TITLE: Diagnosis of Snoring and Obstructive Sleep Apnea: A Review of the Accuracy

DATE: 17 April 2009

CONTEXT AND POLICY ISSUES:

Diagnosis of Snoring and Obstructive Sleep Apnea

Snoring is common in the adult population, and it is associated with varying levels of upper airway resistance.\(^1\) Sleep may not be interrupted at lower levels of airway resistance, but with increased resistance transient arousals from sleep will occur due to the effort required to maintain ventilation. This is known as upper airways resistance syndrome (UARS). With further increases in resistance, patients may experience episodes of hypoventilation and oxygen desaturation, known as obstructive sleep apnea-hypopnea syndrome (OSA).\(^1\)

The 2006 Canadian guidelines for the diagnosis of sleep-disordered breathing in adults define the severity of OSA using the apnea and hypopnea index (AHI).\(^2\) Apnea is defined as an event that lasts 10 seconds or longer, and is a reduction in airflow (nasal pressure or respiratory inductance) greater than 50% from baseline, or a reduction in airflow less than 50% accompanied by arousal or oxygen desaturation of 4% or more. Hypopnea has been defined as an abnormal respiratory event lasting at least 10 seconds with at least 4% oxygen desaturation, and with a minimum reduction in airflow or thoracoabdominal movement of 30%. Mild OSA is five to 15 events of apnea or hypopnea per hour of sleep; moderate OSA is classified as 15 to 30 events per hour; and severe disease is more than 30 events per hour.\(^2\)

OSA is more common in those who snore, in men, and in patients with obesity, hypertension, and physical abnormalities of the upper airways.\(^3\) The definition of primary snoring specifically excludes those symptoms associated with OSA and UARS, including sleep disruption, insomnia, or daytime sleepiness. However, it is difficult to differentiate between primary snoring and OSA from clinical symptoms alone. The diagnosis of snoring is made when OSA and UARS are ruled out.\(^1\)
The 2006 Canadian guidelines for the diagnosis of sleep-disordered breathing in adults state that laboratory polysomnography (PSG) is the accepted standard for evaluation and diagnosis of OSA. However, PSG has limitations, as it requires experienced personnel to evaluate the results, it is not readily available to all patients, and the testing itself can be expensive. Most countries can currently meet approximately 10% of the demand for PSG testing in patients with suspected OSA. Portable monitors to diagnose OSA may be less expensive and more readily accessible alternatives to PSG, and they can be used in the home. However, the ability of the monitors to diagnose OSA is highly variable.

### Imaging Techniques for Snoring and OSA Appliances

Oral appliances are being used more frequently for treatment of patients with OSA. Radiologic imaging techniques, such as cephalometric X-rays, computed tomography (CT), and magnetic resonance imaging (MRI), have been used to assess airway anatomy differences in patients with OSA and other sleep-disordered breathing (not including snoring). Some of these imaging techniques have also been used for the assessment of oral appliances for the treatment of OSA.

### RESEARCH QUESTIONS:

1. What is the accuracy and reliability of devices for diagnosis of snoring and do they accurately rule out obstructive sleep apnea?

2. What is the evidence that imaging techniques are necessary for diagnosis of snoring or obstructive sleep apnea or before treatment with an oral appliance following diagnosis?

### METHODS:

A limited literature search was conducted on key health technology assessment resources, including PubMed, The Cochrane Library (Issue 1, 2008), University of York Centre for Reviews and Dissemination (CRD) databases, ECRI, EuroScan, international health technology agencies, and a focused Internet search. Results include articles published between 2004 and March 2009, and are limited to English language publications only. Filters were applied to limit the retrieval to health technology assessments, systematic reviews, meta-analyses, randomized controlled trials, and guidelines.

### SUMMARY OF FINDINGS FOR RESEARCH QUESTION #1:

The literature search identified two meta-analyses and two evidence-based guidelines. Additional studies reporting on some of the portable devices currently used for the diagnosis of OSA are located in the appendix.

