



Published in final edited form as:

*J Perinatol.* 2014 August ; 34(8): 587–593. doi:10.1038/jp.2014.48.

## Obstructive Sleep Apnea in Pregnancy: Reliability of Prevalence and Prediction Estimates

Kathleen M. Antony, M.D.<sup>1</sup>, Alpna Agrawal, Ph.D.<sup>2</sup>, Melanie E. Arndt, B.S.<sup>1</sup>, Adrienne M. Murphy, B.S.<sup>1</sup>, Philip M. Alapat, M.D.<sup>3</sup>, Kalpalatha K. Guntupalli, M.D.<sup>3</sup>, and Kjersti M. Aagaard, M.D. Ph.D.<sup>1</sup>

<sup>1</sup>Department of Obstetrics and Gynecology, Division of Maternal-Fetal Medicine, Baylor College of Medicine, Houston, TX

<sup>2</sup>School of Medicine, University of Texas Health Science Center, Houston, TX

<sup>3</sup>Department of Pulmonary, Critical Care, and Sleep Medicine, Baylor College of Medicine, Houston, TX

### Abstract

**Objective**—We sought to ascertain the validity of two screening scales for obstructive sleep apnea (OSA) in pregnancy and to establish the prevalence of OSA in pregnancy.

**Study Design**—In this prospective observational study, two screening scales were administered. Screen positive subjects were referred for diagnostic polysomnography (PSG); if admitted for antepartum care, screen positive subjects underwent a modified study with a type 3 device (T3D).

**Result**—1509 subjects underwent OSA screening; 58 completed diagnostic testing. Neither measure was a reliable diagnostic tool for OSA as determined by T3D or PSG (detection rates of 10.3% and 18.0%, respectively). Among screen positive subjects undergoing PSG or T3D testing, 15.5% ultimately met ‘gold standard’ OSA diagnostic criteria for an estimated point prevalence of 4.9%.

**Conclusion**—In this prospective trial, screening positive on the BQ or ESS was poorly predictive of OSA among gravidae and was associated with a high false referral rate.

### Keywords

Obesity; Sleep-Disordered Breathing; Screening

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

**Corresponding author and reprint requests:** Kjersti Aagaard, M.D., Ph.D., Associate Professor, Department of Obstetrics and Gynecology, Division of Maternal-Fetal Medicine, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, Phone: 713-798-8467, Fax: 713-798-4216, [aagaardt@bcm.edu](mailto:aagaardt@bcm.edu).

#### Disclosure Statement:

*ResMed* donated single-use nasal cannulas and oximeter sensors for this study.

#### Paper Presentation Information:

This paper was previously presented as a poster at the Society of Maternal-Fetal Medicine’s 33<sup>rd</sup> Annual Meeting- The Pregnancy Meeting™, San Francisco, California, February 11–16, 2013. Abstract Number 671.

**Conflict of Interest:** The authors report no conflict of interest.

## INTRODUCTION

Obstructive sleep apnea (OSA) is characterized by recurrent cessation of respiratory airflow resulting from upper airway collapse during sleep and is accompanied by oxygen desaturation or arousals(1,2). The diagnosis of OSA is established by polysomnography (PSG) (3). Given the time and expense associated with making this diagnosis, several questionnaire-based scales have been developed and validated as screening tools in the nonpregnant population, namely the Berlin questionnaire (BQ) and Epworth sleepiness scale (ESS)(4–7). Prior attempts to validate these questionnaires in pregnancy have had variable results. Delineating the relationship between OSA screening and diagnosis is paramount given the proposed relationship of OSA with adverse pregnancy outcomes, such as preeclampsia and fetal growth restriction(8–14). Furthermore, the majority of studies to date have failed to adjust for OSA risk factors which also serve as potential outcome confounders, most notably obesity(13,15).

There are both epidemiological and physiological data to suggest that pregnant women may be predisposed to OSA(16–18). Diagnosing this condition in pregnancy, however, remains elusive. Our prior published research suggests that the BQ performs poorly in this population with sensitivity and specificity of only 35% and 63.8%, respectively (8). The goal of this study was to enroll a large, prospective, population-based cohort in order to establish the prevalence of OSA in pregnancy and to ascertain the validity of the BQ and the ESS as screening tools for OSA in pregnancy. Based upon our pilot study of 100 patients undergoing BQ screening followed by a modified four-channel sleep screening test(8), we hypothesized that BQ and ESS would perform poorly as screening measures for OSA in pregnancy.

To accomplish these objectives, subjects were prospectively screened for OSA. Screen positive subjects were referred to a sleep lab for polysomnography (PSG). Subsequent univariate and multivariate logistic regression analysis determined the point prevalence of OSA in pregnancy, as well as established relative estimates of false referral rates.

## MATERIALS AND METHODS

This study was performed in the Harris County Hospital District (now Harris Health System) between May 2010 and September 2012. The study was approved by the Institutional Review Board at Baylor College of Medicine and Harris County Hospital District and written informed consent was obtained from all enrolled subjects. Gravidae presenting to the Casa de Amigos Health Center, People's Health Center (now Vallbona Health Center), and Ben Taub High-Risk Obstetrics Clinic were approached for enrollment; a combination of low and high obstetric-risk patients (as defined by maternal and fetal comorbidities, but not risk of OSA *per se*) were specifically recruited to assess the validity of screening questionnaires in a "real world" heterogeneous clinical setting. Consenting subjects were administered the standard screening measure questionnaire (in English or Spanish) which comprised the ESS, the BQ, and questions collecting basic health and sleep information. Women of all gestational ages were recruited. Inclusion criteria were gravidae of 18–50 years of age. Exclusion criteria were subjects with known sleep-disordered

breathing, multifetal gestation, fatal fetal anomalies, and subjects with significant underlying pulmonary or cardiac comorbidities due to likely confounding that could not be adequately controlled for.

Resource constraints only allowed diagnostic testing to be performed on subjects who were suspected of having OSA based on ESS or BQ as described below. Given the poor sensitivity of the BQ, to capture as many cases as possible, enrollees who screened positive based upon either the ESS or Berlin measures were referred for diagnostic testing(8).

The ESS and BQ are screening measures that are clinically employed to assess the risk of being diagnosed with OSA. The ESS consists of eight questions that evaluate the tendency to fall asleep in certain situations and is a subjective measure of sleepiness(5,19). In nonpregnant subjects, ESS has a sensitivity of 66–93.5% and specificity from 48–100%(6,19). For this study, we considered a patient to screen positive for OSA with an ESS of 10 or higher. The BQ questions were developed to elicit factors or behaviors that consistently predicted the presence of sleep disordered breathing(4). In nonpregnant adults, it has a sensitivity of 86% and specificity of 77%(4). It is divided into three categories that focus on the presence and severity of snoring, witnessed hypopneas, daytime sleepiness, and presence of hypertension or obesity(4). A high risk or “screen positive” individual is reported to have at least two out of the three symptom categories positive in the questionnaire(4).

