https://doi.org/10.5664/jcsm.9388

JCSM | Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

Examining the diagnostic validity of the Berlin Questionnaire in a low-income Black American sample

Stephanie Brooks Holliday, PhD[1]; Ann Haas, MS, MPH[2]; Lu Dong, PhD[1]; Madhumita Ghosh-Dastidar, PhD[1]; Lauren Hale, PhD[3,4]; Daniel J. Buysse, MD[5]; Tamara Dubowitz, ScD[3]; Wendy M. Troxel, PhD[3]

[1]*Behavioral and Policy Sciences, RAND Corporation, Santa Monica, California;* [2]*Economics, Sociology, and Statistics, RAND Corporation, Pittsburgh, Pennsylvania;* [3]*RAND Corporation, Pittsburgh, Pennsylvania;* [4]*Department of Family, Population and Preventive Medicine, Stony Brook University, Stony Brook, New York;* [5]*Department of Psychiatry, University of Pittsburgh, Pittsburgh Pennsylvania*

Study Objectives: Black individuals and individuals of low socioeconomic status are at increased risk for obstructive sleep apnea (OSA). The Berlin Questionnaire is one of the most widely used screening tools for OSA; however, there is limited research on its diagnostic accuracy in low-income Black populations.

Methods: This study analyzed data from an ongoing study taking place among a cohort from 2 predominantly Black neighborhoods in Pittsburgh, Pennsylvania (96.3% Black, 79.6% female). The sample included 269 individuals without a prior diagnosis of OSA who completed the Berlin Questionnaire and also participated in a home sleep apnea test. An apnea-hypopnea index ≥ 15 events/h was used to identify individuals with moderate or severe OSA.

Results: 19.3% of individuals met criteria for moderate to severe OSA based on home sleep apnea test, while 31.2% of participants screened as high risk for OSA based on the overall Berlin index. Using apnea-hypopnea index ≥ 15 events/h as the reference standard, the Berlin Questionnaire had a sensitivity of 46.2%, specificity of 72.4%, positive predictive value of 28.6%, and negative predictive value of 84.9% among this sample. Analyses stratified by sex suggested that the Berlin Questionnaire had better diagnostic validity in women than men.

Conclusions: The Berlin Questionnaire has lower sensitivity and positive predictive value in our sample than those observed in general population samples. The measure performed better among women, though a higher proportion of men fell into the moderate or severe OSA range based on the home sleep apnea test. Given the significant downstream consequences of OSA, utilizing screening tools that better detect OSA in Black communities is key.

Keywords: obstructive sleep apnea, screening, assessment, race, disparities

Citation: Holliday SB, Haas A, Dong L, et al. Examining the diagnostic validity of the Berlin Questionnaire in a low-income Black American sample. *J Clin Sleep Med.* 2021;17(10):1987–1994.

BRIEF SUMMARY

Current Knowledge/Study Rationale: Black Americans with socioeconomic disadvantage are at increased risk for obstructive sleep apnea. Although the Berlin Questionnaire is one of the most widely used screening tools for obstructive sleep apnea and has been validated in a broad range of populations, there is little evidence regarding its diagnostic accuracy in low-income Black populations.

Study Impact: This study reports the sensitivity, specificity, positive predictive value, and negative predictive value of the Berlin Questionnaire in a sample of low-income Black community-dwelling adults from 2 Pittsburgh neighborhoods. An understanding of the diagnostic validity of the measure is key to knowing whether it is an adequate screening tool specifically among low-income Black individuals.

INTRODUCTION

An estimated 6% to 17% of individuals in the United States have moderate to severe obstructive sleep apnea (OSA),[1] a condition associated with a number of downstream health consequences, such as cardiovascular disease, including coronary artery disease and stroke[2]; cognitive dysfunction[3,4]; and reduced quality of life.[5,6] Certain groups are at increased risk for OSA, including men, older adults, and individuals who live with obesity (ie, having a body mass index [BMI] 30+ kg/m^2).[1] There is also recognition that race/ethnicity and socioeconomic disadvantage may increase risk of OSA. For example, a meta-analysis found that African Americans have a higher prevalence and greater severity of OSA,[7] likely influenced by potential confounding factors such as household income and greater BMI.[8] There is also evidence that neighborhood disadvantage may contribute to risk of OSA beyond the influence of age and race, at least among children.[9,10] For example, a study of childhood OSA found that children living in census tracts with lower family income, higher proportions of single parent families, higher population density, and lower levels of education had a higher risk of OSA, controlling for age, race, and obesity.[10] Certain neighborhood conditions have also been associated with OSA risk in adults, including exposure to air pollution,[11] worse walking environment,[12] and greater neighborhood crowding.[13] Given that Black Americans have a higher prevalence of OSA and are disproportionately exposed to both individual and neighborhood-level socioeconomic disadvantage,[14] it is important to evaluate screening measures to better identify individuals who may need further evaluation and treatment for OSA.

