III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28(2):193-213.

(65) Buysse DJ, Reynolds CF, III, Monk TH, Hoch CC, Yeager AL, Kupfer DJ. Quantification of subjective sleep quality in healthy elderly men and women using the Pittsburgh Sleep Quality Index (PSQI). *Sleep* 1991;14(4):331-338.

(66) Weaver TE, Laizner AM, Evans LK et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep* 1997;20(10):835-843.

(67) Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;14(6):540-545.

(68) Johns MW. Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the epworth sleepiness scale: failure of the MSLT as a gold standard. *J Sleep Res* 2000;9(1):5-11.

(69) Epstein LJ, Kristo D, Strollo PJ, Jr. et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med* 2009;5(3):263-276.

(70) Morin CM, Colecchi C, Stone J, Sood R, Brink D. Behavioral and pharmacological therapies for late-life insomnia: a randomized controlled trial. *JAMA* 1999;281(11):991-999.

(71) Riedel BW, Lichstein KL. Strategies for evaluating adherence to sleep restriction treatment for insomnia. *Behav Res Ther* 2001;39(2):201-212.

(72) Harvey AG. A cognitive model of insomnia. *Behav Res Ther* 2002;40(8):869-893.

(73) Hauri P, Fisher J. Persistent psychophysiologic (learned) insomnia. *Sleep* 1986;9(1):38-53.

(74) Perlis ML, Giles DE, Mendelson WB, Bootzin RR, Wyatt JK. Psychophysiological insomnia: the behavioural model and a neurocognitive perspective. *J Sleep Res* 1997;6(3):179-188.

(75) Spielman AJ, Caruso LS, Glovinsky PB. A behavioral perspective on insomnia treatment. *Psychiatr Clin North Am* 1987;10(4):541-553.

(76) Kripke DF, Garfinkel L, Wingard DL, Klauber MR, Marler MR. Mortality associated with sleep duration and insomnia. *Arch Gen Psychiatry* 2002;59(2):131-136.

(77) Kripke DF, Langer RD, Elliott JA, Klauber MR, Rex KM. Mortality related to actigraphic long and short sleep. *Sleep Med* 2011;12(1):28-33.

(78) Stone KL, Ewing SK, Ancoli-Israel S et al. Self-reported sleep and nap habits and risk of mortality in a large cohort of older women. *J Am Geriatr Soc* 2009;57(4):604-611.

(79) Tamakoshi A, Ohno Y. Self-reported sleep duration as a predictor of all-cause mortality: results from the JACC study, Japan. *Sleep* 2004;27(1):51-54.

(80) Vgontzas AN, Liao D, Pejovic S et al. Insomnia with short sleep duration and mortality: the Penn State cohort. *Sleep* 2010;33(9):1159-1164.

(81) Knutson KL, Rathouz PJ, Yan LL, Liu K, Lauderdale DS. Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the CARDIA study. *Sleep* 2007;30(6):793-796.

(82) van Hilten JJ, Braat EA, van der Velde EA, Middelkoop HA, Kerkhof GA, Kamphuisen HA. Ambulatory activity monitoring during sleep: an evaluation of internight and intrasubject variability in healthy persons aged 50-98 years. *Sleep* 1993;16(2):146-150.

(83) Buysse DJ, Cheng Y, Germain A et al. Night-to-night sleep variability in older adults with and without chronic insomnia. *Sleep Med* 2010;11(1):56-64.

(84) Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep* 2003;26(3):342-392.

(85) Cole MG, Dendukuri N. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. *Am J Psychiatry* 2003;160(6):1147-1156.

(86) Blackwell T, Ancoli-Israel S, Redline S, Stone KL. Factors that may influence the classification of sleep-wake by wrist actigraphy: the MrOS Sleep Study. *J Clin Sleep Med* 2011;7(4):357-367.

(87) Blackwell T, Redline S, Ancoli-Israel S et al. Comparison of sleep parameters from actigraphy and polysomnography in older women: the SOF study. *Sleep* 2008;31(2):283-291.

(88) Blackwell T, Ancoli-Israel S, Redline S, Stone KL. Factors that may influence the classification of sleep-wake by wrist actigraphy: the MrOS Sleep Study. *J Clin Sleep Med* 2011;7(4):357-367.

(89) Pahor M, Chrischilles EA, Guralnik JM, Brown SL, Wallace RB, Carbonin P. Drug data coding and analysis in epidemiologic studies. *Eur J Epidemiol* 1994;10(4):405-411.

(90) Sheikh J., Yesavage J. Geriatric Depression Scale:recent evidence and development of a shorter version. *Clin Gerontol* 2012;5:165-173.

