CADTH Health Technology Review

At-Home Polysomnography Versus In-Clinic Polysomnography for Sleep Disorders

Chantelle C. Lachance
Sharon Bailey
# Table of Contents

Abbreviations .................................................................................. 4  
Research Questions .......................................................................... 5  
Key Messages .................................................................................. 5  
Context and Policy Issues ................................................................. 5  
  What Are Sleep Disorders? ............................................................. 5  
  What Is the Current Practice? ......................................................... 6  
  What Is a Level 2 PSG, and How May It Benefit? .......................... 7  
Objective ......................................................................................... 7  
Methods ........................................................................................... 7  
  Literature Search Methods ............................................................. 7  
  Selection Criteria and Methods ..................................................... 8  
  Exclusion Criteria .......................................................................... 8  
  Critical Appraisal of Individual Studies ........................................ 8  
Summary of Evidence ....................................................................... 8  
  Quantity of Research Available .................................................... 8  
  Summary of Study Characteristics ............................................... 9  
  Summary of Critical Appraisal ....................................................... 9  
  Summary of Findings ..................................................................... 10  
Limitations ....................................................................................... 12  
  Overall Completeness of the Evidence ......................................... 12  
  Generalizability of the Findings .................................................... 12  
Conclusions and Implications for Decision- or Policy-Making ......... 13  
References ....................................................................................... 14  
Appendix 1: Selection of Included Studies ....................................... 16  
Appendix 2: Characteristics of Included Publications .................... 17  
Appendix 3: Critical Appraisal of Included Publications ................. 19  
Appendix 4: Main Study Findings .................................................... 21  
Appendix 5: References of Potential Interest .................................. 22
List of Tables

Table 1: Selection Criteria ................................................................. 8
Table 2: Characteristics of Included Diagnostic Test Accuracy Study .......... 17
Table 3: Characteristics of Included Primary Clinical Utility Study .................. 17
Table 4: Strengths and Limitations of Diagnostic Test Accuracy Study Using the QUADAS-2 Checklist 18 .... 19
Table 5: Strengths and Limitations of Clinical Utility Study Using the Downs and Black Checklist 19 .............. 20
Table 6: Summary of Findings by Outcome — Diagnostic Test Accuracy for Apneas and Hypopneas (AHI) Detection ................................................................. 21
Table 7: Summary of Findings by Outcome — Clinical Utility of Level 2 PSG ................................................................. 21

List of Figures

Figure 1: Selection of Included Studies ................................................................. 16
Abbreviations

AHI   apnea-hypopnea index
CPAP  continuous positive airway pressure
EEG   electroencephalogram
ICSD  International Classification of Sleep Disorders
OSA   obstructive sleep apnea
PSG   polysomnography
QUADAS-2 Quality Assessment of Diagnostic Accuracy Studies 2
SF-36  Short Form (36) Health Survey
Research Questions

1. What is the diagnostic test accuracy of Level 2 Polysomnography for screening or diagnosis of any sleep disorders in adults?
2. What is the clinical utility of Level 2 Polysomnography for screening or diagnosis of any sleep disorders in adults?
3. What is the cost-effectiveness of Level 2 Polysomnography for the screening or diagnosis of sleep disorders in adults?

Key Messages

- Level 2 PSG may have moderate accuracy compared to Level 1 PSG for diagnosing obstructive sleep apnea. This conclusion was based on 1 study with limitations in reporting that reduce our certainty in the findings.
- There may be no significant differences in daytime sleepiness, hypertension, treatment adherence, and most quality-of-life measures for patients with OSA who were diagnosed with a Level 2 PSG compared with a Level 1 PSG (1 study).
- We did not find any studies on the diagnostic test accuracy or clinical utility of Level 2 PSG for the screening or diagnosing of other sleep disorders (e.g., central disorders of hypersomnolence, sleep-related movement disorders, parasomnias) that met our inclusion criteria.
- We did not find any studies on the cost-effectiveness of Level 2 PSG for screening or diagnosis of any sleep disorders in adults that met our inclusion criteria.

Context and Policy Issues

What Are Sleep Disorders?
A common complaint in the general population is excessive daytime sleepiness. The International Classification of Sleep Disorders (ICSD) defines excessive daytime sleepiness as the inability to maintain wakefulness and alertness during the major waking episodes of the day, with sleep occurring unintentionally or at inappropriate times almost daily for at least 3 months. A longitudinal study used the validated Epworth Sleepiness Scale to find that 33% of respondents living in Quebec reported excessive daytime sleepiness. Other common complaints related to sleep include snoring and associated events (e.g., gasping, choking), reported by patient and/or bed partner, and morning headaches. These clinical presentations can indicate certain sleep disorders or other medical conditions (e.g., heart disease, lung disease, depression).

