Clinical Practice Guidelines for Diagnostic Testing for Adult Obstructive Sleep Apnea: An Update For 2016

An American Academy of Sleep Medicine Clinical Practice Guideline

1.0 INTRODUCTION

The diagnosis of obstructive sleep apnea (OSA) has previously been addressed in the American Academy of Sleep Medicine (AASM) guidelines including "Practice Parameters for the Indications for Polysomnography and Related Procedures: An Update for 2005" and "Clinical Guidelines for the Use of Unattended Portable Monitors in the Diagnosis of Obstructive Sleep Apnea in Adult Patients (2007)".[1, 2] The AASM commissioned a task force (TF) of content experts to develop an updated clinical practice guideline (CPG) on this topic. The objectives of this CPG are to combine and update information from prior guideline documents regarding diagnosis of OSA including the circumstances under which laboratory polysomnography (PSG) or home sleep apnea testing (HSAT) should optimally be performed.

2.0 BACKGROUND

The term sleep disordered breathing encompasses a range of disorders with most falling into the categories of OSA, central sleep apnea (CSA) or sleep-related hypoventilation. This paper focuses on diagnostic issues related to the diagnosis of OSA, a breathing disorder characterized by narrowing of the upper airway that impairs normal ventilation during sleep. Recent reviews on the evaluation and management of CSA and sleep-related hypoventilation have been published separately by the AASM.[3, 4]

The prevalence of OSA depends significantly on the population studied and how OSA is defined (testing methodology, scoring criteria used, apnea hypopnea index (AHI) threshold). In population-based studies utilizing an AHI cutoff of ≥15 events per hour (hypopneas associated with 4% oxygen desaturations) combined with symptoms to define OSA as a clinical syndrome, the prevalence of OSA has most recently been estimated at 14% of men and 5% of women in population-based data.[5] OSA may impact a larger proportion of the population than indicated by these numbers as the definition of AHI used in this study was restrictive as it did not consider hypopneas that disrupt sleep without oxygen desaturation. In addition, the estimate excludes individuals with an elevated AHI who do not have sleepiness but who may nevertheless be at risk for adverse consequences such as cardiovascular disease.[6-8] In specific populations, the prevalence of OSA is substantially higher, for example, in patients being evaluated for bariatric surgery (range of 70-80%) or who have had a transient ischemic attack or stroke (range of 60-70%).[9] Other disease-specific populations found to have increased rates of OSA include, but are not limited to, patients with coronary artery disease, congestive heart failure, arrhythmias, refractory hypertension, type 2 diabetes, and polycystic ovarian disease.[10, 11]

The consequences of untreated OSA are wide ranging and are postulated to result from the fragmented sleep, intermittent hypoxia and hypercapnea, intrathoracic pressure swings, and increased sympathetic nervous activity that accompanies disordered breathing in sleep. Individuals with OSA often feel unrested, fatigued, and sleepy during the daytime. They may suffer from impairments in vigilance, ability to concentrate, cognitive function, social interactions and quality of life. Declines in daytime function can translate into higher rates of job-related and motor vehicle accidents.[12] Patients with untreated OSA may be at increased risk of developing significant cardiovascular disease, including difficult to control blood pressure and increased risk of coronary artery disease, congestive heart failure, arrhythmias and stroke.[13]

OSA is also associated with metabolic dysregulation, affecting glucose control and risk for diabetes.[14] Undiagnosed and untreated OSA is a significant burden on the healthcare system, with increased healthcare utilization seen in those with untreated OSA,[15] thus, highlighting the importance of early and accurate diagnosis of this common disorder.

Recognizing and treating OSA appears to be important for a number of reasons. The treatment of OSA has been shown to improve quality of life, lower rates of motor vehicle accidents, and reduce the risk of chronic health consequences of untreated OSA as mentioned above.[16] There is also data supporting a decrease in healthcare utilization and cost following the diagnosis and treatment of OSA.[17] However, there are challenges and uncertainties in making the diagnosis and a number of questions remain unanswered.

Individuals with OSA are not immune to other sleep disorders; and they may have these diagnoses as part of, or unrelated to, their OSA. Co-morbid insomnia has been found to be a frequent problem in patients with OSA.[18] There are also instances where undiagnosed OSA may be masquerading as another sleep disorder, such as REM Behavior Disorder.[19] When OSA is suspected, a comprehensive sleep evaluation is important to ensure appropriate diagnostic testing is performed to address OSA, as well as other comorbid sleep complaints. Due to the high prevalence of OSA, there is significant cost associated with sending all patients suspected of having OSA for formal in-lab PSG testing (currently considered the gold standard diagnostic test). Further, there also may be limited accessibility to in-lab testing in some areas. HSAT is an alternative way to diagnose OSA in adults that has limitations but which may be less costly and more efficient in some populations. This guideline document addresses some of these issues through an evidence-based approach.

The diagnosis of OSA involves measuring breathing during sleep. The evolution of measurement techniques and definitions of abnormalities justifies updating the guidelines regarding diagnostic testing, but also complicates the evaluation and summary of evidence gathered from older research studies that have included sleep tests with diverse sensor types and scored respiratory events using different definitions. The third edition of the International Classification of Sleep Disorders (ICSD-3) defines OSA as a PSG-determined obstructive respiratory disturbance index (RDI) $\geq$5 events per hour associated with the typical symptoms of OSA (unrefreshing sleep, daytime sleepiness, fatigue or insomnia, awakening with a gasping or choking sensation, loud snoring, or witnessed apneas), or an obstructive RDI $\geq$15 events per hour (even in the absence of symptoms).[20] In addition to apneas and hypopneas that are included in the AHI, the RDI includes respiratory effort-related arousals (RERAs). The scoring of respiratory events is defined in the AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, version 2.3.[21] However, it should be noted that considerable controversy surrounds the definition of a hypopnea event. The AASM Scoring Manual indicates a preferred definition that requires changes in flow to be associated with a 3% oxygen desaturation or a cortical arousal, but allows an alternative definition that requires association with a 4% oxygen desaturation. Depending on which definition is used, the AHI may be considerably different in a given individual.[22-24] The discrepancy between these and other hypopnea definitions used in research studies introduces complexity in the evaluation of evidence regarding the diagnosis of OSA.

One area of focus was the role of HSAT in the diagnosis of OSA. There are potential disadvantages to HSAT relative to PSG as a diagnostic test related to differences in the physiologic parameters being collected, and the availability of personnel to adjust sensors when needed. With regards to HSAT sensor technology, we refer readers to a classification system has been previously proposed: the SCOPER classification which incorporates Sleep, Cardiovascular, Oximetry, Position, Effort and Respiratory parameters.[25] This categorization convention allows for the inclusion of technologies such as peripheral arterial tonometry (PAT).

Measurement error is inevitable in HSAT versus in-laboratory PSG as standard sleep staging channels are typically not monitored (e.g. EEG, EOG and EMG monitoring are not typically performed) which results in use of the recording time rather than sleep time to define the denominator of the Respiratory Event Index (REI; the term used to represent the frequency of apneas and hypopneas derived from HSAT). HSAT monitors that use conventional sensors are unable to detect hypopneas only associated with cortical arousals which are included in the preferred AHI metric recommended by the AASM scoring manual. Sensor dislodgement and poor quality signal during HSAT are additional contributors to the measurement error of the REI. All these factors can result in the underestimation of the "true" AHI based upon in-laboratory monitoring as well as the need for repeated studies because of inadequate data.

Though the guideline has a diagnostic focus, diagnosis occurs in the context of management of a patient within the healthcare system, and therefore outcomes other than diagnostic accuracy are relevant in the evaluation of management strategies. These include impact on clinical outcomes (e.g., sleepiness, quality of life, morbidity, mortality, adherence to therapy) and efficiency of care (e.g., time to test, time to treatment, costs). Therefore, these important clinical outcomes variables are also included and synthesized in the current document.

Prior AASM guideline documents[1, 2] addressing the diagnosis of OSA included statements that are no longer felt by the TF to be required, thus these statements were not addressed in the current update. Moreover, prior guidelines included consensus statements that have not been specifically evaluated in clinical trials. Following careful qualitative evaluation by the TF, some of these statements that were felt to be of high importance to ensure quality care were adopted in the current guideline as 'good practice' recommendations. The scope of this paper also does not include a comprehensive update of guidelines regarding technical specification for sleep testing. Nevertheless, the paper considers whether currently recommended technology was used in the research studies that were evaluated. In particular, the TF felt that the use of currently AASM recommended flow (nasal pressure transducer and thermistor) and effort sensors (respiratory inductance plethysmography) during PSG and HSAT increased the value of evidence derived from validation studies.[21] As part of the data extraction process, validation studies were classified on the basis of whether the currently recommended respiratory sensors were used for PSG or HSAT.

3.0 METHODS

3.1 Expert Task Force

The AASM commissioned a TF of sleep medicine physicians with expertise in the diagnosis and management of adults with OSA to develop this guideline. These content experts were required to disclose all potential conflicts of interest (COI) according to the AASM's COI policy both prior to being appointed to the TF, and throughout the research and writing of this paper. All relevant conflicts of interest are listed in the Disclosures section.

3.2 PICO Questions

A PICO (**P**atient, **P**opulation or **Pr**oblem, **I**ntervention, **C**omparison, and **O**utcomes) question template was used to develop clinical questions to be addressed in this guideline. The PICO format is an established framework for guiding literature searches targeted at addressing the clinical questions and developing evidence-based clinical practice recommendations. PICO questions were developed based on a review of the existing AASM practice parameters on indications for use of PSG and unattended portable monitoring for the diagnosis of patients with OSA, and a review of systematic reviews, meta-analyses, and guidelines published since 2004. The AASM Board of Directors (BOD) approved the final

list of PICO questions presented in Table 1 before the literature search was performed. To develop the PICO questions, the TF identified the commonly used approaches and devices for the diagnosis of OSA. The TF then developed a list of patient-oriented clinically relevant outcomes that are indicative of whether the diagnostic approach should be recommended for clinical practice. Both diagnostic and clinical outcomes were identified by the TF. The TF rated their relative importance to determine which outcomes are critical for decision-making. A summary of the critical outcomes by PICO is presented in Table 2. A summary of the clinical significance thresholds for the clinical outcome measures is presented in Table 3. The clinical significance thresholds were established based on TF consensus and experience and were used to determine if the changes in outcomes were clinically significant and whether the intervention should be recommended in clinical practice. It should be noted that there was insufficient evidence to directly address PICO question 1, as no studies were identified that compared the efficacy of clinical prediction algorithms to history and physical exam. However, the TF decided to compare the efficacy of clinical prediction algorithms to PSG and/or HSAT.

Table 1 - PICO Questions

1. In adult patients with suspected OSA, do clinical prediction algorithms accurately identify patients with a high pretest probability for OSA compared to history and physical exam? (See Recommendation 4.1)
2. In adult patients with suspected OSA, does HSAT accurately diagnose OSA, improve clinical outcomes and improve efficiency of care compared to attended PSG? (See Recommendation 4.2)
3. In adult patients scheduled for upper airway surgery for snoring or OSA, does attended PSG or HSAT accurately identify patients with OSA and improve clinical outcomes compared to using a history and physical exam or clinical prediction algorithms? (No recommendations, see Future Directions)
4. In adult patients with comorbid conditions (post-stroke, chronic heart failure, chronic obstructive pulmonary disease, opioid use, neuromuscular disease, hypoventilation, insomnia) and suspected OSA, does HSAT accurately diagnose OSA, improve clinical outcomes and efficiency of care compared to attended PSG? (See Recommendation 4.3)
5. In adult patients undergoing attended PSG for suspected OSA, does a split-night study accurately diagnose OSA and improve efficiency of care compared to a full-night study? (See Recommendation 4.4)
6. In adult patients undergoing attended PSG for suspected OSA, do two nights of attended PSG accurately diagnose OSA and improve efficiency of care compared to a single night of attended PSG? (See Recommendation 4.5)
7. In adult patients undergoing HSAT for suspected OSA, is there a minimum number of hours of HSAT that accurately diagnoses OSA and improves efficiency of care? (See Recommendation 4.2)
8. In adult patients undergoing HSAT for suspected OSA, do multiple nights of HSAT accurately diagnose OSA and improve efficiency of care compared to a single night of HSAT? (See Recommendation 4.2)
9. In adult patients with diagnosed OSA, does repeat PSG or HSAT to confirm severity of OSA or efficacy of therapy improve outcomes relative to clinical follow-up without repeat testing? (No recommendations, see Future Directions)

Table 2 – "Critical" Outcomes by PICO

PICO Question	Diagnostic Accuracy*	Subjective Sleepiness	QOL**	CPAP Adherence	AHI	Depression	Cardiovascular Endpoints
1	√						
2	FN only	√	√	√	√	√	√
3	√	√	√				√
4	√						
5	√	√	√	√	√	√	√
6	√	√	√	√	√	√	√
7	√						
8	√						
9		√	√	√		√	√

* Diagnostic Accuracy is determined by the number of True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FM) diagnoses

**Based on SAQLI and FOSQ measures of QOL. SF-36 measure of QOL was determined to be important but not critical for decision-making based on TF consensus.

Table 3 – Summary of Clinical Significance Thresholds for Clinical Outcome Measures

Outcome Measure	Clinical Significance Threshold
Epworth Sleepiness Score (ESS)	2 points
Functional Outcomes of Sleep Questionnaire (FOSQ)	1 point
Sleep Apnea QOL Index (SAQLI)	1 point
CPAP Adherence (hrs/night)	0.5 hrs/night
CPAP Adherence (% nights >4 hrs)	10%
SF-36 (Vitality Score)	14 points
SF-36 (Physical Component Summary Score)	3 points
SF-36 (Mental Component Summary Score)	3 points

3.3 Literature Searches, Evidence Review and Data Extraction

The TF performed an extensive review of the scientific literature to retrieve articles which addressed at least one of the nine PICO questions. Multiple literature searches were performed by the AASM research staff using the PubMed and Embase databases throughout the guideline development process (see Figure 1). The search yielded articles with various study designs however the analysis was limited to randomized controlled trials (RCTs) and observational studies. The articles that were cited in the 2007 AASM clinical practice guideline,[2] 2005 practice parameter,[1] 2003 review,[26] and 1997 review[27] were included for data analysis if they met the study inclusion criteria described below.

