Executive Summary of Respiratory Indications for Polysomnography in Children: An Evidence-Based Review

Merrill S. Wise, MD[1]; Cynthia D. Nichols, PhD[2]; Madeleine M. Grigg-Damberger, MD[3]; Carole L. Marcus, MBBCh[4]; Manisha B. Witmans, MD[5]; Valerie G. Kirk, MD[6]; Lynn A. D'Andrea, MD[7]; Timothy F. Hoban, MD[8]

[1]Methodist Healthcare Sleep Disorders Center, Memphis, TN; [2]Munson Medical Center, Traverse City, MI; [3]University of New Mexico School of Medicine, Albuquerque, NM; [4]Children's Hospital of Philadelphia, Philadelphia, PA; [5]Stollery Children's Hospital and University of Alberta, Edmonton, AB; [6]Alberta Children's Hospital and University of Calgary, Calgary, AB; [7]Medical College of Wisconsin and Children's Hospital of Wisconsin, Milwaukee, WI; [8]University of Michigan, Ann Arbor, MI

Objective: This comprehensive, evidence-based review provides a systematic analysis of the literature regarding the validity, reliability, and clinical utility of polysomnography for characterizing breathing during sleep in children. Findings serve as the foundation of practice parameters regarding respiratory indications for polysomnography in children.

Methods: A task force of content experts performed a systematic review of the relevant literature and graded the evidence using a standardized grading system. Two hundred forty-three evidentiary papers were reviewed, summarized, and graded. The analysis addressed the operating characteristics of polysomnography as a diagnostic procedure in children and identified strengths and limitations of polysomnography for evaluation of respiratory function during sleep.

Results: The analysis documents strong face validity and content validity, moderately strong convergent validity when comparing respiratory findings with a variety of relevant independent measures, moderate-to-strong test-retest validity, and limited data supporting discriminant validity for characterizing breathing during sleep in children. The analysis documents moderate-to-strong test-retest reliability and interscorer reliability based on limited data. The data indicate particularly strong clinical utility in children with suspected sleep related breathing disorders and obesity, evolving metabolic syndrome, neurological, neurodevelopmental, or genetic disorders, and children with craniofacial syndromes. Specific consideration was given to clinical utility of polysomnography prior to adenotonsillectomy (AT) for confirmation of obstructive sleep apnea syndrome. The most relevant findings include: (1) recognition that clinical history and examination are often poor predictors of respiratory polygraphic findings, (2) preoperative polysomnography is helpful in predicting risk for perioperative complications, and (3) preoperative polysomnography is often helpful in predicting persistence of obstructive sleep apnea syndrome in patients after AT. No prospective studies were identified that address whether clinical outcome following AT for treatment of obstructive sleep apnea is improved in association with routine performance of polysomnography before surgery in otherwise healthy children. A small group of papers confirm the clinical utility of polysomnography for initiation and titration of positive airway pressure support.

Conclusions: Pediatric polysomnography shows validity, reliability, and clinical utility that is commensurate with most other routinely employed diagnostic clinical tools or procedures. Findings indicate that the "gold standard" for diagnosis of sleep related breathing disorders in children is not polysomnography alone, but rather the skillful integration of clinical and polygraphic findings by a knowledgeable sleep specialist. Future developments will provide more sophisticated methods for data collection and analysis, but integration of polysomnographic findings with the clinical evaluation will represent the fundamental diagnostic challenge for the sleep specialist.

Keywords: Polysomnography, pediatric, indications, clinical utility, sleep related breathing disorders, obstructive sleep apnea syndrome

Citation: Wise MS; Nichols CD; Grigg-Damberger MM; Marcus CL; Witmans MB; Kirk VG; D'Andrea LA; Hoban TF. Executive Summary of respiratory indications for polysomnography in children: an evidence-based review. *SLEEP* 2011;34(3):389-398.

1.0 INTRODUCTION

Evaluation of children with suspected sleep disorders begins with and is based primarily on a thorough history. In appropriate cases the diagnostic process includes performance of polysomnography (PSG), most commonly for characterization of breathing during sleep. Because PSG requires significant time and health care resources, understanding the strengths, limitations, and clinical utility of PSG is necessary to ensure optimal utilization.

The Indications for Polysomnography in Children task force was established by the AASM Standards of Practice Committee and approved by the AASM Board of Directors. The objectives were to: (1) provide a systematic and comprehensive review of the relevant medical literature regarding respiratory indications for PSG in children; (2) grade the strength of evidence contained in the literature using a standardized grading system; (3) summarize information regarding the validity and reliability, clinical utility, and when available, outcomes associated with use of PSG in children with suspected respiratory disturbance during sleep; and (4) discuss the strengths and limitations of current knowledge about the utility of PSG in children. Findings from this paper will provide the foundation for evidence-based practice parameters regarding respiratory indications for PSG in children. This Executive Summary represents a condensed version of the review paper. The full review paper follows this Executive Summary. The evidence table is available on the AASM web site (www.aasmnet.org).

It is beyond the scope of this review to evaluate standards for how to perform PSG, equipment for PSG in children, methods for scoring respiratory events during sleep in children, or economic cost and cost/benefit analyses of PSG in children. Unattended testing outside the sleep laboratory in children has been

Submitted for publication December, 2010
Accepted for publication December, 2010
Address correspondence to: Sharon L. Tracy, PhD, American Academy of Sleep Medicine, 2510 North Frontage Road, Darien, IL 60561-1511; Tel: (630) 737-9700; Fax: (630) 737-9790; E-mail: stracy@aasmnet.org