(91) Fitti JE, Kovar MG. The Supplement on Aging to the 1984 National Health Interview Survey. *Vital Health Stat 1* 1987;(21):1-115.

(92) Pincus T, Summey JA, Soraci SA, Jr., Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983;26(11):1346-1353.

(93) Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. *J Clin Psychiatry* 1987;48(8):314-318.

(94) Knutson KL, Rathouz PJ, Yan LL, Liu K, Lauderdale DS. Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the CARDIA study. *Sleep* 2007;30(6):793-796.

(95) Durrence HH, Lichstein KL. The sleep of African Americans: a comparative review. *Behav Sleep Med* 2006;4(1):29-44.

(96) Wetter DW, Young TB. The relation between cigarette smoking and sleep disturbance. *Prev Med* 1994;23(3):328-334.

(97) Lustberg L, Reynolds CF. Depression and insomnia: questions of cause and effect. *Sleep Med Rev* 2000;4(3):253-262.

(98) Roth T, Zorick F, Sicklesteel J, Stepanski E. Effects of benzodiazepines on sleep and wakefulness. *Br J Clin Pharmacol* 1981;11 Suppl 1:31S-35S.

(99) Patel SR, Blackwell T, Redline S et al. The association between sleep duration and obesity in older adults. *Int J Obes (Lond)* 2008;32(12):1825-1834.

(100) Simon GE, VonKorff M. Prevalence, burden, and treatment of insomnia in primary care. *Am J Psychiatry* 1997;154(10):1417-1423.

(101) Hayward R, Jordan KP, Croft P. Healthcare use in adults with insomnia: a longitudinal study. *Br J Gen Pract* 2010;60(574):334-340.

(102) Novak M, Mucsi I, Shapiro CM, Rethelyi J, Kopp MS. Increased utilization of health services by insomniacs--an epidemiological perspective. *J Psychosom Res* 2004;56(5):527-536.

(103) Chevalier H, Los F, Boichut D et al. Evaluation of severe insomnia in the general population: results of a European multinational survey. *J Psychopharmacol* 1999;13(4 Suppl 1):S21-S24.

(104) Daley M, Morin CM, LeBlanc M, Gregoire JP, Savard J, Baillargeon L. Insomnia and its relationship to health-care utilization, work absenteeism, productivity and accidents. *Sleep Med* 2009;10(4):427-438.

(105) Leger D, Guilleminault C, Bader G, Levy E, Paillard M. Medical and socio-professional impact of insomnia. *Sleep* 2002;25(6):625-629.

(106) Ozminkowski RJ, Wang S, Walsh JK. The direct and indirect costs of untreated insomnia in adults in the United States. *Sleep* 2007;30(3):263-273.

(107) Kaufmann CN, Canham SL, Mojtabai R et al. Insomnia and health services utilization in middle-aged and older adults: results from the Health and Retirement Study. *J Gerontol A Biol Sci Med Sci* 2013;68(12):1512-1517.

(108) Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA* 1993;269(18):2386-2391.

(109) Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998;36(1):8-27.

(110) Chronic Condition Data Warehouse - Chronic Condition Categories. *Internet* 2014;Available at: URL: http://www.ccwdata.org.

(111) Chevalier H, Los F, Boichut D et al. Evaluation of severe insomnia in the general population: results of a European multinational survey. *J Psychopharmacol* 1999;13(4 Suppl 1):S21-S24.

(112) Daley M, Morin CM, LeBlanc M, Gregoire JP, Savard J, Baillargeon L. Insomnia and its relationship to health-care utilization, work absenteeism, productivity and accidents. *Sleep Med* 2009;10(4):427-438.

(113) Hayward R, Jordan KP, Croft P. Healthcare use in adults with insomnia: a longitudinal study. *Br J Gen Pract* 2010;60(574):334-340.

(114) Leger D, Guilleminault C, Bader G, Levy E, Paillard M. Medical and socio-professional impact of insomnia. *Sleep* 2002;25(6):625-629.

(115) Novak M, Mucsi I, Shapiro CM, Rethelyi J, Kopp MS. Increased utilization of health services by insomniacs--an epidemiological perspective. *J Psychosom Res* 2004;56(5):527-536.

(116) Ozminkowski RJ, Wang S, Walsh JK. The direct and indirect costs of untreated insomnia in adults in the United States. *Sleep* 2007;30(3):263-273.

(117) Simon GE, VonKorff M. Prevalence, burden, and treatment of insomnia in primary care. *Am J Psychiatry* 1997;154(10):1417-1423.

