SciMedCentral

Journal of Sleep Medicine & Disorders

Research Article

The Art of Divination: Do Standard Questionnaires Predict Sleep Apnea in Veterans with Kidney Disease?

Muna Canales[1]*, Shahab Bozorgmehri[2], Areef Ishani[3], I. David Weiner[4], Richard Berry[5], and Rebecca Beyth[6]

[1]Division of Nephrology, Malcom Randall VA Medical Center and University of Florida Department of Medicine, USA
[2]Division of Nephrology, Malcom Randall VA Medical Center and University of Florida Department of Medicine, USA
[3]Minneapolis VA Medical Center and University of Minnesota, Minneapolis, USA
[4]Division of Nephrology, Malcom Randall VA Medical Center and University of Florida Department of Medicine, USA
[5]Division of Pulmonary, Critical Care and Sleep Medicine, Malcom Randall VA Medical Center and University of Florida Department of Medicine, USA
[6]Division of General Medicine, Malcom Randall VA Medical Center and University of Florida Department of Medicine, USA

***Corresponding author**
Muna T. Canales, Division of Nephrology, Malcom Randall VA Medical Center and University of Florida Department of Medicine, Mail Code 111G, 1601 SW Archer Rd, Gainesville, FL 32610, Tel: 352-374-6102; Fax: 352-379-4048; Email: muna.canales@medicine.ufl.edu

Submitted: 25 September 2020

Accepted: 29 September 2020

Published: 30 September 2020

ISSN: 2379-0822

Keywords

- Sleep apnea
- Veterans
- Chronic kidney disease
- predictive validity

Abstract

Background: Sleep apnea is common among Veterans with kidney disease, but the predictive validity of existing questionnaires to identify sleep apnea in this population is unknown.

Methods: We performed a cross-sectional analysis of 248 Veterans selected only for having moderate to severe kidney disease. We collected anthropometrics, demographic data, responses to the Berlin Questionnaire and performed full unattended polysomnography on all participants. We calculated the predictive validity of the Berlin and a constructed STOP-BANG score to predict sleep apnea (apnea-hypopnea index ≥15 events/hour).

Results: Our cohort was 95% male of age 73.2±9.6 years; 78% were Caucasian and 20% were black; 52% had body-mass index ≥30 kg/m^2, 96% had hypertension and 39% had sleep apnea. There was no difference in daytime sleepiness by sleep apnea status (p=0.25). The mean (SD) Berlin score was 1.91 (0.79); the mean STOP-BANG score was 4.5 (1.3). While mean Berlin score did not differ by sleep apnea status (p=0.28), mean STOP-BANG score was higher for those with sleep apnea (p=0.003). The sensitivity, specificity, positive and negative predictive values for Berlin score ≥2 were 75%, 34%, 42% and 68%, respectively; accuracy was 50%, and area under the curve was 0.55(0.49-60). The sensitivity, specificity, positive and negative predictive values for STOP-BANG score ≥3 were 97%, 5%, 40% and 73%, respectively; accuracy was 41%, and area under the curve was 0.51(0.49-54). Alternative cut-points of Berlin and STOP-BANG did not substantially alter our findings.

Conclusion: Among Veterans with kidney disease, the Berlin and STOP-BANG questionnaires do not reliably distinguish sleep apnea disease from health.

ABBREVIATIONS

CKD: Chronic Kidney Disease; SA: Sleep Apnea; PSG: Polysomnography; BQ: Berlin Questionnaire; SNORE: Sleep and Nephrology Outcomes REsearch; MDRD: Modification of Diet in Renal Disease; eGFR: estimated Glomerular Filtration Rate; VISTA: Veterans Information System and Technology Architecture; VAMC: Veterans Affairs Medical Center; BMI: Body-Mass Index; AHI: Apnea-Hypopnea Index; ANOVA: Analysis of Variance; AUC: Area Under The Curve; HTN: Hypertension

INTRODUCTION

Chronic kidney disease (CKD), defined as a reduction in estimated glomerular filtration rate or presence of albuminuria, disproportionately affects United States Veterans. While it is estimated that 1 in 7 Americans has CKD, nearly 1 in 3 Veterans are affected [1]. CKD confers an independent and graded excess risk of numerous negative health consequences including heightened risk of cardiovascular disease and death [2,3]. Beyond renin-angiotensin system inhibition, few therapeutic options exist to slow progression of CKD and mitigate related consequences. As such, novel therapeutic strategies are urgently needed.

Sleep apnea (SA) also disproportionately affects patients with CKD, including Veterans with CKD. Literature to date has found the prevalence of moderate to severe SA in CKD to range from 40-50%, compared to prevalence of 6-13% in the general population [4-7]. Furthermore, dysregulated breathing during sleep defined by presence of SA, or related nocturnal hypoxia has been associated with faster decline in renal function and development of CKD in a spectrum of populations including U.S. Veterans [7-10]. As such, identification of CKD patients with SA

may identify a key group to target for SA screening and treatment as a novel strategy to potentially slow progression of CKD and related consequences.