Table 1—Definitions of reliability and validity

Type of Reliability or Validity	Definition	Polysomnography Example
Test-retest reliability	Stability of a measurement across time	Consistency of PSG data on 2 consecutive nights
Interrater reliability	Consistency of a measurement when used by multiple raters	Agreement between 2 people scoring the same PSG
Intrarater reliability	Consistency of a measurement when used by the same rater	Agreement between 2 scorings of the same PSG by the same person
Types of Construct Validity: (test-retest, convergent, discriminant)	Extent to which explanatory concepts account for performance on the test	
Test-retest validity (or responsiveness)	Change in the expected direction on 2 administrations of a test following a manipulation that is expected to have an impact on the measure	Reduction in the PSG-determined AHI following adenotonsillectomy
Convergent validity	Measures that should be related are in reality related	Positive correlation between oxygen saturation by oximetry and ABG
Discriminant validity	Measures that should not be related are in reality not related	Absence of significant correlation between PLM index and apnea/hypopnea index
Face validity	Agreement by experts or examinees that the test looks like it is measuring what it intends to measure	Agreement between experts that measuring air flow at the nose and mouth is a reasonable assessment of breathing during sleep Agreement between questionnaire data assessing clinical symptoms of OSA and OSA determined by PSG
Content validity	Agreement that the test samples the phenomena about which conclusions will be drawn	Agreement between experts that sleep stages can be determined by using EEG, EOG, and EMG during PSG
Criterion (predictive) validity	Agreement between the test and a direct measure of the behavior or characteristic	Increased signal amplitude on snore sensor when patient has audible snoring Consistent subjective report of sleeping when awakened from a specific sleep stage

used predominantly in research settings, and there is a paucity of research comparing it to traditional in-laboratory attended PSG or other objective clinical outcomes. For this reason, the task force did not address validity, reliability, or clinical utility of unattended testing outside the sleep laboratory in children.

2.0 BACKGROUND

The dramatic growth of pediatric sleep medicine over the past 3 decades is well documented. Expansion of the literature regarding PSG in children creates an opportunity for systematic and comprehensive review and evidence grading. Several professional organizations have produced clinical guidelines or practice parameters that include indications for PSG in children. Earlier publications included expected limitations associated with a less mature literature and less sophisticated sleep technology, and many recommendations were primarily consensus-based rather than evidence-based. Two professional organizations (The American Thoracic Society[1] in 1996 and The American Academy of Pediatrics[2] in 2002) produced clinical guidelines or practice parameters regarding indications for PSG in children. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology, and Specifications[3] provides explicit rules for scoring respiratory events, sleep stages, arousals, and other aspects of pediatric PSG.

Assessment of clinical utility and indications for performing a diagnostic test is often challenging, particularly when the diagnostic test is viewed as a *de facto* "gold standard." The ideal approach is assessment of whether patient outcome is improved in association with performance of the test. There are few published studies that address this issue with regard to diagnostic PSG in children. A second approach involves assessment of the operating characteristics of the diagnostic test in an effort to document validity, reliability, and clinical utility. Validation of a diagnostic test often involves establishment of different types of validity (test-retest, convergent, discriminant, face, content, and criterion) and reliability (test-retest, interrater, intrarater) (see Table 1 for definitions). The validity and reliability of techniques used for collecting and processing data influence the clinical utility of the diagnostic procedure. For this project the task force viewed clinical utility as a multidimensional concept, and the following attributes were considered to define clinical utility: the diagnostic test (PSG) must (1) have acceptable validity and reliability in the clinical populations of interest, (2) be useful for diagnosis and management decisions, and results should inform clinical decision-making, (3) be applied when effective therapies are available (results can influence outcome only when effective treatment is available), and (4) be interpretable by clinicians with necessary skills to use results in a meaningful way and to recognize false signals (artifact). A recurring challenge in this project is comparison of clinical utility of PSG across studies performed using different methods for measurement of respiratory parameters. An-

other challenge in regard to obstructive sleep apnea syndrome (OSAS) is that the explicit diagnostic criteria listed in the International Classification of Sleep Disorders, 2nd Edition,[4] include PSG respiratory findings that must be present to confirm a diagnosis of OSAS. Thus, determination of the clinical utility of PSG for diagnosis of OSAS involves "incorporation bias" since polygraphic diagnostic criteria are incorporated into diagnostic criteria.

Table 2—Levels of evidence

Level	Description
1	Evidence provided by a **prospective** study in a **broad spectrum** of persons with the suspected condition, using a **reference (gold) standard** for case definition, where test is applied in a **blinded fashion**, and enabling the assessment of appropriate test of diagnostic accuracy. All persons undergoing the diagnostic test have the presence or absence of the disease determined. Level 1 studies are judged to have a low risk of bias.
2	Evidence provided by a **prospective** study of a **narrow spectrum** of persons with the suspected condition, or a **well-designed retrospective** study of a **broad spectrum** of persons with an established condition (by **"gold standard"**) compared to a **broad spectrum of controls**, where test is applied in a **blinded** evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy. Level 2 studies are judged to have a moderate risk of bias.
3	Evidence provided by a **retrospective** study where either person with the established condition or controls are of a **narrow spectrum**, and where **the reference standard, if not objective, is applied by someone other than the person that performed (interpreted) the test**. Level 3 studies are judged to have a moderate to high risk of bias.
4	Any study design where **test is not applied in an independent evaluation** or evidence is provided by expert opinion alone or in **descriptive case series without controls**. There is **no blinding or there may be inadequate blinding**. The **spectrum of persons tested may be broad or narrow**. Level 4 studies are judged to have a very high risk of bias.

The composition of the task force includes individuals who are content experts in respiratory and nonrespiratory areas of pediatric sleep medicine, with clinical and research experience in pediatric PSG. All task force members completed AASM conflict of interest forms and were found to have no potential conflicts.

3.0 METHODS

Details of the process employed by the task force are provided in the full review paper available at the end of this Executive Summary. The task force developed a literature search strategy, established methods for selection of relevant papers, developed procedures for extracting data and grading the strength of evidence, and generated successive drafts of the review paper. Descriptions of the levels of evidence are listed in Table 2.

4.0 RESULTS

4.1 Overview of Results

Approximately 3500 candidate papers were identified and screened, and 243 papers were selected for inclusion. Presentation of results is organized into 3 sections: sleep related breathing disorders (SRBD), other chronic respiratory disorders, and clinical utility of PSG for therapeutic intervention.

4.2 Sleep Related Breathing Disorders

The SRBD section is composed of 4 subsections: (1) studies that support validity and/or reliability of PSG for characterization of breathing in children, (2) clinical utility of PSG in children with risk factors for SRBD, (3) clinical utility of PSG prior to adenotonsillectomy (AT) or other surgical procedures, and (4) clinical utility of PSG for assessment of infants less than 12 months of age with suspected SRBD.