The Berlin Questionnaire is one of the most widely used and studied screening tools for obstructive sleep apnea.[15] The Berlin Questionnaire is a brief instrument with 10 questions that can be easily administered at a doctor's office. The measure comprises 3 domains: (1) habitual snoring, (2) sleepiness, and (3) hypertension or BMI > 30 kg/m^2. A meta-analysis of sleep apnea screeners found that the Berlin Questionnaire most accurately predicted diagnosis of OSA,[16] and it has been shown to have specificity in the detection of mild, moderate, and severe OSA higher than other screening tools.[15,17] However, the meta-analysis did not examine moderators of the diagnostic accuracy, such as race or sex, leaving questions about how the questionnaire functions in diverse groups.

The Berlin Questionnaire has been validated in a variety of populations and settings. This includes general population, primary care, and sleep clinic samples (eg, refs. 18–20) in countries across the world (eg, refs. 21–23). Given the brief nature of the instrument, it is easy and low-resource to implement, which is a key consideration when working with low-resource communities with limited access to health services. Despite its widespread use, no prior research to the best of our knowledge has examined the diagnostic validity of the Berlin Questionnaire in a largely Black American sample. A study using data from the Jackson Heart Study conducted a confirmatory factor analysis to validate the structure of the questionnaire and examined its internal consistency, finding relatively low internal consistency (alpha = 0.53), but did not explore diagnostic validity.[24] There has also been limited work examining whether the questionnaire functions differently across Black American women and men. A recent study suggested that women experience fatigue associated with OSA more often than men, whereas men experience sleepiness more often. This difference was especially pronounced for Black men.[25] This finding was consistent with previous research suggesting that women with OSA may present with "nonspecific" symptoms, such as fatigue, depression, and sleep disturbances, compared to symptoms that are seen more traditionally as indicators of sleep-disordered breathing (eg, snoring, gasping).[26]

Our study fills several important gaps in the literature. First, this study provides an opportunity to examine the prevalence of traditional OSA symptoms, as measured by the Berlin Questionnaire, in a predominantly Black, low-income sample. In addition, because the Berlin Questionnaire is a widely used screening measure and Black Americans are a population at increased risk for OSA, it is essential to know how well it functions in this population. Our objective was to explore the diagnostic validity of the Berlin Questionnaire in a community-based sample of predominantly low-income Black individuals. Given the reported sex difference in OSA presentation, we also conducted exploratory analyses to understand whether there were sex differences in the diagnostic validity of the Berlin Questionnaire in this population.

METHODS

Participants

Participants were recruited as part of the Pittsburgh Hill/Homewood Research on Neighborhood Change and Sleep (PHRESH Zzz) study, part of a series of studies leveraging a natural experiment of neighborhood-level change in 2 neighborhoods in Pittsburgh, PA. These neighborhoods are largely low-income, with an estimated median household income of $17,982 and $21,492, and 91% of residents are Black.[27] A random sample of households was initially recruited for enrollment in 2011 and participated in several waves of data collection, consisting of individual in-home interviews. Data for the present analyses were collected in 2016, following the opening of a full-service supermarket and greenspace expansion in 1 of the 2 neighborhoods. This study was approved by the RAND Institutional Review Board.

Procedure

In 2016, our sample included 828 individuals who participated in PHRESH Zzz. The in-home interview completed by participants included the Berlin Questionnaire with interviewer-administered (objective) assessments of height and weight following the survey. In addition, a subsample of individuals were invited to complete an in-home apnea assessment if they had reported never receiving a diagnosis of OSA by a physician (149 participants reported a physician diagnosis and were not eligible for the in-home apnea assessment). Participants were invited to take part during their in-home interview. Given budgetary constraints and equipment limitations, the in-home apnea assessment was limited to the first 291 individuals who agreed to participate. Participants with fewer than 3.5 hours of data were excluded from in-home apnea assessment analyses (n = 22). Our sample for the present analyses, therefore, included the 269 participants with in-home apnea assessment data. Individuals in this subsample were similar to those who did not participate on most sociodemographic variables (including sex and BMI) but were younger than those who did not participate.

Measures

Sociodemographic characteristics

Participants self-reported their sociodemographic characteristics, including race, age, sex, level of education, and household income. We also examined indicators of cardiovascular and general health for individuals in this sample. This included an indicator of hypertension, which included individuals who reported a prior diagnosis of hypertension, those who reported currently taking medication for hypertension, or who had high blood pressure by objective measurement (systolic blood pressure 140 mm Hg or higher or diastolic 90 mm Hg or higher). For blood pressure, 2 measurements were taken 60 seconds apart using a Micro Life automated blood pressure monitor Microlife USA, Inc, Clearwater, FL) after the participant had been seated for 5 minutes, and the average of the measurements was taken (for more detail on the objective assessments, see ref. 28). Self-rated health was assessed with a single item, with responses made on a 5-point scale ranging from "poor" to "excellent."[29]

Berlin Questionnaire

The Berlin Questionnaire was administered as part of the in-home interview. It comprises 10 items and assesses 3 categories of risk factor for OSA: 1) habitual snoring, 2) sleepiness, and 3) hypertension or BMI > 30 kg/m^2. BMI was calculated using objectively measured height and weight; hypertension was based on participant self-report in response to the question, "Do you have high blood pressure?" The Berlin Questionnaire yields a score for each category (positive vs negative), and individuals are classified as high risk for OSA if there are 2 or more categories with a positive score.