(118) Kutty K. Sleep and chronic obstructive pulmonary disease. *Curr Opin Pulm Med* 2004;10(2):104-112.

(119) Sharma B, Owens R, Malhotra A. Sleep in congestive heart failure. *Med Clin North Am* 2010;94(3):447-464.

(120) Stepanski EJ, Rybarczyk B. Emerging research on the treatment and etiology of secondary or comorbid insomnia. *Sleep Med Rev* 2006;10(1):7-18.

Appendices

Appendix A. A Comparison of STOP-BANG Questions and equivalent questions obtained from the MrOS cohort

Component	STOP-BANG **Definition**	MrOS Sleep Study **Definition**
S**nore loudly**	Do you snore loudly (louder than talking or loud enough to be heard through closed doors?) (Yes=Snore loudly)	During the past month, how often have you had trouble sleeping because you cough or snore loudly? (Less than once a week or more = Snore loudly)
T**ired**	Do you often feel tired, fatigued, or sleepy during the daytime? (Yes= Tired)	Do you have difficulty being as active as you want to be in the morning (or afternoon) because you are sleepy or tired? (Extreme/moderate/a little difficulty = Tired)
O**bserved apneas**	Has anyone observed you stop breathing during your sleep? (Yes=Observed apneas)	Please ask your bed partner how often in the past month you have had long pauses between breaths while asleep? (Less than once a week or more = Observed apneas)
P**ressure (Hypertension)**	Do you have, or are you being treated for high blood pressure? (Yes=Pressure)	Has a healthcare provider or doctor ever told you that you have hypertension or high blood pressure? (Yes=Pressure)
B**MI**	BMI >35 kg/m^2 ?(Yes=BMI)	BMI >35 kg/m^2 ?(Yes=BMI)
A**ge**	Age >50 y? (Yes=Age)	Age >50 y? (Yes=Age)
N**eck circumference**	Neck circumference >40 cm? (Yes=Neck circumference)	Neck circumference >40 cm? (Yes=Neck circumference)
Male G**ender**	Male gender? (Yes=Gender)	Male gender? (Yes=Gender)

Appendix B. Predictive parameters of Different STOP-BANG score cut-offs using a 4% oxygen desaturation criteria

STOP-BANG score cut-off	Sensitivity (%)	Specificity (%)	PPV (%)	NPV(%)
All OSA (AHI≥5) vs. None (AHI<5)				
≥3 vs. 2	91.3	16.0	62.7	54.6
≥4 vs. 0-3	67.0	46.6	65.9	47.8
≥5 vs. 0-4	37.7	73.2	68.5	43.2
≥6 vs. 0-5	16.1	92.0	75.5	41.5
≥7 vs. 0-6	4.1	98.0	76.3	39.9
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)				
≥3 vs. 2	92.3	12.9	27.5	82.4
≥4 vs. 0-3	72.0	42.1	30.8	80.8
≥5 vs. 0-4	45.9	71.0	36.2	78.6
≥6 vs. 0-5	22.0	90.3	44.8	76.4
≥7 vs. 0-6	6.5	97.9	52.7	74.6
Severe OSA (AHI≥30) vs. Moderate/Mild/None (AHI<30)				
≥3 vs. 2	94.7	12.3	10.6	95.5
≥4 vs. 0-3	73.1	39.6	11.7	93.1
≥5 vs. 0-4	48.6	68.2	14.4	92.4
≥6 vs. 0-5	25.5	88.5	19.6	91.6
≥7 vs. 0-6	6.4	97.1	19.4	90.4