4.2.1 Studies that assess validity or reliability of PSG in children

The task force's strategy was to evaluate papers that provide useful information about validity or reliability by comparing PSG respiratory findings with other independent measures that are relevant to assessment of SRBD in children. The task force also identified studies that compare PSG findings before and after therapeutic interventions such as AT and studies that include control groups without intervention, which allows assessment of test-retest validity and reliability. The movement of PSG respiratory parameters in the expected direction after surgery supports the validity and reliability of PSG for measurement of breathing during sleep.

4.2.1.1 Correlation of PSG findings with independent measures

4.2.1.1.1 History of snoring and other nocturnal symptoms

Thirty-five papers (2 Level 1, 4 Level 2, 11 Level 3, and 18 Level 4) were identified. Findings provide limited and inconsistent evidence to support the validity of PSG for evaluation of suspected SRBD when using the clinical history as an independent comparison measure. One interpretation of these data is that the clinical history is not sufficiently accurate, reliable, or stable to represent a meaningful comparison with the objective physiological measurements encompassed by PSG.

4.2.1.1.2 Audio or video recordings

Four articles (2 Level 2 and 2 Level 3) addressed the correlation between audio or video recordings and PSG findings. As a group, these studies provide moderate support for the validity of PSG, but investigators found that audio or video records do not provide sufficient specificity to reliably distinguish primary snoring from OSAS in children.

4.2.1.1.3 Questionnaires

Nine papers (2 Level 2, 3 Level 3, and 4 Level 4) reported on the correlation between pediatric sleep questionnaire (SQ) results and PSG findings in children with suspected OSAS. SQ results had a variable correlation with PSG findings, and most studies report relatively weak associations, suggesting that questionnaires do not provide strong evidence to support the validity of PSG respiratory measurements. This observation does not necessarily indicate poor validation of PSG, but

instead, it may suggest that currently available pediatric sleep questionnaires are not able to discriminate between children with primary snoring versus OSAS, nor gauge the severity of OSAS as determined by PSG.

4.2.1.1.4 Subjective and objective measures of sleepiness

Eleven studies (1 Level 1, 3 Level 2, 4 Level 3, and 3 Level 4) addressed the correlation between PSG findings and subjective sleepiness in children with suspected OSAS. Most papers support an association between subjective sleepiness and abnormal PSG or multiple sleep latency test (MSLT) parameters; however, the presence of subjective sleepiness alone does not accurately predict the presence of PSG-defined OSAS. Findings provide support and validation for the role of PSG in children to determine whether subjectively reported sleepiness is related to the presence of underlying OSAS. Five articles (2 Level 2, 1 Level 3, and 2 Level 4) addressed the correlation between PSG findings and objective measures of sleepiness. Findings support consistent but often weak associations between abnormal respiratory PSG parameters and MSLT-defined sleepiness among children with OSAS. Findings provide some limited support for the validity of PSG respiratory measures for characterization of OSAS in children. These findings also suggest that objective sleepiness is less often present and less severe in children with OSAS compared with adults. Several studies suggest that the MSLT may be more sensitive than subjective ratings of sleepiness in children with OSAS. Results provide support for the use of PSG for evaluation of suspected OSAS in children with sleepiness.

4.2.1.1.5 Physical examination

Eleven studies (2 Level 2, 3 Level 3, and 6 Level 4) provided evidence to address validity of PSG for characterization of SRBD in children in comparison to physical examination. The strength of association between physical examination findings and PSG findings is variable. This reflects the observation that PSG is a physiological measurement of breathing during sleep, whereas the physical examination is focused on anatomic structures during wakefulness. Physical findings provide only limited independent validation of PSG for characterization of SRBD in children and cannot take the place of PSG for diagnosis of OSAS.

4.2.1.1.6 Radiographic and endoscopic evaluation

Seven papers (1 Level 2, 3 Level 3, and 3 Level 4) addressed whether radiographic imaging or naso-oro-pharyngeal endoscopy provide independent support for validity of PSG findings in characterizing SRBD. This small number of papers provides consistently positive associations between independent assessments with radiographic or endoscopic methods for imaging the upper airway and PSG findings in children with suspected OSAS. As a group, these studies provide a moderate degree of support for the validity of PSG. An inclusion bias is likely in these studies because the subjects who underwent radiographic studies were suspected of having OSAS on the basis of the history or physical examination. No studies included a broad spectrum of subjects (i.e., subjects not suspected of having OSAS).

4.2.1.1.7 Neurocognitive or psychological assessments

Twenty studies addressed the construct validity of PSG for the evaluation of SRBD utilizing measures of neuropsychological, behavioral, and emotional functioning as the convergent construct. Five papers provided Level 2 evidence, 7 provided Level 3 evidence, and 8 provided Level 4 evidence. The magnitude of association and the nature of the relationships between these measures and sleep disordered breathing during PSG varied across studies. When the studies are viewed collectively, children with SRBD appear to function at lower levels when compared to children without SRBD. Complex interrelated factors such as duration of disease, genetic factors, sociocultural influences, and timing of exposure to SRBD in children probably influence neurocognitive function in this population. It is also possible that current methods for characterization of respiratory disturbance may not reflect subtle alterations in sleep microarchitecture, which may be better predictors of neurobehavioral outcomes. These studies provide moderate support for the construct validity of PSG and suggest that even mild SRBD may be associated with impairments in behavior and neuropsychological functioning.

4.2.1.1.8 Serial or ambulatory BP measurements

Twelve articles addressed whether independent measures of blood pressure (BP) correlate with respiratory PSG findings, including 5 Level 2, 5 Level 3, and 2 Level 4 papers. Two papers (1 Level 2, 1 Level 3) reported findings regarding the correlation between postoperative AHI and remission of hypertension (HTN) or elevated BP in children with OSAS. In summary, findings provide moderate-to-strong evidence for convergent (construct) validity of PSG respiratory measures using various BP measurements as an independent measurement. An AHI $\geq$ 5 in school age children was an *independent* risk factor for elevated systolic and diastolic BP even after adjusting for various confounding factors including BMI**.** Level 2 evidence supports an AHI $\geq$ 5 per hour as the threshold for OSAS severity associated with clinically significant elevations of BP values in children. The positive association between left ventricular remodeling and findings from 24-hour blood pressure monitoring highlights the relationship between PSG respiratory findings and cardiovascular morbidity. Obesity can be a confounding factor for elevated BP, risk of HTN, and OSAS severity among children who snore or have SRBD.