Objective measure of OSA

Participants completed an in-home sleep apnea assessment using the ApneaLink Plus (ResMed, San Diego, CA), which is an FDA-approved, type 3 home sleep apnea testing device.[30,31] Automated scoring was used for scoring of apneas/hypopneas and all desaturations. In addition, each record was reviewed for accuracy and edited as necessary by trained sleep technicians under the supervision of 1 of the study investigators (D.J.B.).

Apnea-hypopnea index (AHI) was calculated as the sum of all apneas and hypopneas divided by total evaluation time. We used a 4% oxygen desaturation threshold to identify hypopneas,[32] which is considered acceptable according to the American Academy of Sleep Medicine scoring manual, Version 2.6,[33] and recommended by the Center for Medicare & Medicaid Services.[34] We examined 2 thresholds for OSA: AHI ≥ 5 events/h, which identified anyone with mild, moderate, or severe OSA, and AHI ≥ 15 events/h, which allowed us to focus on individuals with moderate or severe OSA.[17] Participants received feedback about their AHI results. Those with an elevated AHI were provided information and recommendations for further evaluation.

Statistical analysis

To examine the diagnostic validity of the Berlin Questionnaire, we calculated 4 metrics: sensitivity, which indicates the percentage of those with the condition who are correctly classified as having the condition; specificity, which indicates the percentage of those without the condition who are correctly classified as not having the condition; positive predictive value (PPV), which indicates the percentage of those classified as having the condition who actually have the condition; and negative predictive value (NPV), which indicates the percentage of those classified as not having the condition who actually do not have the condition. We also computed the likelihood ratios for a positive and negative test. Our primary analyses examined AHI ≥ 15 events/h as the reference standard for determining whether an individual has moderate to severe OSA, though we also examined a threshold of AHI ≥ 5 events/h to identify those with mild OSA as well.[17] In addition to examining these metrics for the full sample, we also conducted analyses stratified by sex. We also conducted sensitivity analyses to determine if diagnostic validity varied by level of education. Data were analyzed using SAS 9.4 (SAS Institute, Cary, NC).

RESULTS

Table 1 presents the demographic characteristics of participants. A substantial percentage of the sample had BMI greater than 30 kg/m^2 (46.8%), and about two-thirds had hypertension based on self-report or objective measurement. Based on the overall Berlin Questionnaire index, a total of 31.2% of participants were screened as high risk for OSA. **Table 2** reports results of the home sleep apnea test (HSAT). Approximately 19.3% of participants had an AHI ≥ 15 events/h. A greater proportion of men (30.9%) than women (16.4%) had an AHI ≥ 15 events/h (P = .02). An additional 37.2% of individuals had an AHI between 5 and 15 events/h.

Table 3 presents the diagnostic validity of the Berlin Questionnaire for AHI ≥ 15 events/h. Using a threshold of AHI ≥ 15 events/h, the overall Berlin Questionnaire index was associated with a somewhat higher sensitivity (46.2%) and NPV (74.9%) but lower PPV (28.6%). Specificity was not substantially different (72.4%). There was variation in the sensitivity and specificity of individual categories on the Berlin Questionnaire at this threshold as well. Sensitivity ranged from 9.6% for sleepiness to 84.6% for BMI or high blood pressure. Specificity ranged from 28.1% for BMI or high blood pressure to 85.7% for sleepiness. PPV and NPV were not substantially different for individual Berlin Questionnaire categories. We also examined diagnostic validity using a threshold of AHI ≥ 5 events/h (see **Table S1** in the supplemental material). Based on this threshold, the Berlin Questionnaire yielded a lower sensitivity (36.2%) than when AHI ≥ 15 events/h was examined and similar specificity (75.2%). A similar pattern of findings was observed for the variation in diagnostic validity across the individual categories as was found for the AHI ≥ 15 events/h threshold.

Table 4 presents results stratified by sex for AHI ≥ 15 events/h. The sensitivity of the overall Berlin Questionnaire index was significantly lower for men (17.6%) than women (60.0%) in our sample ($P < .01$), though specificity was not significantly different (71.1% and 72.6%, respectively; $P = .84$). Similarly, both PPV and NPV were higher among women (PPV = 30.0%, NPV = 90.3%) than men (PPV = 21.4%, NPV = 65.9%) in our sample, though this difference was only statistically significant for NPV ($P < .01$). Regarding individual categories on the Berlin Questionnaire, sleepiness was associated with especially low sensitivity in men (11.8%) and women (8.6%). However, sleepiness also yielded the highest PPV in men (40.0%) and the highest specificity in both women (84.4%) and men (92.1%). **Table S2** in the supplemental material presents results stratified by sex for AHI ≥ 5 events/h. No significant sex differences were observed for sensitivity, specificity, PPV, or NPV. Regarding the individual categories, the only significant differences observed were for BMI or high blood pressure, for which specificity and PPV were significantly higher in men vs women ($P = .01$ and $P = .03$, respectively).