Given the high burden and consequencess of SA in CKD, it is imperative to provide tools for the practitioner to be able to determine which CKD patients would benefit from formal testing for SA. Complicating the process of identifying SA in CKD is that CKD patients often do not display typical symptoms of SA [7,11-13]. Furthermore, while simple SA screening tools such as the Berlin Questionnaire (BQ) and the STOP-BANG exist to identify high risk for SA in the general population and in surgical patients, such tools have not been well tested against formal polysomnography (PSG) in Veterans or CKD patients [14]. We therefore formally tested the validity of the BQ to diagnose PSG-defined SA in a cohort of 248 Veterans with moderate to severe CKD who were not selected for sleep diagnoses or complaints. We also examined the validity of the STOP-BANG questionnaire components to predict SA in this cohort.

MATERIALS AND METHODS

We performed this analysis using baseline data from the Sleep and Nephrology Outcomes Research (SNORE) Study. The SNORE Study is a prospective cohort study that enrolled 254 Veterans with CKD. The detailed SNORE Study design has been published previously [15]. Briefly, Veteran enrollees in the North Florida/South Georgia Veterans Health System with stage G3b and G4 CKD (MDRD estimated glomerular filtration rate [eGFR] 15-44 ml/min/1.73m^2) on two occasions in the prior year were identified using the Veterans Information Systems and Technology Architecture (VISTA). We excluded Veterans who were 1) already on treatment for SA or on nocturnal oxygen therapy, 2) had active non-prostate or non-melanoma skin cancer, 3) on dialysis or with kidney or other solid organ transplant, or 4) with life expectancy < 3 years. The study visit was comprised of clinical assessments, completion of a Medical History Questionnaire, BQ, and the Epworth Sleepiness Scale (see Study Measurements) followed by a sleep study set-up and an unattended PSG at a local hotel. The next morning, participants returned to the Malcom Randall Veterans Affairs Medical Center (VAMC) to return equipment and complete blood and urine studies.

Our protocol was approved by the University of Florida Institutional Review Board and by the Malcom Randall VAMC Research and Development Committee. All participants provided written informed consent.

STUDY MEASUREMENTS

Sleep *Apnea Risk* Assessment

The BQ is a self-administered questionnaire with 10 questions (5 related to snoring, 4 related to awake-time sleepiness or drowsy-driving and 1 related to hypertension) and provision of height and weight to calculate body-mass index (BMI) [14]. Scoring yields a "high risk" or "low risk" categorization for mild or worse SA. Those meeting criteria in 2 of the 3 following categories are labeled as "high risk": 1) persistent symptoms (3-4 times per week) in two or more snoring questions, 2) persistent symptoms in awake-time sleepiness or drowsy driving, 3) history of hypertension or BMI > 30 kg/m^2.

The STOP-BANG questionnaire is an 8-item tool developed for SA screening in pre-operative patients(16). This questionnaire consists of 4 questions (STOP component) relating to 1) snoring loudly, 2) feeling tired, 3) observed apneas, 4) presence of high blood pressure and 4 questions (BANG component) regarding presence of: 1) BMI > 35 kg/m^2, 2) age > 50 years, 3) neck circumference > 40 cm, and 4) male gender(16). One point is assigned for each positive response for a maximum point total of 8. In practice, a STOP-BANG score ≥ 3 is considered "high risk" for mild or worse SA [16]. While the STOP-BANG questionnaire was not administered as a stand-alone tool in the SNORE Study, a STOP-BANG "score" was generated from components of the BQ and demographic and anthropometric data as in previous studies [17,18].

Diagnosis of Sleep Apnea

SA was diagnosed with overnight, unattended polysomnography. Trained study personnel attached leads for unattended 25-channel polysomnography recording system to monitor sleep stage, breathing, body movements, oxygen saturation, electrocardiography, leg movements and snoring (AURA PSG System, Grass Technologies, West Warwick, RI). The recording montage was: electroencephalography, electrooculography and submental electromyography to ascertain sleep state; continuous pulse oximetry; nasal-oral thermocouple to assess nasal and oral airflow; nasal pressure transducer to measure nasal airflow; respiratory inductance plethysmography to measure chest and abdominal excursions; 2-lead electrocardiography; and position/activity sensor to monitor body position and movement.

Sleep studies were staged and respiratory events scored by a single American Board of Sleep Medicine-certified, registered sleep technician. Scoring rules for respiratory events were according to the AASM Scoring Manual v 2.0, 2012. Apneas were scored if there was cessation of airflow lasting > 10 seconds. Hypopneas were scored if there was a clear reduction in airflow for > 10 seconds accompanied by a 4% or more oxygen desaturation in primary analysis; in secondary analysis, we used a 3% or more oxygen desaturation parameter. The apnea-hypopnea index (AHI) was defined as the average number of apneas and hypopneas per hour of sleep. Events were defined as obstructive, central or mixed according to the 2012 Scoring Manual.