***Using 4% desaturation criteria**

Appendix C. Predictive parameters of STOP-BANG Component Combinations

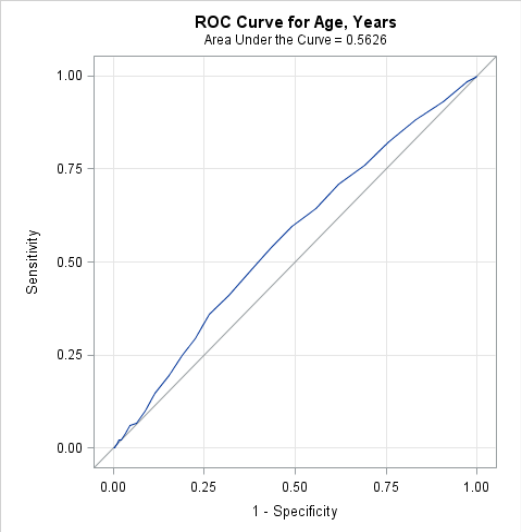
STOP-BANG score cut-off	Prevalence (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV(%)	OR (95% CI)**
Snore loudly + Tired	22.3	25.3	80.0	49.2	58.3	1.35 (1.13-1.62)
Snore loudly + OA	13.8	16.6	88.3	52.0	58.0	1.54 (1.24-1.92)
Snore loudly + HTN	22.7	27.8	81.2	53.1	59.5	1.65 (1.38-1.98)
Snore loudly + BMI	1.9	2.8	98.8	63.0	57.0	2.17 (1.23-3.83)
Snore loudly + Neck	17.6	23.6	87.1	58.3	59.8	2.23 (1.82-2.74)
Tired + OA	11.1	12.8	90.2	49.8	57.5	1.41 (1.11-1.79)
Tired + HTN	23.8	25.8	77.7	46.9	57.8	1.21 (1.02-1.45)
Tired + BMI	1.9	2.8	98.8	64.8	57.0	2.29 (1.29-4.07)
Tired + Neck circum	17.5	22.2	86.2	55.2	59.2	1.85 (1.51-2.27)
OA + HTN	11.3	14.4	91.2	55.5	58.2	1.76 (1.38-2.24)
OA + BMI	1.0	1.1	99.3	54.2	56.7	1.40 (0.61-3.20)
OA + NC	9.2	12.3	93.2	59.0	58.1	2.00 (1.53-2.61)
HTN + BMI	2.7	4.2	98.4	66.7	57.3	2.53 (1.55-4.12)
HTN + NC	21.7	28.3	83.3	56.5	60.3	2.06 (1.71-2.48)
BMI + NC	3.3	4.9	97.8	63.2	57.3	2.23 (1.44-3.44)

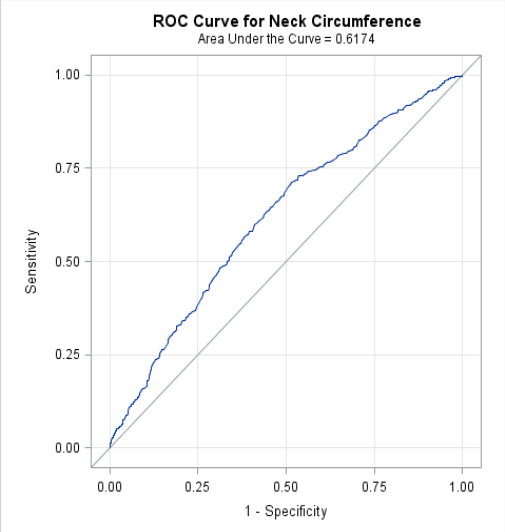
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)

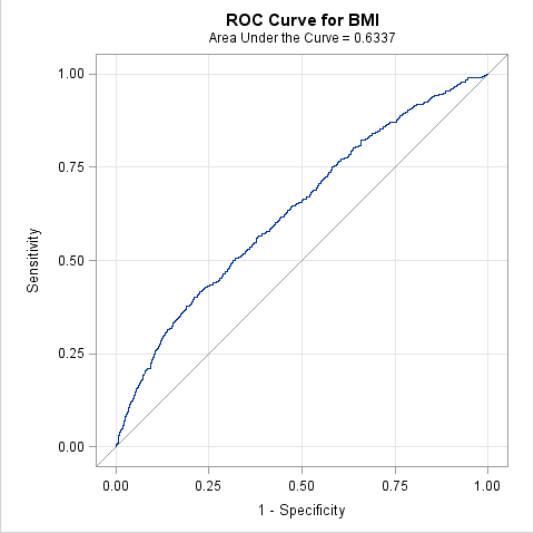
***Using 3% desaturation criteria**

****Age and clinic site adjusted**

Appendix D. ROC Curve Results for Age, Neck circumference, Body Mass Index and Prediction of Severe OSA.







Appendix E. Predictive parameters of Different STOPBANG Score cut offs by severity of OSA (using 3% criteria) after redefining the TIRED component.

Cut points	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
AHI ≥5				
≥3 vs. 0-2	88.0	21.2	80.6	32.2
≥4 vs. 0-3	60.1	56.1	83.6	27.4
≥5 vs. 0-4	30.8	80.5	85.5	23.8
≥6 vs. 0-5	12.0	96.2	92.1	22.7
≥7 vs. 0-6	2.5	99.5	94.8	21.5
AHI ≥15				
≥3 vs. 0-2	90.2	17.1	45.4	69.5
≥4 vs. 0-3	63.0	48.2	48.2	63.0
≥5 vs. 0-4	35.6	77.0	54.3	61.0
≥6 vs. 0-5	14.9	93.3	63.0	58.9
≥7 vs. 0-6	3.4	99.0	72.4	57.3
AHI ≥30				
≥3 vs. 0-2	93.0	15.4	18.9	91.2
≥4 vs. 0-3	69.5	46.1	21.5	87.7
≥5 vs. 0-4	41.5	74.3	25.5	85.7
≥6 vs. 0-5	18.0	91.4	30.8	84.0
≥7 vs. 0-6	4.0	98.4	34.5	82.9

OSA=Obstructive sleep apnea; PPV=Positive predictive value; NPV=Negative predictive value; AHI=Apnea Hypopnea Index.