Finally, we examined diagnostic validity by educational attainment. We compared individuals with less than a high school diploma or a high school diploma/general educational development certificate to those with some college or higher for both AHI ≥ 5 and AHI ≥ 15 events/h. Analyses yielded no significant differences for those groups (data not reported).

DISCUSSION

This paper aimed to examine the diagnostic accuracy of the Berlin Questionnaire in a sample of predominantly low-income Black women, recruited from 2 disadvantaged neighborhoods in

Table 1—Descriptive characteristics (n = 269).

	Overall (n = 269)	Females (n = 214)	Males (n = 55)
Demographic characteristics			
Age (years)	55.0 (14.4)	55.0 (14.9)	55.1 (12.8)
Female	79.6%	n/a	n/a
African American/Black	96.3%	97.7%	90.1%
Education			
< High school	11.2%	11.7%	9.1%
High school or GED	38.3%	40.2%	30.9%
Some college/training	37.2%	36.9%	38.2%
College degree or higher	13.4%	11.2%	21.8%
Annual family income ($)	21,700 (17,500)	21,400 (17,000)	23,100 (19,700)
Health/comorbidities			
BMI ≥ 30 kg/m^2	46.8%	52.3%	25.5%
Mean BMI (kg/m^2)	30.7 (7.3)	31.6 (7.4)	27.1 (5.5)
Hypertensiona	68.8%	70.6%	61.8%
Poor or fair self-reported healthb	30.5%	31.3%	27.3%
Positive screen for OSA based on Berlin Questionnaire			
High risk (based on all 3 categories)	31.2%	32.7%	25.5%
Category 1: snoring	32.0%	32.7%	29.1%
Category 2: sleepiness	13.4%	14.5%	9.1%
Category 3: BMI or high blood pressure	74.3%	78.0%	60.0%

Values are reported as percentage or mean (standard deviation). [a]Includes individuals who reported having a prior diagnosis with hypertension, reported currently taking medication for hypertension, or who had high blood pressure by objective measurement (SBP 140 mmHg or higher or DBP 90 mmHg or higher). [b]Other options are "good," "very good," or "excellent." BMI = body mass index, DBP = diastolic blood pressure, GED = general educational development certificate, n/a = not available, OSA = obstructive sleep apnea, SBP = systolic blood pressure.

Table 2—Home apnea screening results.

	Overall (n = 269)	Females (n = 214)	Males (n = 55)
AHI (events/h)	6.1 (2.7, 12.1)	5.6 (2.7, 11.2)	7.8 (3.4, 17.2)
Normal (< 5)	43.5%	46.3%	32.7%
Mild (5 to < 15)	37.2%	37.4%	36.4%
Moderate (15 to < 30)	12.3%	11.2%	16.4%
Severe (30+)	7.1%	5.1%	14.6%
Duration of evaluation (min)	405.9 (328.6, 473.2)	410.5 (327.6, 476.6)	392.7 (329.1, 448.1)
Duration of recording (min)	446.3 (376.6, 534.1)	453.8 (381.9, 537.2)	427.2 (366.9, 519.5)
Apnea index (events/h)	1.0 (0.0, 4.0)	1.0 (0.0, 3.0)	2.0 (1.0, 8.0)
Hypopnea index (events/h)	4.4 (1.7, 7.9)	4.3 (1.6, 7.7)	4.9 (2.0, 9.2)
Oxygen distress index (events/h)	6.7 (2.3, 12.4)	6.5 (2.2, 11.2)	8.0 (2.8, 16.6)
Minutes saturation ≤ 90%	15.0 (2.0, 109.0)	14.0 (1.0, 101.0)	28.0 (2.0, 122.0)

Values are reported as percentage or median (25th, 75th percentile). Four cases were missing oxygen distress index and minutes saturation ≤ 90%. AHI = apnea-hypopnea index.

Table 3—Diagnostic accuracy of the overall Berlin Questionnaire score and individual categories, AHI ≥ 15 events/h (n = 269).