### Systematic reviews and meta-analyses

**Pre-operative screening tests for OSA**

Obstructive sleep apnea carries a high risk of perioperative morbidity and mortality. Polysomnography is recommended by the American Society of Anesthesiologists for preoperative screening when indicated, but it is not always possible. Anesthetists therefore, require screening tests to rule out OSA in patients with a low-risk of the syndrome. A meta-analysis was performed on pre-operative screening tests for OSA. Two independent
researchers reviewed articles from 1966 to 2008 that measured the diagnostic value of overnight PSG versus questionnaires, clinical models (history and physical exam with or without other measurements or investigations) and prediction equations (regression equations or algorithms). The investigators excluded studies that did not use overnight monitored PSG as the reference standard and that did not provide prevalence of OSA information with raw data in 2 x 2 tables or did not provide sensitivity and specificity or positive or negative likelihood ratios. The authors also excluded studies that did not have a complete description of the clinical methods; however, this was not clearly defined in the meta-analysis. 

There were 26 studies in the final analysis, including eight questionnaires and 18 clinical prediction tests (algorithms, regression models, and neural networks). Each test was evaluated for the frequency of true or false positive or negative diagnoses. From this data, the tests were assigned a diagnostic odds ratio (DOR), based on their specificity and sensitivity. The false negative rate (FN) was calculated as one minus the sensitivity. Tests with a DOR of 10 to 80 were classified as good, those greater than 81 were rated as excellent, and an ideal test had a DOR more than 81 and an FN rate of zero.

The results of the analysis showed that no single questionnaire or clinical model met all the criteria for the ideal preoperative screening tool. The most accurate tools were the Berlin and Sleep Disorders questionnaires, the Kushida index (morphometry), and the Battagel combined clinical-cephalometry clinical model. However, the authors proposed that due to the high degree of heterogeneity and FN rates, a significant proportion of patients with OSA would be missed by all questionnaires and most clinical models. For example, although the Berlin questionnaire was the most accurate with FN rates ranging from 14.5 to 38.2, it would not be able to conclusively rule out OSA pre-operatively. As well, the accuracy of these tests was not readily reproduced in multiple validation studies. Limitations of this meta-analysis include an unclear definition for the clinical methods exclusion criteria and the significant heterogeneity between studies included in the analysis.

**Comparison of Portable Sleep Monitors with Polysomnography**

Ghegan and colleagues performed a meta-analysis of studies comparing portable (home) sleep monitors with laboratory PSG for evaluation of OSA. The primary objective was the difference in the respiratory disturbance index (RDI), with the null hypothesis that there would be no difference between the two methods. Secondary objectives included a comparison of the mean low oxygen saturation index (LSAT), recorded sleep time, percentage of inadequate studies, and cost. The investigators included studies that had simultaneous or sequential portable monitoring and PSG, and also included one of the secondary outcome parameters.

The authors included 18 studies in their analysis, including the Watch_PAT and Apnoescreen-II (further details provided in the appendices). For the RDI, pooled data on 12 studies demonstrated that the portable monitors were 10% lower on average than with lab PSG tests [odds ratio (OR) 0.9; 95% Confidence Interval (CI) 0.87 to 0.92]. Combined data of three studies revealed no significant difference for the mean LSAT between portable and laboratory PSG monitoring (OR 1.0; 95% CI 0.94 to 1.1). In five studies, the laboratory sleep times were on average 13% longer compared with home studies (OR 0.87; 95% CI 0.86 to 0.89). There was a significant degree of heterogeneity (p=0.000) between studies for average sleep time. A cost-comparison from four countries showed that portable sleep monitors were 35 to 88% less than laboratory studies.
The authors concluded that home sleep monitors may underestimate the severity of sleep apnea, although the lower costs make them useful screening tools for OSA.

Guidelines and recommendations

The 2006 Canadian guidelines for the diagnosis of sleep-disordered breathing (refers to OSA, central sleep apnea-hypopnea syndrome, and sleep hypoventilation syndrome) in adults categorize complete laboratory PSG as level I for diagnostic procedures and assigned it a level of evidence of C (case control or cohort studies with a risk of bias). Portable monitoring devices are classified as level III sleep studies and are useful for improving access to diagnosis of sleep-disordered breathing (level C evidence). Portable monitoring devices can also be used to confirm OSA diagnosis in patients with a moderate-high pre-test probability of the illness, but they have limited use with other forms of sleep-disordered breathing and co-morbid illnesses (level C evidence).

These findings are aligned with the American Academy of Sleep Medicine guidelines. In 2007, the Academy issued guidelines for the use of portable monitors in the diagnosis of OSA in 2007. The guidelines state that lab PSG monitoring is the standard of practice for the diagnosis of OSA. They indicated that portable monitors should only be used for the diagnosis of OSA in conjunction with a comprehensive sleep evaluation by a certified sleep practitioner. The guidelines also state that portable monitors may be used an alternative to PSG only in patients with a high pre-test probability of moderate to severe OSA, nor is it appropriate for patients with significant co-morbid illnesses (e.g. congestive heart failure).