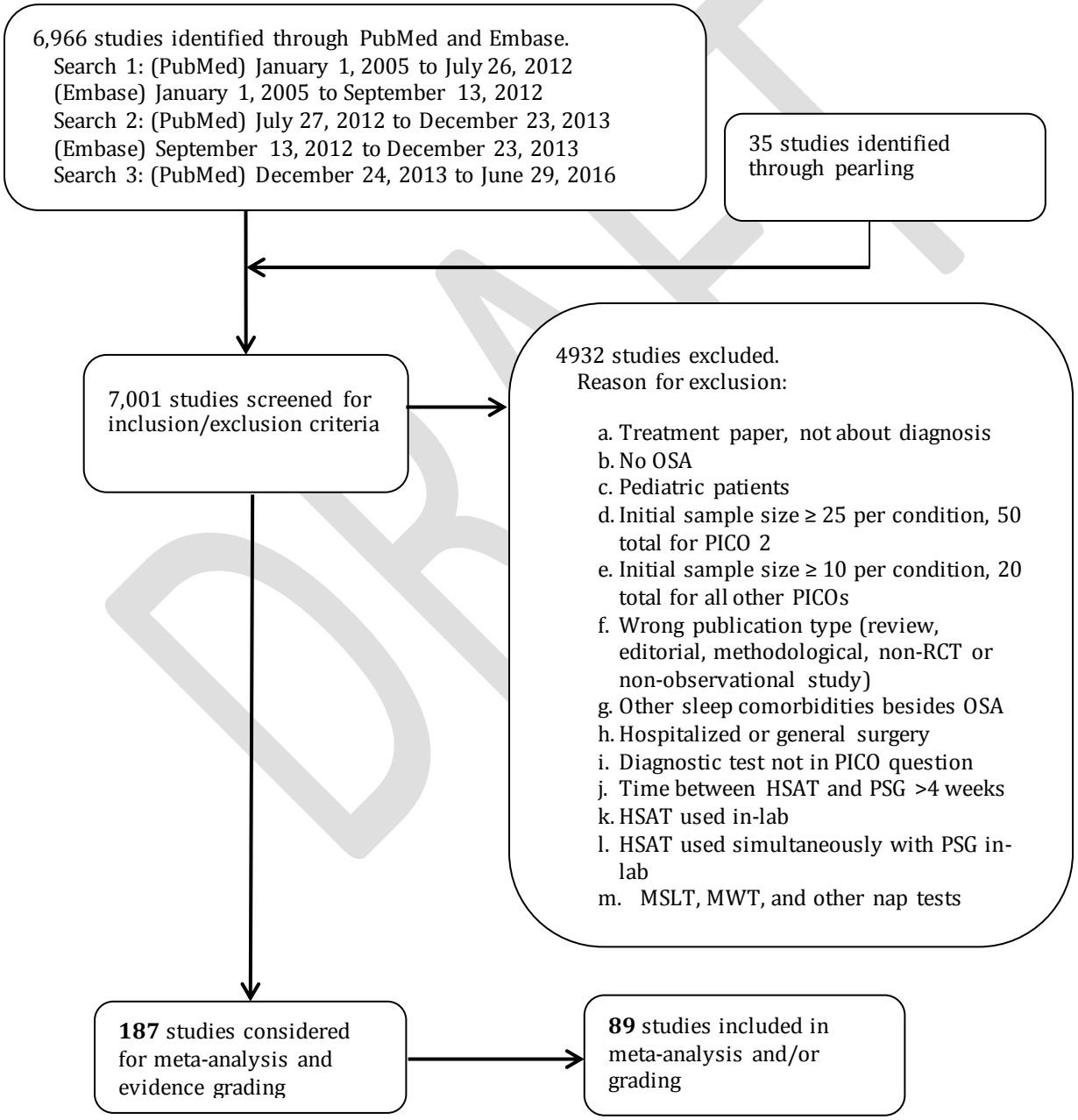
The literature search in PubMed and Embase was conducted using a combination of MeSH terms and keywords as presented in the Supplemental Materials. The PubMed database was searched from January 1, 2005 through July 26, 2012 for any relevant literature published since the last guideline. This search was updated again on December 9, 2013 to capture the latest literature. A total of 3,730 citations were identified in PubMed and supplemented by pearling (i.e., checking the reference sections of search results for articles otherwise missed) which identified an additional 35 citations.

The literature search in Embase was performed using a combination of terms and keywords as presented in the Supplemental Materials. The Embase database was searched from January 1, 2005 through September 13, 2012. This search was updated on December 9, 2013 to capture the latest literature and cross-checked with the results from the PubMed search. A total of 1,300 citations were identified in Embase.

A final search was performed using PubMed on June 29, 2016 to update the guideline prior to submission to the AASM BOD for review and approval for public comment. The final search identified an additional 1,936 articles for inclusion/exclusion assessment. The literature searches yielded a total of 5,065 citations from both databases.

Each abstract was assessed by 2 reviewers to determine whether they met inclusion criteria presented in the Supplemental Materials. Articles were excluded per the criteria listed in the Supplemental Materials and Figure 1. A total of 187 articles met these criteria and were considered for data extraction, meta-analysis, and grading. A total of 89 studies were included in meta-analysis and/or grading.

Figure 1 – Evidence Base Flow Diagram



3.4 Meta-Analysis

Meta-analysis was performed on both diagnostic and clinical outcomes of interest for each PICO question when possible. Diagnostic approaches were categorized into the following: clinical prediction algorithms; history and physical exam; HSAT; attended PSG; split-night attended PSG; two-night attended PSG; single-night HSAT; multiple-night HSAT; follow-up attended PSG; and follow-up HSAT. Adult patients were categorized into the following: suspected OSA; suspected OSA with comorbid conditions; diagnosed OSA; and scheduled for upper airway surgery.

For diagnostic outcomes of interest, data on the pre-test probability for OSA, sensitivity and specificity of the tested diagnostic approach, and number of patients for each study included in the analysis was entered into statistical software to derive pooled sensitivity and specificity estimates for each OSA severity threshold (i.e., AHI ≥5, AHI ≥15, AHI ≥30). For analyses that included more than 5 studies, hierarchical modeling of the sensitivity and specificity data was first performed using STATA software to derive the pooled estimates. The pooled estimates of sensitivity and specificity were entered into Review Manager 5.3 software to generate a forest plot with the sensitivity and specificity represented by a mean value and standard deviation. Pooled estimates of sensitivity and specificity, along with the TF estimates of OSA prevalence, were entered into the GRADE Guideline Development Tool (GDT) to estimate the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) diagnoses per 1,000 patients for both high risk and low risk patients. The TF determined the downstream consequences of an accurate versus inaccurate diagnosis (see Supplemental Materials, Table S1) and used the above estimates to weigh the benefits of an accurate diagnosis versus the harms of an inaccurate diagnosis. This information was used, in part, to assess whether a given diagnostic approach could be recommended when compared against PSG. For analyses of diagnostic outcomes that included less than 5 studies, pooled estimates of sensitivity and specificity could not be derived using hierarchical modeling. For such analyses, the range of sensitivity and specificity across the studies was entered directly into the GDT to estimate the number of TP, TN, FP, and FN per 1,000 patients for both high risk and low risk patients.

For clinical outcomes of interest (e.g., subjective sleepiness, QOL, CPAP adherence etc.), data on change scores before and after treatment were entered into the Review Manager 5.3 software to derive the mean difference and standard deviation between the experimental diagnostic approach and the gold standard or comparator. For studies that did not report change scores, data from post-treatment values taken from the last treatment time-point were used for meta-analysis. All meta-analyses of clinical outcomes were performed using the random effects model with results displayed as a forest plot. Interpretation of clinical significance for the clinical outcomes of interest was conducted by comparing the absolute effects of each diagnostic approach to the Clinical Significance Threshold previously determined by the TF for each clinical outcome of interest (see Table 3).

There was insufficient evidence to perform meta-analyses for PICOs 3 and 9, thus no recommendations are provided.

3.5 Strength of Recommendations

The assessment of evidence quality was performed according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) process.[28] The TF assessed the following four components to determine the direction and strength of a recommendation: quality of evidence, balance of beneficial and harmful effects, patient values and preferences and resource use as described below.

1. Quality of evidence – based on an assessment of the overall risk of bias (randomization, blinding, allocation concealment, selective reporting, and author disclosures), imprecision (clinical significance thresholds), inconsistency (I^2 cutoff of 75%), indirectness (study population), and risk of publication bias (funding sources), the TF determined their overall confidence that the

estimated effect found in the body of evidence was representative of the true treatment effect that patients would see. For diagnostic accuracy studies, the QUADAS-2 tool was used in addition to the quality domains for the assessment of risk of bias in intervention studies. The QUADAS-2 tool was used to consider specific study characteristics for additional risk of bias and applicability related to patient selection, index test (test or device whose performance is being evaluated), reference standard, and flow and timing of patients and tests through the study. The quality of evidence was based on the quality of studies that reported on outcomes that the TF deemed critical for decision-making.

2. Benefits vs. Harms – based on the meta-analysis (if applicable), analysis of any harms/side effects reported within the accepted literature, and the clinical expertise of the TF, the TF determined if the beneficial outcomes of the intervention outweighed any harmful side effects.
3. Patient values and preferences – based on the clinical expertise of the TF members and any data published on the topic relevant to patient preferences, the TF determined if patient values and preferences would be generally consistent, and if patients would use the intervention based on the body of evidence.
4. Resource use – based on the clinical expertise of the TF members. The TF judged resource use to be important for determining whether to recommend the use of HSAT for diagnosis of OSA. Resource use was not considered for other diagnostic approaches such as clinical prediction tools and questionnaires.

Taking these major factors into consideration, each recommendation statement was assigned a direction (For or Against) and strength (Strong or Weak). Additional information is provided in the form of "Remarks" immediately following the recommendation statements, when deemed necessary by the TF. Remarks are based on the evidence evaluated during the systematic review, are intended to provide context for the recommendations and are intended to guide clinicians in implementing the recommendations in daily practice.

Discussions accompany each recommendation to summarize the relevant evidence and explain the rationale leading to each recommendation. These sections are an integral part of the GRADE system and offer transparency to the process.

3.6 Approval and Interpretation of Recommendations

A draft of the guideline was available for public comment for a two-week period on the AASM website. The TF took into consideration all the comments received and made decisions about whether or not to revise the draft based on the comments. The revised guideline was submitted to the AASM BOD who subsequently approved these recommendations.

The recommendations in this guideline define principles of practice that should meet the needs of most patients in most situations. This guideline should not, however, be considered inclusive of all proper methods of care or exclusive of other methods of care reasonably used to obtain the same results. The ultimate judgment regarding the suitability of any specific recommendation must be made by the clinician, in light of the individual circumstances presented by the patient, the available diagnostic tools, the accessible treatment options, and available resources.

The AASM expects this guideline to have an impact on professional behavior, patient outcomes, and possibly, health care costs. This clinical practice guideline reflects the state of knowledge at the time of the literature review and will be reexamined and updated as new information becomes available.

4.0 CLINICAL PRACTICE RECOMMENDATIONS

The following clinical practice recommendations are based on a systematic review and evaluation of evidence following the GRADE methodology. Remarks are provided to guide clinicians in the implementation of these recommendations. All figures, including meta-analyses and Summary of Findings tables are presented in the Supplemental Materials. Table 4 shows a summary of the recommendation statements including the strength of recommendation and quality of evidence. A decision tree for the diagnosis of patients suspected of having OSA is presented in Figure 2.

The subsequent clinical practice recommendations are good practice statements, the implementation of which is deemed necessary for appropriate and effective diagnosis and management of obstructive sleep apnea:

Diagnostic testing for OSA should be performed in conjunction with a comprehensive sleep evaluation and adequate follow-up.

OSA is one of many medical conditions that can cause sleep symptoms, therefore diagnostic testing for OSA is best carried out after comprehensive sleep evaluation. The clinical evaluation for OSA should include a thorough sleep history and a physical examination that includes the respiratory, cardiovascular, and neurologic systems. Although the examiner may pay particular attention to observations regarding snoring, apneas, nocturnal choking or gasping, restlessness, and excessive sleepiness, other aspects of a sleep history should also be elicited as many patients suffer from more than one sleep disorder or present with atypical sleep symptoms. In addition, medical conditions associated with increased risk for OSA, such as obesity, hypertension, stroke, and congestive heart failure should be identified. The general evaluation should serve to establish a differential diagnosis, which can then be used to select the appropriate test(s). Follow up under the supervision of board certified sleep medicine specialist insures that study findings and recommendations are relayed appropriately; and that appropriate expertise in prescribing and/or administering therapy is available to the patient.

The Task force recognizes that there may be specific contexts (e.g., preoperative evaluation of OSA) in which evaluation of OSA needs to occur in an expedited manner when it may not be practical to perform a comprehensive sleep evaluation prior to diagnostic testing. In such situations, the Task Force recommends a clinical pathway be developed and administered by a board certified sleep medicine specialist or appropriately licensed medical staff member designated by the board certified sleep medicine specialist that includes the following elements: a focused sleep apnea evaluation by a clinical provider and/or use of tools/questionnaires that capture clinically important information that is reviewed by a board certified sleep medicine specialist prior to testing; and availability of comprehensive sleep evaluation and follow up under the supervision of a board certified sleep medicine specialist following testing.

Attended polysomnography is the gold standard diagnostic test for the diagnosis of obstructive sleep apnea.

Misdiagnosing patients can lead to significant harm due to loss of benefits of therapy in those with OSA and the prescription of inappropriate therapy in those without OSA. Sleep apnea-focused questionnaires and clinical prediction rules lack sufficient accuracy and therefore direct measurement of sleep disordered breathing is necessary to establish a diagnosis of OSA. PSG is widely accepted as the gold standard test for diagnosis of OSA. Further, this test has typically been used as the gold standard for comparison of

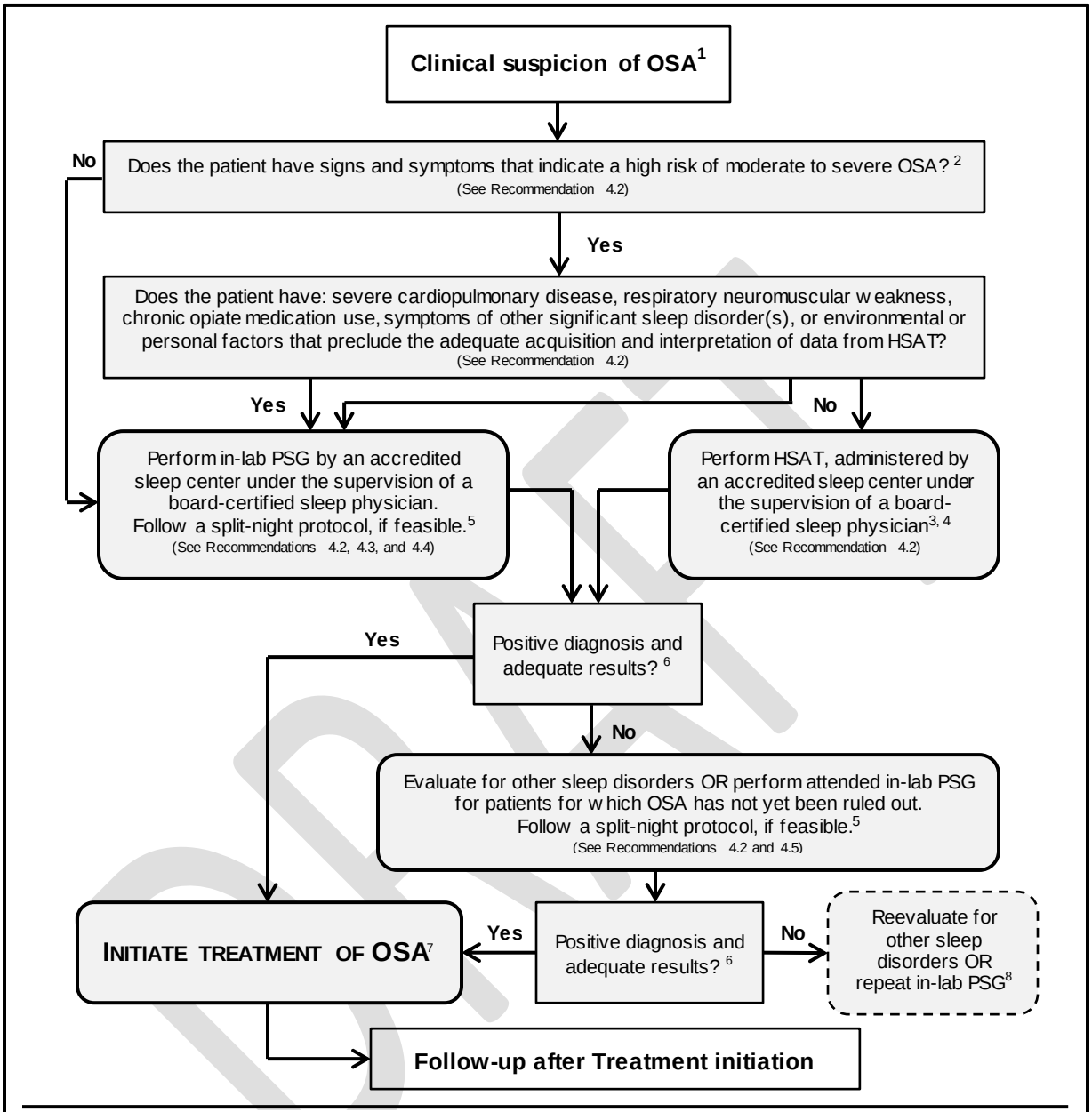
other diagnostic tests such as HSAT, which have lower accuracy. Besides the diagnosis of OSA, attended PSG can identify co-existing sleep disorders including other forms of sleep disordered breathing. In some cases, and within the appropriate context, the use of HSAT as the initial sleep study may be acceptable, though PSG should be used when HSAT results do not provide satisfactory post-test probability of confirming or ruling out OSA.