Screen positive subjects were referred for attended PSG at the Harris County Hospital District Sleep Disorders Center. Subjects were scheduled for PSG during their 26th–28th week of gestation or as soon thereafter as feasible. The rationale for using this gestational age for testing was twofold: first, we did not have adequate resources to test women twice during pregnancy to assess for pre-existing OSA versus new OSA, so we could only perform diagnostic testing once; second, we wanted to capture as many cases of OSA as possible to allow accurate cohort assignment. We chose 26–28 weeks based upon the prior finding that CPAP requirement increased around 24–26 weeks in gravidae with OSA, suggesting that the course of the disease would have worsened or manifested itself by that gestational age(20). Subjects recruited after 28 weeks gestation were scheduled as soon after screening positive as feasible. Screen positive subjects were contacted by mail and by designated phone contact (K.M.A.1) in their native language on at least two occasions. Once a PSG was scheduled, the subject was contacted in their native language by two means on two separate occasions (K.M.A.1 and the Sleep Center). Given subjects’ and the Sleep Center’s constraints, few overnight PSGs were accomplished. The protocol was subsequently amended to allow inpatient T3D testing as a surrogate for PSG. For subjects undergoing T3D testing, this was performed during inpatient admission at any gestational age.

Attended PSG testing was performed using a Carefusion Somnostar polysomnography system (CareFusion, Yorba Linda, California). Via this multi-channel system, the following variables were monitored continuously: electroencephalography, electrocardiography, electro-oculography, submental and leg electromyography, and electrocardiography. Continuous heart rate and pulse oximetry were monitored using a finger probe, and airflow was determined via nasal cannula-pressure transducer and thermistor. Additional variables

measured included body position sensors, snore microphones, and thoracic and abdominal piezoelectric bands. Attended PSG tests were administered by a certified technologist, and studies were reviewed by a physician, board certified in sleep medicine. OSA was diagnosed with Respiratory Disturbance Index (RDI) > 5 using attended PSG. Attended PSG was scored using the AASM Manual for the Scoring of Sleep and Associated Events, 2007(3).

The *ResMed ApneaLink Plus* (ResMed Corp, MAP Medizin Technologie GmbH, Poway, CA) is a type 3 unattended home sleep testing device with a minimum of four channels used to test for OSA. It consists of a pulse oximeter for oximetry data, heart rate recording, respiratory effort belt, and a nasal sensor for detection of flow limitation, apnea hypopnea index (AHI), and snoring. The device generates a report with SpO<sub>2</sub> information and the AHI, among other variables. An AHI  $\geq 5$  was considered diagnostic of OSA. At the time this study was performed, no home sleep testing had been validated in pregnancy, but it had been used in prior obstetric studies as a diagnostic tool(8,21). For the purposes of this study, T3D was considered similarly diagnostic. However, if a T3D was performed, the referral for PSG testing was not cancelled as PSG remains the standard for diagnosis. Had a sufficient number of subjects undergone both T3D and PSG testing, it would have additionally allowed for calculations of the predictive value of T3D in pregnancy.

The diagnostic criteria for OSA differ between PSG and T3D; PSG uses RDI and T3D uses AHI. These measures have been previously compared in a non-gravid population and shown to correlate well; thus, while they measure different components of the sleep study, the precedent exists to utilize the index specific to the study type(22).

Covariates considered in these analyses were maternal age, race/ethnicity, smoking, parity, gestational age, prepregnancy body mass index (where appropriate), pregestational diabetes, and chronic hypertension. BMI categories were defined according to the International Obesity Task Force classification: normal weight 19 to 24.9, overweight 25 to 29.9, and obese >30 kg/m<sup>2</sup>. BMI was calculated using height and weight data (kg/m<sup>2</sup>) that were collected during initial assessment. When prepregnancy weight was not known, the earliest available weight was used.

As there is no sample size estimate calculation to assess the validity of a questionnaire, our sample size estimation was based on having adequate power to examine the relationship between OSA and various maternal and neonatal outcomes, adjusted for the possible confounding influence of obesity(23). In our previous analysis of n=100 subjects, the relationship between standard Berlin measures and each of its components (snoring questions, sleepiness questions) and the T3D test that we used as criteria for diagnosis were analyzed via linear and logistic regression analyses(8). Based upon the screen positive rate for OSA of 36% and overall OSA diagnosis positive rate of 20%, we initially estimated that a minimum of 1100 subjects would need to be screened to detect a 20% difference in our primary perinatal outcomes. After 500 subjects were enrolled the protocol was amended; enrollment was increased to 1600 as only 4 diagnostic PSGs had been completed to that point. Based upon the same pilot study which found a screen positive rate of OSA of 36%, we anticipated performing PSG on 40% of subjects(8). This would give a proximate of the prevalence of OSA in pregnancy.

Descriptive findings of study sample characteristics by OSA screening results from the BQ and ESS and OSA diagnosis based on T3D or PSG were reported. Chi-squared testing or Fisher's exact test was performed to assess differences between OSA groups by study sample characteristics. Scale reliability for the BQ and ESS were assessed by computing Cronbach's  $\alpha$ , which measures internal consistency, and item-test correlation, which is the correlation of each item with the overall scale. Regarding unanswered test items, as no numerical value could be assigned to contribute to the overall sum on the ESS or BQ, they were treated as a "zero" for summation purposes. However, in all other item analyses, they were treated as missing data.

Multivariate logistic regression analysis was used to estimate associations between screening positive on the BQ scale and its items and the ESS scale and its items adjusting for confounders. In item analyses of BQ and ESS, items were dichotomized for ease of interpretation. Results were reported as adjusted odd ratios with 95% confidence intervals.

Due to resource constraints, overall screen negative subjects were not referred for diagnostic testing. As screen negative subjects did not undergo diagnostic testing, the detection rate and false referral rate were calculated(24). These terms represent the proportion of screened subjects who screened positive and had the disease and the proportion of subjects who screened positive who did not have the disease, respectively.

A probability value of  $<0.05$  was considered statistically significant. Stata 10.0 (Stata Corporation, College Station, TX) and SPSS 13.0 (SPSS Incorporated, Chicago, IL) were used for analyses.