The 2006 Canadian guidelines state that clinical predication formulas can be used to assess the probability of sleep-disordered breathing and to prioritize patients for further evaluation, but are insufficient to establish a diagnosis (level C evidence).

SUMMARY OF FINDINGS FOR RESEARCH QUESTION #2:

Guidelines and recommendations

Canadian Thoracic Society 2006 sleep-disordered breathing guidelines for adults state that oral appliances are appropriate first-line treatment for patients with mild to moderate OSAHS who have minimal daytime symptoms (level A evidence – high-quality meta-analysis or single randomized controlled trial with a low risk of bias). The guidelines recommend follow-up treatment to ensure the effectiveness of the appliance, but they do not make any recommendations for the use of imaging techniques.

The American Academy of Sleep Medicine issued practice parameters for treatment of snoring and OSA with oral appliances in 2005. The guidelines state that cephalometric evaluation is not always necessary for patients prescribed an appliance. They recommend that patients undergo PSG testing or attend a cardiorespiratory sleep study after the appliance has been fitted. The guidelines did not make any recommendations for the use of other imaging techniques.

LIMITATIONS

There is a paucity of evidence-based data for the diagnosis of snoring, OSA, and for the use of radiological imaging techniques for OSA diagnosis and oral appliances. There are no current evidence-based studies, analyses, or guidelines for the diagnosis of primary snoring. There is one meta-analysis that reviews portable monitors compared with polysomnography, and one
meta-analysis for preoperative screening tools. The available literature on radiological imaging was limited. The Canadian guidelines do offer levels of evidence ratings for their recommendations, but the American guidelines do not.

CONCLUSIONS AND IMPLICATIONS FOR DECISION OR POLICY MAKING:

There is a lack of evidence-based articles for the diagnosis of primary snoring, therefore the included studies primarily focus on devices used for the diagnosis of OSA (the diagnosis of primary snoring is made when OSA is excluded). Both the American and Canadian guidelines state that polysomnography is the gold standard for the diagnosis of OSA. Portable monitoring devices should only be used in patients with a moderate to high probability of the syndrome, and without any severe co-morbid illnesses.\(^2\)\(^9\) The meta-analysis by Ghegan and colleagues also concluded that portable monitoring devices may underestimate the severity of sleep apnea.\(^4\) Therefore, portable devices should only be used to screen snoring patients who are very likely to have OSA, and the diagnosis of OSA should be confirmed with PSG monitoring.

The routine use of radiologic imaging techniques for use in diagnosis of OSA or when fitting oral appliances is not supported by the Canadian or American guidelines.\(^2\)\(^6\). Further information is required regarding the use of imaging techniques in the diagnosis of obstructive sleep apnea. Additional research is required to develop monitoring devices or screening tools that will accurately rule out OSA in patients who snore.

PREPARED BY:
Verla Chatsis, BA, BSP, ACPR
Carolyn Spry, MLIS, Information Specialist
Health Technology Inquiry Service
Email: htis@cadth.ca
Tel: 1-866-898-8439
REFERENCES:


APPENDIX:
Controlled trials or observational studies were not part of the search criteria for this report. Below are the details of some studies on select portable devices. This does not represent a comprehensive review of portable monitors that are currently used for the diagnosis of OSA.