Sleep disorders are a group of conditions that disturb normal sleep patterns. Sleep disorders among adults are 1 of the most common clinical problems encountered, and can affect the quality, timing, and amount of sleep resulting in daytime distress and impairment in functioning. For example, survey data from the Canadian Longitudinal Study on Aging (n = 27,210) suggests that nearly 20% of adults older than 45 years
old are at risk of obstructive sleep apnea (OSA) in Canada. Certain sleep conditions can also perpetuate adverse outcomes such as mortality, cardiovascular disease, mental and cognitive disorders, accidents, and injuries. The ICSD includes 6 categories of sleep disorders with a spectrum of disorders classified within each category:

- sleep-disordered breathing, including OSA, central sleep apnea, and obesity hypoventilation syndrome
- central disorders of hypersonmolence, including narcolepsy and idiopathic hypersonmia
- sleep-related movement disorders, for example restless legs syndrome
- parasomnias, including nonrapid eye movement-related parasomnias (e.g., sleep walking, sleep terrors), rapid eye movement-related parasomnias (e.g., rapid eye movement sleep behaviour disorder, nightmare disorder), and other parasomnias
- circadian rhythm sleep-wake disorders, including those with environmental factors (e.g., shift work) and those happening when the circadian timing system becomes altered relative to the external environment (e.g., delayed sleep phase syndrome)
- insomnia, including environmental, genetic, psychological, and/or behavioural factors leading to hyperarousal.

**What Is the Current Practice?**

To determine if patients have a sleep disorder, they may be recommended by their health care provider to partake in a sleep study known as polysomnography (PSG). A Level 1 or type I (attended) PSG is the current gold standard for diagnosing sleep-disordered breathing, such as OSA, and other sleep disorders (i.e., certain central disorders of hypersonmolence, sleep-related movement disorders, and parasomnias) when used in conjunction with patients’ clinical history and other tests. Also, to obtain reimbursement for continuous positive airway pressure (CPAP) devices, the typical treatment for OSA, a Level I PSG is currently required for patients living in certain locations within Canada, including Ontario. A Level 1 PSG occurs in a registered sleep laboratory, within a hospital or specialized sleep clinic, under the supervision of trained health care staff to continuously monitor several physiologic signals during sleep:

- sleep stages through eye movement via electrooculogram, brain wave activity via electroencephalography (EEG), muscle activity and/or movement via electromyography
- respiratory airflow to measure the number and depth of respirations, such as episodes of shallow breathing (hypopneas) or episodes of breathing cessation (apneas)
- respiratory effort to measure movements of the chest and abdomen
- oxygen saturation via pulse oximetry
- cardiac dysrhythmias via cardiac monitoring
- any abnormal movements or behaviours via video monitoring
- body position
- snoring via microphone.
When a patient arrives at the clinic, a sleep technician places the equipment on the patient and monitors them throughout the sleep test in the technologist's control room. The test results are later interpreted by a sleep specialist, and a treatment plan is discussed with the patient in a follow-up visit (e.g., no intervention required, CPAP therapy recommended). A Level 1 PSG is largely standard practice in Canada, and capacity is limited to the number of dedicated beds within each registered sleep clinic. However, during the COVID-19 pandemic, in-clinic PSG testing was temporarily paused in Canada to prioritize essential services and surgeries. As a result, wait times for in-clinic PSG testing was further exacerbated by the pausing of in-clinic PSG. The COVID-19 pandemic also illuminated access challenges for patients living in Canada: access to registered sleep clinics varies greatly depending on location (rural and/or remote versus urban settings) as well as other factors (e.g., ability to travel, care responsibilities at nighttime).

What Is a Level 2 PSG, and How May It Benefit?

Level 2 or type II PSGs use the same monitoring sensors as full PSG (Level 1) but are unattended, affording the ability to be performed outside of the sleep clinic and in the patient's home. Using a Level 2 PSG may be a way to help mitigate the long wait times for those awaiting a Level 1 PSG but also improve patient access to care. Studies of Level 3 and 4 PSG devices, home devices that mainly screen for OSA and do not include EEG monitoring, suggest that they are associated with faster diagnoses and treatment times and could be cost-effective; it is unclear if Level 2 PSG devices may result in similar outcomes if patients can be tested from the comfort of their home without a sleep technician present during the test. Before Level 2 PSG devices are used throughout Canada, it is important to understand their accuracy, clinical utility, and cost-effectiveness.

Objective

To support the decision-making about Level 2 PSG for screening and diagnosis of any sleep disorders, we prepared this Rapid Review to summarize and critically appraise the studies available on the diagnostic accuracy, clinical utility, and cost-effectiveness of Level 2 PSG in adults.

Methods

Literature Search Methods

An information specialist conducted a literature search on key resources, including MEDLINE, Embase, the Cochrane Database of Systematic Reviews, the International HTA Database, the websites of Canadian and major international health technology agencies, as well as a focused internet search. The search approach was customized to retrieve a limited set of results, balancing comprehensiveness with relevancy. The search strategy comprised controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. Search concepts were developed based on the elements of the research questions and selection criteria. The main search concept was polysomnography. Conference abstracts were excluded. The search was completed on June 13, 2023, and was limited to English-language documents published since January 1, 2018.
Selection Criteria and Methods
One reviewer screened citations and selected studies. In the first level of screening, 1 reviewer screened titles and abstracts and then retrieved potentially relevant articles to assess for inclusion. Table 1 presents the final selection of full-text articles based on the inclusion criteria.