## RESULTS

1617 BQ or ESS screening measures were completed. 79 women had enrolled on more than one occasion during their pregnancy; in these cases data from the questionnaire completed closest to 26–28 weeks was used for analysis; the duplicate questionnaires these women completed were not included as doing so would introduce co-linearity. Twenty-one subjects were excluded for multi-fetal gestation and eight were excluded for fatal fetal anomalies. One subject participated in the study during two separate pregnancies; her data was retained as the data were unique for each pregnancy. 75 subjects only fully completed one screening measure or the other, but not both; their available data was analyzed. Five subjects did not fully complete either screening measure (BQ or ESS), thus the BQ or ESS measure could not adequately be characterized, but the available individual items of each questionnaire were used for analysis.

The majority of gravid subjects were Hispanic (87.8%) and non-smokers (96.5%), which is consistent with the obstetric population of Harris Health System (Figure 1). Over three-fourths of the subjects were overweight (35.1%) or obese (40.4%). As shown in Figure 1, after excluding the five subjects that did not complete either screening measure, of the 1504 remaining subjects, 456 (30.3%) screened positive on either the BQ or ESS and 1048 (69.4%) screened negative on at least one measure.

Table 1 demonstrates the demographics of subjects who screened positive or negative on the BQ or ESS and shows the characteristics of subjects who screened positive or negative for OSA; the prevalence of subjects who screened positive varied by screening method. Among subjects who adequately completed the respective screening method, the prevalence of those screening positive was 15.1% by BQ, 20.4% by ESS, and 31.9% by either BQ or ESS. The screen positive rate by either BQ or ESS differs in the table as a result of excluding 75 subjects with inadequate completion of the screening measures.

Multivariate logistic regression analyses adjusting for age, ethnicity, smoking status, parity, gestational age at survey, BMI, and pregestational diabetes showed increased odds of screening positive on either the BQ or the ESS among Black subjects (aOR 3.59, 95% CI 2.12–6.07), subjects who smoked during pregnancy (aOR 2.60, 95% CI 1.22–5.54), and subjects with BMI over 30 (aOR 2.72, 95% CI 1.89–3.93). There was no association with gestational age at time of screening.

Figure 2 demonstrates the percentage distribution of BQ and ESS items. Among subjects who screened positive on the BQ scale, 40.4% of the sample had a BMI>30, 21.4% reported snoring, and 24.9% felt tired during waking time. Cronbach's  $\alpha$  for the BQ items was 0.62 and item-test correlations ranged from 0.30 to 0.66. Among subjects who screened positive on the ESS, 22.8%–83.9% of the sample reported some degree of dozing off. Cronbach's  $\alpha$  for the ESS items was 0.82 and item-test correlations ranged from 0.55 to 0.74.

As shown in Table 2, screening positive on the BQ scale was not associated with screening positive on the ESS after controlling for covariates. Specific BQ items positively associated with at least two ESS items after adjusting for covariates were items about snoring, feeling tired, and falling asleep while driving.

58 women were tested for OSA by T3D ( $n$  52 (50 results used)) or PSG ( $n$  8). Two women underwent both T3D testing and PSG; one was negative on both the T3D and PSG and one was negative on T3D and positive on PSG; the PSG results (rather than the T3D results) were used for both women in these analyses. For the purposes of this analysis, PSG and T3D were both considered diagnostic for OSA. 15.5% of women tested were diagnosed with OSA. Fischer's exact test showed no differences between OSA diagnosis by study sample characteristics, including age, race/ethnicity, smoking, BMI, gestational age at testing, and pregestational diabetes. (Data not shown.). As shown in Table 3, no items on the BQ were significantly associated with testing positive on PSG or T3D. Similarly, no items on the ESS were positively associated with testing positive on PSG or T3D, however, the item concerning the likelihood of falling asleep "as a passenger in a car for an hour without a break" was negatively associated with testing positive on PSG or T3D.

Table 4 gives the detection rates for each screening and diagnostic measure. The overall detection rate for screening positive on BQ or ESS was 15.5, and the false referral rate was 84.5. The detection rate for BQ was 10.3 with a false referral rate of 89.7 whereas the detection rate for ESS was 18.0 with a false referral rate of 82.1. Using the PSG, the detection rates were higher than using the T3D. However, very few subjects completed PSG.



As 31.9% of women screened positive on either BQ or ESS and 15.5% of those women tested positive on T3D or PSG, the overall prevalence of OSA in this study would be 4.9%, if the test positive rate among untested women is the same as tested women. Among subjects undergoing diagnostic testing, 8.0% (4/50) undergoing T3D were positive and 62.5% (5/8) undergoing PSG were positive. Thus, if the testing modes were analyzed separately, using T3D the prevalence of OSA would be 2.6% and by PSG the prevalence would be 19.9%.

## DISCUSSION

As we have demonstrated here, the prevalence of OSA in our studied obstetric population remains difficult to determine. Our 4.9% prevalence rate is based on all positive cases having occurred among screen positive women.

We have observed that screening by either BQ or ESS is poorly predictive of OSA among gravidae. The detection rate of BQ and ESS were both poor, and the false referral rates were high, particularly when using T3D as the diagnostic test. No individual item on either questionnaire was positively associated with testing positive by PSG or T3D; one item was negatively associated with screening positive on PSG or T3D, which was the item about falling asleep as a passenger in a car. It is not clear why there was a negative association; it may be that women with true OSA were too tired to go on trips over an hour, even as a passenger, but this is speculation.

By study design we were not able to calculate the true prevalence or the sensitivity and specificity, but the poor predictive value of BQ and ESS during pregnancy has been previously suggested by ourselves and others(9,12). After this study was initiated, Facco, *et al.*, demonstrated that the BQ has a sensitivity and specificity of 39% and 68%, respectively, while the ESS has a sensitivity and specificity of 36% and 77%, respectively(9). One recent study found a much higher sensitivity and specificity of the BQ among gravidae, but only performed diagnostic testing on women who either did not score as "high risk" in any of the three categories of the BQ or scored as "high risk" in all three categories(10). Thus the sensitivity and specificity in their study may have limited clinical utility among a population with more intermediate scores, or prone to confounding by virtue of inclusion of obesity on the BQ.

Regarding the questionnaires themselves, after adjusting for confounding variables, screening positive on BQ was not associated with screening positive on the ESS. Overall, the association between BMI>30 and screening positive on BQ was expected as obesity is a weighted component of the BQ. The Cronbach  $\alpha$  for BQ items of 0.62 reflects that it has relatively low internal consistency as a measure. The Cronbach  $\alpha$  for ESS was 0.82, which reflects a good amount of internal consistency; however, its ability to predict OSA in pregnancy is poor, as noted here and elsewhere(9,12).