Table 4. Summary of Recommendation Statements

Recommendation Statement	Strength of Recommendation	Evidence Quality	Benefits vs. Harms	Patient Values & Preferences
4.1a We recommend that clinical tools/questionnaires not be used to diagnose OSA in the absence of objective sleep testing.	Strong against	Moderate	High certainty that harms outweigh benefits	Vast majority of well-informed patients would most likely not choose clinical tools/questionnaires for accurate diagnosis
4.2ai We recommend that attended PSG or HSAT with a technically adequate device be used to diagnose OSA in uncomplicated adult patients presenting with signs and symptoms that indicate a high risk of moderate to severe OSA.	Strong for	Moderate	High certainty that benefits outweigh harms	Vast majority of well-informed, high risk uncomplicated patients would want option of PSG or HSAT to diagnose suspected OSA
4.2aii We recommend that if a single HSAT is negative or inconclusive that an attended in-lab PSG be performed for the diagnosis of OSA in symptomatic patients.	Strong for		High certainty that benefits outweigh harms	Vast majority of well-informed symptomatic patients would want PSG performed if the initial HSAT is negative
4.2aiii We recommend that if a single HSAT is technically inadequate that an attended in-lab PSG be performed for the diagnosis of OSA in symptomatic patients.	Weak for	Low	Low certainty that benefits outweigh harms	Majority of well-informed symptomatic patients would want PSG performed if the initial HSAT is technically inadequate
4.3a We recommend that PSG, rather than HSAT, be used for the diagnosis of OSA in patients with severe cardiorespiratory disease, neuromuscular disease with respiratory muscle impairment, awake hypoventilation or high risk of sleep related hypoventilation, chronic opioid medication use, or severe insomnia.	Strong for	Very Low	High certainty that benefits outweigh harms	Vast majority of well-informed patients with significant comorbidities would most likely choose PSG to diagnose suspected OSA

4.4a We suggest that if clinically appropriate, a split-night diagnostic protocol be used, rather than a full-night diagnostic protocol for attended PSG for the diagnosis of OSA.	Weak for	Low	Low certainty that benefits outweigh harms	Majority of well-informed patients would most likely choose a split-night diagnostic protocol to diagnose suspected OSA
4.5a We suggest that when the initial attended PSG is negative and there is still clinical suspicion for OSA, a second attended PSG be considered for the diagnosis of OSA in symptomatic patients.	Weak for	Very low	Low certainty that benefits outweigh harms	Majority of well-informed symptomatic patients would most likely choose a second attended PSG to diagnose suspected OSA when the initial PSG is negative and there is still a suspicion that OSA is present

Figure 2 – Flow-chart for implementation of clinical practice guidelines



1. Clinical suspicion based on a comprehensive sleep evaluation
2. High risk of moderate to severe OSA is indicated by the presence of daytime sleepiness and two or more of the following: diagnosed hypertension, habitual loud snoring, and witnessed apnea, gasping or choking.
3. This recommendation is based on conducting: one or more consecutive nights of HSAT recording.
4. This recommendation is based on HSAT devices that incorporate a minimum of the following sensors: nasal pressure, chest and abdominal respiratory inductance plethysmography (RIP) and oximetry; or peripheral arterial tonometry (PAT) with oximetry and actigraphy. For additional information, refer to the AASM Manual for the Scoring of Sleep and Associated Events.
5. This recommendation is based on a split-night protocol conducted only when the following criteria are met: 1) A moderate to severe degree of OSA is observed during a minimum of 2 hours of diagnostic PSG; AND 2) At least 3 hours are available to complete CPAP titration. If these criteria are not met, a full-night diagnostic protocol should be followed.
6. Technically adequate results include a minimum of 4 hours of oximetry and flow data, obtained during a recording attempt that encompasses the habitual sleep period. For additional information, refer to the AASM Manual for the Scoring of Sleep and Associated Events.
7. Treatment of OSA should be initiated based on technically adequate in-lab PSG or HSAT results which lead to PAP therapy initiation that includes titration using APAP or attended PSG, early follow up after initiation of therapy, and PSG titration studies for failed APAP therapy.
8. Repeat in-lab PSG when the results of the repeated PSG are not technically adequate.

4.1 Clinical tools/questionnaires for the diagnosis of OSA in adults

4.1a We recommend that clinical tools/questionnaires not be used to diagnose OSA in the absence of objective sleep testing. (STRONG AGAINST)

4.1b *Summary*

There were no studies that evaluated whether the use of clinical tools/questionnaires or prediction rules improve a sleep clinician's ability to decide whether a sleep study is warranted beyond a typical sleep focused history and physical exam. However, there were several studies that compared the ability of clinical tools/questionnaires and clinical prediction algorithms to predict high pre-test probability of obstructive sleep apnea to PSG and HSAT. The overall quality of evidence was downgraded to moderate due to inconsistency and/or imprecision of findings.

In the clinic-based setting, clinical tools/questionnaires and prediction rules have a low level of accuracy for the diagnosis of OSA at any threshold of AHI consideration.

Clinician prediction algorithms may be used in sleep clinic patients with suspected OSA but are not necessary to establish the need for objective sleep testing. These tools may be more helpful to identify patients who are at high risk for OSA in non-sleep clinic settings but this is a question that was not addressed by our evidence review.

Evaluation with a clinical tool/questionnaire or prediction rule may be less burdensome to the patient with possible OSA than HSAT or PSG. However, this must be balanced with the enhanced diagnostic accuracy of the PSG or HSAT. On the other hand, PSG and HSAT, while involving more resources, greater cost and potentially more burden to the patient, provide greater value in terms of the more accurate diagnosis of OSA compared to OSA-targeted questionnaires or clinical prediction rules. Based on its clinical judgement, the task force determined that the vast majority of patients would not favor the use of clinical questionnaires or prediction tools alone for the diagnosis of OSA.

4.1c *Discussion*

The literature search did not identify publications which directly compared the performance of clinical prediction algorithms relative to history and physical exam to accurately identify patients with a high pretest probability for OSA. Validation studies as described below were identified involving comparison of clinical prediction models to sleep study results. The data from these validation studies are summarized in the Supplemental Materials, Tables S2-S38. Because of uncertainty regarding clinical outcomes for patients misclassified by the prediction rules, the TF was unable to establish a cut-off for number of misclassified patients that would be acceptable. Nevertheless, all the clinical prediction models evaluated resulted in upper ranges of predicted false negatives per 1000 patients that exceeded 100, a number that was felt by the TF to be clearly excessive for stand-alone diagnostic test for OSA. The TF was not able to establish a value for false negative. The following discussion has been organized to review the data based upon OSA questionnaire or clinical prediction rule type, as it compares to objective testing.

BERLIN QUESTIONNAIRE:

Twenty two studies compared the ability of the Berlin Questionnaire to predict OSA compared to attended PSG.[29-50] The Berlin questionnaire involves 11 questions divided into 3 categories to classify the patient as high or low risk for OSA.[51] The studies were characterized by a wide geographic variation including Brazil,[33] Canada,[29, 37] Greece,[32] Iran,[31] Korea,[35] Turkey,[38] and the United States[36, 39, 40] and various patient populations considered including primary care clinics, Veteran population, sleep clinics and studies focused on patients with cardiac disease. The patients included in these studies were suspected

to have OSA and involved mostly men (50% or greater in the majority of studies), were overweight/obese, and were middle aged. Overall, the Berlin Questionnaire had a large number of false negative results compared to an attended PSG sleep study, thereby limiting its utility as an instrument to diagnose patients with OSA. Specifically, when assessing the ability of the Berlin Questionnaire to identify subjects with an AHI cutoff of ≥5, the estimate for sensitivity ranged from 0.25-0.95, while the specificity ranged from 0.06-0.86 (See Supplemental Materials, Table S2). The result was an unacceptably high number of false negative results (range of 43 to 652; assuming a prevalence of 87%). Lower specificity was observed with increasing cutoff of AHI category considered. Furthermore, the questionnaire had suboptimal accuracy, ranging from 50-77%; levels which became progressively more compromised with consideration of higher OSA severity thresholds (See Supplemental Materials, Tables S3S4).

Five studies evaluating the performance of the Berlin Questionnaire in comparison to HSAT were identified .[52-56] When considering an AHI cutoff of ≥15, the pooled estimate for sensitivity was 0.74 (95% CI: 0.64 to 0.82), for specificity was 0.44 (95% CI: 0.32 to 0.58) and for accuracy was 63%, and were comparable to attended PSG results. However, when considering an AHI cutoff of ≥5, there were higher numbers of false negative results (range of 357 to 679 per 1000 patients; assuming a prevalence of 87%) (See Supplemental Materials, Tables S5-S8). The overall quality of evidence for the use of the Berlin questionnaire was downgraded to moderate due to either heterogeneity, indirectness, or imprecision in meta-analyses that included different AHI cut-offs.

EPWORTH SLEEPINESS SCALE:

The Epworth Sleepiness Scale (ESS) is a self-reported questionnaire involving 8 questions to assess the propensity for daytime sleepiness or dozing.[57] Eight studies comparing ESS scores relative to attended PSG were performed in China and other countries including Brazil, Croatia, Turkey and the United States thus reflecting a wide geographic sampling.[33, 38, 40, 48, 49, 58-60] Participants were those suspected of OSA and included mainly male, middle-aged and overweight or obese individuals. The overall results indicate that the ESS had a large number of false negative results compared to PSG, limiting its utility for the diagnosis of OSA. When considering an AHI of ≥5, the characteristics of the ESS relative to attended PSG revealed a range of sensitivity of 0.26-0.94 and specificity of 0.50-0.89. (See Supplemental Materials, Table S9) The ESS demonstrated an accuracy ranging from 47% to 74% for the AHI ≥5 cut-off. The ESS had a high number of false negative results (range of 52 to 644 per 1000 patients; assuming a prevalence of 87%) When considering a higher level cutoff of AHI ≥15, the number of false negative results increased and ranged from 269 to 506 per 1000 patients. (See Supplemental Materials, Table S10) Findings from one study comparing ESS relative to HSAT showed low sensitivity of 0.36 (95% CI: 0.19 to 0.57) and high specificity of 0.77 (95% CI: 0.66 to 0.86).[56] (See Supplemental Materials, Table S12) The overall quality of evidence for the use of the ESS was downgraded to low due to heterogeneity, indirectness, or imprecision in meta-analyses that included different AHI cut-offs.

STOP-BANG QUESTIONNAIRE

Nine studies involving primarily middle-aged, obese males suspected of OSA compared the STOP-Bang questionnaire to attended PSG.[37, 47, 48, 50, 61-65] The STOP-Bang questionnaire is an OSA screening tool consisting of 4 yes/no questions and 4 physical attributes.[66] Overall the findings reveal that the STOP-Bang questionnaire had high sensitivity, but low specificity for the detection of OSA; findings which became more pronounced with higher levels of OSA category cutoffs considered. The number of false negative studies results compared to the sleep study limits its use as an instrument to diagnose patients with OSA. Specifically, when considering an AHI ≥5, the sensitivity in the studies was high (range of 0.85-0.98), but specificity was lower (range of 0.19-0.48) with a range of accuracy of 79 to 90% in high-risk patients. The number of false negatives was lower than what was found in studies using the Berlin and ESS questionnaires (range of 17 to 131 per 1000 patients; assuming a prevalence of 87%) when

compared with attended PSG (See Supplemental Materials, Figure S13). The sensitivity further improved and specificity was further compromised with progressively higher level of AHI cutoffs considered (See Supplemental Materials, Tables S14,S15). In general, these findings were very similar when examining the data of STOP-Bang relative to HSAT[54, 67] (See Supplemental Materials, Tables S16-S18) or compared to either test[54, 66, 67] (See Supplemental Materials, Tables S19-S21). The overall quality of evidence for the use of the STOP-Bang questionnaire was downgraded to moderate due to either indirectness or imprecision in meta-analyses that included different AHI cut-offs.

STOP QUESTIONNAIRE:

STOP questionnaire relative to attended PSG data showed a relatively moderate to high sensitivity, fair specificity, and low accuracy[47-49, 60, 65, 68] (See Supplemental Materials, Tables S22-S24). When considering an AHI ≥5, the sensitivity ranged from 0.65 to 0.96, the specificity ranged from 0.60-0.80, and the accuracy for high risk OSA ranged from 64-94%. The number of false negatives was relatively high (range of 35-304 per 1000 patients; assuming a prevalence of 87%) when compared to attended PSG (See Supplemental Materials, Table S22). When considering an AHI cutoff of ≥15, the sensitivity ranged from 0.62-0.98, the specificity ranged from 0.10-0.63, and the accuracy for high risk OSA ranged from 60-79%), with the number of false negatives ranging from 13 to 243 per 1000 patients; assuming a prevalence of 64%) (See Supplemental Materials, Table S23). The quality of evidence for the use of the STOP questionnaire was downgraded to low due to heterogeneity and imprecision in meta-analyses that included different AHI cut-offs.

MORPHOMETRIC MODELS:

Morphometric models have been developed to predict OSA based upon sleep study data. For example, in a group of hypertensive patients, a multivariable apnea prediction score which combined symptoms, body mass index, age and sex was used to assess OSA risk.[69] In another study involving primarily middle-aged males, those with OSA were compared to those without OSA by using a morphometric clinical prediction formula incorporating measures of craniofacial anatomy (e.g. palatal height, maxillary and mandibular intermolar distances).[70] While these studies support relatively high sensitivity (range of 0.88-0.98) to predict AHI ≥5, the specificity was quite low (range of 0.11-0.31). (See Supplemental Materials, Table S25) Similar findings have been observed considering adjusted neck circumference in both the hypertensive and chronic kidney disease populations in terms of relatively high sensitivity, but poor specificity with improvement in specificity with higher thresholds of OSA.[54, 69] (See Supplemental Materials, Tables S26, S27) The overall quality of evidence for the use of morphometric models was downgraded to low due to heterogeneity and imprecision in meta-analyses that included different AHI cut-offs.

MULTIVARIABLE APNEA PREDICTION QUESTIONNAIRE:

The Multivariable Apnea Prediction questionnaire has been assessed relative to PSG in those with suspected OSA,[30, 71-73] a sample of hypertensive patients,[69] and also a sample of older adults[74] with lower levels of specificity and high false positive values. (See Supplemental Materials, Tables S28, S29). The overall quality of evidence for the use of the multivariable apnea prediction questionnaire was downgraded to moderate due to imprecision.

CLINICAL PREDICTION MODELS:

Four studies have evaluated clinical prediction models relative to PSG in terms of OSA diagnosis[59, 75-77] and 4 have evaluated these models against HSAT.[74, 78-80] Two of the studies evaluated respiratory parameters (i.e., a study involving a Chinese cohort that evaluated snoring while sitting and another single study assessing respiratory conductance and oximetry relative to PSG demonstrated a sensitivity ranging from 0.59-1.00), and a specificity ranging from 0.64-0.84).[75, 77] Other studies assessing clinical prediction rules including age, waist circumference, ESS score and minimum oxygen saturation, and

another evaluating gender, nocturnal choking, snoring and body mass index relative to PSG show reasonably high sensitivity (range of 0.72-0.94) and specificity (range of 0.75-0.91) considering different AHI thresholds.[59, 76] Clinical prediction rules evaluated relative to HSAT have been studied in select populations, i.e. the elderly,[74] bariatric surgery candidates,[78] and commercial drivers.[79] (See Supplemental Materials, Table S30 -S32)

The overall quality of evidence for the use of clinical prediction models was downgraded to moderate due to either heterogeneity or imprecision across meta-analyses at different AHI cut-offs.

OTHER OSA PREDICTION TOOLS:

Other OSA prediction tools include the OSA50 (see Supplemental Materials, Tables S33,S34), the clinical decision support system (see Supplemental Materials, Table S35), the OSAS score (see Supplemental Materials, Table S36), and the Kushida Index (see Supplemental Materials, Table S37). The OSA50 questionnaire involves four components including age ≥50, snoring, witnessed apneas and waist circumference.[81] A study involving Turkish bus drivers[82] and a validation study for the OSA50 in the primary care setting[81] showed a sensitivity ranging from 0.63-0.82 and a specificity of 0.82 in both studies. (See Supplemental Materials, Tables S33, S34) A hand-held clinical decision support system (assessing sleep behavior, breathing during sleep and daytime functioning) relative to PSG studied in veterans with ischemic heart disease showed a high sensitivity of 0.98 (95% CI: 0.92 to 1.00) and a high specificity of 0.87 (95% CI: 0.66 to 0.97).[36] (See Supplemental Materials, Table S35) The OSAHS score involves assessment of the Friedman tongue position, tonsil size, and body mass index. In a sample of individuals suspected to have OSA, the sensitivity of the OSAHS score was 0.86 (95% CI: 0.80 to 0.91) relative to PSG in terms of ascertaining AHI >5, however; specificity was lower at 0.47 (95% CI: 0.34 to 0.56) with a high number of false positives.[34] (See Supplemental Materials, Table S36) The Kushida Index showed a high sensitivity of 0.98 (95% CI: 0.95 to 0.99) and high specificity of 1.00 (95% CI:0.92 to 1.00) to detect AHI ≥5 from PSG in a single study. (See Supplemental Materials, Table S37) The quality of the evidence was downgraded to moderate due to either heterogeneity or imprecision.