Table 1: Selection Criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Adult individuals suspected of having any sleep disorders</td>
</tr>
<tr>
<td>Intervention/Index test</td>
<td>Level 2 PSG (at-home or unattended)</td>
</tr>
<tr>
<td>Reference standard</td>
<td>Q1. Level 1 PSG (in-clinic or attended)</td>
</tr>
<tr>
<td></td>
<td>Q2 and Q3. Not applicable</td>
</tr>
<tr>
<td>Comparators</td>
<td>Q1. Not applicable</td>
</tr>
<tr>
<td></td>
<td>Q2 and Q3. Level 1 PSG (in-clinic or attended) or Level 3 PSG (at-home or unattended)</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Q1. Diagnostic test accuracy (e.g., sensitivity, specificity, positive predictive value, negative predictive value)</td>
</tr>
<tr>
<td></td>
<td>Q2. Clinical utility (e.g., time to diagnosis, cardiovascular outcomes, new cardiovascular events, cerebrovascular outcomes, mortality, quality of life, motor vehicle accidents)</td>
</tr>
<tr>
<td></td>
<td>Q3. Cost-effectiveness (e.g., cost per QALY gained, ICER)</td>
</tr>
<tr>
<td>Study designs</td>
<td>Health technology assessments, systematic reviews, randomized controlled trials, nonrandomized studies, economic evaluations</td>
</tr>
</tbody>
</table>

ICER = incremental cost-effectiveness ratio; PSG = polysomnography; QALY = quality-adjusted life-year.

Exclusion Criteria
We excluded articles if they did not meet the selection criteria outlined in Table 1, were duplicate publications, or were published before 2018.

Critical Appraisal of Individual Studies
One reviewer critically appraised the included publications using the following tools as a guide: Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) checklist\textsuperscript{18} for the diagnostic test accuracy study and the Downs and Black checklist\textsuperscript{19} for clinical utility study. Summary scores were not calculated for the included studies; rather, the strengths and limitations of each included publication were described narratively.

Summary of Evidence
Quantity of Research Available
Appendix 1 presents the study selection details. We identified 1 crossover study\textsuperscript{20} that addressed the diagnostic test accuracy of Level 2 PSG (research question 1) and 1 observational study\textsuperscript{21} that examined the clinical utility of Level 2 PSG (research question 2). We did not identify any relevant health technology assessments or randomized controlled trials. Further, we did not identify any relevant economic evaluations
regarding Level 2 PSG for screening or diagnosing adult sleep disorders. Appendix 5 provides additional references of potential interest that did not meet the inclusion criteria.

Summary of Study Characteristics
Appendix 2 provides detailed characteristics of included publications.

Included Study for Question 1: Diagnostic Test Accuracy
We included 1 crossover study that examined the diagnostic test accuracy of Level 2 PSG for the diagnosis of OSA. This 2018 study was conducted in Brazil. The study authors used a single-gate approach with 1 set of eligibility criteria (i.e., 1 group of participants with suspected OSA according to the study protocol) for admission. The study authors considered a clinical diagnosis of OSA was possible when the patient had at least 1 of the following symptoms: witnessed apneas; snoring; snoring associated with excessive daytime sleepiness; or snoring and daytime tiredness. The index test was a home monitoring sleep test using Level 2 PSG: Embletta X100 system (Embla, Natus Inc., Middleton, US) whereby the PSG technician went to the participant's home to hook up and later disassemble the equipment (i.e., the test itself was unattended). The reference standard was an in-clinic sleep test, reported by the author as the "gold standard," (p. 119) using the same equipment (i.e., Embletta X100 system) conducted in the standard way tests are performed in the clinic. The adult participants from 1 study site were consecutively included and randomly assigned to 1 of 2 groups according to the test sequence. Group 1 (n = 20) had the at-home PSG first, followed by a second, in-clinic PSG. Group 2 (n = 20) had the in-clinic PSG first, followed by a second, at-home PSG. The second PSG always happened on the night after the first PSG. The relevant outcomes for this review were sensitivity, specificity, positive predictive value, and negative predictive value in relation to the apnea-hypopnea index (AHI, used for determining the presence and severity of OSA).

Included Study for Question 2: Clinical Utility
We included 1 observational study that examined the clinical utility of Level 2 PSG for diagnosing OSA. This 2022 study included 225 participants living in Portugal with OSA, whereby a diagnosis was made using PSG. The study compared patients that received a Level 2 ambulatory, at-home PSG (intervention, n = 114) with patients that received a Level 1 attended, in-clinic PSG (comparator, n = 111) for their diagnosis of OSA. The study included a retrospective cohort component (for collecting general demographic data, for example) and a prospective, cross-sectional component to evaluate key outcomes using a patient questionnaire. The relevant outcomes for this review were quality of life, daytime sleepiness, hypertension, and treatment adherence (hours per night, percentage of days using CPAP). The follow-up period was not clearly defined or reported.