There are several factors which may contribute to the BQ and ESS poor performance in pregnancy. Both questionnaires query daytime sleepiness, which is a common complaint during pregnancy(25–27); thus this symptom may not distinguish women with and without OSA. The BQ queries snoring, as well. The frequency of snoring increases during

pregnancy, with 37% reporting snoring often or every night in the last week (27), so it also may not serve as a very discriminating symptom. As shown in Table 3, no items were positively associated with testing positive for OSA by PSG or T3D.

Strengths of our study include its robust subject accrual, and dedicated means to minimize barriers to diagnosis and follow-up. In addition, because this study was performed in a “real-world” setting, it reveals the challenges of employing a broad-based screening program for this condition which has garnered increased attention in recent years. As we have shown here, questionnaire-based “diagnosis” of OSA in pregnancy is prone to disease misclassification and suggest that among gravid ESS and BQ should only be regarded as screening measures.

This study was designed with the limited availability of PSG in mind, thus only screen positive subjects were referred for PSG. As a prevalence study should refer all subjects for diagnostic testing, we attempted to refer as many subjects as possible and thus referred subjects screening positive on either the BQ or ESS.

Limitations of this study include the relatively small number of subjects who underwent attended PSG testing and the use of often used but as of yet unvalidated T3D for diagnosis of OSA in pregnancy: 456 subjects were referred for PSG, and 58 completed diagnostic testing. Although our ratio of screen positive to completion of diagnostic testing was poor, it is still the largest population based study performed to date and thus imparts valuable data.

There were several challenges to completing diagnostic testing in our population based cohort, and all are reflective of the realities of clinical practice among those at greatest risk for undiagnosed OSA. Notably, the PSGs themselves were technically feasible when performed, which is consistent with prior findings(15,28–31). PSG completion was solely hindered by the systematic availability of this limited resource and also by subject-based impediments inherent to any gravid population. There were a number of factors impeding the completion of PSGs, all of which are worthy of consideration for investigators and clinicians alike. First, despite collaboration with the Department of Pulmonary, Critical Care, and Sleep Medicine at the onset of the study design and throughout the study period, there were significant obstacles to completing PSG. The waiting period for PSGs often exceeded subjects’ gestation; thus, we coordinated with Sleep Center staff and when appointments became available due to cancellations, study subjects were prioritized. Second, mid-way through the study a co-pay was initiated for attended PSGs. In order to relieve this financial burden, all co-pays were covered for enrolled study subjects. Third, while attended PSGs were covered by a county-based health care plan, they were not covered by emergency pregnancy Medicaid. Fourth, when appointments at the Sleep Disorders Center were made, subjects often lacked transportation and childcare which led to cancelations or truancy in over 95% of appointments. Despite these recognized obstacles, either T3D or PSG diagnostic testing was completed in 12.7% of subjects; this represents the largest population-based outpatient validation testing undertaken and reported to date (PubMed and Ovid, 1986-present).



An additional limitation is the fact that the unattended T3D testing device used in this study had not been validated in pregnancy. T3D testing is known to have a high false negative rate, especially among patients with low AHI(32)(26), and women have a lower AHI than men(33)(27). In this study, among screen positive subjects undergoing sleep studies, the number of subjects completing both T3D and PSG was insufficient to draw conclusions regarding the accuracy of T3D. However, the percentage of patients tested who were diagnosed as positive via each method was discrepant (8.0% via T3D versus 62.5% via PSG), which may suggest non-equivalence. Since the completion of this study, two ambulatory sleep testing devices have demonstrated validity in pregnancy (34,35)(28, 29). Further studies of the validity of other T3D are needed, especially as they are commonly used as a surrogate for PSG in the literature. Regarding bias, as the majority of subjects were tested during antepartum admissions for non-study related indications, it is unlikely that subjects' concern regarding OSA introduced bias in subjects undergoing T3D testing. However, while all screen positive women were referred for PSG, women who were themselves concerned about a sleep disorder may have had higher motivation to undergo attended PSG, resulting in a higher rate of positive diagnosis. A final limitation is its limited external validity given that the majority of subjects were Hispanic.

In sum, because of the inherent challenges in diagnosis and lack of pregnancy-specific screening measures, the prevalence of obstructive sleep apnea in pregnancy remains poorly characterized. Our current estimate of 4.9% in the general obstetric population likely represents the most robust point prevalence data to date. This analysis suggests that the BQ and ESS as they are currently administered and scored are poorly predictive of OSA in pregnancy. This study also elucidated possibly pervasive barriers to PSG. Based on our data presented herein, in the interval until screening and alternate diagnostic measures are fully developed and validated in pregnancy, cautious use of these screening tools in clinical obstetrical practice is warranted.

## Acknowledgments

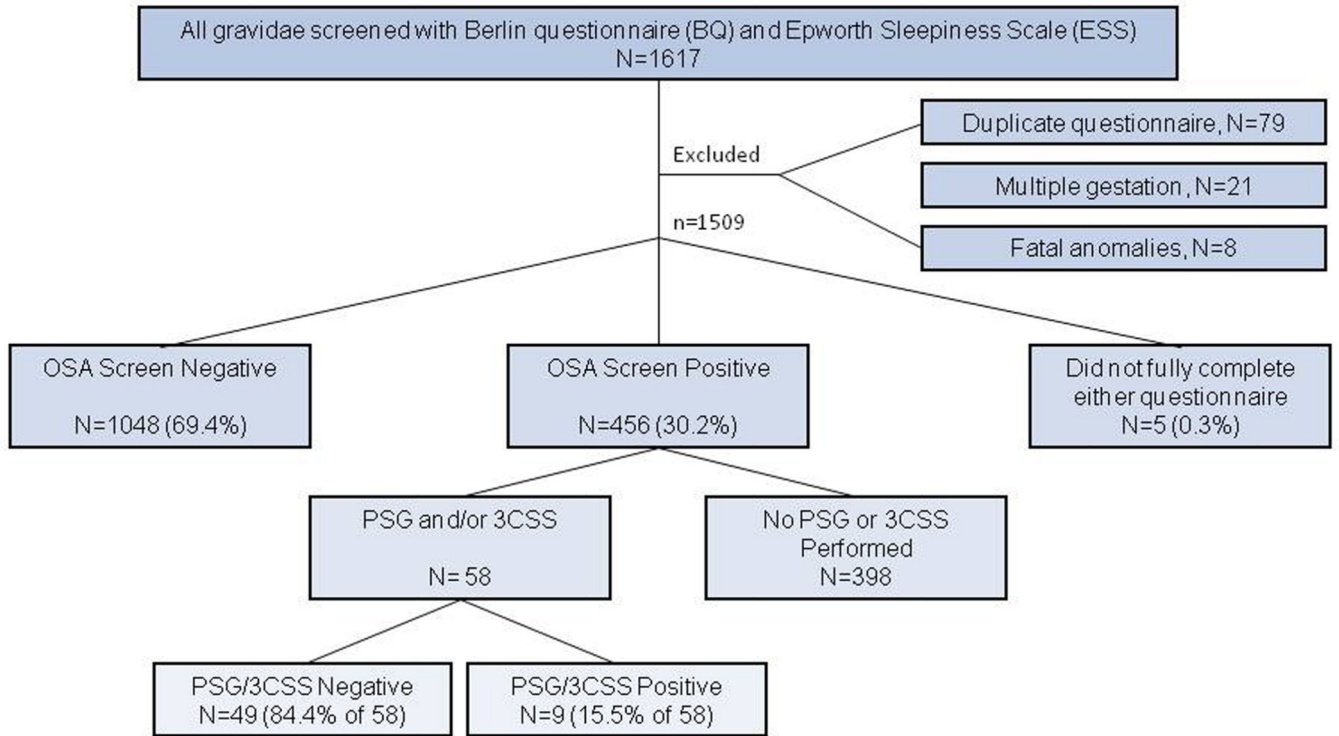
We thank Lata Casturi, RPGST at Baylor College of Medicine; *ResMed* for the use of the *ResMed ApneaLink* devices; and staff and nurses at Casa de Amigos, People's Health Center, and Ben Taub General Hospital Obstetrics Clinic and Antepartum Unit. KM Aagaard, M.D., Ph.D. was supported by a DP2 award (DP21DP2OD001500-01 NIH Director New Innovator Award). This study was financially supported by the Baylor College of Medicine Department of Obstetrics and Gynecology Resident Research Grant 2011–2012 (to KM Antony). *ResMed* donated single-use nasal cannulas and oximeter sensors for this study.