Appendix F. Predictive parameters of Different STOPBANG Score cut offs by severity of OSA (using 3% criteria) after redefining the Pressure component

Cut points	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
AHI ≥5				
≥3 vs. 0-2	93.4	15.1	80.4	38.1
≥4 vs. 0-3	70.5	43.5	82.3	28.4
≥5 vs. 0-4	40.7	72.2	84.5	24.7
≥6 vs. 0-5	16.8	92.6	89.3	23.0
≥7 vs. 0-6	4.2	98.5	91.3	21.6
AHI ≥15				
≥3 vs. 0-2	94.4	10.5	44.7	71.1
≥4 vs. 0-3	73.6	37.1	47.3	64.7
≥5 vs. 0-4	45.0	67.4	51.4	61.6
≥6 vs. 0-5	20.4	89.5	60.0	59.5
≥7 vs. 0-6	5.3	97.7	63.1	57.4
AHI ≥30				
≥3 vs. 0-2	96.6	9.4	18.4	92.9
≥4 vs. 0-3	80.0	35.1	20.7	89.2
≥5 vs. 0-4	51.5	64.9	23.7	86.3
≥6 vs. 0-5	24.7	87.3	29.2	84.5
≥7 vs. 0-6	6.6	97.0	32.0	83.1

OSA=Obstructive sleep apnea; PPV=Positive predictive value; NPV=Negative predictive value; AHI=Apnea Hypopnea Index.

Appendix G. Predictive parameters of different STOP-BANG score cut-offs including all apneic events

STOP-BANG score cut-off	Prevalence of OSA (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
All OSA (AHI≥5) vs. None (AHI<5)*					
≥3 vs. 2	97.6	88.5	13.0	97.6	2.7
≥4 vs. 0-3		61.8	44.9	97.8	2.8
≥5 vs. 0-4		33.5	68.1	97.7	2.5
≥6 vs. 0-5		13.0	91.3	98.4	2.5
≥7 vs. 0-6		3.3	98.6	98.9	2.5
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)*					
≥3 vs. 2	78.9	89.6	15.8	79.9	28.8
≥4 vs. 0-3		63.8	46.4	81.7	25.5
≥5 vs. 0-4		35.4	73.7	83.4	23.3
≥6 vs. 0-5		14.2	92.0	87.0	22.3
≥7 vs. 0-6		3.8	98.7	91.4	21.5
Severe OSA (AHI≥30) vs. Moderate/Mild/None (AHI<30)*					
≥3 vs. 2	43.4	91.8	14.1	45.1	69.1
≥4 vs. 0-3		68.0	43.3	47.9	63.7
≥5 vs. 0-4		40.7	72.1	52.8	61.3
≥6 vs. 0-5		17.8	90.9	60.1	59.0
≥7 vs. 0-6		4.8	97.9	63.4	57.3

*AHI defined using PORDI0P variable that includes all apneic events, including subtle snoring.

PPV=Positive predictive value; NPV=Negative predictive value; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index

Appendix H. Predictive parameters of Different STOP-BANG score cut-offs excluding central sleep apnea events from the definition of OSA

STOP-BANG score cut-off	Prevalence of OSA (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
All OSA (AHI≥5) vs. None (AHI<5)*					
≥3 vs. 2	79.6	90.1	17.9	81.1	31.5
≥4 vs. 0-3		64.3	48.9	83.1	25.9
≥5 vs. 0-4		35.5	74.7	84.6	22.9
≥6 vs. 0-5		14.6	93.8	90.2	21.9
≥7 vs. 0-6		3.9	99.3	95.7	20.9
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)*					
≥3 vs. 2	43.9	91.6	14.1	45.5	68.2
≥4 vs. 0-3		67.7	43.2	48.2	63.1
≥5 vs. 0-4		40.3	71.9	52.8	60.6
≥6 vs. 0-5		18.1	91.1	61.4	58.7
≥7 vs. 0-6		5.0	98.1	67.7	56.9
Severe OSA (AHI≥30) vs. Moderate/Mild/None (AHI<30)*					
≥3 vs. 2	17.1	93.8	12.7	18.1	90.9
≥4 vs. 0-3		75.4	41.2	20.9	89.0
≥5 vs. 0-4		48.5	69.7	24.7	86.8
≥6 vs. 0-5		23.6	89.3	31.3	85.0
≥7 vs. 0-6		7.0	97.5	36.6	83.6

*AHI defined using POORDI3 variable that excludes central sleep apnea events.