Participants with studies with less than 1 hour of recorded sleep time were offered the opportunity to repeat the study or were dis-enrolled from the study if they declined repeat study or if repeat study was still less than 1 hour of sleep time (n=6). A subset of sleep studies (15%) had between 1 hour and 4 hours of recorded sleep time. Although we included these studies, we accounted for the reduced sleep time in our analyses (see Statistical Analysis). Finally, while the majority of our studies utilized both nasal pressure transducer and thermistor for measurement of airflow only the thermistor was functional for a subset (31%); this was also accounted for in our analyses (see Statistical Analysis).

Other Measurements

Each participant underwent measurement of blood pressure

at baseline (average of triplicate), and anthropometrics including height and weight for BMI calculation. Serum creatinine and urinary albumin excretion were measured at baseline in order to define the level of renal function. EGFR was calculated using the Modification of Diet in Renal Disease formula [19]. Urinary albumin excretion was estimated from a spot urine sample for albumin and creatinine. We administered a Medical History Questionnaire to determine demographics and identify co-morbid conditions. Diagnoses were defined by self-report with the exception of hypertension and diabetes that were defined by self-report or use of anti-hypertensive or anti-diabetic medication, respectively. Finally, each participant completed the 8-question Epworth Sleepiness Scale in order to gauge daytime sleepiness. A score ≥ 10 suggests excessive daytime sleepiness.

STATISTICAL ANALYSIS

Baseline characteristics including demographics, renal function, co-morbidities, PSG data and SA-risk questionnaire data for those with and without SA were compared using ANOVA for continuous variables, and chi-square test for categorical variables. Continuous variables that were not normally distributed were compared using non-parametric tests. We examined the performance of the BQ and STOP-BANG to predict SA defined at the threshold of AHI ≥ 15 similar to previous study of SA in older men and given this threshold likely portends higher risk of cardio-metabolic consequences of SA [20]. We determined the sensitivity, specificity, positive predictive value, negative predictive value, accuracy and area under the curve of 1) the BQ at the standard threshold score of ≥ 2 and 2) the STOP-BANG at the threshold of ≥ 3 used in surgical populations [14,16]. In sensitivity analyses, we examined alternative thresholds of ≥ 3 for the BQ and at escalating cut-points of the STOP-BANG to the maximum score of 8. We also repeated the above analyses using a 3% oxygen desaturation for hypopneas. Finally, we repeated the above analyses excluding those with 1) < 4 hours total sleep time and then 2) those with thermistor only.

RESULTS

Of the 254 SNORE participants recruited, 248 had sleep studies adequate for inclusion in the SNORE Study and constitute the analytic cohort for this analysis. Table 1 displays baseline demographics, anthropometrics, renal function and co-morbid conditions overall and by presence or absence of SA. Our cohort consisted of predominantly older Caucasian men though 20% were Black. Nearly all our cohort was hypertensive. Veterans with SA were more likely to be male, non-white, with higher BMI and neck circumference and higher prevalence of diabetes mellitus. Table 2 represents objective PSG data. The prevalence of moderate to severe SA in this cohort was 39%. Only 8% had central sleep apnea. Those with SA spent more time in Stage 1 and less time in Stage 3 and R sleep, compared to those without SA. Table 3 shows the results of SA risk and sleepiness questionnaires overall and by components for those with and without SA. There was no statistical difference in prevalence of daytime sleepiness between those with and without SA. Based upon BQ score ≥ 2, 70% of our cohort were considered "high risk" for SA. There was no difference in mean BQ score or proportion identified as "high risk" by BQ score at all 3 possible cutpoints between those with and without SA. A STOP-BANG score ≥ 3 identified 96% of our cohort as "high risk" for SA. Those with SA had a higher mean STOP-BANG Score and were more likely to report that someone observed them stop breathing. However, the proportion identified as "high risk" by STOP-BANG did not vary by SA status.

Validity of the BQ Questionnaire to Predict SA

The median AHI and proportion with SA by BQ score are shown in Figure 1. Median AHI did not vary by BQ score, but proportion with SA trended up with higher BQ score but was not statistically significant (Figure 1b, p trend 0.27). At the threshold used in the general population to predict "high risk" for SA (BQ ≥ 2), the BQ was moderately sensitive with moderately high negative predictive value (Table 4). However, specificity and positive predictive value were low and accuracy was only 50% reflecting only 50% proper assignment of disease or no disease status. When we examined a more extreme threshold for "high risk" of BQ ≥ 3, sensitivity dropped dramatically to 25% but specificity was improved and moderate. Accuracy was minimally improved. The AUC's at both thresholds were low. When we alternatively defined SA using a 3% oxygen desaturation threshold for hypopneas, predictive validity did not change. Finally, in sensitivity analyses, performance of the BQ was not meaningfully changed with we excluded those with total sleep time < 4 hour or those with only thermistor to detect air flow (data not shown).