Summary of Critical Appraisal
Appendix 3 presents additional details regarding the strengths and limitations of included publications.

Diagnostic Test Accuracy Study
Zancanella and colleagues (2021) recruited participants using a convenience sample and randomly assigned them into 2 groups, according to the test sequence. The study authors avoided a case-control study
design (i.e., all participants had a clinical diagnosis of OSA according to the study protocol). The authors did not include all participants in the analysis; it is unclear if it was appropriate to exclude 6 participants that had lost EEG sensors during home monitoring (e.g., should a participant’s loss of EEG sensors be a data point for interpretation instead of criteria for exclusion?), which may also be a concern for its applicability. By excluding these participants, the estimates of diagnostic accuracy may also be overoptimistic and could have introduced bias. The study population, index test, and reference standard seem to match those targeted by the review questions. The study authors described their methods to conduct and interpret the index test and reference standard. All participants received the same index test and reference standard and the timing between tests was within 1 day of each other, ensuring the patients’ clinical condition would not have changed between tests. It is unclear whether the authors interpreted the index test and reference standard results without knowing the opposing test results. The study authors did not report measures of uncertainty. Zancanella and colleagues declared that they have no conflicts of interest, including “no significant financial support for this work that could have influenced its outcome” (p. 121). The authors did not report any additional context leaving it unclear what significant means (e.g., no funding versus a specific monetary value) and if it had the potential to introduce any reporting bias.

Clinical Utility Study

The study by Andrade and Paiva (2018) clearly reported the objectives, main outcomes, participant inclusion criteria, participant characteristics, interventions, findings, estimates of random variability, and actual P values. The authors recruited study participants from the same population over the same time, but they did not report any potential confounders (nor describe any adjustment of potential confounding factors such as OSA severity and pre-existing health conditions [e.g., cardiovascular and pulmonary disease]). The authors did not report any details about characteristics of participants lost to follow-up or adverse events. For example, they did not describe why the quality-of-life outcome had data from 148 participants when they included 225 participants overall. When considering external validity, the setting appeared representative of that in the population; though, it was unclear if the population invited and/or agreed to participate is representative of the entire population. When considering bias, there was no apparent data dredging, the statistical tests appeared appropriate, and authors clearly described the outcome measures. Because the study design was retrospective, the investigators were not able to blind participants or clinicians to exposures. It is unclear if study authors blinded exposures to the investigator(s) that analyzed the data. The authors did not clearly define or report the allowable length of time between when a patient had the PSG and when their outcome data were collected (e.g., 2 months versus 3 years). Moreover, the authors did not describe how they adjusted for losses of participants to follow-up for key outcomes. The authors did not state whether they conducted any power calculations, so it is unknown if the study had the appropriate number of patients required to avoid a type II error (e.g., accept the null hypothesis incorrectly and report that there is no difference between the 2 groups). Andrade and Paiva (2018) reported no conflicts of interest.

Summary of Findings

Appendix 4 presents the main study findings.
Diagnostic Accuracy of Level 2 PSG
We identified 1 study that examined the diagnostic test accuracy of Level 2 PSG for apneas and hypopneas using the AHI. When using “AHI <15>” (p. 120, interpreted as 15 events per hour, representing moderate OSA), the Level 2 (at-home) PSG, compared to the Level 1 (in-clinic) PSG, had:

- 80% sensitivity
- 83% specificity
- 91% positive predictive value
- 67% negative predictive value.

This suggests that Level 2 PSG tests have moderate sensitivity to detect OSA. This means that, out of every 100 people with OSA, Level 2 PSG will:

- detect 80 people with OSA
- miss up to 20 people (i.e., false negatives).

This also suggests that Level 2 PSG tests have moderate specificity in detecting those who do not have OSA. This means that, out of every 100 individuals who do not have OSA:

- 83 people will test negative
- 17 people will be wrongly diagnosed with OSA (i.e., false positives).

Clinical Utility of Level 2 PSG
We identified 1 study that examined the clinical utility of Level 2 PSG and provided results for the following outcomes: (i) quality of life; (ii) daytime sleepiness; (iii) hypertension; and (iv) treatment adherence.

Quality of Life
After OSA diagnosis and noninvasive ventilation treatment, the study authors compared mean scores for each specific scale of the Short Form (36) Health Survey (SF-36) between groups (i.e., Level 2 PSG versus Level 1 PSG). There were no statistically significant differences in mean scores between groups for 7 of the 8 scales of the SF-36, but there was a statistically significant increase in the Role-Physical scale for the Level 2 PSG group (P = 0.042).

However, whether these differences are clinically important is unknown, as the Minimal Clinically Important Difference was not reported for this scale.

Daytime Sleepiness
This study compared Epworth Sleepiness Scale scores, a validated scale to measure excessive daytime sleepiness in adults between groups, and found no statistically significant differences between the Level 2 PSG group and the Level 1 PSG group (P = 0.111).