Table 1: Portable monitors for diagnosis of OSA – observational studies

<table>
<thead>
<tr>
<th>Device / Manufacturer / Author(s) / Patient Population / Setting</th>
<th>Interventions and Comparators</th>
<th>Outcome Measures</th>
<th>Main Findings</th>
</tr>
</thead>
</table>
| Apnoescreen II (AP-II) with actigraphy (Erich Jaeger GMBH & CoKg, Germany) García-Díaz et al (2007)                              | - Patients had simultaneous overnight PSG and AP-II monitoring in the lab  
- Within 15 days of the first test, patients had AP-II monitoring at home  
- Apnea (for PSG & AP-II): Cessation of airflow for ≥ 10 seconds  
- PSIG hypopnea: A reduction in airflow by ≥ 50% for ≥ 10 seconds and with at least 4% O₂ desaturation or the presence of an arousal  
- AP-II hypopnea: A reduction in airflow by ≥ 50% for ≥ 10 seconds and with at least 4% O₂ desaturation  
- Respiratory disturbance index (RDI): Number of apneic and hypopneic episodes per hour of sleep estimated by AP-II actigraphy (vs AHI for PSG) | - To compare sleep parameters between AP-II and PSG  
- The sensitivity/specificity and other parameters for diagnosis of SAHS  
- The agreement between PSG AHI and AP-II RDI                                                                                                                                 | - Three patients were excluded from the results  
- See table 2 for sensitivity and specificity data  
- Mean difference between PSG AHI and home AP-II RDI: Observer A: 3.1 ± 17 (95% CI: -1.1 to 7.5)  
Observer B: 1.6 ± 16.4 (95% CI: -2.5 to 5.8)  
- Mean difference between PSG AHI and lab AP-II RDI: Observer A: 2.8 ± 10.5 (95% CI: 0.13 to 5.5)  
Observer B: 1.03 ± 10.3 (95% CI: -1.6 to 3.6) |
| ARES Unicorder Ayappa et al (2008)                                                                                             | - All patients were provided with written instructions (no in-person training) for two nights of home-based Ares monitoring, with  
- Within two weeks, patients had overnight simultaneous PSG and ARES monitoring in a sleep lab.  
- The lab PSG monitoring was either a full night or split into a PSG diagnostic study and CPAP titration  
- Apnea: Reduction in airflow amplitude to 10% of baseline | To compare AHI and RDI for:  
- Lab ARES versus PSG  
- Home ARES versus lab PSG  
- Automated analysis versus tech-edited data from ARES lab and home                                                                 | - There were 88/102 successful home ARES recordings, and 96/102 patients returned for the lab monitoring  
- AHI 4% was higher in the lab than home (p=0.05), as was the AHI 1% (p=0.01)  
- Technical review showed little change in diagnostic accuracy  
- See table 2 for sensitivity and specificity data for AHI |
Diagnosis of Snoring and Obstructive Sleep Apnea

- Exclusion criteria included the inability to read English, or inability to wear the device on the forehead.
- Technician analyzing PSG data was blinded to ARES results.

### SleepStrip
(Roxon Medi-Tech Inc.)
Pang et al (2006)\(^{12}\)
- 39 adults with suspected OSA enrolled (17 M, 22 F)
- The patients’ results were read by two blinded investigators

All patients had overnight attended level 1 PSG in the hospital, and then used the Sleep Strip at home the night after the PSG.

**SleepStrip definitions**
- Apnea: Cessation of airflow for ≥ 10 seconds
- Hypopnea: A reduction in airflow for ≥ 10 seconds and reduction in thoraco-abdominal movement or airflow, and with at least 4% O\(_2\) desaturation or the presence of an arousal

**Outcome Measures**
- To assess the level of agreement between SleepStrip™ scores and lab-based AHI
- To calculate the sensitivity/ specificity for SleepStrip™ diagnosis
- Whether the duration of sleep differed between the two tests

**Main Findings**
- 37 patients completed the study
- Five SleepStrips™ reported an error
- Correlation between AHI and SleepStrip™ results was 0.139 (P = 0.19)
- See table 2 for sensitivity and specificity data
- Patient-reported sleep time was higher with the SleepStrip™ than PSG (P = 0.052)

### Somnocheck
(Weinmann GmbH, Germany)
Tonelli de Oliveira et al (2009)\(^{5}\)
157 patients with suspected OSA (113 M, 44 F)
- Pregnant women and those with severe coexisting diseases (e.g. cancer) were excluded

**Outcome Measures**
- Concordance between PSG and Somnocheck monitoring
- Concordance between lab and home Somnocheck monitoring
- Accuracy of Somnocheck as described by sensitivity and specificity and other parameters

**Main Findings**
- 8 patients were excluded from the lab Somnocheck analysis, and 36 patients from the home analysis
- The overall concordance between PSG and Somnocheck was 0.64 (95% CI: 0.55 to 0.74) for the lab test and 0.53 (0.42 tp 0.63) for the home test (P < 0.001)
- See table 2 for sensitivity and specificity data
- Correlation between lab and home
Diagnosis of Snoring and Obstructive Sleep Apnea

Device / Manufacturer / Author(s) / Patient Population / Setting

Interventions and Comparators

Outcome Measures

Main Findings


− All patients had overnight attended level 1 PSG in the hospital, and wore the WatchPAT simultaneously
− Apnea: Cessation of airflow for ≥ 10 seconds
− Hypopnea: A reduction in airflow for ≥ 10 seconds and reduction in thoraco-abdominal movement or airflow, and with at least 4% O₂ desaturation or the presence of an arousal

Correlation between PSG and WatchPAT for AHI, sleep time, lowest O₂ saturation (LSAT), and REM and O₂ percentage.