In summary, morphometric models, clinical prediction rules considering a variety of variables including symptoms, exam findings and subject characteristics have been examined in terms of ability to predict OSA relative to PSG and HSAT. Some of these studies were conducted in focused populations (e.g. commercial drivers, elderly, bariatric surgery patients, etc.) thus limiting interpretation in terms of generalizability. Overall, while sensitivity appears to be in the higher range, specificity tends to be lower with a higher number of false positives thereby limiting utility of these clinical/morphometric rules and models in the diagnosis of OSA.

OVERALL QUALITY OF EVIDENCE:

The quality of evidence for specific clinical prediction algorithms and models ranged from very low to high after being downgraded due to imprecision, indirectness, and heterogeneity. However due to the heterogeneity of the questionnaires and prediction tools described above, and the low likelihood that future research would result in change of the accuracy of these questionnaires and tools, the task force felt that the overall quality of evidence for the recommendation against using clinical questionnaires or predictive tools alone was moderate.

BENEFITS VS HARMS

Given the downstream effects of misclassification in terms of high false negatives for most of these clinical tools/questionnaires, there is the possibility of not capturing the diagnosis of OSA when indeed the diagnosis exists. As a result, this would translate into high levels of OSA-related decrements in quality of life, morbidity, and mortality due to undiagnosed and untreated OSA. On the other hand, false positive results from sleep apnea clinical questionnaires or prediction tools would result in unnecessary testing

and/or treatment for sleep apnea which is incorrectly identified. Therefore, the TF determined that the potential harms of using clinical questionnaires of prediction tools alone to diagnose OSA outweigh the potential benefits.

PATIENTS' VALUES AND PREFERENCES

Evaluation with a clinical tool/questionnaire or prediction rule may be less burdensome to the patient with possible OSA than HSAT or PSG. However, this must be balanced with the enhanced diagnostic accuracy of the PSG or HSAT. On the other hand, PSG and HSAT, while involving more resources, greater cost and potentially more burden to the patient, provide greater value in terms of the more accurate diagnosis of OSA compared to OSA-targeted questionnaires or clinical prediction rules. Based on its clinical judgement, the TF determined that the vast majority of patients would not favor the use of clinical questionnaires or prediction tools alone for the diagnosis of OSA.

4.2 HSAT for the diagnosis of OSA in adults

4.2ai We recommend that attended polysomnography or home sleep apnea testing with a technically adequate device, be used to diagnose OSA in uncomplicated adult patients presenting with signs and symptoms that indicate a high risk of moderate to severe OSA. (STRONG FOR)

4.2aii We recommend that if a single technically adequate home sleep apnea test is negative or inconclusive that an attended in-lab polysomnogram be performed for the diagnosis of OSA in symptomatic patients. (STRONG FOR)

4.2aiii We recommend that if a single home sleep apnea test is technically inadequate, that an attended in-lab polysomnogram be performed for the diagnosis of OSA in symptomatic patients. (WEAK FOR)

Remarks: The following remarks are based on specifications used by studies that support these recommendation statements:

An uncomplicated patient is defined by the absence of severe cardiopulmonary disease, respiratory neuromuscular weakness, chronic opiate medication use, concern for other significant sleep disorder(s), and environmental or personal factors that preclude the adequate acquisition and interpretation of data from HSAT.

HSAT is administered by an accredited sleep center under the supervision of a board certified sleep medicine physician or clinician with equivalent training.

High risk of moderate to severe OSA is indicated by the presence of daytime sleepiness and two or more of the following: diagnosed hypertension, habitual loud snoring, witnessed apnea, and/or gasping or choking at night.

Single HSAT recording is conducted over at least one night.

A technically adequate HSAT device incorporates a minimum of the following sensors: nasal pressure, chest and abdominal respiratory inductance plethysmography (RIP) and oximetry; or peripheral arterial tonometry (PAT) with oximetry and actigraphy. For additional information regarding HSAT sensor requirements, refer to the AASM Manual for the Scoring of Sleep and Associated Events.[21]

A technically adequate study includes a minimum of 4 hours of technically adequate oximetry and flow HSAT data, obtained during a recording attempt that encompasses the habitual sleep period.

In patients diagnosed with OSA using HSAT, PAP therapy can be initiated with APAP (with attended PSG titration used for patients failing APAP therapy) or attended PSG titration. There should be early follow up after initiation of therapy.

4.2b *Summary*

In RCTs, PSG in comparison to HSAT has failed to show superior outcomes for the patient population specified above when HSAT is performed using the specifications listed in the remarks section of the recommendation. Therefore in this context, either PSG or HSAT are recommended for diagnosis of OSA. A clinician's choice of study type for a particular patient should be guided by clinical judgement and incorporate consideration of patient preferences. The overall quality of evidence was downgraded to moderate due to imprecision, inconsistency, or indirectness.

Diagnosis of OSA via HSAT has been demonstrated to result in positive effects on clinical outcomes and efficient care when specific types of HSAT devices are used in an appropriate patient population and management pathway and performed with adequate clinical/technical expertise. Use of HSAT in other contexts may not provide similar benefit. On the other hand, there may be unstudied or understudied contexts in which HSAT has not been evaluated that may provide benefit to a specific patient.

HSAT is less sensitive than PSG in establishing a diagnosis of OSA. A false negative result in a symptomatic patient could result in harm to the patient because the patient would be denied a beneficial therapy. Therefore, after a single negative HSAT result, the performance of PSG is recommended. Performing a repeat HSAT is not recommended in this situation because of the increased risk of the patient not completing the diagnostic process prior to a definitive diagnosis. The TF had a high level of concern that with repeated HSAT use in this situation, patient drop out would compromise patient outcomes. The TF had similar concerns regarding performing a second HSAT when the initial HSAT is technically inadequate but recognized there may be specific situations where this may be appropriate.

The majority of RCTs most generalizable to clinical practice that compared HSAT to PSG required PSG to be performed when HSAT did not confirm a diagnosis of OSA.[83-85] There is significant concern that clinical outcomes relative to use of PSG may not be acceptable if PSG is not performed in patients that have a negative HSAT result in symptomatic patients. In addition, because of an increased likelihood that repeat HSAT testing will also lead to a negative result and concern for risk of patient drop out from the diagnostic process with repeated testing, PSG rather than repeat HSAT is recommended when the initial HSAT result is negative. There are similar concerns that following a technically inadequate HSAT study that repeat testing may be associated with higher rate of technical failure on the second study and of increased risk of drop out from the diagnostic process. But the TF also recognizes that there may be specific cases in which repeat HSAT study is appropriate after an initial failed HSAT. These would include cases in which both of the following are present; the clinician judges that there is a high likelihood of successful recording on a second attempt and the patient expresses a preference for this approach.

Patients value accurate diagnosis and good clinical outcomes and would want to use HSAT in a context that it would provide these benefits.

4.2c *Discussion*

The formulation of this recommendation statement was guided by evidence from 28 HSAT validation studies that measured diagnostic accuracy relative to PSG,[30, 52, 65, 66, 81, 86-108] as well as 7 RCTs that

compared clinical outcomes from management pathways.[83-85, 109-112] Four of these RCTs that did not require oximetry testing as a criterion for inclusion and used conventional methods for determination of PAP pressures, i.e. APAP or attended titration, were felt to be most relevant to clinical practice.[83-85, 112] This subset of studies will be referred to as "RCTs most generalizable to clinical practice" for the remainder of this discussion section.

ACCURACY:

The data from validation studies are summarized in the Supplemental Materials, Tables S38-S58. Briefly, when Type 2 HSAT monitors are compared to attended PSG using an AHI ≥5 cut-off, accuracy in the high-risk population (prevalence 87%) ranged from 84-91% based on two studies (See Supplemental Materials, Table S38).[65, 86] Using a cut-off of AHI ≥15, the accuracy of the monitors in the high-risk population was 88% based on the same studies (See Supplemental Materials, Table S39).[65, 86] When Type 3 HSAT monitors are compared to attended PSG using an AHI ≥5 cut-off, accuracy in a high- risk population (prevalence 87%) ranged from 84-91% and lower risk (prevalence 55%) ranged from 70-78% based on eight studies (see Supplemental Materials, Table S40).[87-90, 92, 93, 113] Using a cut-off of AHI ≥15, the accuracy of the monitors in the high risk population ranged from 65-91% based on six studies (See Supplemental Materials, Table S41).[87, 89-92, 94] Using a cut-off of AHI ≥30, the accuracy of the monitors in the high risk population ranged from 62-96% based on eight studies (See Supplemental Materials, Table S42).[87-92, 94, 95]

Peripheral arterial tonometry (PAT) with oximetry and actigraphy when evaluated simultaneously with simultaneous in-home PSG had a sensitivity of 0.88, specificity of 0.87 and accuracy 88% using AHI ≥5 cut-off and would result in an estimated average of 121 misdiagnosed patients out of 1000 tested in a high-risk group and 125 misdiagnosed patients out of 1000 tested in a low-risk group (See Supplemental Materials, Table S55).[105] Specificity was appreciably lower (0.43-0.47) in 2 studies that compared this PAT device to in lab PSG using AHI ≥5 cut-off[107, 108] but found to be better (0.77-1.0) in studies that evaluated the PAT device at higher AHI cut-offs (See Supplemental Materials, Tables S56-S58).[106-108]

For 2-3 channel monitors using cut-offs of AHI ≥5,[96-98] AHI ≥15,[96-100] and AHI ≥30,[97, 98] accuracy ranged from 81-93%, 72-87%, and 71-90%, in a high-risk population, respectively, and ranged from 77-88%, 68-95%, and 88-91% in a low-risk population, respectively (see Supplemental Materials, Tables S43-S45). When 2-3 channel HSAT was compared with in-home PSG, the accuracy was 86% (range of 76-93%) in the high-risk group using a cut-off of AHI ≥15,[52] and ranged from 83-91% using a cut-off of AHI ≥30 (See Supplemental Materials, Tables S46, S47).[52, 81]

A single study that evaluated oximetry relative to in-home PSG using a cut-off of AHI ≥5 in the high-risk group demonstrated an accuracy of 73% (range of 68-78%) and in the low-risk group of 79% (range of 74-84%).[66] Using oximetry to diagnose OSA at an AHI ≥5 cut-off would result in an estimated average of 274 misdiagnosed patients out of 1000 tested in a high-risk group and 210 misdiagnosed patients out of 1000 tested in a low-risk group.[114] This same study using a cut-off of AHI ≥15 and AHI ≥30, demonstrated an accuracy of 86% (range of 83-91%) and 74% (range of 71-76%) for the high-risk group and an accuracy of 80% (range of 75-84%) and 63% (range of 59-67%) for the low-risk group, respectively (See Supplemental Materials, Tables S52-S54).[114]. Data from studies that evaluated other forms of single channel HSAT are summarized in the supplemental materials (Tables S48-S51)

The potential consequences for patients classified in true and false positive/negative categories were enumerated and are summarized in the Supplemental Materials, Table S1. For false positive results consequences included unnecessary treatment and utilization of resources, unnecessary use of patient time and money for treatment and follow up, psychological distress, delay in diagnosis of true condition, and exposure to side effects of therapy, For false negative results consequences included denial of necessary treatment, reduced quality of life, psychological distress, need for retesting in patients with high

risk of OSA, risk of consequences of untreated OSA including death and increased costs due to medical conditions caused by untreated OSA. The TF concluded that the numbers of patients diagnostically misclassified by HSAT compared to PSG was high enough to be of clinical concern. In a population that has a high prevalence of moderate to severe OSA, both the increased likelihood of false negatives and the significant impact of missed diagnoses on patient outcomes would cause significant harm. This reasoning supports the use of a sleep test with higher sensitivity (i.e., attended PSG) in this population if HSAT provides a negative or non-diagnostic result.

The TF also concluded that the assessment of the impact of suboptimal diagnostic accuracy noted on validation studies on clinical outcomes is very complicated to assess for a number of reasons including: night to night variability that contributes to discordance among tests, the use of a different AHI cut-offs reflecting inconsistent definitions for HSAT and PSG scoring, and the uncertainty regarding clinical outcomes for patients misclassified by HSAT.

For these reasons, studies that compared clinical outcomes in patients randomized to management pathways using either attended PSG or HSAT for diagnosis provided greater clarity regarding acceptability of clinical outcomes using HSAT.

SUBJECTIVE SLEEPINESS:
A meta-analysis of seven RCTs that represented high quality evidence compared ESS change in patients diagnosed by HSAT vs attended PSG, and then followed by PAP titration /initiation (see Supplemental Materials, Figure S59, and Table S67).[83-85, 109-112] The meta-analysis showed a 0.38 unit (95% CI: -1.07 to 0.32) greater improvement in patients randomized to the HSAT pathway that was not statistically significant and below the level of difference (2) that was felt by the TF to be clinically significant. The quality of evidence for subjective sleepiness was high.

QUALITY OF LIFE:
The quality of evidence available for six RCTs utilizing various validated quality of life specific instruments (i.e., FOSQ, SAQLI, and SF-36) ranged from moderate to high and the pooled effects were not statistically or clinically different between pathways (see Supplemental Materials, Figures S60 – S64, and Table S67).[83, 84, 109-112] The TF determined that the quality of evidence for QOL was high.

CPAP ADHERENCE:
Six RCTs representing high quality of evidence evaluated CPAP adherence (mean hours of use per night) and found no statistical difference between the two pathways (see Supplemental Materials, Figure S65, and Table S67).[83-85, 110-112] Five RCTs representing high quality evidence found a statistically non-significant trend towards increased CPAP adherence in the HSAT arm versus the PSG arm as determined by number of nights with greater than 4 hours use that the TF did not deem to be clinically significant (see Supplemental Materials, Figure S66, and Table S67).[85, 109, 83, 84, 112]

OVERALL QUALITY OF EVIDENCE:
The TF determined the critical outcomes assessed in the studies to be false negative results, subjective sleepiness, adherence, cardiovascular events, and QOL as assessed by the FOSQ and SAQLI. The overall quality of evidence is moderate. The evidence was downgraded due to imprecision.

The majority of the studies identified by the literature search were validation studies that provided data on diagnostic accuracy but do not directly assess the impact of misdiagnosis on patient outcomes. Seven RCTs were identified that evaluated patient outcomes.[83-85, 109-112] Four of these RCTs that did not require oximetry testing as a criterion for inclusion and used conventional methods for determination of PAP pressures, i.e. APAP or attended titration, were felt to be most relevant to clinical practice.[83-85, 112] The quality of evidence that failed to demonstrate a difference in outcomes with regards to sleepiness, quality

of life, and CPAP adherence was high. The quality of evidence regarding cost effectiveness was low as only one of these studies addressed this outcome.

RESOURCE USE:
Though a single night of HSAT is less resource intensive than a single night of laboratory based full PSG, the relative cost-effectiveness of management pathways that incorporate each of these diagnostic strategies is less clear. Economic analyses have been published that compare the cost effectiveness of management pathways that incorporate diagnostic strategies including HSAT and lab PSG.[115-117] All have concluded that PSG is the preferred diagnostic strategy from an economic perspective for adults suspected to have moderate to severe OSA. An important factor in these analyses is the favorable cost effectiveness of OSA treatment in patients with moderate to severe OSA, particularly when longer time horizons are considered. As a result, diagnostic strategies that lead to increased false negatives and leave patients untreated or increase false positives and unnecessarily treat patients have less favorable cost effectiveness. It is important to note that these economic analyses are susceptible to error because of imprecision of modelling of management pathways and limitations in the quality of data available to estimate parameters. The impact of errors can be magnified over when extrapolated over long time horizons.