	Sensitivity	Specificity	PPV	NPV	Positive Likelihood Ratio	Negative Likelihood Ratio
Overall Berlin Questionnaire index	46.2% [32.2%, 60.5%]	72.4% [65.9%, 78.2%]	28.6% [19.2%, 39.5%]	84.9% [78.9%, 89.7%]	1.67 [1.16, 2.40]	0.74 [0.57, 0.97]
Category 1: snoring	46.2% [32.2%, 60.5%]	71.4% [64.9%, 77.3%]	27.9% [18.8%, 38.6%]	84.7% [78.7%, 89.6%]	1.62 [1.13, 2.32]	0.75 [0.58, 0.98]
Category 2: sleepiness	9.6% [3.2%, 21.0%]	85.7% [80.3%, 90.1%]	13.9% [4.7%, 29.5%]	79.8% [74.1%, 84.8%]	0.67 [0.28, 1.65]	1.05 [0.95, 1.17]
Category 3: BMI or high blood pressure	84.6% [71.9%, 93.1%]	28.1% [22.2%, 34.6%]	22.0% [16.5%, 28.4%]	88.4% [78.4%, 94.9%]	1.18 [1.02, 1.36]	0.55 [0.28, 1.07]

Numbers in brackets reflect 95% confidence intervals for each estimate. Sensitivity indicates percentage of those with the condition who are correctly classified as having the condition. Specificity indicates the percentage of those without the condition who are correctly classified as not having the condition. PPV indicates the percentage of those classified as having the condition who actually have the condition. NPV indicates the percentage of those classified as not having the condition who actually do not have the condition. The positive likelihood ratio indicates the probability of a person who has the disease testing positive divided by the probability of a person who does not have the disease testing positive. The negative likelihood ratio indicates the probability of a person who has the disease testing negative divided by the probability of a person who does not have the disease testing negative. BMI = body mass index, NPV = negative predictive value, PPV = positive predictive value.

Table 4—Diagnostic accuracy of the overall Berlin Questionnaire score and individual categories by sex, AHI ≥ 15 events/h.

	Sensitivity	Specificity	PPV	NPV	Positive Likelihood Ratio	Negative Likelihood Ratio
Women (n = 214)						
Overall Berlin Questionnaire index	60.0%* [42.1%, 76.1%]	72.6% [65.5%, 79.0%]	30.0% [19.6%, 42.1%]	90.3%* [84.2%, 94.6%]	2.19 [1.53, 3.14]	0.55 [0.36, 0.84]
Category 1: snoring	60.0%* [42.1%, 76.1%]	72.6% [65.5%, 79.0%]	30.0% [19.6%, 42.1%]	90.3%* [84.2%, 94.6%]	2.19 [1.53, 3.14]	0.55 [0.36, 0.84]
Category 2: sleepiness	8.6% [1.8%, 23.1%]	84.4% [78.2%, 89.3%]	9.7% [2.0%, 25.8%]	82.5% [76.2%, 87.7%]	0.55 [0.18, 1.70]	1.08 [0.96, 1.22]
Category 3: BMI or high blood pressure	94.3%* [80.8%, 99.3%]	25.1%* [19.0%, 32.2%]	19.8% [14.0%, 26.6%]	95.7%* [85.5%, 99.5%]	1.26 [1.12, 1.42]	0.23 [0.06, 0.89]
Men (n = 55)						
Overall Berlin Questionnaire index	17.6% [3.8%, 43.4%]	71.1% [54.1%, 84.6%]	21.4% [4.7%, 50.8%]	65.9% [49.4%, 79.9%]	0.61 [0.20, 1.91]	1.16 [0.86, 1.56]
Category 1: snoring	17.6% [3.8%, 43.4%]	65.8% [48.6%, 80.4%]	18.8% [4.05%, 45.6%]	64.1% [47.2%, 78.8%]	0.52 [0.17, 1.58]	1.25 [0.91, 1.72]
Category 2: sleepiness	11.8% [1.46%, 36.4%]	92.1% [78.6%, 98.3%]	40.0% [5.3%, 85.3%]	70.0% [55.4%, 82.1%]	1.49 [0.27, 8.12]	0.96 [0.79, 1.17]
Category 3: BMI or high blood pressure	64.7% [38.3%, 85.8%]	42.1% [26.3%, 59.2%]	33.3% [18.0%, 51.8%]	72.7% [49.8%, 89.3%]	1.12 [0.72, 1.74]	0.84 [0.40, 1.76]

*Indicates that the value is significantly different for females compared to males, $P < .05$. Numbers in brackets reflect 95% confidence intervals for each estimate. Sensitivity indicates percentage of those with the condition who are correctly classified as having the condition. Specificity indicates the percentage of those without the condition who are correctly classified as not having the condition. PPV indicates the percentage of those classified as having the condition who actually have the condition. NPV indicates the percentage of those classified as not having the condition who actually do not have the condition. The positive likelihood ratio indicates the probability of a person who has the disease testing positive divided by the probability of a person who does not have the disease testing positive. Then negative likelihood ratio indicates the probability of a person who has the disease testing negative divided by the probability of a person who does not have the disease testing negative. BMI = body mass index, NPV = negative predictive value, PPV = positive predictive value.

Pittsburgh, PA. About 31% of participants were identified as high-risk for OSA based on the Berlin Questionnaire. This is somewhat higher than estimates in other similar samples, such as the Jackson Heart Study cohort, in which 11.9% of participants were identified as high-risk (3.5% of men and 16.8% of women).[24] However, the Jackson Heart Study sample has a broader income distribution, whereas our sample is largely low-income (with an average income of $21,700), which may contribute to this discrepancy.