Validity of the STOP-BANG Questionnaire to Predict SA

The median AHI and proportion with SA by STOP-BANG score are shown in Figure 2. Median AHI inreased monotonically with higher STOP-BANG score. Also, a progressively higher STOP-BANG score identified greater proportion with SA through a score of 7; the highest score of 8 identified a lower proportion with SA but only 2 participants had this score making estimates of SA prevalence in this category unstable (p trend 0.003). A STOP-BANG score of ≥ 3 performed with excellent sensitivity to identify SA defined as AHI ≥ 15 and consequently moderate negative predictive value (Table 5). However, the specificity for SA at this threshold was very poor as essentially the entire cohort (96%) had a score of ≥ 3. Accuracy was also poor at 41% for this threshold as was the AUC. With progressively higher thresholds to identify SA (up to maximum score of 8), as expected, sensitivity declined and specificity increased with no particular threshold demonstrating both moderate to high sensitivity and specificity. Accuracy and the AUC did not change with escalating threshold. In sensitivity analyses, results were unchanged with a 3% oxygen desaturation threshold for hypopneas, or after excluding those with total sleep time < 4 hours or those with thermistor only (data not shown).

DISCUSSION

Among Veterans with moderate to severe CKD, both the BQ and the STOP-BANG performed no better than a coin flip to predict the presence of moderate to severe SA at any threshold. Both tests were sensitive at low thresholds, but not specific. As the threshold for "high risk" increased, specificity increased with an unacceptable expense in sensitivity. Finally, there was

Table 1. Characteristics of the SNORE Study Cohort

	Overall (n=248)	No Sleep Apnea (n=151)	Sleep Apnea (n=97)	p
Demographics				
Age, years, mean (SD)	73.2 (9.6)	72.3 (9.8)	74.6 (9.2)	0.07
Male Gender, n (%)	236 (95.2)	140 (93)	96 (99)	0.03
Race, n (%)				
White	194 (78)	122 (80.8)	72 (74.2)	0.03
Black	50 (20)	29 (19.2)	21 (21.7)	
Hispanic	4 (1.6)	0 (0)	4 (4.1)	
Anthropometrics				
Body-mass index, kg/m^2, mean (SD)	30.3 (4.8)	29.8 (4.6)	31.2 (4.9)	0.02
Neck circumference, cm, mean (SD)	43.0 (3.7)	42.5 (3.6)	43.9 (3.7)	<0.01
Waist to hip ratio, mean (SD)	0.995 (0.07)	0.990 (0.07)	1.003 (0.07)	0.15
Laboratory Data				
MDRD eGFR, mean (SD)	34.9 (8.7)	34.7 (9.0)	35.3 (8.2)	0.59
Baseline Urinary Albumin to Creatinine Ratio, median [IQR]	33.7 [9.5-219.2]	25.0 [7.4-200.3]	48.4 [12.7-344.9]	0.07
Co-morbid Conditions*				
Health Status Excellent/Very Good, n (%)	53 (21)	32 (21)	21 (22)	0.93
Systolic Blood Pressure, mmHg, mean (SD)	136.2 (19.3)	134.4 (18.8)	138.9 (19.8)	0.07
Diastolic Blood Pressure, mmHg, mean (SD)	73.2 (10.0)	72.7 (10.3)	74.0 (9.7)	0.31
Hypertension, n (%)	237 (96)	142 (94)	95 (98)	0.21
Diabetes Mellitus, n (%)	136 (55)	72 (48)	64 (66)	0.005
Chronic Obstructive Pulmonary Disease, n (%)	31 (13)	19 (13)	12 (12)	0.96
Cardiovascular disease, n (%)	96 (39)	61 (40)	35 (36)	0.50
Peripheral Vascular Disease, n (%)	16 (6)	8 (5)	8 (8)	0.36
Congestive Heart Failure, n (%)	46 (19)	26 (17)	20 (21)	0.50
Stroke, n (%)	33 (13)	18 (12)	15 (15)	0.42

*All are self-report except for hypertension and diabetes mellitus which are self-report and/or use of associated medications and systolic, diastolic blood pressure which are direct measurements

Table 2. Baseline Objective Sleep Data.

Polysomnography Data	Overall (n=248)	No Sleep Apnea (n=151)	Sleep Apnea (n=97)	p
Total Sleep Time, minutes, mean (SD)	348.2 (98.8)	353.3 (94.7)	340.2 (104.8)	0.31
Stage 1, % mean (SD)	12.9 (7.3)	10.6 (3.9)	16.6 (9.5)	<0.001
Stage 2, % mean (SD)	60.9 (9.8)	60.6 (9.5)	61.5 (10.3)	0.47
Stage 3, % mean (SD)	11.6 (7.8)	13.2 (7.3)	9.1 (7.9)	<0.001
Stage R, % mean (SD)	14.6 (7.1)	15.7 (6.8)	12.8 (7.2)	0.002
Stage W, % mean (SD)	30.8 (14.5)	29.7 (13.6)	32.5 (15.7)	0.014
Sleep efficiency, % mean (SD)	69.2 (14.5)	70.3 (13.6)	67.5 (15.7)	0.15
Sleep Latency, minutes, median [IQR]	10.5 [3.5-21.5]	11.0 [3.5-23.5]	17.861 [3.5-19.0]	0.58
AHI*, events/hours, median [IQR]	10.3 [3.9-23.6]	5.4 [1.9-9.2]	30.6 [20.8-39.6]	<0.001
No Sleep Apnea (AHI < 5), n (%)	73 (29)	---	---	---
Mild Sleep Apnea (AHI 5-14.9), n (%)	78 (31)	---	---	---
Moderate Sleep Apnea (AHI 15-29.9), n (%)	47 (19)	---	---	---
Severe Sleep Apnea (AHI ≥ 30), n (%)	50 (20)	---	---	---
Central Apnea Index, median [IQR]	0.3 [0.0-1.5]	0.1 [0.0-0.4]	1.3 [0.4-9.8]	<0.001
Mean oxygen saturation, mean (SD)	94.1 (1.8)	94.5 (1.8)	93.6 (1.8)	<0.001
Minimum oxygen saturation, mean (SD)	81.7 (7.2)	84.0 (6.0)	78.0 (7.4)	<0.001
ODI**, events/hours, median [IQR]	10.0 [3.6-23.2]	4.4 [1.8-8.6]	29.2 [18.7-41.0]	<0.001