Relative cost-effectiveness of management pathways that use HSAT vs PSG for diagnosis can also be assessed in the context of a RCT if resource utilization is measured. Among the four RCTs most generalizable to clinical practice[83-85, 112] only one provided this information.[84] Rosen et al, 2012 reported that in–trial costs estimated based on Medicare Fee Schedule for the various study procedures including office visits and diagnostic testing, taking into account the need to repeat studies, were 25% less in the home arm than the lab arm.[84] A subsequent cost minimization analysis of this RCT also considered costs from a provider perspective. [118] While provider costs (capital, labor, overhead) were generally less for the home program, this was not true for all modelled scenarios. The provider perspective highlighted the large number of cost components necessary to ensure high quality home-based OSA management, which narrowed the cost difference relative to lab management.

The available studies indicate that potential cost advantages of HSAT over PSG are not as high as reflected by the cost difference of a single night of testing. When HSAT is used appropriately there is a need for additional HSAT and PSG testing for patients with failed and non-diagnostic studies in order to achieve an accurate diagnosis. In addition, if a home management pathway is used in a manner that results in reduced effectiveness relative to PSG, use of HSAT could in fact be less cost effective than using PSG. The TF believes that if HSAT is used in the recommended context and management pathway it will be more cost-effective than if it is used outside this framework.

BENEFITS VS HARMS:
Use of HSAT may provide potential benefits to patients with suspected OSA including convenience, comfort, increased access to testing, and decreased cost. HSAT can be performed in the home environment with less attached sensors during sleep. The availability of HSAT for diagnosis may improve access to diagnostic testing in resource limited settings or when the patient is unable to leave the home environment or healthcare setting for testing. In addition, HSAT may be less costly when used appropriately. These potential benefits must be weighed against the potential for harm resulting from need for additional diagnostic testing for patients with non-diagnostic HSAT findings as well as misdiagnosis and the subsequent inappropriate therapy or lack of therapy. As summarized above, the use of HSAT has not been demonstrated to provide inferior benefit to attended PSG in several RCTs. The TF judged that if HSAT was used in a similar context with regards to patient population and management pathway; clinical/technical expertise and HST device type, the risk of harm is minimized and the

probability of potential economic benefits increased. For this reason the appropriate context in which HSAT is recommended was carefully specified in the recommendation. In addition, a second recommendation is made regarding HSAT use that requires PSG for symptomatic patients if the initial HSAT result is negative to avoid misdiagnosis. In particular, requiring a repeat HSAT is not recommended in this situation because of an increased risk of the patient not completing the diagnostic process. A similar but weaker recommendation is made regarding technically inadequate initial HSAT for symptomatic patients: a repeat HSAT is not recommended in this situation because of the increased risk of the patient not completing the diagnostic process; though the TF recognizes that in specific cases the clinician may elect to perform a second HSAT if there is a high likelihood of successful recording on a second attempt and the patient expresses a preference for this approach.

The TF recognizes that HSAT may have value to patients in some other contexts than covered by this recommendation but has limited its recommendation to apply to situations where there is sufficient evidence to guide evaluation of benefits versus harms.

PATIENTS' VALUES AND PREFERENCES:

Individual patients will differ on whether they prefer PSG vs HSAT depending on their circumstances and values. One of the four RCTs, most generalizable to clinical practice performed both types in each subject and found 76% preferred HSAT.[85] This means that a significant percentage (24%) still preferred PSG. There is insufficient data about diagnostic testing preferences in clinical practice where preferences may differ from what is seen in the RCT setting. The availability of different options for diagnosis may increase satisfaction if patient preferences are considered as a part of the decision process in choosing the diagnostic test type. If HSAT was used, the TF felt that patients would value accurate diagnosis and good clinical outcomes and would want to use HSAT in a context that would provide these benefits. The TF determined that patients would prefer not having a repeat HSAT test if an initial HSAT result is negative as repeated HSAT would be less likely to produce a definitive result and would unnecessarily inconvenience the patient. In this situation, proceeding directly to PSG which has greater sensitivity to detect OSA would be preferred. The TF also felt that patients would prefer not to have a repeat HSAT test if the initial HSAT was technically inadequate to avoid inconvenience but that in specific cases in which there was high likelihood of an adequate result with repeat testing some patients may desire this option.

SPECIAL CONSIDERATIONS:

CLINICAL POPULATION:

A review of RCTs indicated that the following criteria to establish the presence of high risk of moderate to severe OSA for suitability of HSAT use was reasonable: the presence of two or more of the following symptoms: habitual loud snoring, witnessed apnea/gasping/choking, or diagnosed hypertension; and daytime sleepiness. Seven RCTs comparing HSAT to PSG that evaluated PAP adherence and/or patient reported outcomes included medically uncomplicated adult clinical populations at higher risk of moderate to severe OSA.[83-85, 109-112] Among these were four studies that the TF judged to be the most relevant to clinical practice.[83-85, 112] Relevant details regarding the clinical populations included in these four studies are provided: Two of four studies[84, 85] required ESS >12 as an entry criterion; one[112] required at least two out of three criteria (sleepiness (ESS >10), witnessed apnea, snoring) for participation; and one which was performed in a Veteran's Administration population that did not specify any specific entry criteria besides suspected OSA (though the average ESS for participants was elevated at >12 and 95% were men).[85] In the latter study, 9.9% of individuals in the PSG arm were found to have AHI <5.[85] In addition to sleepiness, studies in this subset had specific inclusion criteria such as snoring, witnessed apnea, gasping/choking at night, and/or hypertension that were considered as part of the entry criteria in at least 2 studies.[83, 85] One study incorporated neck circumference in the determination of high risk of OSA.[84]

EXCLUDED PATIENT POPULATIONS:

Based on information from 3 of the RCTs most generalizable to clinical practice which also specified exclusion criteria, the TF determined that HSAT should be used in an uncomplicated clinical population as defined by the absence of severe cardiopulmonary disease (e.g. heart failure, chronic obstructive pulmonary disease (COPD)), respiratory neuromuscular weakness, chronic opiate medication use, concern for other significant sleep disorder (e.g. parasomnia, narcolepsy, severe insomnia), and environmental/personal factors that preclude the adequate acquisition and interpretation of data from HSAT.[84, 85, 112]

All three of these studies excluded patients with significant cardiopulmonary disease and other significant sleep disorders.[84, 85, 112] Two studies excluded patients taking opioids, having uncontrolled psychiatric disorder, neuromuscular disease, and patients with significant safety related issues related to driving or work. Other notable exclusion criteria specified by at least one of the studies included lack of an appropriate living situation, pregnancy and alcohol abuse. The single study that did not mention exclusion criteria noted that 3/148 individuals in the HSAT arm were diagnosed with central sleep apnea and 4/148 individuals required supplemental oxygen or bilevel PAP and left the study. [83] In the lab arm of the study 6/148 individuals were diagnosed with central sleep apnea and 12/148 required supplemental oxygen or bilevel PAP. Studies outside the subset summarized here generally had similar inclusion/exclusion criteria.

FOLLOW-UP:

Based on information from the four RCTs most generalizable to clinical practice,[83-85, 112] the TF determined that HSAT should be used in the context of an OSA management pathway which incorporates a PAP therapy initiation protocol that includes APAP or attended PSG titration, early follow up after initiation of therapy, and PSG titration studies for patients failing APAP therapy. All four RCTs incorporated early follow up of APAP titration (within 2-7 days subsequent to HSAT) by skilled technical staff.[83-85, 112] Three of the studies[83-85] required use of PSG if HSAT did not provide adequate data or showed a low AHI after 1 or 2 unsuccessful attempts; and after 1 or 2 failed APAP trials (e.g. insufficient use, elevated residual AHI, persistent large leak). There was evidence of significant subject dropout from the management pathway in 2 of the studies: Rosen et al. 2012[84] reported that 17% of subjects randomized to the home arm and Kuna et al., 2011[83] reported that 25.5% of subjects with inconclusive home studies failed to complete the diagnostic protocol. The TF was concerned that in clinical practice there would be higher levels of drop out compared to the RCT setting from diagnostic testing among patients with initial study attempt that did not result in a diagnosis of OSA. In particular, there was concern that patients with false negative HSAT results may not complete PSG testing after they had been informed of a negative result. For this reason, the TF recommends that if the initial HSAT shows a negative result, PSG rather than repeat HSAT be performed. There are similar concerns that following a technically inadequate HSAT study that repeat testing may be associated with higher rate of technical failure on the second study and with increased risk of drop out from the diagnostic process. But the TF also recognizes that there may be specific cases in which repeat HSAT study is appropriate after an initial failed HSAT. These would include cases in which both of the following are present; the clinician judges that there is a high likelihood of successful recording on a second attempt and the patient expresses a preference for this approach.

This TF's concern in the situation of an initial inadequate recording (technical failure) was less because the reason for this was more readily explained to the patient. The TF felt that in some situations a second HSAT study in a patient with initial inadequate recording may be reasonable depending on the likelihood of success with a second attempt as judged by the clinician and patient preference.

CLINICAL EXPERTISE:

Based on information from the four studies most generalizable to clinical practice, the TF determined that HSAT should be administered by an accredited sleep center under the supervision of a board certified sleep medicine physician or clinician with equivalent training. All four RCTs were performed at academic or tertiary sleep centers with highly skilled Sleep Medicine providers and technical staff.[83-85, 112] HSAT recordings were reviewed by a sleep medicine specialist. One RCT that was not included in this subset (because an overnight oximetry was used as entry criteria) used a simplified nurse-led model of care involving nurse specialists experienced in management of sleep disorders (mean of 8.3 years of experience with CPAP therapy).

HOME SLEEP APNEA TESTING DEVICE:

Based on information from the four RCTs most generalizable to clinical practice,[83-85, 112] the TF determined that testing should be performed using a technically adequate HSAT device. Among the four RCTs, three used conventional Type 3 monitors (nasal pressure, thoracic and abdominal excursion using RIP technology, oxygen saturation, EKG, body position, and oral thermistor in some cases) [83, 84, 112] and one used a 4 channel device[85] based on peripheral arterial tone with 3 additional channels (heart rate, pulse oximetry, and actigraphy). Additional guidance on technical specifications regarding HSAT testing is provided in The AASM Manual for the Scoring of Sleep and Associated Events.

ADEQUACY OF RECORDING:

The TF determined that a protocol that requires a minimum of 4 hours of good quality data from HSAT recording performed during the habitual sleep period is adequate to diagnose OSA. In the four RCTs most generalizable to clinical practice, the minimum requirement for an acceptable study was 4 hours of adequate flow and oximetry signals.[83-85, 112] While one HSAT study[85] used PAT as a surrogate of flow, two studies recorded nasal pressure flow[83, 112] and one study recorded thermistor in addition to nasal pressure flow[84]; the latter three studies also recorded thoracic and abdominal movements.[83, 84, 112] All of these studies showed at least equivalence of outcomes in the home vs in-laboratory management pathways in terms of adherence to PAP therapy and functional improvement.[83, 84, 112]

A number of non-RCT validation studies also reported minimum requirements for duration of acceptable signal quality.[30, 52, 81, 86, 88, 93, 97, 119] The required signals and minimum durations included nasal pressure flow and oximetry for at least 3 hours[88, 93, 119] or 4 hours[52, 81, 86, 97] and single-channel nasal airflow recording for a minimum of 3 hours[30] or only 2 hrs.[53] The diagnostic accuracy of the cardiorespiratory devices compared with in-lab PSG for the detection of OSA at different AHI cutoff points was relatively high. One study reported a sensitivity and specificity of 88% and 84%, respectively for HSAT AHI cut-point of ≥9/h.[52] In a separate study, the sensitivity and specificity for home vs in-lab PSG were 90.5% and 88.9% for AHI cut-point of >10, but 88% and 55%, respectively for AHI cut-point of >5/hr.[86] Similarly, at AHI cut-off of >10, HSAT had a sensitivity of 87%, a specificity of 86% and a positive likelihood ratio of 6.25.[88]

The overall body of evidence investigating the minimum number of hours of adequate data on HSAT required to accurately diagnose OSA is very limited. Therefore, based on available indirect evidence, the TF weighed the 'risk' of undergoing less than the required duration of good quality HSAT with resultant false negative (or false positive) results against the 'benefit' of potentially increasing the accuracy of study results by repeating the HSAT in order to attain the minimum required duration of recording. However, repeating a less-than-4-hour study in the scenario of a 'positive' diagnosis of OSA is less likely to alter clinical decision-making and may, in fact, result in unnecessary delays in care with increased cost. Conversely, a 'negative' HSAT in the scenario of a high pre-test probability of OSA will likely need to be repeated even when the test is of adequate quality and duration. Given that there are no data to suggest

that a criteria that allows fewer than 4 hours of technically adequate recording compromises the accuracy of test results, and in the absence of direct evidence on the impact of a minimum number of hours of adequate HSAT on efficiency of care and clinical outcomes, the TF arrived at the above stated recommendation with the caveat that a shorter study duration of HSAT that allows a diagnosis of OSA may be clinically acceptable rather than repeating the study. The TF believes that the goals of establishing an accurate diagnosis while minimizing patient inconvenience and cost distinctly align with patient preferences.

The adequacy of a single night HSAT test performed for the diagnosis of OSA in the context of appropriate clinical population and management pathway is supported by published evidence. Our literature search only identified 2 studies relevant to the question of whether multiple night recording is superior to a single night.[30, 72] These studies compared multiple nights (3) of single channel HSAT (nasal pressure transducer or oximetry) to the first night of recording. Utilizing PSG as the reference, the studies found that recording over 3 consecutive nights may decrease the probability of insufficient data and marginally improve accuracy when compared to a single night of recording. The TF considered the evidence inadequate to establish the superiority of a 3 nights over a single night protocol as the studies only included a single channel recordings and did not compare clinically meaningful outcomes and/or efficiency of care (e.g., time to treatment, costs) between management pathways that used single versus multiple night protocol for HSAT.

A single HSAT recording encompassing multiple nights may have potential advantages or drawbacks relative to a study that includes only a single night of recording. For example, if multiple-night HSAT improved accuracy or resulted in fewer inconclusive or inadequate studies, there may be benefits in patient outcomes or costs. On the other hand, the potential of multiple-night recordings to increase cost and patient inconvenience must be considered. There is a lack of sufficient evidence to routinely recommend more than one night's duration for HSAT.

4.3 HSAT for the diagnosis of OSA in adults with comorbid conditions

4.3a We recommend that polysomnography, rather than home sleep apnea testing, be used for the diagnosis of OSA in patients with severe cardiorespiratory disease, neuromuscular disease with respiratory muscle impairment, awake hypoventilation or high risk of sleep related hypoventilation, chronic opioid medication use, or severe insomnia. (STRONG FOR)

4.3b *Summary*

Based on the limited data available regarding validity of HSAT in patients with severe cardiorespiratory disease, neuromuscular disease with respiratory muscle impairment, high risk of hypoventilation, opioid medication use, or severe insomnia, both in terms of accuracy of detection of OSA and other forms of sleep disordered breathing that can occur in these populations (central sleep apnea and hypoventilation), the TF recommends that HSAT should not routinely be used for evaluation of sleep disordered breathing in these patient populations. The overall quality of evidence was very low based on imprecision, and a limited number of modest sized studies with fair study design.

PSG is the gold standard method for diagnosis of OSA and other forms of sleep disordered breathing. HSAT has not been adequately validated or demonstrated to provide favorable clinical outcomes and

efficient care when incorporated into specific management pathways in these patient populations and may result in harm through inaccurate assessment of sleep disordered breathing.

Patients value accurate OSA diagnosis and favorable clinical outcomes and therefore would want to be evaluated in a context that would provide these benefits.