4.2.1.1.9 Quality of life measures

Health-related quality-of-life (HRQOL) measures are validated questionnaires completed by the subject or caregiver that identify the quality-of-life (QOL) impact of a medical disorder on different domains of a patient's life. Studies evaluating QOL in children with suspected or confirmed SRBD have used either a generic HRQOL instrument such as the Child Health Questionnaire (CHQ), or a disease-specific QOL tool developed to evaluate children with SRBD such as the OSA-18 or the OSD-6. Pediatric otolaryngologists have developed OSA-specific QOL surveys to assess outcome following AT.

The task force posed 2 questions to assess whether QOL measurements provide independent validation of PSG for characterization of SRBD in children: (1) do caregiver-rated QOL scores correlate with the severity of SRBD on PSG? (2) does improvement in QOL measures following AT for treatment of SRBD correlate with resolution of SRBD on PSG? Our search identified 8 studies (2 Level 2, 3 Level 3, and 3 Level 4) that evaluated the correlation between QOL scores

and PSG findings in children with SRBD. Four studies (1 Level 2, 1 Level 3, and 2 Level 4) correlated QOL and preoperative PSG findings in children or adolescents with suspected SRBD. Four studies compared PSG respiratory findings and QOL before and after AT (1 Level 2, 2 Level 3, and 1 Level 4). In summary, results from generic and disease-specific QOL instruments show generally low, and rarely moderate, correlation with objective respiratory PSG data in children or adolescents with primary snoring and OSAS. QOL scores could not differentiate primary snorers from those who had OSAS on PSG. QOL scores most often could not differentiate mild from severe OSAS. QOL scores showed improvement, even when postoperative PSG showed mild to even severe residual OSAS. Discrepancies between QOL measures and PSG respiratory findings may reflect the different types of measurements between a physiological study (PSG) and the issues probed by QOL instruments. Findings indicate that QOL measures alone do not provide significant independent validation of PSG respiratory measures.

4.2.1.1.10 Therapeutic intervention studies that provide evidence of test-retest validity

Therapeutic intervention studies provide an opportunity to evaluate test-retest validity when PSG is performed on the same group of subjects before and after an intervention known to improve respiratory function during sleep. When the test values change in the expected direction following the intervention, test-retest validity is demonstrated. Our search identified 45 studies. In 23 papers (5 Level 2, 9 Level 3, and 9 Level 4) PSG was performed before and after AT. In 8 studies (1 Level 3 and 7 Level 4) PSG was performed before and after other surgical procedures. In 11 studies (4 Level 2, 4 Level 3, and 3 Level 4), PSG was performed before and after nonsurgical intervention such as orthodontic treatment, and in 3 studies (1 Level 2 and 2 Level 4) PSG was performed before and after mixed surgical and nonsurgical interventions.

All studies provided data that support test-retest validity of PSG. The interventional studies often differed regarding definitions of apnea and hypopnea. Although in some circumstances this would be considered a measurement weakness, it also provides an opportunity to evaluate convergent validity when studies with similar designs use different operational definitions for the same construct. In the pre- and post-AT studies, for example, convergent validity for the measurement of SRBD with PSG was demonstrated by the observation that multiple face-valid yet slightly different definitions of SRBD yielded similar results. The overall consistency of results provides moderate-to-strong evidence for test-retest validity of PSG for characterization of SRBD in children.

4.2.1.1.11 Other measures

The task force identified 13 studies that assess construct, face, or convergent validity of PSG for characterization of SRDB through correlations with other independent measures in addition to those discussed above. Investigators used surrogates of end-organ dysfunction in SRDB such as hormone levels, inflammatory markers, markers of cardiovascular dysfunction, and biochemical markers of neurocognitive dysfunction as independent measures. Six studies provided Level 2 evidence, 4 provided Level 3 evidence, and 3 papers provided Level 4 evidence. These independent measures provide low-to-moderate strength of evidence to support construct and convergent validity for PSG.

4.2.1.1.12 Summary of Section 4.2.1.1

Collectively, the comparison of PSG to other independent measures documents strong face validity and content validity, moderately strong convergent validity, moderate-to-strong test-retest validity, and limited data to support discriminant validity for characterizing breathing during sleep in children.

4.2.1.2 Test-retest reliability and scoring reliability

Reliability testing evaluates the consistency and stability of a measurement across time or determines the accuracy of a measurement when used by multiple raters. Our search identified 6 papers (1 Level 1, 1 Level 2, 2 Level 3, and 2 Level 4) that address the issue of test-retest reliability for PSG in infants and children, including 1 Level 1 and 1 Level 2 study that reported interscorer reliability data. Findings provide good-to-excellent support for test-retest reliability for respiratory PSG parameters in infants and children.

4.2.1.3 Daytime nap PSG compared with full night PSG

Three studies (all with Level 4 evidence) compared daytime nap studies with overnight PSG. Findings provide very limited support for the potential role of nap PSG as a screening method or a diagnostic procedure in children with suspected SRBD. Nap PSG is not as sensitive as overnight PSG in identifying SRBD. Nap studies tend to underestimate the severity of SRBD when compared to overnight PSG.

4.2.1.4 Nocturnal home oximetry compared with PSG

Three papers compared diagnostic utility of home oximetry with PSG in children. Home oximetry findings may be relatively specific for OSAS in certain settings when positive, but findings are insensitive and no studies provided support that home oximetry alone offers acceptable diagnostic accuracy to replace PSG in children.

4.2.2 Clinical utility of PSG in children with risk factors for SRBD

The task force reviewed and summarized the literature with respect to a series of clinical attributes that are thought to represent varying levels of risk for SRBD. This approach evaluates clinical utility through a "risk stratification" strategy in order to support optimal clinical decision-making regarding indications for PSG in children.