Hypertension
This study compared hypertension between groups and found no statistically significant differences between the Level 2 PSG group and the Level 1 PSG group (P = 0.721).
Treatment Adherence
The study compared treatment adherence between Level 2 and Level 1 PSG groups. Overall, 88.8% of patients used CPAP for more than 4 hours on more than 70% of nights, with no statistically significant differences in adherence between groups (P = 0.915). There were also no statistically significant differences between the Level 2 group and Level 1 PSG group in mean hours of CPAP use per night (P = 0.884) and in percentage of days of CPAP use (P = 0.193). The study authors concluded that overall CPAP use in both groups was “considered within clinically efficient limits” (p. 1327); however, this was not described further.

Cost-Effectiveness of Level 2 PSG
We did not identify any relevant evidence regarding the cost-effectiveness of Level 2 PSG for the screening or diagnosis of sleep disorders in adults; therefore, no summary can be provided.

Limitations
Overall Completeness of the Evidence
The findings in this review are limited by the quantity of relevant evidence we identified. For all research questions, no HTA, systematic reviews, or randomized controlled trials met the inclusion criteria for this review.

We did not find any evidence on the following; therefore, no conclusions can be formed on these aspects of the research questions:

- the diagnostic test accuracy of Level 2 PSG for screening or diagnosis of sleep disorders other than OSA
- the clinical utility of Level 2 PSG for screening or diagnosis of sleep disorders other than OSA
- the cost-effectiveness of Level 2 PSG for the screening or diagnosis of any sleep disorders

Notwithstanding, we found little evidence about the diagnostic test accuracy (1 study) and clinical utility (1 study) of Level 2 PSG for screening and diagnosis of OSA. It is unclear whether the lack of evidence is from the true paucity of available evidence regarding Level 2 PSGs versus a limitation of the rapid methodology used for this review (i.e., a limited literature search from the past 5 years). However, this methodological approach balances comprehensiveness with timeliness. Furthermore, when screening the literature for this review, we did see more literature regarding Level 3 and 4 PSG devices, suggesting that the literature search strategy was likely sufficient to capture recent, relevant evidence about at-home PSG devices.

Generalizability of the Findings
The included studies were conducted in Brazil and Portugal and some of the devices used (e.g., Domino by SOMNOmedics) do not have active license listings in Canada; therefore, it is unclear whether the results summarized in this review are generalizable to the health care context in Canada.

These limitations warrants taking caution when interpreting the findings of this review.
Conclusions and Implications for Decision- or Policy-Making

This review identified and summarized the evidence available on the diagnostic test accuracy (1 crossover study)\(^{20}\) and clinical utility (1 observational study)\(^{21}\) of Level 2 PSG for screening or diagnosing sleep disorders in adults.

The limited evidence on diagnostic test accuracy from 1 study\(^{20}\) focused on the accuracy of Level 2 PSG for diagnosing OSA. Comparing Level 2 PSG with Level 1 PSG, the authors reported 80% sensitivity, 83% specificity, 91% positive predictive value, and 67% negative predictive value. These findings suggest moderate sensitivity to detect OSA and moderate specificity to detect those who do not have OSA. Though, the study authors did not provide context about whether moderate accuracy is acceptable in the field of sleep medicine.\(^{20}\)

The limited evidence on the clinical utility of Level 2 PSG from 1 study\(^{21}\) focused on a few relevant outcomes that are downstream of testing and diagnosis of OSA, after resulting treatment, including quality of life, daytime sleepiness, hypertension, and treatment adherence. The authors did not report significant differences in daytime sleepiness, hypertension, treatment adherence, and most quality of life measures for patients with OSA who were diagnosed with a Level 2 PSG compared with a Level 1 PSG.\(^{21}\) These findings suggest that patients may have similar clinical outcomes with an in-home PSG versus an in-clinic PSG, which may be an important implication for policy or decision-making if future studies support these clinical findings and observe Level 2 PSGs to be cost-effective or preferred by patients. However, this study had certain limitations to consider when interpreting the overall findings: study authors did not describe or adjust for any confounding factors (e.g., OSA severity), which may have introduced bias (i.e., internal validity); and the study authors excluded 6 participants after having a Level 2 PSG because EEG sensors fell off during the study. If a PSG fails in the real-world setting, it is presumed that the patient would have to retake the test either at home again or in a sleep clinic. Further research could be explored to determine if these situations would result in a slower time to diagnosis or a reduction of quality-of-life outcomes.

Overall, we found few studies\(^{20,21}\) addressing our research questions that met our inclusion criteria. Both included studies focused on patients with OSA\(^{20,21}\) and we did not find any studies that examined other sleep disorders. We also did not identify any relevant studies addressing the cost-effectiveness of screening or diagnosis of sleep disorders in adults. Though not eligible for this report, 1 study used a theoretical economic decision model to explore Level 2 PSGs within the British Columbia context.\(^{15}\) This study found that Level 2 studies may provide substantial cost advantages versus in-clinic PSGs, but the study authors stated that further empirical studies need to be conducted to test their algorithms.\(^{15}\) Further, decision-makers may wish to consider whether Level 2 PSG impacts other important clinical utility outcomes, including time to diagnosis, direct cardiovascular and cerebrovascular outcomes, mortality, and motor vehicle accidents. We require research focused on sleep disorders other than OSA and the cost-effectiveness of Level 2 PSG.