4.3c *Discussion*

No RCTs were identified involving patients with significant co-morbidities as outlined above that compared clinical outcomes from HSAT to PSG. Three studies regarding validity of HSAT for the diagnosis of OSA in patient populations with significant cardiorespiratory morbidity met our inclusion criteria.[120-122]

PATIENTS WITH COMORBID HEART FAILURE:

Three studies included patients with heart failure.[113, 120, 122] A study of 50 patients with stable heart failure (Class 2-4; left ventricular ejection fraction < 40%) compared home oximetry to lab PSG.[120] Home oximetry was considered positive if the 2% ODI $\geq$10, and the lab PSG was considered positive if AHI $\geq$15 using a hypopnea criteria that did not require oxygen desaturation or arousal. Home oximetry data was not obtained in 3 patients and studies had to be repeated in 2 patients. This study found a sensitivity of 0.85 and specificity of 0.93 in identifying sleep disordered breathing.[120] The specificity was poor for identifying central sleep apnea based on desaturation/resaturation pattern (specificity 0.17; sensitivity 1.0) with 10 of 12 patients with OSA identified as having CSA. A study of 50 patients with heart failure (Class 3; LVEF ≤35%) comparing an HSAT monitor that included ECG (2 leads), oximetry, and respiratory impedance sensors to lab PSG, was able to obtain valid data in 44 patients in the home setting. Sensitivity, specificity and accuracy at AHI $\geq$5 and AHI $\geq$15 cut-offs were 0.92, 0.52, 0.73 and 0.67, 0.78, 0.75 respectively.[122] Data regarding methods used to distinguish central from obstructive events were not provided. A study of 100 patients with stable heart failure (mean LVEF (SD) 34.6% +/- 11) compared simultaneous 2-channel HSAT (nasal pressure flow and oximetry) and home PSG. Ten patients did not have adequate HSAT recordings. [113] In the 90 patients with valid HSAT recordings, the sensitivity and specificity was 0.98 and 0.60 respectively using an AHI ≥5 cut off (hypopneas required 4 % oxygen desaturation for both HSAT and PSG) and 0.93 and0.92% using an AHI ≥15 cut-off. 29% of patients had central sleep apnea, 19% had OSA and 13% had both based on PSG. The type of sleep apnea could not be determined using the HSAT device. The meta-analysis (see Supplemental Materials, Table S68) found that in a population of 1000 patients at high risk of OSA (64% prevalence) 45-230 more false negative and 18-79 more false positives would result from the use of HSAT.[113, 120, 122] The quality of evidence for accuracy outcomes was downgraded to very low due to imprecision and indirectness.

PATIENTS WITH COMORBID COPD:

Only one study addressed the validity of HSAT (nasal pressure, respiratory excursion (piezoelectric sensor), body position and pulse oximetry) in patients with COPD.[121] Of 72 patients with stable COPD (GOLD stage II and III) and symptoms of OSA, only 26 patients (36%) had HSAT studies of reasonable quality.[121] When comparing PSG to HSAT, the intraclass correlation coefficient was 0.47 (accuracy not provided).[121] Data regarding detection of hypoventilation was not provided. Evidence was downgraded to very low based on imprecisions (only one study with limited sample size) and risk of bias (significant data loss).

PATIENTS WITH OTHER COMORBIDITIES:

No studies were identified that met our inclusion criteria that specifically evaluated the use of HSAT for diagnosis of OSA in patients with history of stroke, opioid use, neuromuscular disease with respiratory muscle impairment, high risk of hypoventilation, or insomnia.

OVERALL QUALITY OF EVIDENCE:
There was no evidence regarding suitability of HSAT for diagnosis of OSA in neuromuscular disease with respiratory muscle weakness, hypoventilation, chronic opioid disease and severe insomnia. The evidence regarding suitability of HSAT for diagnosis of OSA in patients with comorbid heart failure was very low based on 3 studies with small sample size and inadequate study design.[113, 120, 122] The evidence regarding suitability of HSAT for diagnosis of OSA in patients with COPD was very poor based on the availability of a single small size study with a majority of subjects with unavailable HSAT data because of recording failure. The overall quality of evidence regarding patients with comorbid conditions was very low based on imprecision, and a limited number of small sized studies with inadequate study design.

BENEFITS VS HARMS:
Certain patient populations are at increased risk of having forms of SDB other than OSA (central sleep apnea, hypoventilation, and hypoxemia). These forms of SDB can cause significant morbidity and/or mortality if left untreated. Therapies used to treat these non-OSA breathing disorders can be different from what is used to treat OSA and can be costly. HSAT has not been validated to diagnose some of these types of SDB (CSA, hypoventilation). Use of HSAT in populations at increased risk for non-OSA breathing disorders increases the likelihood of not detecting these breathing disorders (false negative) or falsely identifying these breathing disorders (false positive) which could lead to inadequate or unnecessary treatments, increasing costs, morbidity and/or mortality. Though the cost of a PSG study is higher than a HSAT, the TF assessed the benefits of increased accuracy, use of appropriate therapy and improved clinical outcomes to outweigh this factor. There are instances where PSG cannot be performed for practical reasons (hospitalization, inability of patient to leave home setting or participate in PSG), and in these instances use of HSAT may be reasonable as the alternative is to not address SDB at all.

PATIENTS' VALUES AND PREFERENCES:
Based on their clinical judgement, the TF judged that patients with high likelihood of non-OSA SDB would want these breathing disorders to be adequately diagnosed and treated as therapy of these disorders can result in significant improvement in health and well-being. If the optimal diagnostic test (PSG) was not feasible, then they would desire to have other diagnostic tests (HSAT) available that may aid their clinical provider in providing care for sleep disordered breathing.

4.4 Diagnosis of OSA in adults using a split-night vs a full-night PSG

4.4a We suggest that if clinically appropriate, a split-night diagnostic protocol, rather than a full-night diagnostic protocol for attended polysomnography be used for the diagnosis of OSA. (WEAK FOR)

Remarks: Clinically appropriate is defined as the absence of conditions identified by the clinician that are likely to interfere with successful diagnosis and treatment using a split night protocol

This recommendation is based on a split-night protocol that initiates CPAP titration only when the following criteria are met:

A moderate to severe degree of OSA is observed during a minimum of 2 hours of diagnostic PSG; **AND** At least 3 hours are available for CPAP titration.

4.4b *Summary*

In the context of an appropriate protocol, a split-night study has acceptable accuracy to diagnose OSA in the uncomplicated adult patient and may improve efficiency of care when performed in the context of adequate clinical and technical expertise. The recommendation is based on evidence from studies that included typical sleep clinic patients studied for symptoms of OSA. Many studies were retrospective case series in which it is possible that patients deemed clinically inappropriate for split night study were not included. There may be specific patient characteristics that are not well suited to conducting a shorter diagnostic evaluation or titration period characteristic of the split night study (i.e severe insomnia, claustrophobia, concern for other forms of sleep disordered breathing or non-breathing related sleep disorders). The quality of evidence was determined to be low.

The split-night protocol potentially provides enhanced efficiency of care in terms of ability to diagnose OSA and establish CPAP treatment pressure needs within a single night recording. Potential disadvantages include insufficient diagnostic sampling and insufficient time to ascertain appropriate treatment CPAP pressure.

A split night study may be preferred relative to separate PSG and PAP titration studies due to the convenience and cost savings of completing a diagnostic and titration study during one versus two separate sleep studies; however, this needs to be balanced with consequences of potentially inconclusive diagnostic or titration portions of the sleep study. If the diagnostic portion is inconclusive, a second sleep study is needed. If the titration portion is inconclusive, a second titration sleep study or the use of autotitrating CPAP may be needed. Based on their clinical judgement, the Task Force determined that the majority of well-informed patients would choose the split-night protocol over a full-night protocol.

4.4c *Discussion*

Our literature search yielded 8 studies that met inclusion criteria.[123-130] Three focused on OSA diagnostic accuracy of the initial portion of the standard PSG recording compared to the full night diagnostic recording[123, 124, 129] and one compared the diagnostic accuracy of the diagnostic portion of a split night study to a separate full night study.[125] Two studies focused on comparison of CPAP success while on CPAP in those undergoing a split-night study versus a full night sleep recording[127, 128] and one examined CPAP adherence in those undergoing split night studies versus full night studies.[126] One additional study that did not meet inclusion criteria for the literature search examined cost-effectiveness of the split night study versus the full night study.[115]

DIAGNOSTIC ACCURACY:

The studies that examined diagnostic accuracy and performance characteristics used the initial truncated PSG to serve as a representative surrogate of the initial diagnostic portion of a split-night study, and compared the first 2-3 hours of the recording versus the full night of sleep recording. One study by Chou et al, 2011 assessed the diagnostic validity of a 2-hour recording and identified an optimal AHI cut-off of 30/hour as providing the highest accuracy (90.9%). Also, a 2-hour diagnostic period had an AUC =0.97 which was similar to that of a 3-hour period.[124] Of the studies that considered 3 hours, excellent consistency of the AHI (concordance correlation coefficient adjusted for REM and supine sleep =0.96 and AUC =0.93), even in those with a milder degree of OSA (AUC for AHI cutoffs of 5, 10 and 15 were 0.95, 0,97 and 0.995 respectively), was observed and demonstrated no statistically significant difference in the AHI derived from the first 3 hours of total sleep time versus the total sleep time.[123, 129] Finally, another study by Chuang et al, 2008 showed an AHI Pearson correlation coefficient between a full night study and the diagnostic portion of the split night study of 0.63 ($p < 0.0001$) when the split-night study recording time was ≥90 minutes.[125] The quality of evidence for diagnostic accuracy was downgraded to low due to indirectness, imprecision, and inconsistency of findings.

CPAP OUTCOMES:

One of the studies that examined CPAP success in the split-night versus full night CPAP titration recordings focused on upper airway resistance syndrome and found no difference of success rate of CPAP titration defined as respiratory effort related arousal (RERA) index <5 on the final CPAP setting.[127] Similarly, another study involving comparison of split-night CPAP recordings versus CPAP titration recordings in patients with OSA, showed no statistically significant difference of the AHI, arousal index and percentage sleep time less than 90% oxygen saturation performed on consecutive nights in the same individuals.[128] One study identified a comparable CPAP adherence at 4-6 weeks after starting CPAP in those undergoing a split night study (78.7%) versus a full-night study with follow-up titration (77.5%).[126] The quality of evidence for CPAP outcomes was downgraded to low due to the imprecision associated with a limited number of studies and small sample size.

RESOURCE USE:

The only available cost effectiveness analysis demonstrated that there was no significant difference in cost or effectiveness when comparing pathways incorporating split night studies versus full night studies.[115] In probabilistic terms, split-night studies (and full-night studies) were most cost effective at higher amounts of third-party willingness to pay compared to HSAT.[115] There was a low level of confidence in the certainty regarding resource use given the lack of high quality evidence to inform cost effectiveness analysis.

OVERALL QUALITY OF EVIDENCE:

The available studies were methodologically limited due to a number of issues: use of the initial portion of a full-night PSG recording as a surrogate for the baseline portion of a split-night study,[123, 129] lack of consistent use of standard monitoring (e.g. nasal pressure transducer),[123] and ambiguity in generalizability to populations with co-morbidity such as underlying cardiopulmonary disease. Given the small number of studies and limited interpretability due to methodological limitations, the TF considered the level of evidence to be low. The overall quality of evidence was determined to be low due to the limited number of studies and sample size leading to imprecision, the indirectness of the studies, and the inconsistency of findings. Given the lack of definitive data, the TF elected not to designate a specific AHI threshold to inform decision to initiate PAP titration during a split night study protocol.

BENEFITS VS HARMS:

Split-night studies have benefits as well as shortcomings which are worthwhile to note. For instance, the split-night protocol has the potential advantage of providing enhanced efficiency of care in terms of ability to diagnose OSA and establish CPAP treatment pressure needs within the same recording compared to a protocol requiring a full-night sleep recording followed by a separate CPAP titration study. Potential disadvantages of the split-night study include insufficient diagnostic sampling (e.g., limited REM time and limited supine time in those with difficulty initiating sleep), and insufficient time to ascertain appropriate CPAP treatment pressure. Based on their clinical judgement, the TF determined that there is low certainty of evidence that the benefits of a split-night study exceed the harms when compared against a full-night study.

PATIENTS' VALUES AND PREFERENCES:

In terms of patient preferences, a split-night study may be preferred due to the convenience and cost and the benefits of completing a diagnostic and titration study during one versus two separate sleep studies; however, this needs to be balanced with the possible need to return for a second sleep study if the diagnostic or titration portions of the split study are inconclusive. When comparing the split-night study to the full-night study, limited existing data are consistent and demonstrate a high level of reproducibility of the standard AHI metric, success of CPAP titration in effectively identifying the optimal CPAP pressure, and suggestion of similar follow-up CPAP adherence and cost effectiveness. Based on their clinical

judgement, the Task Force determined that the majority of well-informed patients would choose the split-night protocol over a full-night protocol based on clinical experience.

4.5 Diagnosis of OSA in adults using a single-night vs. a two-night PSG study

4.5a We suggest that when the initial attended polysomnogram is negative and there is still clinical suspicion for OSA, a second attended polysomnogram be considered for the diagnosis of OSA in symptomatic patients. (WEAK FOR)

4.5b *Summary*

There was limited evidence from which to assess the efficacy of performing a repeat PSG; the recommendation is based on evidence for performing a single-night PSG versus two-nights of PSG for the diagnosis of OSA. The quality of evidence for this recommendation was very low. While the certainty of evidence regarding night to night variability of AHI from the meta-analysis was considered to be high, there was limited evidence from which to assess the efficacy of single-night PSG versus two-night PSG in terms of diagnostic accuracy and clinical outcomes. This lead to a downgrading of the overall quality of evidence to very low to reflect the low certainty of the task force that a repeat PSG would improve patient outcomes.

A meta-analysis of four observational studies compared the AHI scores from 2 consecutive nights of PSG and found no clinically or statistically significant difference (mean difference in the AHI of 0.14; 95% CI -1.86 to 2.15).[29, 131-133] (See Supplemental Materials, Table S69) However, 8-25% of patients were diagnosed with clinically significant OSA only on the second study, which could have therapeutic implications.

Proceeding with a second PSG in patients with a negative initial PSG in order to establish a diagnosis of OSA must be balanced against the possibility of a false positive diagnosis, inconvenience to the patient and added cost of a second study. Given the night to night variability in the AHI, a high AHI on a second study could conceivably represent a false positive diagnosis in a patient without OSA. The added financial burden and inconvenience of a second night of testing in a sleep lab must also be considered, though cost-analysis and patient satisfaction data are not available regarding this issue.

Discussion of a repeat attended PSG with a patient who has had a negative initial PSG is warranted to ensure further testing is in-line with the patient's values and preferences, given the potential benefits and harms associated with additional testing. For these reasons, the TF recommends use of a second night PSG in symptomatic patients with ongoing clinical suspicion of OSA and a negative initial PSG only after careful consideration of the clinical situation and the potential utility for the patient.

4.5c *Discussion*

The meta-analysis of four observational studies[29, 131-133] mentioned above compared AHI data from 2 consecutive nights of PSG (See Supplemental Materials, Table S69 and Figure S70) and found the mean difference in the AHI between the 2 nights was 0.14, which is neither statistically nor clinically significant. There was a wide range of OSA severity in the populations included in the 4 studies (AHI range: 7-34). None of the studies included data on body position during the 2 nights of PSG. One of two studies that reported on sleep architecture changes[132, 133] found a statistically significant increase in REM sleep on the second PSG.[133] Only 1 of the studies indicated that PSG scorers were blinded to the other PSG result.[133]

Despite the minor difference in mean AHI between study nights, a subset of individuals had considerable night to night variability in their AHIs, which could have potential clinical implications if the AHI crosses a treatment threshold only on the second PSG. Using an AHI cutoff of ≥5 to diagnose OSA, 3 of the studies[29, 132, 133] identified that 9.9-25% of subjects had an AHI <5 on the first PSG but an AHI ≥5 only on the second PSG. Likewise, using an AHI cutoff of ≥15 or 20 as a potential treatment threshold, 2 of the studies[29, 132] observed that 7.6% and 25% of subjects crossed this threshold only on the second study. OSA severity was also noted to vary in a subset of subjects with 26% to 35% changing the severity classification of their OSA (in either direction) on the 2 nights, though the majority were a shift of a single category (e.g., mild to moderate).[29, 132]

OVERALL QUALITY OF EVIDENCE:

The overall quality of evidence based on the GRADE criteria for the 4 studies was considered to be high for comparing night-to-night AHI variability.[29, 131-133] However, the available literature did not address other clinically meaningful outcomes (e.g. impact on costs, quality of life, comorbidities and long-term outcomes) resulting from undergoing a second night of PSG testing. As such, the TF downgraded the overall quality of evidence addressing this question to very low to reflect the likelihood that future research could result in different estimates of effect for the outcomes of interest, many of which were not available in the current literature.