*Apnea-hypopnea index, based upon 4% oxygen desaturation; **oxygen desaturation index

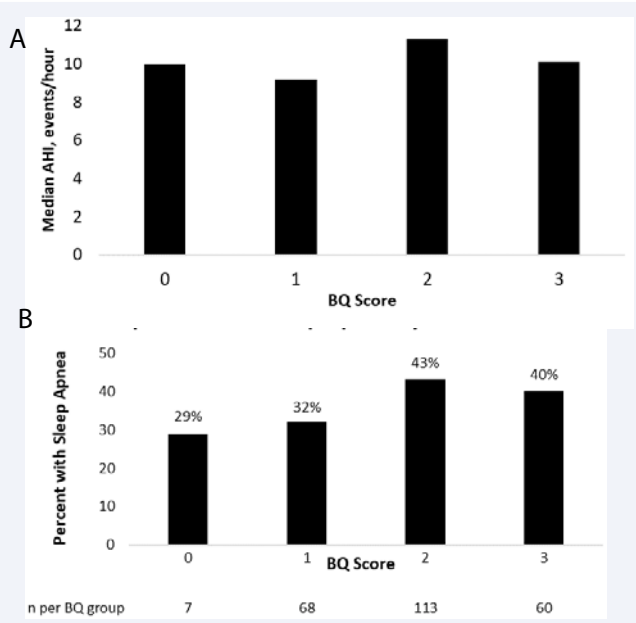

Figure 1 Medlan apnea-hypopnea Index (AHI) (a) and proportion with Sleep pnea (b) by Berlin Questionnalre Score.

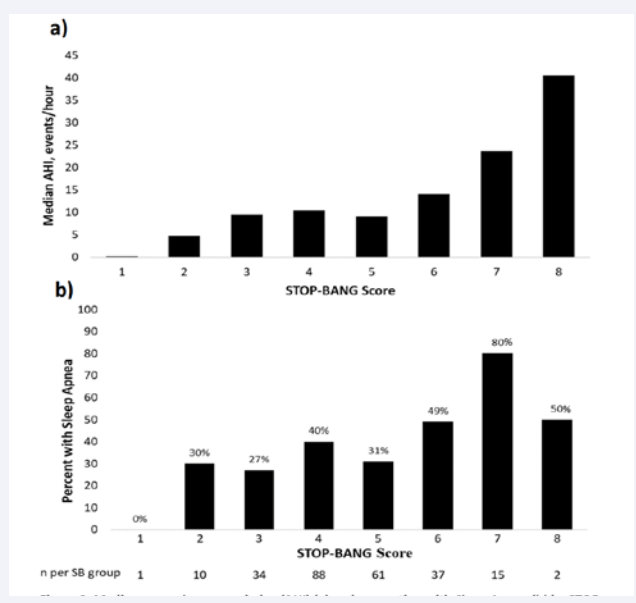

Figure 2 Medlan apnea-hypopnea Index (AHI) (a) and proportion with Sleep pnea (b) by STOPBANG Score.

no difference in sleepiness among those with and without SA. Taken together, our findings suggest that the BQ and STOP-BANG questionnaires are not sufficient stand-alone tools to either rule in or rule out SA among Veterans with CKD. This may be related to lack of traditional SA symptoms in CKD.

