
Sushmita Pamidi, Susan Redline, David Rapoport, Indu Ayappa, Luciana Palombini, Ramon Farre, Jason Kirkness, Jean-Louis Pépin, Olli Polo, Andrew Wellman, and R. John Kimoff; on behalf of the American Thoracic Society Ad Hoc Committee on Inspiratory Flow Limitation

This official Workshop Report of the American Thoracic Society (ATS) was approved by the ATS Board of Directors, March 2017

Abstract

This report summarizes the proceedings of the American Thoracic Society Workshop on the Noninvasive Identification of Inspiratory Flow Limitation in Sleep Studies held on May 16, 2015, in Denver, Colorado. The goal of the workshop was to discuss methods for standardizing the scoring of flow limitation from nasal cannula pressure tracings. The workshop began with presentations on the physiology underlying flow limitation, existing methods of scoring flow limitation, the effects of signal acquisition and filtering on flow shapes, and a review of the literature examining the adverse outcomes related to flow limitation. After these presentations, the results from online scoring exercises, which were crowdsourced to workshop participants in advance of the workshop, were reviewed and discussed. Break-out sessions were then held to discuss potential algorithms for scoring flow limitation. Based on these discussions, subsequent online scoring exercises, and webinars after the workshop, a consensus-based set of recommendations for a scoring algorithm for flow limitation was developed. Key conclusions from the workshop were: (1) a standardized and automated approach to scoring flow limitation is needed to provide a metric of nonepisodic elevated upper airway resistance, which can then be related to clinical outcomes in large cohorts and patient groups; (2) at this time, the most feasible method for standardization is by proposing a consensus-based framework, which includes scoring rules, developed by experts; (3) hardware and software settings of acquisition devices, including filter settings, affect the shape of the flow curve, and should be clearly specified; and (4) a priority for future research is the generation of an open-source, expert-derived training set to encourage and support validation of automated flow limitation scoring algorithms.

Supported and developed by the American Thoracic Society and National Institutes of Health grant R24HL114473.

Correspondence and requests for reprints should be addressed to Sushmita Pamidi, M.D., McGill University Health Centre—Glen Site, Respiratory Division, Block D–D05.2506, 1001 Decarie Boulevard, Montreal, PQ H4A 3J1 Canada. E-mail: sushmita.pamidi@mcgill.ca.

This article has an online supplement, which is accessible from this issue’s table of contents at www.atsjournals.org

Contents

Overview
Methods
Workshop Findings and Recommendations
  Signal Analysis and Flow Measurement Techniques in Assessment of Flow Limitation
  Recommendations
  Physiologic Basis for Flow Limitation with Respect to Flow Shape
  Characteristics Described in the Literature
  Flattening
  NED
  Increased inspiratory time
  Snoring
  Recommendations
  Current Automated Approaches for Detecting Flow Limitation in the Literature
  Recommendations
  The Utility of Resistance Measurements and Transcutaneous CO₂ in Validating Manual Scoring Methods
  Transcutaneous CO₂ Recommendations
  Relationship of IFL to Adverse Outcomes
  Recommendations
  Results of the Crowdsourcing Online Challenges for the Visual Detection of Flow Limitation
  Recommendations
  Summary of Key Findings and Future Directions
Overview

Although the sleep literature has described various methodologies for manual and automated detection of noninvasive inspiratory flow limitation (IFL) using the nasal cannula flow signal, these methods have not been standardized or validated. This workshop was organized to provide a forum for discussion among international experts on airflow limitation regarding an approach to standardizing visual analysis of IFL that could be used across sleep centers. The specific objectives of the workshop were as follows:

1. **Signal analysis**: To understand how methodological differences in flow measurement techniques influence the visual identification of flow limitation.

2. **Physiology**: To review the conceptual basis of shape characteristics of the flow/time curve as a marker of upper airway dysfunction.

3. **Existing algorithms**: To understand current approaches for detecting flow limitation from the shape of the flow/time curve alone, as well as using esophageal pressure (Pes) and carbon dioxide (CO₂) measurements.

4. **Adverse effects**: To review the available data on IFL in a normal population, and the relationship of IFL to adverse clinical outcomes.

5. **Expert consensus**: To generate community input on the visual detection of flow limitation using online crowdsourcing exercises, and use this information to develop a consensus-based algorithm for scoring.

6. **Future directions**: To discuss approaches to developing a consensus-based training set to encourage and support validation of automated algorithms.