## REFERENCES

1. Dempsey JA, Veasey SC, Morgan BJ, O'Donnell CP. Pathophysiology of sleep apnea. *Physiol Rev.* 2010; 90(1):47–112. 2010/01/21 ed. [PubMed: 20086074]
2. Lurie A. Obstructive sleep apnea in adults: epidemiology, clinical presentation, and treatment options. *Adv Cardiol.* 2011; 46:1–42. 2011/10/19 ed. [PubMed: 22005188]
3. Iber, C.; Ancoli-Israel, S.; Chesson, A.; Quan, SF. *The AASM Manual for the Score of Sleep and Associated Events: Rules, Terminology, and Technical Specifications.* 1st ed.. Westchester, Illinois: American Academy of Sleep Medicine; 2007.
4. 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. [PubMed: 10507956]

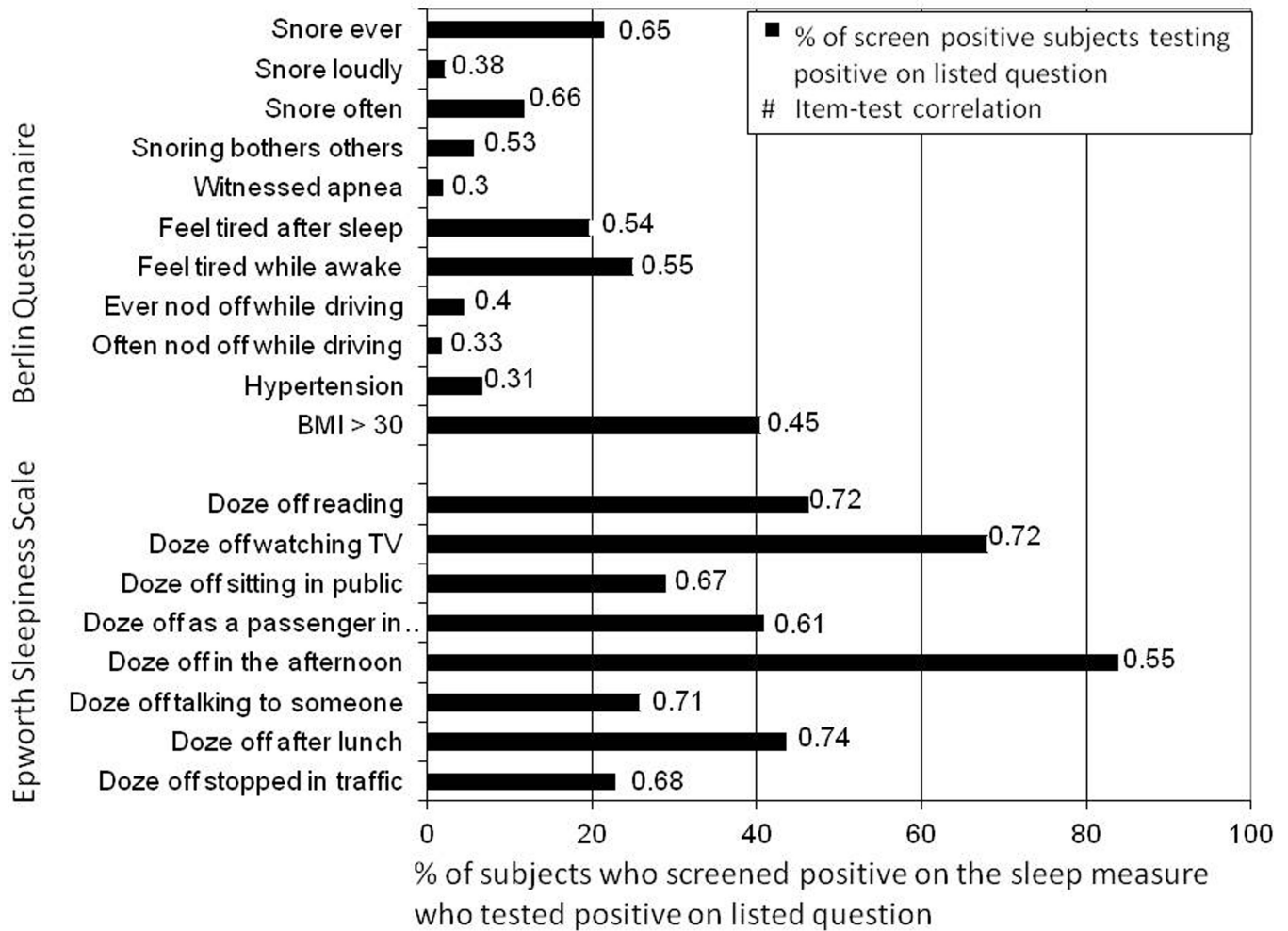
5. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991; 14(6):540–545. 1991/12/01 ed. [PubMed: 1798888]
6. Johns MW. Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the epworth sleepiness scale: failure of the MSLT as a gold standard. *J Sleep Res* [Internet]. 2000 Mar; 9(1):5–11. 2000/03/25 ed. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10733683>.
7. Sharma SK, Vasudev C, Sinha S, Banga A, Pandey RM, Handa KK. Validation of the modified Berlin questionnaire to identify patients at risk for the obstructive sleep apnoea syndrome. *Indian J Med Res*. 2006; 124(3):281–290. 2006/11/07 ed. [PubMed: 17085831]
8. Olivarez SA, Maheshwari B, McCarthy M, Zacharias N, van den Veyver I, Casturi L, et al. Prospective trial on obstructive sleep apnea in pregnancy and fetal heart rate monitoring. *Am J Obs Gynecol*. 2010; 202(6):552 e1–552 e7. 2010/02/23 ed.
9. Facco FL, Ouyang DW, Zee PC, Grobman Wa. Development of a pregnancy-specific screening tool for sleep apnea. *J Clin Sleep Med* [Internet]. 2012 Jan; 8(4):389–394. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3407257&tool=pmcentrez&rendertype=abstract>.
10. Wilson DL, Walker SP, Fung AM, O'Donoghue F, Barnes M, Howard M. Can we predict sleep-disordered breathing in pregnancy? The clinical utility of symptoms. *J Sleep Res*. 2013 2013/06/12 ed.
11. Bourjeily G, Raker CA, Chalhoub M, Miller MA. Pregnancy and fetal outcomes of symptoms of sleep-disordered breathing. *Eur Respir J*. 2010; 36(4):849–855. 2010/06/08 ed. [PubMed: 20525714]
12. Olivarez SAA, Ferres M, Antony K, Mattewal A, Maheshwari B, Sangi-hagheykar H, et al. Obstructive sleep apnea screening in pregnancy, perinatal outcomes, and impact of maternal obesity. *Am J Perinatol*. 2011; 28(8):651–658. 2011/04/12 ed. [PubMed: 21480159]