The predictive validity of the BQ or STOP-BANG to identify SA varies widely in the reported literature, in part due to variability of population studied, the type of reference sleep study, and the definition of SA used(21). While there have been no studies testing the validity of such tools in Veterans with CKD, reports among Veterans or CKD patients align with our findings [22-24]. For example, Nicholl et al studied 109 CKD patients (eGFR < 60, mean age 65 ± 12 years), not selected for sleep complaints, and found that, similar to our findings, a BQ score ≥ 2 had a 83% sensitivity and 34% specificity for moderate to severe SA [23]. Also in line with our findings, a STOP-BANG score ≥ 3 or ≥ 4 had a sensitivities of 93% and 97%, respectively; specificity of STOP-BANG at these thresholds was only slightly better than in our cohort (31% and 53%, respectively) [23]. Notably, this cohort had a similar prevalence of SA as our cohort (38%) though limited channel PSG was performed which may underestimate the severity of SA [23]. Another study of over 1000 Veterans not selected for CKD who were referred for unattended sleep study (primarily via Watchpat 200) found that progressivly higher STOP-BANG score (BQ not studied) was associated with monotonically higher AHI [22]. However, STOP-BANG score ≥ 3 or ≥ 4 was 99% or 97% sensitive and only 5% or 13% specific for moderate to severe SA, respectively [22]. With progressively more stringent STOP-BANG cutpoints, specificity rose at the expense of sensitivity [22]. Though this was a referral population of Veterans with younger mean age (55 years) and higher prevalence of SA (67%), overall the performance of the STOP-BANG was similar to what we observed. In short, our findings and those of others who studied Veterans or CKD patients suggest that the BQ or STOP-BANG do not adequately allow for ruling in or ruling out disease without an unacceptable rate of either high false positive or false negative rate.

The potential reasons for the poor performance of the BQ and STOP-BANG in Veterans with CKD are unclear. One hypothesis relates to the observation that both tools utilize factors that are ubiquitous in CKD or among Veterans, such as the presence of HTN. Incorporating a non-discriminating factor into a screening tool dilutes the specificity of the tool. To overcome this, we examined higher thesholds for "high risk" and found that while specificity did improve, there was a marked decline in sensitivity. Also, other studies have modified the BQ for CKD (by removing HTN) or the STOP-BANG for Veterans without improvement in performance of the tools [22,24]. We then considered that if certain components of these tools are "diluting" the discriminant power of these tools, perhaps examination of each tool by component would reveal key predictors. However, in our study, only observing the participant stop breathing was statitically different between those with and without SA-all others did not distinguish disease state. This, coupled with our observation that CKD patients with SA are not sleepier, suggests that we may need to look to other metrics to predict SA in this population.

Another hypothesis is that the poor performance of these tools in our population may be related to the fact that we used Type II monitoring instead of Type I monitoring as our reference standard. Indeed, the STOP-BANG questionnaire was developed against Type I monitoring. However, the sensitivity and specificity of Type II monitoring versus Type I monitoring is very high at 85-94% and 77-95%, respectively for AHI ≥ 15 with AUC 89-94 [25]. In addition, Type II monitoring has also become the standard for large NIH funded prospective trials of sleep outcomes [26,27]. Notably, the BQ was developed by consensus and not against PSG [14]. Thereafter, the BQ was tested in a variety of populations against a spectrum of levels of PSG, all of

Table 3. Baseline Sleep Apnea Risk Questionnaire Data from the SNORE Cohort

	Overall (n=248)	No Sleep Apnea (n=151)	Sleep Apnea (n=97)	p
Berlin Questionnaire Data				
Category 1 Positive (Snoring), n (%)	158 (64)	91 (60)	67 (69)	0.16
Category 2 Positive (Tiredness), n (%)	84 (34)	53 (35)	31 (32)	0.31
Category 3 Positive (HTN or BMI), n (%)	232 (94)	138 (91)	94 (97)	0.11
BQ score, mean (SD)	1.91 (0.79)	1.87 (0.81)	1.98 (0.75)	0.28
BQ score ≥ 1, n (%)	241 (97)	146 (97)	95 (98)	0.71
BQ score ≥ 2, n (%)	173 (70)	100 (66)	73 (75)	0.13
BQ score ≥ 3, n (%)	60 (24)	36 (24)	24 (25)	0.87
STOP-BANG Data				
Do you **S**NORE loudly[1], n (%)	58 (23)	30 (20)	28 (29)	0.10
Do you feel **T**IRED, fatigued or sleepy during the day[2], n (%)	103 (42)	62 (41)	41 (42)	0.85
Has someone **O**BSERVED you stop breathing during sleep[3], n (%)	43 (17)	17 (11)	26 (27)	0.002
Have or being treated for high blood **P**RESSURE[4] , n (%)	204 (82)	121 (80)	83 (86)	0.27
BMI > 35, n (%)	33 (13)	18 (12)	15 (15)	0.42
Age > 50, n (%)	242 (98)	146 (97)	96 (99)	0.25
Neck circumference > 40 cm, n (%)	204 (82)	121 (80)	83 (86)	0.27
Gender: male, n (%)	236 (95)	140 (93)	96 (99)	0.03
STOP-BANG Score, mean (SD)	4.5 (1.3)	4.3 (1.2)	4.8 (1.3)	0.003
STOP-BANG Score 0-2 (low risk), n (%)	11 (4)	8 (5)	3 (3)	0.36
STOP-BANG Score 3-4 (intermediate risk), n (%)	122 (49)	78 (52)	44 (45)	
STOP-BANG Score 5-8 (high risk), n (%)	115 (46)	65 (43)	50 (52)	
STOP-BANG Score ≥ 3, n (%)	237 (96)	143 (95)	94 (97)	0.54
Daytime Sleepiness				
Epworth Sleepiness Score > 10, n (%)	84 (34)	47 (31)	37 (38)	0.25