BENEFITS VS HARMS:

A second night of an attended PSG in symptomatic patients allows for the diagnosis of OSA in 8-25% of patients with initial false negative studies. Establishing a diagnosis of OSA in these patients allows for treatment that could lead to improved outcomes in a variety of areas, including symptom control (e.g. less daytime sleepiness), better quality of life, and potentially decreased cardiovascular morbidity over time. However, routinely repeating a PSG in patients with an initial negative PSG has potential downsides. There is a risk that a repeat testing could lead to false positive cases being identified. In addition, the routine use of a 2 night study protocol would cause inconvenience to the patient, increased utilization of resources and healthcare costs, and perhaps even delays in the care of other patients awaiting PSG. Based on their clinical judgement, the TF determined that the benefits of a second PSG outweigh the harms; however the certainty that the benefits outweigh the harms is low.

PATIENTS' VALUES AND PREFERENCES:

Patient preference were also considered when weighing the values and trade-offs of a repeat PSG in a patient suspected of having OSA with an initial false negative study. The patient's desire and motivation for further testing can be affected by a variety of factors from the patient's perspective (e.g. quality of life, costs, etc.) and thus a discussion with the patient is warranted prior to pursuing repeat testing. Based on their clinical judgement, the TF determined that the majority of well-informed symptomatic patients would choose a second PSG to diagnose suspected OSA when the initial PSG is negative.

5.0 DISCUSSION AND FUTURE DIRECTIONS

This extensive literature search and review identified many areas that warrant additional study to better inform clinical decision-making and improve patient outcomes.

More accurate and user-friendly clinical screening tools and models are needed to better predict diagnosis and severity of OSA, as well as to improve risk stratification and efficiency of patient management. Identification of biomarkers that detect and quantitate exposure to obstructive sleep disordered breathing and predict likelihood of adverse clinical outcomes could provide novel information that may improve the diagnosis and management of OSA. These advancements could also improve the efficiency by which conventional sleep apnea tests that measure the physiology of breathing during sleep

are used. In addition, these approaches may be useful in situations where conventional tests may not be readily available or logistically feasible to conduct in a timely fashion (e.g. inpatient settings, preoperative clinics).

For patients scheduled for upper airway surgery for snoring, there is currently insufficient evidence to determine if the diagnosis of OSA through sleep testing can decrease peri-operative risk and improve surgical outcomes. Because it has been established that questionnaires cannot be used to diagnose OSA, many sleep experts have followed previous guidelines recommending sleep testing to evaluate for OSA prior to performing surgery for snoring. Further research designs to evaluate this protocol would be useful.

While PSG remains the gold standard for the diagnosis of OSA, it involves cumbersome sensors and monitors that, if were minimized and less invasive, could make PSG more tolerable for patients. Newer technology that is less intrusive and more comfortable may influence patient preferences regarding diagnostic approaches. Split-night PSG testing, which may improve the efficiency of PSG, has not been adequately studied. The quality of evidence regarding split-night sleep studies is low and additional research is needed to better determine its overall impact on patient outcomes. Past research often utilized outmoded testing methodology (e.g. they did not use nasal pressure cannulas) and/or outdated scoring criteria, limiting its relevance. There is also a lack of data on the utility of split night testing in patients with significant underlying cardiopulmonary disease. Finally, the cost-effectiveness of split night studies warrants further exploration.

Significant progress has been made in better understanding the accuracy and clinical utility of HSAT, however more is needed. Future research should focus on evaluating HSAT devices in patients with different pretest probabilities for OSA as well as in more diverse patient populations, especially those routinely excluded (e.g. at risk for hypoventilation and CSA) from past studies, and those unable to be studied in the sleep laboratory environment (e.g. due to critical illness, immobility, safety). In addition, the types and numbers of HSAT sensors necessary to adequately diagnose OSA require elucidation. Research should focus on how to better define the optimal physiologic parameters to be measured, particularly with regards to the minimal number of parameters necessary and how devices measuring different parameters compare with one another and in different clinical situations. Furthermore, a better understanding of factors associated with inadequate or failed HSAT testing could help to optimize efficiency of care with regards to choosing the most appropriate diagnostic method for a given patient and/or clinical situation. Greater study of the cost-effectiveness of home-based management is needed to better define situations in which it may or may not offer value to the healthcare system relative to laboratory-based management. Finally, there is a paucity of data on how patient preferences currently influence clinical decision-making regarding the type of diagnostic testing. The role of patient preference regarding diagnostic pathways (i.e. HSAT vs. attended PSG) and how this may impact outcomes remains to be explored.

More work is needed to determine the duration and number of nights that are optimal for sleep testing. For example, when is a second night of PSG testing indicated in patients suspected of having OSA but who have a negative initial study. Future studies should attempt to determine factors that may predict which patients may benefit from a second night of PSG and measure the impact on clinically meaningful outcomes (e.g., impact on costs, quality of life and medical morbidity). Likewise, the duration and number of testing nights required to accurately diagnose or exclude a diagnosis OSA with HSAT is in need of further study. In terms of the minimal duration of HSAT recording time, future comparative effectiveness studies should consider the impact of HSAT duration on clinical accuracy, clinical efficiency and functional outcomes. Comparative effectiveness studies should also consider the impact of the number of nights of HSAT testing on clinically meaningful outcomes and/or efficiency of care (e.g. time to treatment, costs).

Finally, there is a need for controlled trials to determine the role of repeat testing during chronic clinical management. There was insufficient evidence to determine whether and under what scenarios repeat PSG or HSAT to confirm severity of OSA or efficacy of therapy improves outcomes relative to clinical follow-up without retesting.

REFERENCES

1. Kushida CA, Littner MR, Morgenthaler T et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep* 2005;28:499-521.
2. Collop NA, Anderson WM, Boehlecke B et al. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2007;3:737-47.
3. Aurora RN, Chowdhuri S, Ramar K et al. The treatment of central sleep apnea syndromes in adults: practice parameters with an evidence-based literature review and meta-analyses. *Sleep* 2012;35:17-40.
4. Berry RB, Chediak A, Brown LK et al. Best clinical practices for the sleep center adjustment of noninvasive positive pressure ventilation (NPPV) in stable chronic alveolar hypoventilation syndromes. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2010;6:491-509.
5. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM Increased prevalence of sleep-disordered breathing in adults. *American journal of epidemiology* 2013;177:1006-14.
6. Barbe F, Duran-Cantolla J, Sanchez-de-la-Torre M et al. Effect of continuous positive airway pressure on the incidence of hypertension and cardiovascular events in nonsleepy patients with obstructive sleep apnea: a randomized controlled trial. *Jama* 2012;307:2161-8.
7. Punjabi NM, Caffo BS, Goodwin JL et al. Sleep-disordered breathing and mortality: a prospective cohort study. *PLoS medicine* 2009;6:e1000132.
8. Marin JM, Carrizo SJ, Vicente E, Agusti AG Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet* 2005;365:1046-53.
9. Ravesloot MJ, van Maanen JP, Hilgevoord AA, van Wagensveld BA, de Vries N Obstructive sleep apnea is underrecognized and underdiagnosed in patients undergoing bariatric surgery. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies* 2012;269:1865-71.
10. Punjabi NM The epidemiology of adult obstructive sleep apnea. *Proceedings of the American Thoracic Society* 2008;5:136-43.
11. Franklin KA, Lindberg E Obstructive sleep apnea is a common disorder in the population-a review on the epidemiology of sleep apnea. *Journal of thoracic disease* 2015;7:1311-22.
12. Jackson ML, Howard ME, Barnes M Cognition and daytime functioning in sleep-related breathing disorders. *Progress in brain research* 2011;190:53-68.
13. Budhiraja R, Budhiraja P, Quan SF Sleep-disordered breathing and cardiovascular disorders. *Respiratory care* 2010;55:1322-32; discussion 30-2.
14. Lam JC, Mak JC, Ip MS Obesity, obstructive sleep apnoea and metabolic syndrome. *Respirology* 2012;17:223-36.
15. Kapur V, Blough DK, Sandblom RE et al. The medical cost of undiagnosed sleep apnea. *Sleep* 1999;22:749-55.
16. Kakkar RK, Berry RB Positive airway pressure treatment for obstructive sleep apnea. *Chest* 2007;132:1057-72.
17. Kapur VK Obstructive sleep apnea: diagnosis, epidemiology, and economics. *Respiratory care* 2010;55:1155-67.
18. Lichstein KL, Justin Thomas S, Woosley JA, Geyer JD Co-occurring insomnia and obstructive sleep apnea. *Sleep medicine* 2013;14:824-9.
19. Iranzo A, Santamaria J Severe obstructive sleep apnea/hypopnea mimicking REM sleep behavior disorder. *Sleep* 2005;28:203-6.
20. American Academy of Sleep M. International classification of sleep disorders, 3rd ed. Darien, IL: American Academy of Sleep Medicine, 2014.
21. Berry RB BR, Gamaldo CE, Harding SM, Lloyd RM, Marcus CL and Vaughn BV for the American Academy of Sleep Medicine. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.3 In: Editor, ed.^eds. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.3. City: American Academy of Sleep Medicine, 2016.
22. Ruehland WR, Rochford PD, O'Donoghue FJ, Pierce RJ, Singh P, Thornton AT The new AASM criteria for scoring hypopneas: impact on the apnea hypopnea index. *Sleep* 2009;32:150-7.
23. Redline S, Kapur VK, Sanders MH et al. Effects of varying approaches for identifying respiratory disturbances on sleep apnea assessment. *American journal of respiratory and critical care medicine* 2000;161:369-74.
24. Duce B, Milosavljevic J, Hukins C The 2012 AASM Respiratory Event Criteria Increase the Incidence of Hypopneas in an Adult Sleep Center Population. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2015;11:1425-31.
25. Collop NA, Tracy SL, Kapur V et al. Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2011;7:531-48.
26. Flemons WW, Littner MR, Rowley JA et al. Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society. *Chest* 2003;124:1543-79.
27. Chesson AL, Jr., Ferber RA, Fry JM et al. The indications for polysomnography and related procedures. *Sleep* 1997;20:423-87.
28. Morgenthaler TI, Deriy L, Heald JL, Thomas SM The Evolution of the AASM Clinical Practice Guidelines: Another Step Forward. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2016;12:129-35.
29. Ahmadi N, Shapiro GK, Chung SA, Shapiro CM Clinical diagnosis of sleep apnea based on single night of polysomnography vs. two nights of polysomnography. *Sleep & breathing = Schlaf & Atmung* 2009;13:221-6.
30. Rofail LM, Wong KK, Unger G, Marks GB, Grunstein RR Comparison between a single-channel nasal airflow device and oximetry for the diagnosis of obstructive sleep apnea. *Sleep* 2010;33:1106-14.

31. Amra B, Nouranian E, Golshan M, Fietze I, Penzel T Validation of the persian version of berlin sleep questionnaire for diagnosing obstructive sleep apnea. *International journal of preventive medicine* 2013;4:334-9.
32. Bouloukaki I, Komninos ID, Mermigkis C et al. Translation and validation of Berlin questionnaire in primary health care in Greece. *BMC pulmonary medicine* 2013;13:6.
33. Danzi-Soares NJ, Genta PR, Nerbass FB et al. Obstructive sleep apnea is common among patients referred for coronary artery bypass grafting and can be diagnosed by portable monitoring. *Coronary artery disease* 2012;23:31-8.
34. Friedman M, Wilson MN, Pulver T et al. Screening for obstructive sleep apnea/hypopnea syndrome: subjective and objective factors. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 2010;142:531-5.
35. Kang K, Park KS, Kim JE et al. Usefulness of the Berlin Questionnaire to identify patients at high risk for obstructive sleep apnea: a population-based door-to-door study. *Sleep & breathing = Schlaf & Atmung* 2013;17:803-10.
36. Laporta R, Anandam A, El-Solh AA Screening for obstructive sleep apnea in veterans with ischemic heart disease using a computer-based clinical decision-support system. *Clinical research in cardiology : official journal of the German Cardiac Society* 2012;101:737-44.
37. Pereira EJ, Driver HS, Stewart SC, Fitzpatrick MF Comparing a combination of validated questionnaires and level III portable monitor with polysomnography to diagnose and exclude sleep apnea. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2013;9:1259-66.
38. Ulasli SS, Gunay E, Koyuncu T et al. Predictive value of Berlin Questionnaire and Epworth Sleepiness Scale for obstructive sleep apnea in a sleep clinic population. *The clinical respiratory journal* 2014;8:292-6.
39. Sert Kuniyoshi FH, Zellmer MR, Calvin AD et al. Diagnostic accuracy of the Berlin Questionnaire in detecting sleep-disordered breathing in patients with a recent myocardial infarction. *Chest* 2011;140:1192-7.
40. Subramanian S, Hesselbacher SE, Aguilar R, Surani SR The NAMES assessment: a novel combined-modality screening tool for obstructive sleep apnea. *Sleep & breathing = Schlaf & Atmung* 2011;15:819-26.
41. Best MW, Fitzpatrick M, Milev R, Bowie CR, Jokic R Utility of the Berlin questionnaire for predicting obstructive sleep apnea in individuals with treatment-resistant depression. *Sleep & breathing = Schlaf & Atmung* 2013;17:1221-7.
42. Karakoc O, Akcam T, Genc H, Yetkin S, Piskin B, Gerek M Use of the Berlin Questionnaire to screen at-risk patients for obstructive sleep apnea. *B-Ent* 2014;10:21-5.
43. Suksakorn S, Rattanaumpawan P, Banhiran W, Cherakul N, Chotinaiwattarakul W Reliability and validity of a Thai version of the Berlin questionnaire in patients with sleep disordered breathing. *Journal of the Medical Association of Thailand = Chotmaihet thangphaet* 2014;97 Suppl 3:S46-56.
44. Margallo VS, Muxfeldt ES, Guimaraes GM, Salles GF Diagnostic accuracy of the Berlin questionnaire in detecting obstructive sleep apnea in patients with resistant hypertension. *Journal of hypertension* 2014;32:2030-6; discussion 37.
45. Popevic MB, Milovanovic A, Nagorni-Obradovic L, Nesic D, Milovanovic J, Milovanovic AP Screening commercial drivers for obstructive sleep apnea: translation and validation of Serbian version of Berlin Questionnaire. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2016;25:343-9.
46. Khaledi-Paveh B, Khazaie H, Nasouri M, Ghadami MR, Tahmasian M Evaluation of Berlin Questionnaire Validity for Sleep Apnea Risk in Sleep Clinic Populations. *Basic and clinical neuroscience* 2016;7:43-8.
47. Cowan DC, Allardice G, Macfarlane D et al. Predicting sleep disordered breathing in outpatients with suspected OSA. *BMJ open* 2014;4:e004519.
48. Pataka A, Daskalopoulou E, Kalamaras G, Fekete Passa K, Argyropoulou P Evaluation of five different questionnaires for assessing sleep apnea syndrome in a sleep clinic. *Sleep medicine* 2014;15:776-81.
49. Luo J, Huang R, Zhong X, Xiao Y, Zhou J STOP-Bang questionnaire is superior to Epworth sleepiness scales, Berlin questionnaire, and STOP questionnaire in screening obstructive sleep apnea hypopnea syndrome patients. *Chinese medical journal* 2014;127:3065-70.
50. Kim B, Lee EM, Chung YS, Kim WS, Lee SA The utility of three screening questionnaires for obstructive sleep apnea in a sleep clinic setting. *Yonsei medical journal* 2015;56:684-90.
51. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. *Annals of internal medicine* 1999;131:485-91.
52. Gantner D, Ge JY, Li LH et al. Diagnostic accuracy of a questionnaire and simple home monitoring device in detecting obstructive sleep apnoea in a Chinese population at high cardiovascular risk. *Respirology* 2010;15:952-60.
53. Simpson L, Hillman DR, Cooper MN et al. High prevalence of undiagnosed obstructive sleep apnoea in the general population and methods for screening for representative controls. *Sleep & breathing = Schlaf & Atmung* 2013;17:967-73.
54. Nicholl DD, Ahmed SB, Loewen AH et al. Diagnostic value of screening instruments for identifying obstructive sleep apnea in kidney failure. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2013;9:31-8.
55. Sforza E, Chouchou F, Pichot V, Herrmann F, Barthelemy JC, Roche F Is the Berlin questionnaire a useful tool to diagnose obstructive sleep apnea in the elderly? *Sleep medicine* 2011;12:142-6.
56. Facco FL, Ouyang DW, Zee PC, Grobman WA Development of a pregnancy-specific screening tool for sleep apnea. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2012;8:389-94.
57. Johns MW A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;14:540-5.
58. Chen NH, Chen MC, Li HY, Chen CW, Wang PC A two-tier screening model using quality-of-life measures and pulse oximetry to screen adults with sleep-disordered breathing. *Sleep & breathing = Schlaf & Atmung* 2011;15:447-54.
59. Zou J, Guan J, Yi H et al. An effective model for screening obstructive sleep apnea: a large-scale diagnostic study. *PloS one* 2013;8:e80704.
60. Pecotic R, Dodig IP, Valic M, Ivkovic N, Dogas Z The evaluation of the Croatian version of the Epworth sleepiness scale and STOP questionnaire as screening tools for obstructive sleep apnea syndrome. *Sleep & breathing = Schlaf & Atmung* 2012;16:793-802.
61. Ong TH, Raudha S, Fook-Chong S, Lew N, Hsu AA Simplifying STOP-BANG: use of a simple questionnaire to screen for OSA in an Asian population. *Sleep & breathing = Schlaf & Atmung* 2010;14:371-6.