13. Louis JM, Auckley D, Sokol RJ, Mercer BM. Maternal and neonatal morbidities associated with obstructive sleep apnea complicating pregnancy. *Am J Obs Gynecol*. 2010; 202(3):261 e1–261 e 5. 2009/12/17 ed.
14. Chen YH, Kang JH, Lin CC, Wang IT, Keller JJ, Lin HC. Obstructive sleep apnea and the risk of adverse pregnancy outcomes. *Am J Obs Gynecol*. 2012; 206(2):136 e1–136 e5. 2011/10/18 ed.
15. Sahin FK, Koken G, Cosar E, Saylan F, Fidan F, Yilmazer M, et al. Obstructive sleep apnea in pregnancy and fetal outcome. *Int J Gynaecol Obstet* [Internet]. 2008 Feb; 100(2):141–146. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17976624>.
16. Sahota PK, Jain SS, Dhand R. Sleep disorders in pregnancy. *Curr Opin Pulm Med*. 2003; 9(6): 477–483. 2003/10/10 ed. [PubMed: 14534398]
17. Izci B, Riha RL, Martin SE, Vennelle M, Liston WA, Dundas KC, et al. The upper airway in pregnancy and pre-eclampsia. *Am J Respir Crit Care Med*. 2003; 167(2):137–140. 2002/11/02 ed. [PubMed: 12411285]
18. Pilkington S, Carli F, Dakin MJ, Romney M, De Witt KA, Dore CJ, et al. Increase in Mallampati score during pregnancy. *Br J Anaesth*. 1995; 74(6):638–642. 1995/06/01 ed. [PubMed: 7640115]
19. Rosenthal LD, Dolan DC. The Epworth sleepiness scale in the identification of obstructive sleep apnea. *J Nerv Ment Dis*. 2008; 196(5):429–431. 2008/05/15 ed. [PubMed: 18477888]
20. Guilleminault C, Kreutzer M, Chang JL. Pregnancy, sleep disordered breathing and treatment with nasal continuous positive airway pressure. *Sleep Med* [Internet]. 2004 Jan; 5(1):43–51. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S138994570300193X>.
21. Connolly G, Razak AR, Hayanga A, Russell A, McKenna P, McNicholas WT. Inspiratory flow limitation during sleep in pre-eclampsia: comparison with normal pregnant and nonpregnant women. *Eur Respir J*. 2001; 18(4):672–676. 2001/11/22 ed. [PubMed: 11716173]
22. Nigro CA, Serrano F, Aimaretti S, Gonzalez S, Codinardo C, Rhodius E. Utility of ApneaLink for the diagnosis of sleep apnea-hypopnea syndrome. *Med (B Aires)*. 2010; 70(1):53–59. 2010/03/17 ed.
23. Antony KM, Agrawal A, Arndt ME, Murphy AM, Alapat PM, Guntupalli KK, et al. Association of adverse perinatal outcomes with screening measures of obstructive sleep apnea. *J Perinatol Press*. 2014

24. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst.* 2001; 93(14):1054–1061. 2001/07/19 ed. [PubMed: 11459866]
25. Franklin KA, Holmgren PA, Jonsson F, Poromaa N, Stenlund H, Svanborg E. Snoring, pregnancy-induced hypertension, and growth retardation of the fetus. *Chest.* 2000; 117(1):137–141. 2000/01/13 ed. [PubMed: 10631211]
26. Facco FL, Kramer J, Ho KH, Zee PC, Grobman WA. Sleep disturbances in pregnancy. *Obs Gynecol.* 2010; 115(1):77–83. 2009/12/23 ed.
27. Hutchison BL, Stone PR, McCowan LME, Stewart AW, Thompson JMD, Mitchell Ea. A postal survey of maternal sleep in late pregnancy. *BMC Pregnancy Childbirth* [Internet]. *BMC Pregnancy and Childbirth.* 2012 Jan. 12(1):144. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3541269&tool=pmcentrez&rendertype=abstract>. [PubMed: 23228137]
28. Schiza SE, Bouloukaki I, Mermigkis C, Bourjeily G, El Sabbagh R, Sawan P, et al. Epworth sleepiness scale scores and adverse pregnancy outcomes. *Sleep Breath.* 2013 2013/02/20 ed.
29. Edwards N, Blyton DM, Kirjavainen TT, Sullivan CE. Hemodynamic responses to obstructive respiratory events during sleep are augmented in women with preeclampsia. *Am J Hypertens.* 2001; 14(11 Pt 1):1090–1095. 2001/11/29 ed. [PubMed: 11724205]
30. Guilleminault C, Palombini L, Poyares D, Takaoka S, Huynh NT, El-Sayed Y. Pre-eclampsia and nasal CPAP: part 1. Early intervention with nasal CPAP in pregnant women with risk-factors for pre-eclampsia: preliminary findings. *Sleep Med.* 2007; 9(1):9–14. 2007/07/24 ed. [PubMed: 17644420]
31. Poyares D, Guilleminault C, Hachul H, Fujita L, Takaoka S, Tufik S, et al. Pre-eclampsia and nasal CPAP: part 2. Hypertension during pregnancy, chronic snoring, and early nasal CPAP intervention. *Sleep Med.* 2007; 9(1):15–21. 2007/07/24 ed. [PubMed: 17644475]
32. Oktay B, Rice TB, Atwood CW Jr, Passero M Jr, Gupta N, Givelber R, et al. Evaluation of a single-channel portable monitor for the diagnosis of obstructive sleep apnea. *J Clin Sleep Med* [Internet]. 2011 Aug 15; 7(4):384–390. 2011/09/08 ed. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3160740&tool=pmcentrez&rendertype=abstract>.
33. Ralls FM, Grigg-Damberger M. Roles of gender, age, race/ethnicity, and residential socioeconomic in obstructive sleep apnea syndromes. *Curr Opin Pulm Med.* 2012; 18(6):568–573. 2012/09/20 ed. [PubMed: 22990656]
34. O'Brien LM, Bullough AS, Shelgikar AV, Chames MC, Armitage R, Chervin RD. Validation of Watch-PAT-200 against polysomnography during pregnancy. *J Clin Sleep Med.* 2012; 8(3):287–294. 2012/06/16 ed. [PubMed: 22701386]
35. Louis J, Auckley D, Miladinovic B, Shepherd A, Mencin P, Kumar D, et al. Perinatal Outcomes Associated With Obstructive Sleep Apnea in Obese Pregnant Women. *Obs Gynecol.* 2012 2012/09/26 ed.



**Figure 1.**

Flow through study. 1509 subjects were eligible for inclusion. 5 subjects did not fully complete either questionnaire. 1048 subjects completed at least one questionnaire and were screen negative for both the BQ and the ESS, including 75 subjects who fully completed one questionnaire but incompletely completed the other. 456 subjects screened positive for OSA on either the BQ, the ESS, or both. All 456 subjects were referred to the Sleep Center for PSG. If admitted to antepartum during gestation, a T3D was performed. 398 subjects did not complete PSG or T3D. 58 subjects completed PSG and/or T3D. Of the 58 subjects who completed sleep testing via PSG or T3D, 15.5% were diagnosed with obstructive sleep apnea.



**Figure 2.**

Item-test correlations and percent of subjects who screened positive on the sleep measure screening positive on the listed item. The bar graph shows the percent of screen positive subjects on the Berlin Questionnaire or Epworth Sleepiness Scale, respectively, who tested positive on the listed item. The item-test correlation for each item is shown to the right of the bar. Cronbach's  $\alpha$  for BQ items was 0.62, and itemized correlation ranged from 0.30 to 0.66. Cronbach's  $\alpha$  for ESS items was 0.82, and itemized correlations ranged from 0.55 to 0.74.

**Table 1**

Demographic characteristics by the presence of screening positive for obstructive sleep apnea on the Berlin Questionnaire (BQ) or the Epworth Sleepiness Scale (ESS).

	Total	n	BQ -	BQ +	ESS -	ESS +	Screen -	Screen +	aOR	aOR
	%	%	%	%	%	%	%	%	%	%
	100.0	1509	n=1275 84.9	n=227 15.1	n=113 5	n=290 79.6	n=973// 68.1	n=456 31.9		
Maternal age groups		1509								
<20	8.3	125	8.6	6.6	6.9	15.5	7.2	11.6	Ref	Ref
20-29	50.2	758	51.5	44.1	50.1	52.8	51.4	48.7	NS	NS
30-39	37.2	562	36.4	41.4	38.9	26.9	38.0	33.8	0.43 <sup>†</sup>	NS
40-46	4.2	64	3.5	7.9	4.1	4.8	3.4	5.9	3.56 <sup>*</sup>	NS
$\chi^2$ test			<b>p&lt;0.001</b>		<b>p&lt;0.001</b>		<b>p&lt;0.001</b>			
Race/Ethnicity		1509								
Hispanic	87.8	1325	89.3	78.9	90.2	76.2	91.3	79.2	Ref	Ref
Black	8.6	130	7.3	16.3	7.0	16.6	6.0	15.1	3.06 <sup>‡</sup>	3.59 <sup>*</sup>
Caucasian	1.1	17	0.9	2.2	1.1	1.7	0.8	2.0	NS	NS
Other	2.5	37	2.4	2.6	1.8	5.5	2.0	3.7	2.98 <sup>*</sup>	2.63 <sup>*</sup>
$\chi^2$ test			<b>p&lt;0.001</b>		<b>p&lt;0.001</b>		<b>p&lt;0.001</b>			
Smoking		1496								
Nonsmoker in pregnancy	96.5	1443	97.2	92.4	96.8	94.1	97.5	93.6	Ref	Ref
Smoked during pregnancy	3.5	53	2.8	7.6	3.2	5.9	2.5	6.4	2.67 <sup>*</sup>	2.60 <sup>*</sup>
$\chi^2$ test			<b>p&lt;0.001</b>		<b>p&lt;0.05</b>		<b>p&lt;0.001</b>			
Parity group		1509								
P0	26.2	395	26.1	27.3	24.6	33.4	24.8	29.8	Ref	Ref
P1	27.1	409	27.7	23.8	27.3	25.9	27.7	25.7	NS	NS
P2+	46.7	705	46.2	48.9	48.1	40.7	47.5	44.5	NS	NS