*Based upon 4% oxygen desaturation; [1,2,3,4]Based upon positive response to the following Berlin Questionnaire items: [1]Question 1; [2]Question 7; [3]Question 5; [4]Hypertension Question

Predictive Validity of BQ to Detect AHI ≥ 15

	≥ 1 (n=241)	≥ 2 (n=173)	≥ 3 (n=60)
Sensitivity	98%	75%	25%
Specificity	3%	34%	76%
Positive Predictive Value	39%	42%	40%
Negative Predictive Value	71%	68%	61%
Accuracy	40%	50%	56%
AUC (95% CI)	0.51 (0.49-0.53)	0.55 (0.49-60)	0.50 (0.45-0.56)

Predictive Validity of STOP-BANG to Detect AHI ≥ 15

	≥3 (n=237)	≥4 (n=203)	≥5 (n=115)	≥6 (n=54)	≥7 (n=17)	≥8 (n=2)
Sensitivity	97%	88%	52%	32%	13%	1%
Specificity	5%	22%	57%	85%	97%	99%
Positive Predictive Value	40%	42%	43%	57%	76%	50%
Negative Predictive Value	73%	73%	65%	66%	64%	61%
Accuracy	41%	48%	55%	64%	65%	61%
AUC (95% CI)	0.51 (0.49-0.54)	0.55 (0.50-0.59)	0.54 (0.48-0.60)	0.58 (0.53-0.64)	0.55 (0.52-0.59)	0.50 (0.49-0.51)

which showed significant variability in performance even within type of PSG [21,28]. Regardless, a recent systematic review found that among validation studies of BQ or STOP-BANG, use of Type I vs Type II monitoring did not materially affect point estimates for sensitivity and specificity across a spectrum of SA severity [29].

Key strengths of our study include performance of full PSG in all of our participants with 100% concomitant completion of the BQ. In addition, we provide the first look at the predictive validity of both the BQ and STOP-BANG in a cohort of Veterans with CKD not selected for sleep disorders or complaints. Key limitations of our study include the reconstruction of the STOP-BANG in lieu of administration of the complete questionnaire to each participant. However, other studies have used a similar approach in order to study validity of the STOP-BANG retrospectively [17,18]. Also, it was not ideal to have a proportion of our cohort with <4 hours total sleep time or thermistor only, however sensitivity analyses excluding those participants showed no difference in results. Finally, our results are largely generalizable to male Veterans with CKD given the low percentage of females represented in our cohort. However, our population accurately reflects the race distribution within the Veterans Affairs system.

CONCLUSION

Among Veterans selected only for presence of CKD, moderate to severe SA is common and is not associated with excess daytime sleepiness. Standard SA risk assessment tools perform poorly to detect SA with an acceptable compromise in specificity, likely related to lack of excess sleepiness and high background of co-morbid conditions among those with SA. Future studies should investigate other low cost, easy to administer screening tools that will allow providers to reliably forecast the need for PSG. National sleep, cardiology and nephrology societies, in concert with available evidence, should determine the target balance between sensitivity and specificity for novel SA-risk tools in CKD.

ACKNOWLEDGEMENTS

This work was supported through funding from a Veterans Affairs Clinical Science Research and Development Career Development Award (CX000533-01A1) and by support from the Division of Nephrology, Hypertension and Renal Transplantation, Department of Medicine, University of Florida. The funding source for this work played no role in the study design, collection, analysis and interpretation of data or in the writing of the report and the decision to submit the article for publication. We also thank the Geriatric Research Education and Clinical Center at the Malcom Randall VA Medical Center for their support.