62. Sadeghniiat-Haghighi K, Montazeri A, Khajeh-Mehrizi A et al. The STOP-BANG questionnaire: reliability and validity of the Persian version in sleep clinic population. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2015;24:2025-30.
63. Alhouqani S, Al Manhali M, Al Essa A, Al-Houqani M Evaluation of the Arabic version of STOP-Bang questionnaire as a screening tool for obstructive sleep apnea. *Sleep & breathing = Schlaf & Atmung* 2015;19:1235-40.
64. BaHammam AS, Al-Aqeel AM, Alhedyani AA, Al-Obaid GI, Al-Owais MM, Olaish AH The Validity and Reliability of an Arabic Version of the STOP-Bang Questionnaire for Identifying Obstructive Sleep Apnea. *The open respiratory medicine journal* 2015;9:22-9.
65. Banhiran W, Durongphan A, Saleesing C, Chongkolwatana C Diagnostic properties of the STOP-Bang and its modified version in screening for obstructive sleep apnea in Thai patients. *Journal of the Medical Association of Thailand = Chotmaihet thangphaet* 2014;97:644-54.
66. Chung F, Subramanyam R, Liao P, Sasaki E, Shapiro C, Sun Y High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *British journal of anaesthesia* 2012;108:768-75.
67. Chung F, Chau E, Yang Y, Liao P, Hall R, Mokhlesi B Serum bicarbonate level improves specificity of STOP-Bang screening for obstructive sleep apnea. *Chest* 2013;143:1284-93.
68. Chung F, Yegneswaran B, Liao P et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008;108:812-21.
69. Gurubhagavatula I, Fields BG, Morales CR et al. Screening for severe obstructive sleep apnea syndrome in hypertensive outpatients. *Journal of clinical hypertension* 2013;15:279-88.
70. Kushida CA, Efron B, Guilleminault C A predictive morphometric model for the obstructive sleep apnea syndrome. *Annals of internal medicine* 1997;127:581-7.
71. Gurubhagavatula I, Maislin G, Pack AI An algorithm to stratify sleep apnea risk in a sleep disorders clinic population. *American journal of respiratory and critical care medicine* 2001;164:1904-9.
72. 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;33:1097-105.
73. Wilson G, Terpening Z, Wong K et al. Screening for sleep apnoea in mild cognitive impairment: the utility of the multivariable apnoea prediction index. *Sleep disorders* 2014;2014:945287.
74. Morales CR, Hurley S, Wick LC et al. In-home, self-assembled sleep studies are useful in diagnosing sleep apnea in the elderly. *Sleep* 2012;35:1491-501.
75. Chang ET, Yang MC, Wang HM, Lai HL Snoring in a sitting position and neck circumference are predictors of sleep apnea in Chinese patients. *Sleep & breathing = Schlaf & Atmung* 2014;18:133-6.
76. Sharma SK, Malik V, Vasudev C et al. Prediction of obstructive sleep apnea in patients presenting to a tertiary care center. *Sleep & breathing = Schlaf & Atmung* 2006;10:147-54.
77. Zerah-Lancner F, Lofaso F, d'Ortho MP et al. Predictive value of pulmonary function parameters for sleep apnea syndrome. *American journal of respiratory and critical care medicine* 2000;162:2208-12.
78. Kolotkin RL, LaMonte MJ, Walker JM, Cloward TV, Davidson LE, Crosby RD Predicting sleep apnea in bariatric surgery patients. *Surgery for obesity and related diseases : official journal of the American Society for Bariatric Surgery* 2011;7:605-10.
79. Platt AB, Wick LC, Hurley S et al. Hits and misses: screening commercial drivers for obstructive sleep apnea using guidelines recommended by a joint task force. *Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine* 2013;55:1035-40.
80. Tsai WH, Remmers JE, Brant R, Flemons WW, Davies J, Macarthur C A decision rule for diagnostic testing in obstructive sleep apnea. *American journal of respiratory and critical care medicine* 2003;167:1427-32.
81. Chai-Coetzer CL, Antic NA, Rowland LS et al. A simplified model of screening questionnaire and home monitoring for obstructive sleep apnoea in primary care. *Thorax* 2011;66:213-9.
82. Firat H Comparison of four established questionnaires to identify highway bus drivers at risk for obstructive sleep apnea in Turkey. *Sleep and Biological Rhythms* 2012;10:231-36.
83. Kuna ST, Gurubhagavatula I, Maislin G et al. Noninferiority of functional outcome in ambulatory management of obstructive sleep apnea. *American journal of respiratory and critical care medicine* 2011;183:1238-44.
84. Rosen CL, Auckley D, Benca R et al. A multisite randomized trial of portable sleep studies and positive airway pressure autotitration versus laboratory-based polysomnography for the diagnosis and treatment of obstructive sleep apnea: the HomePAP study. *Sleep* 2012;35:757-67.
85. Berry RB, Hill G, Thompson L, McLaurin V Portable monitoring and autotitration versus polysomnography for the diagnosis and treatment of sleep apnea. *Sleep* 2008;31:1423-31.
86. Campbell AJ, Neill AM Home set-up polysomnography in the assessment of suspected obstructive sleep apnea. *Journal of sleep research* 2011;20:207-13.
87. Gjevre JA, Taylor-Gjevre RM, Skomro R, Reid J, Fenton M, Cotton D Comparison of polysomnographic and portable home monitoring assessments of obstructive sleep apnea in Saskatchewan women. *Canadian respiratory journal* 2011;18:271-4.
88. Masa JF, Corral J, Pereira R et al. Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome. *Thorax* 2011;66:567-73.
89. Polese JF, Santos-Silva R, de Oliveira Ferrari PM, Sartori DE, Tufik S, Bittencourt L Is portable monitoring for diagnosing obstructive sleep apnea syndrome suitable in elderly population? *Sleep & breathing = Schlaf & Atmung* 2013;17:679-86.
90. Santos-Silva R, Sartori DE, Truksinas V et al. Validation of a portable monitoring system for the diagnosis of obstructive sleep apnea syndrome. *Sleep* 2009;32:629-36.
91. Yin M, Miyazaki S, Ishikawa K Evaluation of Type 3 Portable Monitoring in Unattended Home Setting for Suspected Sleep Apnea: Factors That May Affect Its Accuracy. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 2006;134:204-09.
92. Planes C, Leroy M, Bouach Khalil N et al. Home diagnosis of obstructive sleep apnoea in coronary patients: validity of a simplified device automated analysis. *Sleep & breathing = Schlaf & Atmung* 2010;14:25-32.

93. Masa JF, Corral J, Pereira R et al. Effectiveness of sequential automatic-manual home respiratory polygraphy scoring. *The European respiratory journal* 2013;41:879-87.
94. Garcia-Diaz E, Quintana-Gallego E, Ruiz A et al. Respiratory polygraphy with actigraphy in the diagnosis of sleep apnea-hypopnea syndrome. *Chest* 2007;131:725-32.
95. Oliveira MG, Treptow EC, Fukuda C et al. Diagnostic accuracy of home-based monitoring system in morbidly obese patients with high risk for sleep apnea. *Obesity surgery* 2015;25:845-51.
96. Ayappa I, Norman RG, Seelall V, Rapoport DM Validation of a self-applied unattended monitor for sleep disordered breathing. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2008;4:26-37.
97. Tonelli de Oliveira AC, Martinez D, Vasconcelos LF et al. Diagnosis of obstructive sleep apnea syndrome and its outcomes with home portable monitoring. *Chest* 2009;135:330-6.
98. Ward KL, McArdle N, James A et al. A Comprehensive Evaluation of a Two-Channel Portable Monitor to "Rule in" Obstructive Sleep Apnea. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2015;11:433-44.
99. Baltzan MA, Verschelden P, Al-Jahdali H, Olha AE, Kimoff RJ Accuracy of Oximetry with Thermistor (OxiFlow) for Diagnosis of Obstructive Sleep Apnea and Hypopnea. *Sleep* 2000;33:1-9.
100. Masdeu MJ, Ayappa I, Hwang D, Mooney AM, Rapoport DM Impact of clinical assessment on use of data from unattended limited monitoring as opposed to full-in lab PSG in sleep disordered breathing. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2010;6:51-8.
101. Nakano H, Tanigawa T, Ohnishi Y et al. Validation of a single-channel airflow monitor for screening of sleep-disordered breathing. *The European respiratory journal* 2008;32:1060-7.
102. Ozmen OA, Tuzemen G, Kasapoglu F et al. The reliability of SleepStrip as a screening test in obstructive sleep apnea syndrome. *Kulak burun bogaz ihtisas dergisi : KBB = Journal of ear, nose, and throat* 2011;21:15-9.
103. Pang KP, Dillard TA, Blanchard AR, Gourin CG, Podolsky R, Terris DJ A comparison of polysomnography and the SleepStrip in the diagnosis of OSA. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery* 2006;135:265-8.
104. Watkins MR, TaImage JB, Thiese MS, Hudson TB, Hegmann KT Correlation between screening for obstructive sleep apnea using a portable device versus polysomnography testing in a commercial driving population. *Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine* 2009;51:1145-50.
105. O'Brien LM, Bullough AS, Shelgikar AV, Chames MC, Armitage R, Chervin RD Validation of Watch-PAT-200 against polysomnography during pregnancy. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2012;8:287-94.
106. Pittman SD, Ayas NT, MacDonald MM, Malhotra A, Fogel RB, White DP Using a wrist-worn device based on peripheral arterial tonometry to diagnose obstructive sleep apnea: in-laboratory and ambulatory validation. *Sleep* 2004;27:923-33.
107. Garg N, Rolle AJ, Lee TA, Prasad B Home-based diagnosis of obstructive sleep apnea in an urban population. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2014;10:879-85.
108. Pittman SD, Pillar G, Berry RB, Malhotra A, MacDonald MM, White DP Follow-up assessment of CPAP efficacy in patients with obstructive sleep apnea using an ambulatory device based on peripheral arterial tonometry. *Sleep & breathing = Schlaf & Atmung* 2006;10:123-31.
109. Andreu AL, Chiner E, Sancho-Chust JN et al. Effect of an ambulatory diagnostic and treatment programme in patients with sleep apnoea. *The European respiratory journal* 2012;39:305-12.
110. Antic NA, Buchan C, Esterman A et al. A randomized controlled trial of nurse-led care for symptomatic moderate-severe obstructive sleep apnea. *American journal of respiratory and critical care medicine* 2009;179:501-8.
111. Mulgrew AT, Fox N, Ayas NT, Ryan CF Diagnosis and initial management of obstructive sleep apnea without polysomnography: a randomized validation study. *Annals of internal medicine* 2007;146:157-66.
112. Skomro RP, Gjevre J, Reid J et al. Outcomes of home-based diagnosis and treatment of obstructive sleep apnea. *Chest* 2010;138:257-63.
113. de Vries GE, van der Wal HH, Kerstjens HA et al. Validity and Predictive Value of a Portable Two-Channel Sleep-Screening Tool in the Identification of Sleep Apnea in Patients With Heart Failure. *Journal of cardiac failure* 2015;21:848-55.
114. Chung F, Liao P, Elsaid H, Islam S, Shapiro CM, Sun Y Oxygen desaturation index from nocturnal oximetry: a sensitive and specific tool to detect sleep-disordered breathing in surgical patients. *Anesthesia and analgesia* 2012;114:993-1000.
115. Deutsch PA, Simmons MS, Wallace JM Cost-effectiveness of split-night polysomnography and home studies in the evaluation of obstructive sleep apnea syndrome. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2006;2:145-53.
116. Chervin RD, Murman DL, Malow BA, Totten V Cost-utility of three approaches to the diagnosis of sleep apnea: polysomnography, home testing, and empirical therapy. *Annals of internal medicine* 1999;130:496-505.
117. Pietzsch JB, Garner A, Cipriano LE, Linehan JH An integrated health-economic analysis of diagnostic and therapeutic strategies in the treatment of moderate-to-severe obstructive sleep apnea. *Sleep* 2011;34:695-709.
118. Kim RD, Kapur VK, Redline-Bruch J et al. An Economic Evaluation of Home Versus Laboratory-Based Diagnosis of Obstructive Sleep Apnea. *Sleep* 2015;38:1027-37.
119. Masa JF, Corral J, Sanchez de Cos J et al. Effectiveness of three sleep apnea management alternatives. *Sleep* 2013;36:1799-807.
120. Series F, Kimoff RJ, Morrison D et al. Prospective evaluation of nocturnal oximetry for detection of sleep-related breathing disturbances in patients with chronic heart failure. *Chest* 2005;127:1507-14.
121. Oliveira MG, Nery LE, Santos-Silva R et al. Is portable monitoring accurate in the diagnosis of obstructive sleep apnea syndrome in chronic pulmonary obstructive disease? *Sleep medicine* 2012;13:1033-8.
122. Abraham WT, Trupp RJ, Phillilps B et al. Validation and clinical utility of a simple in-home testing tool for sleep-disordered breathing and arrhythmias in heart failure: results of the Sleep Events, Arrhythmias, and Respiratory Analysis in Congestive Heart Failure (SEARCH) study. *Congestive heart failure* 2006;12:241-7; quiz 48-9.

123. Khawaja IS, Olson EJ, van der Walt C et al. Diagnostic accuracy of split-night polysomnograms. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 2010;6:357-62.
124. Chou KT, Chang YT, Chen YM et al. The minimum period of polysomnography required to confirm a diagnosis of severe obstructive sleep apnoea. *Respirology* 2011;16:1096-102.
125. Chuang ML, Lin IF, Vintch JR, Liao YF Predicting continuous positive airway pressure from a modified split-night protocol in moderate to severe obstructive sleep apnea-hypopnea syndrome. *Internal medicine* 2008;47:1585-92.
126. Collen J, Holley A, Lettieri C, Shah A, Roop S The impact of split-night versus traditional sleep studies on CPAP compliance. *Sleep & breathing = Schlaf & Atmung* 2010;14:93-9.
127. Kristo DA, Shah AA, Lettieri CJ et al. Utility of split-night polysomnography in the diagnosis of upper airway resistance syndrome. *Sleep & breathing = Schlaf & Atmung* 2009;13:271-5.
128. Yamashiro Y, Kryger MH CPAP titration for sleep apnea using a split-night protocol. *Chest* 1995;107:62-6.
129. Ciftci B, Ciftci TU, Guven SF [Split-night versus full-night polysomnography: comparison of the first and second parts of the night]. *Archivos de bronconeumologia* 2008;44:3-7.
130. Elshaug AG, Moss JR, Southcott AM Implementation of a split-night protocol to improve efficiency in assessment and treatment of obstructive sleep apnoea. *Internal medicine journal* 2005;35:251-4.
131. Gouveris H, Selivanova O, Bausmer U, Goepel B, Mann W First-night-effect on polysomnographic respiratory sleep parameters in patients with sleep-disordered breathing and upper airway pathology. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies* 2010;267:1449-53.
132. Ma J, Zhang C, Zhang J et al. Prospective study of first night effect on 2-night polysomnographic parameters in adult Chinese snorers with suspected obstructive sleep apnea hypopnea syndrome. *Chinese medical journal* 2011;124:4127-31.
133. Selwa LM, Marzec ML, Chervin RD et al. Sleep staging and respiratory events in refractory epilepsy patients: Is there a first night effect? *Epilepsia* 2008;49:2063-8.