In conclusion, while limited evidence suggests that Level 2 PSG may be moderately accurate for diagnosing OSA and may not lead to different clinical outcomes for patients compared with Level 1 PSG, we require more comprehensive research with rigorous methodological approaches to understand this topic better.
References


Appendix 1: Selection of Included Studies

Figure 1: Selection of Included Studies

- 557 citations identified from electronic literature search and screened
- 531 citations excluded
- 26 potentially relevant articles retrieved for scrutiny (full text, if available)
- 0 potentially relevant reports retrieved from other sources (grey literature, handsearch)
- 26 potentially relevant reports
- 24 reports excluded:
  - irrelevant population (1)
  - irrelevant intervention (17)
  - irrelevant comparator (1)
  - irrelevant outcomes (4)
  - other (review articles, editorials) (1)
- 2 reports included in review
### Table 2: Characteristics of Included Diagnostic Test Accuracy Study

<table>
<thead>
<tr>
<th>Study citation, country, funding source</th>
<th>Study design, target condition(s)</th>
<th>Population characteristics</th>
<th>Index test and reference standard</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zancanella et al. (2022)20 Brazil</td>
<td>Study design: Crossover study Target condition: OSA Adult participants with a clinical diagnosis of OSA (according to the study protocol) from 1 site consecutively included and randomly assigned to 1 of 2 groups according to test sequence. Group 1: At-home PSG then in-clinic PSG Group 2: In-clinic PSG then at-home PSGa</td>
<td>34 adult participants who never had a PSG, were not pregnant or taking medication with proven interference with sleep phases and/or other PSG parameters: Group 1: 14 participants (64% male, 36% female, mean age = 39.9 ± 8.9 years). Group 2: 20 participants (75% male, 25% female, 40.2 ± 9.0 years).</td>
<td>Index Test: Level 2 PSG (i.e., at-home sleep study) using Embletta X100 system (Embla, Natus Inc., Middleton, Wisconsin, US). Reference Standard: Level 1 PSG (i.e., in-clinic sleep study) reported by author as the “gold standard” (p. 119) using Embletta X100 system (Embla, Natus Inc., Middleton, Wisconsin, US).</td>
<td>Sensitivity, specificity, positive predictive value, negative predictive value</td>
</tr>
</tbody>
</table>

NR = not reported; OSA = obstructive sleep apnea; PSG = polysomnography.  
*aThe second PSG always happened on the night after the first PSG.

### Table 3: Characteristics of Included Primary Clinical Utility Study

<table>
<thead>
<tr>
<th>Study citation, country, funding source</th>
<th>Study design</th>
<th>Population characteristics</th>
<th>Intervention and comparator(s)</th>
<th>Clinical outcomes, length of follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrade and Paiva (2018)21 Portugal</td>
<td>Observational study with a retrospective cohort component and a prospective, cross-sectional component.</td>
<td>225 adult participants with an OSA diagnosis via PSG.</td>
<td>Intervention: Level 2 ambulatory PSG (i.e., at-home sleep study) using Domino (SOMNOmedics GmbH, Randersacker, Germany) or Embla 7000 (Embla Systems, Inc., Broomfield, Colorado, US). Comparator: Level 1 attended PSG (i.e., in-clinic sleep study) using Alice 5 (Philips Respironics, Murrysville, Pennsylvania, US), Domino (SOMNOmedics</td>
<td>Outcomes: quality of life, daytime sleepiness (ESS score), hypertension, treatment adherence (hours per night, percentage of days) Follow-up: undefined</td>
</tr>
<tr>
<td>Study citation, country, funding source</td>
<td>Study design</td>
<td>Population characteristics</td>
<td>Intervention and comparator(s)</td>
<td>Clinical outcomes, length of follow-up</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-------------</td>
<td>---------------------------</td>
<td>-------------------------------</td>
<td>----------------------------------------</td>
</tr>
</tbody>
</table>

ESS = Epworth Sleepiness Scale; NR = not reported; OSA = obstructive sleep apnea; PSG = polysomnography.
Appendix 3: Critical Appraisal of Included Publications

Table 4: Strengths and Limitations of Diagnostic Test Accuracy Study Using the QUADAS–2 Checklist