4.2.2.1 Obesity

Thirty-four papers addressed the potential clinical utility of PSG in obese or overweight children with suspected SRBD. Multiple groups of investigators report that obesity in children correlates strongly with the presence of SRBD and moderately with the severity of SRBD on PSG. However, the effect of obesity on the risk for SRBD in children is probably not as strong as that observed in adults. There is relatively strong evidence that obese children 8 years or older are at significant risk for obstructive SRBD. The presence of even a modest degree of tonsillar hypertrophy and/or narrow velopharyngeal space potentiates the risk of SRBD in obese children. Obese children

are more likely to have residual OSAS following AT compared with non-obese children, which suggests the need for careful clinical follow-up and possibly repeat PSG after surgery. OSAS in obese children is associated with increased risk for hypertension, metabolic syndrome, and fatty liver disease. In summary, PSG has significant clinical utility for the diagnosis and management of SRBD in obese children and adolescents and for following clinical course after therapeutic intervention.

4.2.2.2 Prematurity

Four papers (2 Level 3 and 2 Level 4) addressed prematurity as a risk factor for PSG-confirmed SRBD. Two papers suggest an association between prematurity and abnormal respiratory PSG parameters, and 1 study stratified risk factors for OSAS among children born prematurely. Findings suggest that prematurity is an independent risk factor for SRBD in children.

4.2.2.3 Race/Ethnicity

Six papers (3 Level 2, 1 Level 3, and 2 Level 4) addressed race or ethnicity as a risk factor for PSG-confirmed SRBD. Most but not all papers support an association between African American race/ethnicity and increased risk for SRBD, and higher risk for residual OSAS following AT.

4.2.2.4 Family history of SRBD

Two papers (1 Level 3 and 1 Level 4) evaluated the correlation between PSG findings and family history of SRBD. Findings suggest that children with a family history of SRBD are at increased risk for SRBD. Data are too limited to support an indication for PSG based solely on a positive family history, but findings suggest that family history of SRBD represents a significant modifier for the expression of SRBD or severity of SRBD in children.

4.2.2.5 Allergic rhinitis or recurrent sinusitis

The association between allergic rhinitis or recurrent sinusitis and SRBD was addressed in 3 papers (1 Level 3 and 2 Level 4). These studies provide limited evidence that allergic rhinitis is independently associated with PSG-confirmed SRBD in children.

4.2.2.6 Systemic hypertension

The task force identified 6 papers (3 Level 2 and 3 Level 4) that addressed the potential clinical utility of PSG for evaluation of hypertension in children. Findings consistently support the clinical utility of PSG for identification of SRBD in children with systemic hypertension, in association with and independent of obesity.

4.2.2.7 Unexplained pulmonary hypertension

No articles were identified that provide data regarding an association between unexplained pulmonary hypertension and SRBD or the clinical utility of PSG in children with unexplained pulmonary hypertension.

4.2.2.8 Other risk factors and special populations

4.2.2.8.1 Chromosomal and neurogenetic disorders

Five studies (all Level 4) addressed the clinical utility of PSG in children with Down syndrome. Findings were uniform in showing a high prevalence of OSAS, with OSAS occurring in at least half the patients evaluated in each study. Snoring was not uniformly present by history or on PSG in children with Down syndrome and OSAS. Seven papers (1 Level 3 and 6 Level 4) addressed the clinical utility of PSG in children with Prader-Willi syndrome. Most investigators reported increased prevalence of SRBD, including 2 papers showing SRBD even in the absence of sleep complaints. No clinical variables were predictive of the severity of SRBD in Prader-Willi syndrome. One paper (Level 4) regarding Rett syndrome demonstrated that unless there is a clinical concern for SRBD, the diagnostic yield associated with PSG is low in this population.

4.2.2.8.2 Disorders with craniofacial anomalies

Six studies (1 Level 2 and 5 Level 4) involving Pierre Robin sequence suggest that significant OSAS is often present during infancy, and PSG is clinically useful in evaluating breathing in this population. Two papers (1 Level 3 and 1 Level 4) suggest a high prevalence of SRBD in children with achondroplasia. The presence and severity of SRBD may not be predicted by history, indicating an important role for PSG in this population. Two Level 4 studies suggest that children with craniofacial dysostosis have a high prevalence of OSAS that may be missed on history. Pharyngeal flap surgery is performed to correct velopharyngeal incompetence, particularly in patients who have had cleft palate repair. OSAS is a known complication of this procedure. Three Level 4 studies included PSG before and after surgery for velopharyngeal incompetence, and findings suggest a high prevalence of SRBD following surgery. The prevalence of SRBD prior to surgery was low, suggesting that PSG following surgery is more useful than before.

4.2.2.8.3 Sickle cell disease (SCD)

Six studies addressed the potential clinical utility of PSG for characterization of SRBD in children with SCD (1 Level 2 and 5 Level 4). The precise incidence of OSAS in children with SCD is not known and there is inconsistency in the literature about whether SRBD occurs more commonly in this population. Although children with SCD often experience nocturnal oxygen desaturations, it is not clear that they are more likely to have OSAS than children without SCD. However, children with SCD and OSAS appear to have more severe nocturnal oxygen desaturations compared with SCD subjects without OSAS. The task force recognizes a number of limitations in the literature in this area, and it is likely that future investigations will provide greater clarity regarding the clinical utility of and indications for PSG in children with SCD. Pulse oximetry may not provide an accurate measurement of SpO_2 values in SCD, and oximetry alone is probably not useful as a screening method for OSAS in children with SCD.

4.2.2.8.4 Neurological disorders

Twenty-four papers addressed the potential clinical utility of PSG for characterization of SRBD in children with neurological disorders including neuromuscular disorders (NMD). Findings showed that PSG was clinically useful in identifying SRBD and managing children with multiple neurological disorders including Duchenne muscular dystrophy (1 Level 3 and 5 Level 4); cerebral palsy (2 Level 4); meningomyelo-

cele, spina bifida and/or Chiari malformation (1 Level 3 and 3 Level 4); other neuromuscular disorders (2 Level 3 and 2 Level 4), and epilepsy (5 Level 4), including several studies that document an association between vagal nerve stimulator (VNS) implantation and OSAS in children. One Level 1 paper involving a variety of neurological disorders also documented clinical utility of PSG. Two other Level 4 studies found that children with neurological comorbidities were likely to have more postoperative complications, more severe preoperative PSG abnormalities, and less optimal responses to AT or other upper airway surgery for OSAS compared with neurologically normal children.