Regarding the prevalence of OSA based on HSAT, 56.5% of our sample had AHI ≥ 5 events/h (53.7% of women and 67.3% of men) and 19.4% had AHI ≥ 15 events/h (16.3% of women and 31.0% of men). This suggests a higher prevalence of OSA in our primarily Black sample compared to the general population. For example, the Wisconsin Sleep Cohort study found that an estimated 11% of participants had an AHI ≥ 15 events/h.[35] Among individuals with age and BMI similar to our sample, 11% to 29% of women and 5% to 14% of men had an AHI ≥ 15 events/h.[35] The substantial rate of OSA in our sample underscores the importance of a screening tool with high diagnostic validity.

However, our findings suggest that the Berlin Questionnaire may not perform adequately as a screening tool for OSA in this population—specifically, non–treatment seeking, predominantly Black adults from disadvantaged neighborhoods. In part, this is because the Berlin Questionnaire appears to be better at identifying individuals who do not have OSA than those who do have OSA, based on the relatively higher specificity than sensitivity in our sample. This was observed when both AHI ≥ 5 and ≥ 15 events/h were used as the criterion for OSA. This means that many people who *do* have OSA are not adequately identified, which is problematic given the downstream consequences of OSA and higher prevalence of OSA in Black Americans. A good screening measure should have a high sensitivity. That said, maximizing specificity has certain benefits—from a cost and resources perspective, it means fewer people are being referred to an unnecessary sleep study, which is a consideration in populations that experience obstacles to accessing quality health care.[36]

An important consideration when interpreting the findings our study is that we focused on residents from randomly selected households in 2 neighborhoods, rather than individuals who were seeking treatment for sleep-related concerns. Previous studies have generally found that the Berlin Questionnaire has better psychometric properties in specialty care settings.[37] This is not surprising, particularly for PPV and NPV, given that the base rate of OSA would be expected to be higher in a specialty care setting. Even so, there have been other validation studies that have found somewhat better psychometric properties for the Berlin Questionnaire in general-population samples than what we found in our sample. For example, a study using a randomly selected general-population sample of adults in Norway found the overall Berlin Questionnaire index resulted in a higher PPV for AHI ≥ 15 events/h than observed in our sample, though sensitivity, specificity, and NPV were similar.[38] Another study in a Greek primary care sample reported a substantially higher sensitivity and PPV when AHI thresholds of both ≥ 5 and ≥ 15 events/h were examined.[22,37] However, it is important to note that these studies were conducted in international samples with translated versions of the Berlin Questionnaire (see **Table 5** for a summary of results from this study and these other 2 studies).

Regarding the role of sex, HSAT results revealed that a substantially higher proportion of men fell into the moderate or severe range (30.9% vs 16.4%), which is consistent with work in other Black samples.[39] However, somewhat more women than men screened positive for OSA on the Berlin Questionnaire (32.7% vs 25.5%), and we found that the Berlin Questionnaire had poorer sensitivity and NPV in men than women (based on AHI ≥ 15 events/h). Research has suggested that women may present with more nonspecific symptoms, rather than the traditional "hallmarks" of OSA.[26] Therefore, one might expect that the Berlin Questionnaire would have better diagnostic validity in men than women, as the measure assesses the more common symptoms of OSA. However, we found that the reverse was true in this sample. Across the Berlin Questionnaire domains, the sleepiness category had the highest PPV in men. This may reflect that men—especially Black men—seem to experience sleepiness as a symptom of OSA more often than women.[25]

The primary strength of this study is that it engages a sample of largely low-income Black individuals—a population that is particularly at risk for OSA. Despite the increased risk, this population has been under-represented in OSA research.[12] In addition, it is uncommon for studies in Black American populations to report on objective sleep apnea measurements alongside the Berlin Questionnaire, which is widely used to screen for the condition. This presented a unique opportunity to explore whether this is an appropriate measure for this population.

Table 5—Diagnostic validity of the Berlin Questionnaire in the present study and other validation studies focusing on a non-specialty-care–seeking population.

	PHRESH Sample (Present Study)		Hrubos-Strøm et al, 2011[38]		Bouloukaki et al, 2013[22]	
	AHI ≥ 5 events/h	AHI ≥ 15 events/h	AHI ≥ 5 events/h	AHI ≥ 15 events/h	AHI ≥ 5 events/h	AHI ≥ 15 events/h
Sensitivity	36%	46%	37%	43%	76%	86%
Specificity	75%	72%	84%	80%	45%	61%
PPV	66%	29%	61%	34%	94%	88%
NPV	48%	85%	66%	86%	15%	58%

Included studies were selected as a point of comparison because they used an oxygen desaturation threshold of 4%. AHI = apnea-hypopnea index, NPV = negative predictive value, PHRESH = Pittsburgh Hill/Homewood Research on Neighborhood Change and Sleep, PPV = positive predictive value.