PPV=Positive predictive value; NPV=Negative predictive value; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index

Appendix I. Predictive parameters of Different STOP-BANG score cut-offs excluding men with unusual occurrences of periodic breathing (i.e. Cheynes-Strokes)

STOP-BANG score cut-off	Prevalence of OSA (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
All OSA (AHI≥5) vs. None (AHI<5)*					
≥3 vs. 2	77.7	90.2	17.9	79.3	34.4
≥4 vs. 0-3		64.6	49.8	81.8	28.8
≥5 vs. 0-4		36.1	75.4	83.7	25.3
≥6 vs. 0-5		15.2	93.9	89.7	24.1
≥7 vs. 0-6		4.1	99.3	95.5	22.9
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)*					
≥3 vs. 2	40.7	92.0	14.0	42.4	71.8
≥4 vs. 0-3		67.9	43.1	45.1	66.2
≥5 vs. 0-4		41.4	71.9	50.3	64.1
≥6 vs. 0-5		19.2	91.0	59.4	62.1
≥7 vs. 0-6		5.5	98.2	67.4	60.2
Severe OSA (AHI≥30) vs. Moderate/Mild/None (AHI<30)*					
≥3 vs. 2	15.1	94.5	12.7	16.2	92.9
≥4 vs. 0-3		75.2	41.1	18.5	90.3
≥5 vs. 0-4		49.1	69.3	22.2	88.4
≥6 vs. 0-5		24.8	88.9	28.6	86.9
≥7 vs. 0-6		7.2	97.4	32.6	85.5

*AHI defined using 3% desaturation Criteria.

Men with unusual occurrences of periodic breathing lasting >10 minutes are excluded.

PPV=Positive predictive value; NPV=Negative predictive value; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index

Appendix J. Predictive parameters of different STOP-BANG score cut-offs restricted to 2,173 men who reported having a bed partner

STOP-BANG score cut-off	Prevalence of OSA (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
All OSA (AHI≥5) vs. None (AHI<5)*					
≥3 vs. 2	78.2	91.3	16.8	80.7	33.8
≥4 vs. 0-3		66.6	46.7	82.6	26.8
≥5 vs. 0-4		38.5	72.1	84.0	23.5
≥6 vs. 0-5		16.7	92.3	89.1	22.5
≥7 vs. 0-6		4.9	99.1	95.5	21.5
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)*					
≥3 vs. 2	40.2	92.4	12.4	44.1	68.4
≥4 vs. 0-3		68.8	39.9	46.1	63.1
≥5 vs. 0-4		42.7	68.6	50.4	61.6
≥6 vs. 0-5		20.3	89.3	58.7	60.0
≥7 vs. 0-6		6.4	97.7	67.1	58.3
Severe OSA (AHI≥30) vs. Moderate/Mild/None (AHI<30)*					
≥3 vs. 2	14.5	94.0	11.2	17.7	90.2
≥4 vs. 0-3		73.8	38.2	19.5	87.8
≥5 vs. 0-4		48.4	66.2	22.5	86.4
≥6 vs. 0-5		25.1	87.3	28.6	85.2
≥7 vs. 0-6		7.9	96.7	33.0	83.8

*AHI defined using 3% desaturation criteria

PPV=Positive predictive value; NPV=Negative predictive value; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index

Appendix K. Predictive Parameters of STOPBANG cut points stratified by Excessive Daytime Sleepiness

	Sensitivity	Specificity	PPV	NPV
AHI≥5				
ESS 0-10				
≥3 vs. 2	89.3	18.3	79.9	31.9
≥4 vs. 2-3	62.9	51.0	82.4	27.4
≥5 vs. 2-4	33.9	76.6	84.1	24.2
≥6 vs. 2-5	13.2	95.0	90.5	23.1
≥7 vs. 2-6	3.2	99.4	95.5	22.0
ESS>10				
≥3 vs. 2	96.4	17.4	83.6	52.2
≥4 vs. 2-3	77.2	43.5	85.7	30.3
≥5 vs. 2-4	48.0	65.2	85.8	22.3
≥6 vs. 2-5	24.8	85.5	88.2	20.6
≥7 vs. 2-6	8.6	98.6	96.3	19.8
AHI≥15				
ESS 0-10				
≥3 vs. 2	90.8	14.8	44.2	68.4
≥4 vs. 2-3	65.6	44.4	46.8	63.5
≥5 vs. 2-4	38.2	73.2	51.5	61.4
≥6 vs. 2-5	16.2	92.2	60.8	59.7
≥7 vs. 2-6	4.1	98.4	65.2	58.0
ESS>10				
≥3 vs. 2	97.2	9.3	49.7	78.3
≥4 vs. 2-3	79.2	32.1	51.8	62.6
≥5 vs. 2-4	53.4	61.7	56.2	58.9
≥6 vs. 2-5	28.7	82.4	60.0	55.6
≥7 vs. 2-6	11.2	96.4	74.1	54.1
AHI≥30				