− Five patients excluded with an inadequate sleep time
− AHI correlation between PSG and WatchPAT was 0.9288 (95% CI: 0.8579 to 0.9650; P < 0.0001)
− LSAT correlation between PSG and WatchPAT was 0.9891 (95% CI: 0.9773 to 0.9947; P < 0.0001)
− Sleep time correlation was 0.5815 (P = 0.005)
− See table 2 for sensitivity and specificity data
− The outcomes for REM and O₂ percentage were not reported

Legend: AHI = apnea-hypopnea index (sum of the apneic + hypopneic events / hour of sleep); CPAP = continuous positive airway pressure; F = females; M = men; O₂ = oxygen; p = probability; SAHS = sleep apnea-hypopnea syndrome

Table 2: Sensitivity and Specificity of Portable OSA Monitoring Devices versus PSG

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apnoescreen II (Garcia-Diaz et al, 2007)10</td>
<td>≥ 10: Lab: 94.6 (87.3 to 100) Home: 91.9 (83.1 to 100)</td>
<td>Lab: 100 Home: 93.8 (86.9 to 100)</td>
</tr>
<tr>
<td></td>
<td>≥ 15: Lab: 100</td>
<td>Lab: 96.7 (90.2 to 100) Home: 95.8 (87.8 to 100)</td>
</tr>
<tr>
<td></td>
<td>≥ 30: Lab: 95.8 (87.8 to 100)</td>
<td>Home: 95.8 (87.8 to 100)</td>
</tr>
<tr>
<td>ARES (Ayappa et al, 2008)11</td>
<td>≥ 5: Lab: 0.98 (0.89 to 1.00) Home: 0.90 (0.78 to 0.96)</td>
<td>Lab: 0.84 (0.67 to 0.93) Home: 0.79 (0.62 to 0.91)</td>
</tr>
<tr>
<td></td>
<td>≥ 10: Lab: 0.97 (0.85 to 1.00) Home: 0.86 (0.70 to 0.95)</td>
<td>Lab: 0.85 (0.72 to 0.93) Home: 0.82 (0.67 to 0.91)</td>
</tr>
<tr>
<td></td>
<td>≥ 15: Lab: 0.92 (0.76 to 0.98) Home: 0.74 (0.56 to 0.87)</td>
<td>Lab: 0.95 (0.84 to 0.99) Home: 0.88 (0.75 to 0.95)</td>
</tr>
<tr>
<td>SleepStrip (Pang et al, 2006)12</td>
<td>&gt; 15: 0.55</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>&gt; 25: 0.44</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>&gt; 40: 0.33</td>
<td>0.95</td>
</tr>
<tr>
<td>Somnocheck at home (Tonelli de Oliveira et al, 2009)5</td>
<td>≥ 5: 96.15 (92.5 to 99.8)</td>
<td>64.7 (42 to 87.4)</td>
</tr>
<tr>
<td></td>
<td>≥ 10: 90.7 (82.7 to 95.2)</td>
<td>82.9 (67.3 to 91.9)</td>
</tr>
<tr>
<td></td>
<td>≥ 15: 81.3 (71.1 to 88.5)</td>
<td>82.6 (69.3 to 90.9)</td>
</tr>
<tr>
<td></td>
<td>≥ 30: 80.6 (68.3 to 91.7)</td>
<td>92.1 (86 to 98.2)</td>
</tr>
<tr>
<td>WatchPAT (Pang et al, 2007)13</td>
<td>&gt; 5: 0.94, P = 0.26</td>
<td>0.8, P = 0.26</td>
</tr>
<tr>
<td></td>
<td>&gt; 15: 0.96, P = 0.25</td>
<td>0.79, P = 0.25</td>
</tr>
<tr>
<td></td>
<td>&gt; 35: 0.83, P = 0.05</td>
<td>0.72, P = 0.05</td>
</tr>
</tbody>
</table>

Legend: AHI = apnea-hypopnea index; CI = confidence interval