That said, there are certain limitations of this work. First, only a subsample of participants completed the HSAT, which was not a random sample of all eligible participants given limitations related to budget and equipment. Though it is not possible to fully account for the possibility of bias introduced by this recruitment strategy, descriptive analyses suggested that there are few differences between participants who completed HSAT and those who did not. That said, individuals in the HSAT sample were significantly younger, which may have resulted in an underestimate of OSA prevalence in this sample. Second, our sample included a higher proportion of females, by design, because the original study recruited the primary grocery shopper in each household. It may be that the men included in this sample are not representative of the overall population of men in the neighborhoods where this study was based. Finally, we defined moderate to severe OSA based on AHI only and did not include other symptoms to make this determination.

Regarding the implications of these findings, as previously noted, research suggests that the Berlin Questionnaire is more accurate in patient populations (eg, patients of sleep clinics)—those already known to be at risk for the condition.[37] For PPV and NPV, this likely reflects in part the higher base rate of OSA in specialty care populations than the general population. Our participants were not treatment-seeking, which may partially explain our findings. The US Preventive Services Task Force has not found enough evidence to determine that OSA should be screened for in asymptomatic adults,[40] and our results are consistent with that recommendation. That said, Black individuals are at an increased risk for OSA, and these findings may also reflect the fact that the Berlin Questionnaire may not assess symptoms of OSA that are more common in a Black population. For example, Johnson and colleagues[39] found that BMI, habitual snoring, and neck circumference were associated with higher odds of OSA in an African American population, whereas sleepiness was not. It may be that screening for neck circumference rather than sleepiness would improve the predictive validity of the Berlin Questionnaire. Although the Berlin Questionnaire is an easy tool, which can be widely implemented, knowing that the predictive validity varies across populations is essential. In this way, these findings also underscore the need to identify and/or develop screening tests that may be more appropriate for a Black population.

ABBREVIATIONS

AHI, apnea-hypopnea index
BMI, body mass index
HSAT, home sleep apnea test
NPV, negative predictive value
OSA, obstructive sleep apnea
PPV, positive predictive value

REFERENCES

1. Senaratna CV, Perret JL, Lodge CJ, et al. Prevalence of obstructive sleep apnea in the general population: A systematic review. *Sleep Med Rev.* 2017;34:70–81.
2. Franklin KA, Lindberg E. Obstructive sleep apnea is a common disorder in the population—a review on the epidemiology of sleep apnea. *J Thorac Dis.* 2015;7(8): 1311–1322.
3. Olaithe M, Bucks RS, Hillman DR, Eastwood PR. Cognitive deficits in obstructive sleep apnea: Insights from a meta-review and comparison with deficits observed in COPD, insomnia, and sleep deprivation. *Sleep Med Rev.* 2018;38:39–49.
4. Stranks EK, Crowe SF. The cognitive effects of obstructive sleep apnea: an updated meta-analysis. *Arch Clin Neuropsychol.* 2016;31(2):186–193.
5. Engleman HM, Douglas NJ. Sleep · 4: sleepiness, cognitive function, and quality of life in obstructive sleep apnoea/hypopnoea syndrome. *Thorax.* 2004;59(7):618–622.
6. Gonsalves MA, Paiva T, Ramos E, Guilleminault C. Obstructive sleep apnea syndrome, sleepiness, and quality of life. *Chest.* 2004;125(6):2091–2096.
7. Ruiter ME, DeCoster J, Jacobs L, Lichstein KL. Sleep disorders in African Americans and Caucasian Americans: a meta-analysis. *Behav Sleep Med.* 2010;8(4):246–259.
8. Scharf SM, Seiden L, DeMore J, Carter-Pokras O. Racial differences in clinical presentation of patients with sleep-disordered breathing. *Sleep Breath.* 2004;8(4): 173–183.
9. Spilsbury JC, Storfer-Isser A, Kirchner HL, et al. Neighborhood disadvantage as a risk factor for pediatric obstructive sleep apnea. *J Pediatr.* 2006;149(3):342–347.
10. Brouillette RT, Horwood L, Constantin E, Brown K, Ross NA. Childhood sleep apnea and neighborhood disadvantage. *J Pediatr.* 2011;158(5):789–795.e1.
11. Billings ME, Gold D, Szpiro A, et al. The association of ambient air pollution with sleep apnea: the Multi-Ethnic Study of Atherosclerosis. *Ann Am Thorac Soc.* 2019;16(3): 363–370.
12. Billings ME, Johnson DA, Simonelli G, et al. Neighborhood walking environment and activity level are associated with OSA: the Multi-Ethnic Study of Atherosclerosis. *Chest.* 2016;150(5):1042–1049.
13. Johnson DA, Drake C, Joseph CL, Krajenta R, Hudgel DW, Cassidy-Bushrow AE. Influence of neighbourhood-level crowding on sleep-disordered breathing severity: mediation by body size. *J Sleep Res.* 2015;24(5):559–565.
14. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Rep.* 2001;116(5):404–416.
15. Amra B, Rahmati B, Soltaninejad F, Feizi A. Screening questionnaires for obstructive sleep apnea: an updated systematic review. *Oman Med J.* 2018;33(3):184–192.
16. Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. *Anesthesiology.* 2009;110(4):928–939.
17. Abrishami A, Khajehdehi A, Chung F. A systematic review of screening questionnaires for obstructive sleep apnea. *Can J Anaesth.* 2010;57(5):423–438.
18. Tan A, Yin JD, Tan LW, van Dam RM, Cheung YY, Lee C-H. Using the Berlin questionnaire to predict obstructive sleep apnea in the general population. *J Clin Sleep Med.* 2017;13(3):427–432.
19. Ahmadi N, Chung SA, Gibbs A, Shapiro CM. The Berlin questionnaire for sleep apnea in a sleep clinic population: relationship to polysomnographic measurement of respiratory disturbance. *Sleep Breath.* 2008;12(1):39–45.
20. 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 (7):485–491.
21. Sforza E, Chouchou F, Pichot V, Herrmann F, Barthélémy 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.
22. Bouloukaki I, Komninos ID, Mermigkis C, et al. Translation and validation of Berlin questionnaire in primary health care in Greece. *BMC Pulm Med.* 2013;13(1):6.
23. Saleh ABM, Ahmad MA, Awadalla NJ. Development of Arabic version of Berlin questionnaire to identify obstructive sleep apnea at risk patients. *Ann Thorac Med.* 2011;6(4):212–216.
24. Fülöp T, Hickson DA, Wyatt SB, et al. Sleep-disordered breathing symptoms among African-Americans in the Jackson Heart Study. *Sleep Med.* 2012;13(8):1039–1049.
25. Eliasson AH, Kashani MD, Howard RS, Vernalis MN, Modlin RE; Integrative Cardiac Health Project Registry. Fatigued on Venus, sleepy on Mars—gender and racial differences in symptoms of sleep apnea. *Sleep Breath.* 2015;19(1):99–107.
26. Lozo T, Komnenov D, Badr MS, Mateika JH. Sex differences in sleep disordered breathing in adults. *Respir Physiol Neurobiol.* 2017;245:65–75.