	Total	n	BQ -	BQ +	ESS -	ESS +	Screen -	Screen +	aOR	aOR	aOR
	%		%	%	%	%	%	%	%	%	%
$\chi^2$ test			p=0.48		p<0.01		p=0.13				
<b>Body Mass Index (BMI)</b>											
		1508									
BMI<25	24.5	369	27.9	5.3	Ref	23.3	29.4	Ref	26.6	19.8	Ref
BMI 25-30	35.1	530	38.9	13.2	NA	34.9	34.3	NS	38.7	26.2	NS
BMI>30	40.4	609	33.1	81.5	NA	41.8	36.3	NS	24.6	54.1	2.72 <sup>‡</sup>
$\chi^2$ test			p<0.001		p=0.08		p<0.001				
<b>Gestational age at survey</b>											
		1509									
1st trimester	15.7	237	15.9	15.0	Ref	15.9	17.9	Ref	16.2	16.2	Ref
2nd trimester	45.3	684	46.3	40.1	NS	45.2	43.8	NS	45.6	43.4	NS
3rd trimester	39.0	588	37.8	44.9	NS	38.9	38.3	NS	38.1	40.4	NS
$\chi^2$ test			p=0.12		p=0.69		p=0.69				p=0.69
<b>Pregestational diabetes</b>											
		1182 <sup>¶</sup>									
No	94.5	1117	95.0	91.8	Ref	94.1	94.8	Ref	94.3	94.1	Ref
Yes	5.5	65	5.0	8.2	NS	5.9	5.2	NS	5.7	5.9	NS
$\chi^2$ test			p=0.08		p=0.69		p=0.86				

NA=not applicable, NS=not significant

\* p<0.05,

<sup>†</sup> p<0.01,

<sup>‡</sup> p<0.001

<sup>§</sup> = "Screen -" signifies subjects who screened negative on both the BQ and ESS. "Screen +" signifies subjects who screened positive on either one or the other.

<sup>¶</sup> =Logistic regressions adjusted for maternal age, ethnicity, smoking during pregnancy, parity, gestational age at survey, BMI, and pregestational diabetes (where appropriate).

<sup>||</sup> =Excluded samples with incomplete BQ or ESS, n=75 (in addition to the 5 missing both BQ and ESS measures)

<sup>¶</sup> =Many subjects did not know their diabetes status prior to pregnancy

**Table 2** Multivariate logistic regression of Berlin Questionnaire and Epworth Sleepiness Scale overall and item-by-item analyses.

	BQ scale	Snoring (Ever)	Snore loudly	Snore often	Snoring bothers others	Witnessed apneas	Feel tired after sleep	Feel tired during waking time	Nod off or fall asleep while driving	Often nod off or fall asleep while driving	HTN	BMI > 30
		aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)	aOR (95% CI)
<b>ESS scale</b>	NS											
Doze off reading		NS	2.60* (1.1-6.3)	NS	1.72* (1.01-2.93)	NS	2.36‡ (1.70-3.30)	1.90‡ (1.40-2.58)	3.49‡ (1.75-6.95)	7.19‡ (2.06-25.15)	NS	NS
Doze off watching TV		1.43* (1.02-2.02)	NS	NS	NS	NS	1.92‡ (1.31-2.81)	2.14‡ (1.51-3.04)	3.30‡ (1.36-8.00)	9.49* (1.25-72.33)	NS	NS
Doze off sitting in a public place		1.50* (1.07-2.11)	NS	1.61* (1.05-2.45)	NS	NS	2.10‡ (1.50-2.95)	2.10‡ (1.53-2.88)	4.10‡ (2.18-7.72)	8.12‡ (2.87-23.01)	NS	NS
Doze off as passenger in car		NS	NS	NS	NS	NS	2.37‡ (1.70-3.30)	2.17‡ (1.60-2.95)	2.93‡ (1.55-5.56)	NS	NS	NS
Doze off lying down in afternoon		NS	NS	NS	NS	-	NS	1.74* (1.09-2.77)	NS	NS	NS	1.44* (1.01-2.06)
Doze off talking to someone		NS	NS	NS	NS	NS	1.66‡ (1.17-2.36)	NS	2.12* (1.12-4.00)	2.56* (1.01-6.50)	NS	NS
Doze off after lunch		NS	NS	NS	NS	NS	1.89‡ (1.36-2.63)	2.05‡ (1.51-2.78)	2.30* (1.21-4.38)	3.17* (1.17-8.58)	NS	NS
Doze off in car stopped in traffic		NS	NS	1.66* (1.06-2.59)	NS	NS	1.57* (1.09-2.27)	1.91‡ (1.36-2.70)	6.63‡ (3.50-12.57)	10.14‡ (3.75-27.38)	NS	NS

NA=not applicable, NS=not significant

\* p<0.05,

‡ p<0.01,

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

100  
\* $p < 0.0001$

Logistic regressions adjusted for maternal age, ethnicity, smoking, parity, BMI (where appropriate), hypertension (where appropriate), gestational age at survey, pregestational diabetes

**Table 3**

Association of BQ and ESS items with testing positive on PSG or T3D. By study design, all subjects who completed PSG or T3D screened positive on either the BQ or ESS; all subjects shown here screened positive for OSA by one measure or the other.

	PSG or T3D Negative n= 49 (% shown)	PSG or T3D Positive n=9 (% shown)	p
<b>Berlin Questionnaire</b>			
Snore ever	46.9	44.4	0.89
Snore loudly	4.2	11.1	0.39
Snore often	36.7	33.3	0.84
Snoring bothers others	12.2	33.3	0.11
Witnessed apnea	6.4	0	0.46
Feel tired after sleep	46.9	44.4	0.89
Feel tired while awake	52.1	33.3	0.30
Ever nod off while driving	14.6	11.1	0.78
Often nod off while driving	6.4	0	0.46
Hypertension	32.6	11.1	0.19
BMI>30	47.9	77.8	0.10
<b>Epworth Sleepiness Scale</b>			
Doze off reading	70.8	87.5	0.32
Doze off watching TV	89.6	88.9	0.95
Doze off sitting in public	54.2	44.4	0.59
Doze off as a passenger in a car	67.4	25.0	<b>0.02</b>
Doze off in the afternoon	91.7	88.9	0.79
Doze off talking to someone	50.0	55.6	0.76
Doze off after lunch	77.1	66.7	0.50
Doze off while stopped in traffic	43.8	50.0	0.74

**Table 4**

Detection rate and false referral rate for screen positive population.

	Detection Rate			False referral rate			Total n
	PSG+ % (n/total)	T3D+ % (n/total)	PSG+ or T3D+ % (n/total)	PSG- % (n/total)	T3D- % (n/total)	PSG- or T3D- % (n/total)	
BQ+	100 (1/1) <sup>f</sup>	7.1 (2/28)	10.3 (3/29)	-(0/1) <sup>f</sup>	92.9 (26/28)	89.7 (26/29)	29
ESS+	57.1 (4/7)	9.4 (3/32)	18.0 (7/39)	42.9 (3/7)	90.6 (29/32)	82.1 (32/39)	39
BQ+ or ESS+	62.5 (5/8)	8.0 (4/50)	15.5 (9/58)	37.5 (3/8)	92.0 (46/50)	84.5 (9/58)	58

<sup>f</sup> One woman was positive on BQ and received PSG testing, and she tested positive on PSG