REFERENCES

1. Patel N, Golzy M, Nainani N, Nader ND, Carter RL, Lohr JW, et al. Prevalence of various comorbidities among veterans with chronic kidney disease and its comparison with other datasets. Ren Fail. 2016; 38: 204-8.
2. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med. 2004; 351: 1296-305.
3. O'Hare AM, Bertenthal D, Covinsky KE, Landefeld CS, Sen S, Mehta K, et al. Mortality risk stratification in chronic kidney disease: one size for all ages? J Am Soc Nephrol. 2006; 17: 846-53.
4. Markou N, Kanakaki M, Myrianthefs P, Hadjiyanakos D, Vlassopoulos D, Damianos A, et al. Sleep-disordered breathing in nondialyzed patients with chronic renal failure. Lung. 2006; 184: 43-9.
5. Ogna A, Forni Ogna V, Haba Rubio J, Tobback N, Andries D, Preisig M, et al. Sleep Characteristics in Early Stages of Chronic Kidney Disease in the HypnoLaus Cohort. Sleep. 2016; 39: 945-53.
6. Nicholl DD, Ahmed SB, Loewen AH, Hemmelgarn BR, Sola DY, Beecroft JM, et al. Declining kidney function increases the prevalence of sleep apnea and nocturnal hypoxia. Chest. 2012; 141: 1422-30.
7. Sakaguchi Y, Hatta T, Hayashi T, Shoji T, Suzuki A, Tomida K, et al. Association of nocturnal hypoxemia with progression of CKD. Clin J Am Soc Nephrol. 2013; 8: 1502-7.
8. Ahmed SB, Ronksley PE, Hemmelgarn BR, Tsai WH, Manns BJ, Tonelli M, et al. Nocturnal hypoxia and loss of kidney function. PLoS One. 2011; 6: e19029.
9. Molnar MZ, Mucsi I, Novak M, Szabo Z, Freire AX, Huch KM, et al. Association of incident obstructive sleep apnoea with outcomes in a large cohort of US veterans. Thorax. 2015; 70: 888-95.
10. Tahrani AA, Ali A, Raymond NT, Begum S, Dubb K, Altaf QA, et al. Obstructive sleep apnea and diabetic nephropathy: a cohort study. Diabetes Care. 2013; 36: 3718-25.
11. Nicholl DD, Ahmed SB, Loewen AH, Hemmelgarn BR, Sola DY, Beecroft JM, et al. Clinical presentation of obstructive sleep apnea in patients with chronic kidney disease. J Clin Sleep Med. 2012; 8: 381-7.
12. Chan GC, Lam B, Yap DY, Ip MS, Lai KN, Tang SC. Proteinuria is associated with sleep apnea in chronic kidney disease. Nephrol Dial Transplant. 2016; 31: 772-9.
13. Adams RJ, Appleton SL, Vakulin A, Hanly PJ, McDonald SP, Martin SA, et al. Chronic Kidney Disease and Sleep Apnea Association of Kidney Disease With Obstructive Sleep Apnea in a Population Study of Men. Sleep. 2017; 40.
14. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med. 1999; 131: 485-91.
15. Canales MT, Bozorgmehri S, Ishani A, Weiner ID, Berry R, Beyth R. Prevalence and correlates of sleep apnea among US Veterans with chronic kidney disease. J Sleep Res. 2020: e12981.
16. Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. Anesthesiology. 2008; 108: 812-21.
17. Farney RJ, Walker BS, Farney RM, Snow GL, Walker JM. The STOP-Bang equivalent model and prediction of severity of obstructive sleep apnea: relation to polysomnographic measurements of the apnea/hypopnea index. J Clin Sleep Med. 2011; 7: 459-65B.
18. 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: 467-72.
19. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Ann Intern Med. 2006; 145: 247-54.
20. Mehra R, Stone KL, Blackwell T, Ancoli Israel S, Dam TT, Stefanick ML, et al. Prevalence and correlates of sleep-disordered breathing in older men: osteoporotic fractures in men sleep study. J Am Geriatr Soc. 2007; 55: 1356-64.
21. Senaratna CV, Perret JL, Matheson MC, Lodge CJ, Lowe AJ, Cassim R, et

al. Validity of the Berlin questionnaire in detecting obstructive sleep apnea: A systematic review and meta-analysis. Sleep Med Rev. 2017; 36: 116-24.

22. Kunisaki KM, Brown KE, Fabbrini AE, Wetherbee EE, Rector TS. STOP-BANG questionnaire performance in a Veterans Affairs unattended sleep study program. Ann Am Thorac Soc. 2014; 11: 192-7.

23. Nicholl DD, Ahmed SB, Loewen AH, Hemmelgarn BR, Sola DY, Beecroft JM, et al. Diagnostic value of screening instruments for identifying obstructive sleep apnea in kidney failure. J Clin Sleep Med. 2013; 9: 31-8.

24. Huang HC, Walters G, Talaulikar G, Figurski D, Carroll A, Hurwitz M, et al. Sleep apnea prevalence in chronic kidney disease - association with total body water and symptoms. BMC Nephrol. 2017; 18: 125.

25. Jonas DE, Amick HR, Feltner C, Weber RP, Arvanitis M, Stine A, et al. Screening for Obstructive Sleep Apnea in Adults: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2017; 317: 415-33.

26. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, 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: 569-85.

27. Quan SF, Howard BV, Iber C, Kiley JP, Nieto FJ, O'Connor GT, et al. The Sleep Heart Health Study: design, rationale, and methods. Sleep. 1997; 20: 1077-85.

28. Hrubos-Strøm H, Randby A, Namtvedt SK, Kristiansen HA, Einvik G, Benth J, et al. A Norwegian population-based study on the risk and prevalence of obstructive sleep apnea. The Akershus Sleep Apnea Project (ASAP). J Sleep Res. 2011; 20: 162-70.

29. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, et al. Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: An American Academy of Sleep Medicine Clinical Practice Guideline. J Clin Sleep Med. 2017; 13: 479-504.

Cite this article

Canales M, Bozorgmehri S, Ishani A, Weiner ID, Berry R, et al. (2020) The Art of Divination: Do Standard Questionnaires Predict Sleep Apnea in Veterans with Kidney Disease? J Sleep Med Disord 6(3): 1109.