27. US Census Bureau. Comparing *2009 American Community Survey Data*. 2018. https://www.census.gov/programs-surveys/acs/guidance/comparing-acs-data/2009.html. Accessed May 29, 2021.
28. Dong L, Dubowitz T, Haas A, et al. Prevalence and correlates of obstructive sleep apnea in urban-dwelling, low-income, predominantly African-American women. *Sleep Med.* 2020;73:187–195.
29. Ware JE Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol.* 1998;51(11):903–912.
30. Rofail LM, Wong KK, Unger G, Marks GB, Grunstein RR. The utility of single-channel nasal airflow pressure transducer in the diagnosis of OSA at home. *Sleep.* 2010;33(8):1097–1105.
31. Oktay B, Rice TB, Atwood CW Jr, et al. Evaluation of a single-channel portable monitor for the diagnosis of obstructive sleep apnea. *J Clin Sleep Med.* 2011;7(4):384–390.
32. Kapur VK, Auckley DH, Chowdhuri S, 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(3):479–504.
33. Berry R, Quan SF, Abreu AR, et al; for the American Academy of Sleep Medicine. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Version 2.6. Darien, IL: American Academy of Sleep Medicine; 2020.
34. Centers for Medicare and Medicaid Services. *Decision Memo for Continuous Positive Airway Pressure (CPAP) Therapy for Obstructive Sleep Apnea* (OSA)(CAG-00093N). 2001. https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=19&fromdb=true. Accessed May 21, 2021.
35. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol.* 2013;177(9):1006–1014.
36. Riley WJ. Health disparities: gaps in access, quality and affordability of medical care. *Trans Am Clin Climatol Assoc.* 2012;123:167–172, discussion 172–174.
37. Senaratna CV, Perret JL, Matheson MC, et al. Validity of the Berlin Questionnaire in detecting obstructive sleep apnea: a systematic review and meta-analysis. *Sleep Med Rev.* 2017;36:116–124.
38. Hrubos-Strøm H, Randby A, Namtvedt SK, 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(1 Pt 2):162–170.
39. Johnson DA, Guo N, Rueschman M, Wang R, Wilson JG, Redline S. Prevalence and correlates of obstructive sleep apnea among African Americans: the Jackson Heart Sleep Study. *Sleep.* 2018;41(10):zsy154.
40. Bibbins-Domingo K, Grossman DC, Curry SJ, et al; US Preventive Services Task Force. Screening for obstructive sleep apnea in adults: US Preventive Services Task Force recommendation statement. *JAMA.* 2017;317(4):407–414.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication November 23, 2020
Submitted in final revised form April 6, 2021
Accepted for publication April 7, 2021

Address correspondence to: Stephanie Brooks Holliday, PhD, 1776 Main Street, Santa Monica, CA 90407; Email: holliday@rand.org

DISCLOSURE STATEMENT

All authors have seen and approved the final manuscript. This study was funded by the National Heart, Lung, and Blood Institute (R01HL122460 and R01HL131531) and the National Cancer Institute (R01CA164137). The authors report no conflicts of interest.