ESS 0-10				
≥3 vs. 2	93.3	13.5	17.9	90.9
≥4 vs. 2-3	71.5	42.5	20.1	88.1
≥5 vs. 2-4	43.7	70.8	23.2	86.2
≥6 vs. 2-5	20.1	90.4	29.7	84.9
≥7 vs. 2-6	5.8	98.0	36.4	83.7
ESS>10				
≥3 vs. 2	97.6	7.3	23.0	91.3
≥4 vs. 2-3	84.2	29.8	25.4	86.9
≥5 vs. 2-4	61.0	58.8	29.6	84.2
≥6 vs. 2-5	31.7	79.6	30.6	80.4
≥7 vs. 2-6	9.8	93.4	29.6	78.5

*3% Desaturation Criteria

PPV=Positive predictive value; NPV=Negative predictive value; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index

Appendix L. Predictive Parameters of STOPBANG cut points by AHI severity and stratified by AGE

	Sensitivity	Specificity	PPV	NPV
AHI≥5				
Age <80 yrs				
≥3 vs. 2	91.1	18.0	80.3	35.6
≥4 vs. 2-3	66.7	50.1	83.1	29.1
≥5 vs. 2-4	39.9	74.7	85.3	25.3
≥6 vs. 2-5	16.9	93.2	90.1	23.4
≥7 vs. 2-6	4.7	99.1	95.0	22.1
Age ≥80 yrs				
≥3 vs. 2	88.0	18.8	80.8	28.7
≥4 vs. 2-3	59.9	50.3	82.4	24.5
≥5 vs. 2-4	25.4	77.0	81.0	21.0
≥6 vs. 2-5	9.2	95.8	89.4	21.4
≥7 vs. 2-6	2.0	100.0	100.0	20.9
AHI≥15				
Age <80 yrs				
≥3 vs. 2	92.7	13.4	43.3	72.1
≥4 vs. 2-3	70.2	42.0	46.3	66.4
≥5 vs. 2-4	45.4	69.3	51.3	64.0
≥6 vs. 2-5	20.9	89.6	58.9	61.4
≥7 vs. 2-6	6.2	97.7	66.3	59.4
Age ≥80 yrs				
≥3 vs. 2	89.6	16.2	49.4	63.0
≥4 vs. 2-3	61.7	45.7	51.0	56.6
≥5 vs. 2-4	29.2	79.1	56.0	55.0
≥6 vs. 2-5	11.7	95.0	68.2	54.1
≥7 vs. 2-6	2.6	99.3	76.9	52.7
AHI≥30				

Age <80 yrs				
≥3 vs. 2	94.7	11.9	16.5	92.3
≥4 vs. 2-3	76.2	39.3	18.8	90.0
≥5 vs. 2-4	51.4	65.9	21.8	88.0
≥6 vs. 2-5	26.0	87.3	27.5	86.5
≥7 vs. 2-6	8.2	96.9	32.5	85.1
Age ≥80 yrs				
≥3 vs. 2	92.8	15.2	24.0	88.0
≥4 vs. 2-3	68.9	45.4	26.7	83.5
≥5 vs. 2-4	37.8	78.9	34.0	81.5
≥6 vs. 2-5	15.0	93.8	40.9	79.3
≥7 vs. 2-6	3.3	98.9	46.2	78.0

*3% Desaturation Criteria

PPV=Positive predictive value; NPV=Negative predictive value; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index