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participant selection</strong></td>
<td><strong>Participant selection</strong></td>
</tr>
<tr>
<td>• Participants recruited using a convenience sample and randomly assigned into 2 groups according to test sequence.</td>
<td>• It is unclear if it was appropriate to exclude the 6 participants that had a loss of EEG sensors during home monitoring. By excluding these participants, the estimates of diagnostic accuracy may be overoptimistic and may have introduced bias.</td>
</tr>
<tr>
<td>• Case-control study design avoided (i.e., all participants had a clinical diagnosis of OSA according to the study protocol).</td>
<td><strong>Index test and reference standard</strong></td>
</tr>
<tr>
<td><strong>Index test and reference standard</strong></td>
<td><strong>Index test and reference standard</strong></td>
</tr>
<tr>
<td>• The study population, index test, and reference standard seem to match those targeted by the review questions.</td>
<td>• It is unclear whether study authors interpreted the index test and reference standard results without knowledge of the opposing test results.</td>
</tr>
<tr>
<td>• Methods to conduct and interpret the index test and reference standard well described.</td>
<td>• It is not clear if study authors should have considered the participant’s loss of EEG sensors a data point for interpretation vs. criteria for exclusion.</td>
</tr>
<tr>
<td>• All participants received the same index test and reference standard.</td>
<td><strong>Flow and timing</strong></td>
</tr>
<tr>
<td><strong>Flow and timing</strong></td>
<td><strong>Flow and timing</strong></td>
</tr>
<tr>
<td>• The second PSG always happened on the night after the first PSG.</td>
<td>• Not all participants included in the analysis.</td>
</tr>
</tbody>
</table>

EEG = electroencephalogram; NA = not applicable; OSA = obstructive sleep apnea; PSG = polysomnography; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2.

Note that this appendix has not been copy-edited.
Table 5: Strengths and Limitations of Clinical Utility Study Using the Downs and Black Checklist\textsuperscript{19}

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Andrade and Paiva (2018)\textsuperscript{21}</strong></td>
<td></td>
</tr>
</tbody>
</table>
| **Reporting** | • Potential confounders not reported.  
• Adverse events and characteristics of participants lost to follow-up not reported. |
| • Objectives, main outcomes, participant inclusion criteria, participant characteristics, interventions, findings, estimates of random variability, and actual P values all clearly reported. | |
| • Setting appeared to be representative of that in the population. | **External validity** |
| • No data dredging (i.e., unreported/post hoc analyses) apparent. | • Population invited/agreed to participate may not be representative of the entire population. |
| • Statistical tests appear appropriate. | **Internal validity (bias)** |
| • Outcome measures clearly described. | • Study participants and clinicians not blinded to exposures. |
| **Internal validity (confounding)** | • Unclear if investigators blinded to exposures. |
| • Study participants recruited from the same population over the same time. | **Internal validity (confounding)** |
| **Power** | • Adjustment for potential confounders not reported. |
| • NA | **Power** |
| | • Unclear how authors adjusted for losses of participants to follow-up for key outcomes. |
| | • No acknowledgement of power calculations reported. |

NA = not applicable.

Note that this appendix has not been copy-edited.
Appendix 4: Main Study Findings

Note that this appendix has not been copy-edited.

Table 6: Summary of Findings by Outcome — Diagnostic Test Accuracy for Apneas and Hypopneas (AHI) Detection

<table>
<thead>
<tr>
<th>Study citation</th>
<th>Index test</th>
<th>Reference standard</th>
<th>Number of participants</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>PPV (95% CI)</th>
<th>NPV (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zancanella et al. (2022)²⁰</td>
<td>At-home PSG (Level 2)</td>
<td>In-clinic PSG (Level 1)</td>
<td>34</td>
<td>0.8 (NR)</td>
<td>0.83 (NR)</td>
<td>0.91 (NR)</td>
<td>0.67 (NR)</td>
</tr>
</tbody>
</table>

AHI = Apnea-Hypopnea Index; CI = confidence interval; NPV = negative predictive value; OSA = obstructive sleep apnea; PPV = positive predictive value; PSG = polysomnography.