4.2.3 Clinical utility of PSG prior to adenotonsillectomy for confirmation of OSA diagnosis

AT is considered a first line of treatment for children with OSAS. Assessment of the clinical utility of PSG prior to AT is challenging for several reasons including significant variations in practice patterns, different pathological cut-offs for AHI values, and different methods for scoring respiratory events in children. Many ENT specialists do not routinely request PSG in children with suspected OSAS prior to AT,[5,6] while some request PSG selectively and others request PSG routinely before AT.

Thirty papers addressed 1 or more aspects of clinical utility of PSG prior to AT. The majority of the studies reviewed provided Level 3 or 4 evidence; only 5 Level 2 and no Level 1 papers were identified. The majority of studies were not designed specifically to assess the clinical utility of PSG prior to AT. However, these papers often provided data to address clinical utility indirectly. Several Level 3 and Level 4 studies show that symptoms, physical examination and certain laboratory tests are poor predictors of respiratory PSG findings in children for whom AT is being considered. This observation supports the clinical utility of PSG prior to AT in order to confirm the diagnosis of OSAS and to provide objective characterization of severity of respiratory disturbance during sleep.

Our search regarding the clinical utility of PSG for assessment of perioperative risk related to AT in children with SRBD identified 11 papers (1 Level 2, 1 Level 3 and 9 Level 4). The literature provides significant documentation to support the clinical utility of preoperative PSG to predict the likelihood of perioperative respiratory compromise in children with OSAS. Findings also suggest that a preoperative PSG with evidence of mild or minimal respiratory disturbance during sleep is associated with very low risk for perioperative complications.

Eight studies of children with adenotonsillar hypertrophy and/or obesity were identified. The preponderance of studies, including 2 Level 2 studies using pediatric scoring criteria, showed that OSAS improved dramatically postoperatively, but that a substantial minority of children experience residual OSAS. The AHI tended to predict those children with persistent OSAS after AT. This would support the utility of both preoperative PSG to determine high-risk patients, and/or postoperative PSG to determine the need for further treatment. However, the task force did not identify any prospective studies that specifically address whether clinical outcome following AT is improved in association with routine performance of PSG prior to AT in otherwise healthy children.

4.2.4 Clinical utility of PSG for assessment of infants less than 12 months of age with suspected SRBD or related conditions

The task force focused on SRBD and related conditions that typically present during infancy including primary sleep apnea of infancy, congenital central hypoventilation syndrome (CCHS), suspected SRBD and gastroesophageal reflux (GER) disease, apparent life threatening events (ALTEs), laryngotracheomalacia, and assessment of risk of sudden infant death syndrome (SIDS). There is topical overlap between some papers discussed in this section and those covered in other sections.

One paper with Level 2 evidence employed a prospective, blinded, controlled crossover design and demonstrated with full PSG that otherwise healthy premature infants at or near term and almost ready for hospital discharge experience frequent, unsuspected adverse cardiorespiratory events. Three articles with Level 4 evidence provided support for the clinical utility of daytime nap PSG or nocturnal PSG in infants born either preterm or at term, for differentiation between normal and abnormal breathing, and cardiorespiratory differences of heart rate and blood pressure, and sleep position. In 1 paper with Level 3 evidence, investigators concluded that full PSG provides the physiological data for proper diagnosis in young infants and that limited cardiorespiratory studies can be misleading in this population. Another article with Level 2 evidence evaluated 14 infants with cyanotic breath-holding spells, and all subjects were found to have PSG abnormalities consistent with SRBD. Four infants had an AHI < 1, and the SRBD would not have been detected without esophageal pressure monitoring. This is a small exploratory study, but findings suggest that infants who present with cyanotic breath holding spells may require PSG to evaluate for SRBD.

4.2.4.1 Suspected primary sleep apnea of infancy

We found no articles that specifically addressed the clinical utility of PSG for establishing a diagnosis of primary sleep apnea of infancy. It is likely that most infants with this entity are diagnosed based on the clinical history and observations in the nursery setting. Clinically, these infants experience recurrent apneas with or without bradycardia and a variety of potential etiologies or comorbid conditions exist including prematurity, GER and other medical disorders, and neurological disorders.

4.2.4.2 Suspected congenital central hypoventilation syndrome

Two papers (1 Level 2 and 1 Level 4) provide limited data that addressed the potential clinical utility of PSG for evaluation of suspected CCHS. Further investigations may clarify the clinical utility and timing of PSG in suspected CCHS, the role of PSG in assessment of asymptomatic carriers of the PHOX2b mutation, and when periodic reevaluation may be necessary.

4.2.4.3 Suspected SRBD and gastroesophageal reflux

Our search regarding the clinical utility of PSG, including the simultaneous recording of lower esophageal pH monitoring, in infants with suspected GER and SRBD identified 7 papers (1 Level 2, 2 Level 3, and 4 Level 4). A Level 2 study reported that in subjects with respiratory dysfunction, GER was present in 75%; conversely, in subjects with GER, respiratory dysfunction was present in 45%. In other studies, findings were variable, and there were a variety of methodological limitations. Further

investigations are needed to understand the diagnostic yield and clinical utility of lower esophageal pH monitoring during overnight PSG in infants.

4.2.4.4 Apparent life-threatening events

Thirteen papers addressed the potential clinical utility of PSG in this population (1 Level 1, 5 Level 2, 4 Level 3, and 3 Level 4). These studies suggest that GER, as well as subtle or nonspecific abnormalities may be identified during PSG in this population, but it was not possible to estimate the diagnostic yield of PSG based on these results. Altered cardiovascular regulation is most likely present but routine clinical PSG parameters do not measure this. It is possible that PSG may be clinically useful in selected populations, particularly when there is clinical concern for upper airway obstruction or other forms of SRBD. In general the prognosis for recurrence of ALTE could not be predicted based on PSG findings, and a significant proportion of infants who experience an ALTE have a normal PSG.