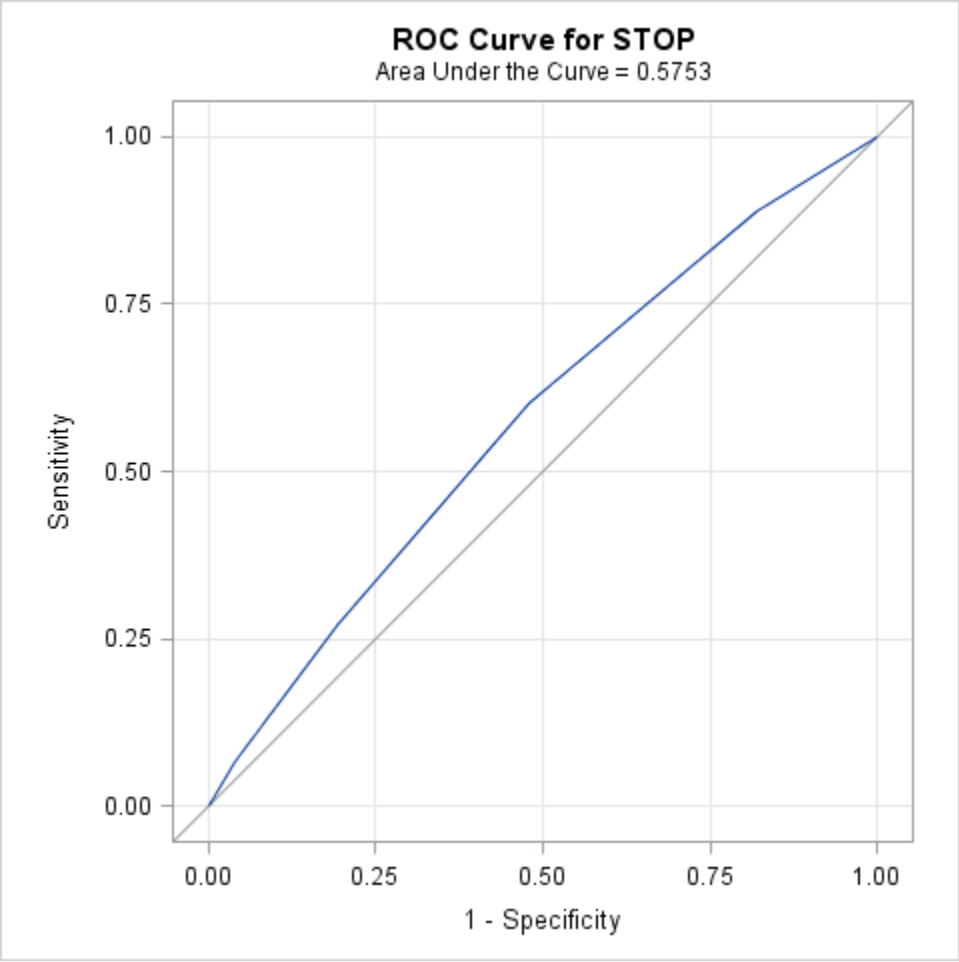
Appendix M. Predictive parameters of Different STOP score cut-offs

	Pre-Test Pr of OSA (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR-	AUC (95% CI)
All OSA (AHI≥5) vs. None (AHI<5)								
≥2 vs. 0-1	78.8	52.5	58.9	82.7	25.0	1.28	0.81	0.56 (0.54-0.58)
≥3 vs. 0-2		21.8	83.6	83.2	22.3	1.33	0.93	0.53 (0.51-0.54)
4 vs. 0-3		5.0	97.2	86.8	21.6	1.77	0.98	0.51 (0.50-0.52)
Moderate/Severe OSA (AHI≥15) vs. Mild/None (AHI<15)								
≥2 vs. 0-1	43.4	55.5	54.0	48.0	61.3	1.21	0.82	0.55 (0.53-0.57)
≥3 vs. 0-2		24.0	81.9	50.3	58.5	1.32	0.93	0.53 (0.51-0.55)
4 vs. 0-3		6.4	96.9	61.2	57.5	2.06	0.97	0.52 (0.50-0.53)
Severe OSA (AHI≥30) vs. Moderate/Mild/None (AHI<30)								
≥2 vs. 0-1	17.5	60.1	52.0	21.0	86.0	1.25	0.77	0.56 (0.54-0.59)
≥3 vs. 0-2		27.1	80.7	22.9	83.9	1.40	0.90	0.54 (0.52-0.56)
4 vs. 0-3		6.6	95.9	25.6	82.9	1.62	0.97	0.56 (0.54-0.58)

*AHI defined using ≥3% oxygen desaturation criteria.

PPV=Positive predictive value; NPV=Negative predictive value; OSA=Obstructive Sleep Apnea; AHI=Apnea Hypopnea Index

Appendix N. ROC Curve Results for STOP Questionnaire and Prediction of Severe OSA.



Straight line depicts no discriminative ability (i.e. Area under the ROC Curve= 0.5). In this model the area under the ROC curve was 0.575 (0.549-0.602).