Table 7: Summary of Findings by Outcome — Clinical Utility of Level 2 PSG

<table>
<thead>
<tr>
<th>Study citation</th>
<th>Outcome</th>
<th>Intervention: Level 2 PSG</th>
<th>Control: Level 1 PSG</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrade and Paiva (2018)²¹</td>
<td>QoL (SF-36 survey): Physical Functioning, mean (SD)</td>
<td>85.7 (18.9) n = 90</td>
<td>83.6 (17.7) n = 93</td>
<td>0.204</td>
</tr>
<tr>
<td></td>
<td>QoL (SF-36 survey): Role-Physical, mean (SD)</td>
<td>84.3 (29.8) n = 90</td>
<td>83.6 (25.8) n = 93</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>QoL (SF-36 survey): Bodily Pain, mean (SD)</td>
<td>74.9 (27.8) n = 90</td>
<td>71.3 (29.4) n = 93</td>
<td>0.414</td>
</tr>
<tr>
<td></td>
<td>QoL (SF-36 survey): General Health, mean (SD)</td>
<td>65.8 (19.3) n = 90</td>
<td>65.3 (21.6) n = 93</td>
<td>0.780</td>
</tr>
<tr>
<td></td>
<td>QoL (SF-36 survey): Vitality, mean (SD)</td>
<td>66.3 (24.1) n = 90</td>
<td>65.9 (24.3) n = 93</td>
<td>0.943</td>
</tr>
<tr>
<td></td>
<td>QoL (SF-36 survey): Social Functioning, mean (SD)</td>
<td>87.6 (22.3) n = 90</td>
<td>85.9 (22.2) n = 93</td>
<td>0.694</td>
</tr>
<tr>
<td></td>
<td>QoL (SF-36 survey): Role - Emotional, mean (SD)</td>
<td>86.6 (26.5) n = 90</td>
<td>83.1 (28.4) n = 93</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td>QoL (SF-36 survey): Mental Health, mean (SD)</td>
<td>75.4 (20.2) n = 90</td>
<td>75.3 (22.7) n = 93</td>
<td>0.687</td>
</tr>
<tr>
<td></td>
<td>Daytime sleepiness: ESS score</td>
<td>NR n = 110</td>
<td>NR n = 100</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>Hypertension</td>
<td>NR</td>
<td>NR</td>
<td>0.721</td>
</tr>
<tr>
<td></td>
<td>Treatment adherence</td>
<td>NR</td>
<td>NR</td>
<td>0.915</td>
</tr>
<tr>
<td></td>
<td>Hours per night of CPAP use, mean (SD)</td>
<td>5.8 (1.4) n = 93</td>
<td>5.6 (1.3) n = 89</td>
<td>0.884</td>
</tr>
<tr>
<td></td>
<td>Percentage of days of CPAP use, mean (SD)</td>
<td>79% (25.4%) n = 93</td>
<td>77.6% (25.3%) n = 89</td>
<td>0.193</td>
</tr>
</tbody>
</table>

CPAP = continuous positive airway pressure; ESS = Epworth Sleepiness Scale; OSA = obstructive sleep apnea; PSG = polysomnography; QoL = quality of life; SF-36 = Short Form (36) Health Survey.

*Mean score reported by participants after OSA diagnosis and noninvasive ventilation treatment.
Appendix 5: References of Potential Interest

Previous CADTH Reports


Diagnosis of snoring and obstructive sleep apnea: A review of the accuracy. Ottawa (ON): CADTH; 2009: https://www.cadth.ca/sites/default/files/pdf/L0092_Diagnosis_Snoring_Obstructive_Sleep_Apnea.pdf


Level 2 PSG for Sleep Bruxism
At-Home Polysomnography Versus In-Clinic Polysomnography for Sleep Disorders

Authors: Chantelle C. Lachance, Sharon Bailey
Contributors: Elizabeth Carson
ISSN: 2563-6596

Disclaimer: The information in this document is intended to help Canadian health care decision-makers, health care professionals, health systems leaders, and policy-makers make well-informed decisions and thereby improve the quality of health care services. While patients and others may access this document, the document is made available for informational purposes only and no representations or warranties are made with respect to its fitness for any particular purpose. The information in this document should not be used as a substitute for professional medical advice or as a substitute for the application of clinical judgment in respect of the care of a particular patient or other professional judgment in any decision-making process. The Canadian Agency for Drugs and Technologies in Health (CADTH) does not endorse any information, drugs, therapies, treatments, products, processes, or services.

While care has been taken to ensure that the information prepared by CADTH in this document is accurate, complete, and up-to-date as at the applicable date the material was first published by CADTH, CADTH does not make any guarantees to that effect. CADTH does not guarantee and is not responsible for the quality, currency, propriety, accuracy, or reasonableness of any statements, information, or conclusions contained in any third-party materials used in preparing this document. The views and opinions of third parties published in this document do not necessarily state or reflect those of CADTH.

CADTH is not responsible for any errors, omissions, injury, loss, or damage arising from or relating to the use (or misuse) of any information, statements, or conclusions contained in or implied by the contents of this document or any of the source materials.

This document may contain links to third-party websites. CADTH does not have control over the content of such sites. Use of third-party sites is governed by the third-party website owners' own terms and conditions set out for such sites. CADTH does not make any guarantee with respect to any information contained on such third-party sites and CADTH is not responsible for any injury, loss, or damage suffered as a result of using such third-party sites. CADTH has no responsibility for the collection, use, and disclosure of personal information by third-party sites.

Subject to the aforementioned limitations, the views expressed herein are those of CADTH and do not necessarily represent the views of Canada's federal, provincial, or territorial governments or any third-party supplier of information.

This disclaimer and any questions or matters of any nature arising from or relating to the use (or misuse) of this document will be governed by and interpreted in accordance with the laws of the Province of Ontario and the laws of Canada applicable therein, and all proceedings shall be subject to the exclusive jurisdiction of the courts of the Province of Ontario, Canada.

The copyright and other intellectual property rights in this document are owned by CADTH and its licensors. These rights are protected by the Canadian Copyright Act and other national and international laws and agreements. Users are permitted to make copies of this document for noncommercial purposes only, provided it is not modified when reproduced and appropriate credit is given to CADTH and its licensors.

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to Requests@CADTH.ca