Two Level 3 and 1 Level 4 studies suggest that infants who experience an ALTE are at increased risk for SRBD because of facial dysmorphology, or other risk factors for SRBD. However, further evaluation is needed to assess the clinical utility of PSG in this population.

4.2.4.5 Laryngotracheomalacia and suspected SRBD

One Level 4 paper was identified that addressed the clinical utility of PSG for assessment of infants with laryngomalacia and suspected SRBD. Findings suggest that PSG may have clinical utility in evaluating SRBD before and after surgical intervention, particularly if there is clinical concern for moderate to severe respiratory disturbance. However, it is not possible to confirm the clinical utility of PSG in this population of infants based on a single paper.

4.2.4.6 Assessing risk of sudden infant death syndrome (SIDS)

Seven papers (each with Level 3 evidence) addressed the potential clinical utility of PSG for assessment of risk for SIDS. All papers were case-control studies with performance of full PSG. Although a variety of PSG findings have been reported in subjects who later died due to SIDS, PSG does not provide sufficiently distinctive or predictive findings to support a routine clinical indication for PSG to determine risk of death due to SIDS. This is an area of active investigation and future work with more sophisticated techniques that may lead to greater clinical utility of PSG.

4.3 Other Chronic Respiratory Disorders

4.3.1 Clinical utility of PSG in children with chronic obstructive lung disease: asthma, cystic fibrosis, and bronchopulmonary dysplasia

Two studies (1 Level 3 and 1 Level 4) evaluated the clinical utility of PSG in children with asthma to identify OSAS. Because of limited data and variable findings, no conclusions can be made regarding whether PSG is routinely indicated in children with asthma. However, clinical screening for signs and symptoms of OSAS, particularly those with suboptimal control or those with multiple risk factors for OSAS, appears warranted. Two Level 4 studies addressed the clinical utility of PSG in children and young adults with cystic fibrosis (CF), and one study demonstrated that PSG can be used to initiate and titrate noninvasive ventilation (NIV) in patients with CF. These limited data suggest that PSG may have clinical utility in managing CF patients with SRBD. No papers were identified that address bronchopulmonary dysplasia. Given that infants with bronchopulmonary dysplasia have significant medical or neurodevelopmental co-morbidities that may confer a higher risk for SRBD, clinical screening for SRBD is warranted.

4.3.2 Clinical utility of PSG in children with chronic restrictive lung disease: kyphoscoliosis and other chest wall abnormalities; restrictive parenchymal lung disease, including diaphragmatic hernia; and neuromuscular weakness and progressive respiratory insufficiency

Two Level 4 papers addressed kyphoscoliosis and other chest wall abnormalities. The papers provide limited evidence to support clinical utility of PSG in identifying SRBD in this population, and there is no evidence to support routine performance of PSG prior to surgical intervention. No papers were identified that address restrictive lung disease, including diaphragmatic hernia. The clinical utility of PSG in children with neuromuscular weakness and progressive respiratory insufficiency is discussed in section 4.2.2.9.4 (Clinical utility of PSG in neurological disorders).

4.4 Clinical Utility of PSG for Therapeutic Intervention

Seven papers, including 1 Level 2 study, addressed the clinical utility of PSG for titration of positive airway pressure (PAP) in children. Published reports suggest there is significant regional variation in practice patterns, and there is general acceptance of PSG as a useful procedure for PAP titration in children. The task force also identified 5 studies (1 Level 3 and 4 Level 4) that evaluated or described the clinical utility of introducing, titrating and reassessing nocturnal intermittent positive pressure ventilation (NIPPV) in children with SRBD and neuromuscular disorders (NMD). Repeat sleep studies were often needed to adjust PPV settings or switch treatment modalities.

The task force identified 1 Level 4 paper that addresses the clinical utility of repeat PSG in children on chronic PAP support. The authors concluded that PSG provides important information for optimizing long term management with PAP.

A search was performed on PSG following AT or other procedures (including rapid maxillary expansion [RME], oral appliances, pharyngeal flap surgery for velopharyngeal incompetence, and supraglottoplasty) to assess response to intervention. In 1 Level 3 and 1 Level 4 study, OSAS on PSG was shown to improve after RME; however, significant residual disease remained, and PSG was useful in determining whether additional treatment was necessary. Two studies (1 Level 2 and 1 Level 4) examined oral appliances in children. These studies support the clinical utility of PSG for evaluation of respiratory function intervention. Pharyngeal flap surgery is performed to correct velopharyngeal incompetence, particularly in patients who have had cleft palate repairs. OSAS is a known complication of this procedure. Three Level 4 studies performed PSG before and after surgery for velopharyn-

geal incompetence. In general, these studies support the use of PSG following pharyngeal flap surgery, but do not support the routine use of PSG preoperatively. Two Level 4 studies used PSG to evaluate efficacy of supraglottoplasty in infants with severe laryngomalacia. One Level 4 study indicated that PSG was useful in assessing the response to surgical procedures such as AT or posterior fossa decompression in children with myelomeningocele; in most cases, SRBD did not resolve postoperatively. One Level 4 study showed that 6 infants with micrognathia who underwent mandibular distraction had improvements in OSAS postoperatively on PSG, although no details were provided.

4.4.4 Consideration of decannulation of tracheostomy

One paper with Level 3 evidence demonstrated that PSG is a useful supplement to airway endoscopy in the evaluation of readiness for decannulation in children with long-term tracheostomy.

4.4.5 PSG for management of mechanical ventilator settings or weaning from ventilator support

One Level 4 paper addressed the potential clinical utility of PSG for management of patients who require mechanical ventilation or weaning of mechanical ventilator support. This paper was of limited usefulness because PSG was not used to adjust ventilation settings, nor did every child have a complete PSG. It is likely that PSG is used selectively to assist with management of ventilator settings or weaning from ventilator support, but the task force found no evidence to support or not support the clinical use of PSG in this situation.

4.4.6 Titration of supplemental oxygen

No papers were identified that addressed the potential clinical utility of PSG for titration of supplemental oxygen for treatment of sleep related hypoxia in children. This may reflect the common clinical practice of using overnight oximetry to assist with supplemental oxygen titration or empirical clinical decisions using caregiver observations.

4.4.7 PSG in relation to use or discontinuation of infant apnea monitors

One paper with Level 4 evidence demonstrated how PSG can be used to evaluate infants who present with an ALTE who were monitored at home with an infant apnea monitor following the event. Although nonspecific abnormalities may be present on PSG in infants being monitored with infant apnea monitors, the task force identified no papers that provide guidance regarding PSG as a predictor for the use or discontinuation of infant apnea monitors or that provide data that predict recurrent apnea or death.

4.4.8 PSG for assessment and monitoring of children with Prader-Willi syndrome being considered for or receiving growth hormone supplementation

Reports of sudden death in children with Prader-Willi syndrome (PWS) who are receiving growth hormone supplementation have raised the issue of whether clinicians should monitor for physiological abnormalities during sleep that may predict risk for SRBD or sudden death in this population. The task force identified 3 papers (1 Level 2, 1 Level 3, and 1 Level 4) that address this issue. Findings do not provide sufficient support for the routine use of PSG to predict risk of death or to monitor for development of significant cardiorespiratory abnormalities in this population. Additional studies with larger numbers of subjects and longitudinal data are needed to develop a more complete profile of risk for sudden death in this population. This profile will most likely require integration of clinical factors and PSG findings.

5.0 SUMMARY

This analysis documents strong face validity and content validity, moderately strong convergent validity when comparing respiratory PSG findings with a variety of relevant independent measures, moderate-to-strong test-retest validity, and limited data that support discriminant validity for PSG for characterization of breathing during sleep in children. The analysis documents moderate-to-strong test-retest reliability and interscorer reliability based on limited data.

Findings indicate particularly strong clinical utility in children with obesity, evolving metabolic syndrome, neurological, neurodevelopmental, or genetic disorders (for example, Down syndrome and Prader-Willi syndrome), and children with certain craniofacial syndromes and clinical features of SRBD. The task force gave specific consideration to the clinical utility of PSG prior to AT for confirmation of OSA, and for assessment of perioperative risk. The most relevant findings included: (1) recognition that the clinical history and physical examination are often poor predictors of respiratory PSG findings, (2) preoperative PSG is helpful in predicting risk of perioperative complications, and (3) preoperative PSG is often helpful in predicting persistence of OSA in a substantial minority of patients after AT. The latter issue is important because it may help identify children who require further treatment. However, the task force did not identify any prospective studies that specifically address whether clinical outcome following AT for treatment of OSA in children is improved in association with *routine* performance of PSG before surgery in otherwise healthy children.

Limited data are available regarding the clinical utility of PSG (1) in infants less than 12 months of age with suspected SRBD; (2) for evaluation of children with chronic respiratory disorders such as chronic obstructive or restrictive lung disease, and suspected SRBD; and (3) for therapeutic purposes including PAP titration, repeat PSG following AT or other surgical procedures, consideration of changes in mechanical ventilator management, decannulation of tracheostomy, and other uses. A small but useful group of papers confirmed the usefulness of PSG for initiation and titration of PAP in children with OSAS. However, the data do not address the optimal timing for repeat studies in children on PAP.

Based on assessment and integration of findings from over 240 evidentiary papers, it is the consensus of the task force that pediatric PSG shows validity, reliability, and clinical utility that is commensurate with most other routinely employed diagnostic clinical tools or procedures. It is apparent that the "gold standard" for diagnosis of SRBD in children is not PSG alone, but rather the skillful integration of clinical and PSG findings by a knowledgeable sleep specialist. Future develop-

ments will provide more sophisticated methods for data collection and analysis, but integration of PSG findings with the clinical evaluation will represent the fundamental diagnostic challenge for the sleep specialist.

6.0 FUTURE DIRECTIONS

This review highlights the need for well-designed, well-powered studies that evaluate the operating characteristics of PSG in a broad range of populations. The most pressing need involves investigation of special populations including children with obesity and other risk factors for cardiovascular disease, neurodevelopmental and neuromuscular disorders, sickle cell disease, and certain craniofacial syndromes. More studies that involve children with metabolic syndrome and children with overt or evolving hypertension are needed, and the clinical utility of PSG in infants less than 12 months of age is not well understood. Whether PSG is routinely indicated prior to AT in otherwise healthy children is not fully resolved, but it is clear that preoperative PSG is useful in identification of children at increased risk for perioperative complications. Postoperative PSG is useful in assessment of response to AT and determination of whether addition treatment is necessary for residual OSAS. Finally, the feasibility and clinical utility of unattended testing outside the sleep laboratory requires investigation in children.

ACKNOWLEDGMENTS

The task force would like to thank Sharon Tracy, PhD and Christine Stepanski, MS for their efforts in the development of this manuscript.

DISCLOSURE STATEMENT

This was not an industry supported study. The authors have indicated no financial conflicts of interest.

REFERENCES

1. American Thoracic Society. Standards and indications for cardiopulmonary studies in children. Am J Respir Crit Care Med 1996;153:866-78.
2. American Academy of Pediatrics. Clinical practice guideline: Diagnosis and management of childhood obstructive sleep apnea syndrome. Pediatrics 2002;109:704-12.
3. Iber C, Ancoli-Israel S, Chesson A, Quan SF, for the American Academy of Sleep Medicine. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, 1st ed. Westchester, IL: American Academy of Sleep Medicine; 2007.
4. American Academy of Sleep Medicine. International classification of sleep disorders, 2nd ed.: Diagnositic and coding manual. Westchester, IL: American Academy of Sleep Medicine; 2005.
5. Weatherly RA, Mai EF, Ruzicka DL, Chervin RD. Identification and evaluation of obstructive sleep apnea prior to adenotonsillectomy in children: a survey of practice patterns. Sleep Med 2003;4:297-307.
6. Mitchell RB, Pereira KD, Friedman NR. Sleep-disordered breathing in children: Survey of current practice. Laryngoscope 2006;116:956-8.