After presentations by experts on these topics and break-out sessions at the workshop, recommendations on an approach to standardizing IFL were made as follows:

1. Establish criteria for recognizing the presence of flow limitation, derived by a group of experts.

2. Develop an algorithm based on shapes of the inspiratory flow curve to score flow limitation and agree on optimal signal and filter settings.

3. Improve agreement in flow limitation scoring and refine a consensus-based scoring algorithm through crowdsourcing of flow tracings.

4. Develop a training set of signals and IFL metrics based on a consensus-based algorithm and final scores from experts in scoring flow limitation.

5. Develop an open-source platform to use in the development and validation of automated algorithms based on this training set.

6. Validate the automation of scoring by comparing with manual scoring by experts.

IFL is characterized by flattening on the flow/time tracing. The most commonly used surrogate for flow in diagnostic sleep studies is currently the tracing from a nasal pressure transducer, which reflects a lack of increase in airflow despite increasing respiratory effort (1). However, to date, standardized criteria for scoring of IFL have not been established. There are some data demonstrating an association between IFL and transient arousals, disturbed sleep architecture, and daytime hypersomnolence (2). Continuous positive airway pressure (CPAP) studies have also demonstrated that treatment of IFL, over and beyond suppressing apneas and hypopneas, may improve neurocognitive function (3). Scoring IFL would likely be of greatest clinical importance in sleep studies with a low apnea–hypopnea index (AHI). Examples include pregnant women and pediatric or symptomatic adult patients with an AHI in the normal to mild range.

IFL, independent of arousals, is currently omitted from standard sleep scoring, and there are no accepted rules for quantifying its presence or severity. However, scoring of respiratory effort–related arousals (RERAs), a subtle form of recurrent upper airway obstruction where flow-limited events are followed by EEG microarousals, is accepted by the American Academy of Sleep Medicine (4). RERAs that are associated with daytime impairment define upper airway resistance syndrome (UARS) (5). As an increasing number of ambulatory sleep studies without EEG are being performed, an important research objective is to determine whether IFL, independent of EEG microarousals, has significant clinical implications.

To date, IFL has been identified by varying visual recognition criteria. Although the concept and gold standard for detecting IFL involves the evaluation of pneumotachographic airflow simultaneous with esophageal manometry as a measure of inspiratory effort, these tools are poorly tolerated by patients, and infrequently used in clinical sleep studies. With movement of sleep data acquisition out of the laboratory and into the home environment, where there is minimal technician involvement, it is unlikely that Pes measurement will become more widespread. Thus, the workshop focused on the indirect detection of IFL, through the use of inspiratory pressure transducer shape alone as a surrogate for the relationship of respiratory effort to flow (i.e., IFL). In addition, even if Pes and true flow were collected, there is a consistent algorithm in the literature for quantifying IFL. Montserrat and colleagues (6) have demonstrated that the nasal pressure airflow signal correlates well with pneumotachographic flow. Thus, the standardized visual detection of IFL from a nasal cannula pressure signal could enhance the evaluation of sleep-disordered breathing in situations where a low AHI fails to capture suspected risk for health outcomes. The overarching goal of this workshop was to develop an approach to standardizing visual analysis of IFL in the clinical setting, which can be applied across sleep centers and in epidemiological datasets without adding unacceptable patient burden, while improving information content.

**Methods**

The workshop proposal was developed by an ad hoc group involved in collaborative work on standardization of IFL scoring (S.P., J.K., S.R., D.R., I.A., and L.P.). The workshop proposal was peer reviewed and funded through the American Thoracic Society (ATS) Assembly project funding mechanism under the auspices of the Assembly on Sleep and Respiratory Neurobiology. A 1-day workshop was held at the ATS International Conference in Denver, Colorado, in May 2015. Participants included international researchers working on IFL with expertise across populations (e.g., pediatrics, pregnancy, mild sleep apnea) and individuals with biomedical engineering expertise. All workshop participants disclosed potential conflicts of interest, and these were managed in accordance with standard ATS policies.

The workshop started with a series of presentations delivered by participants with
expertise on topics of IFL addressing the predefined objectives. For the afternoon, results of a preworkshop online scoring challenge were presented. This challenge, consisting of a test set of records, was crowdsourced to workshop participants and other members of the sleep research community in the weeks preceding the workshop, and was made available through the NHLBI National Sleep Research Resource (NSRR) website (7) (hosted at the Brigham and Women’s hospital; sleepdata.org).

A break-out session format with small groups was used to develop a plan for creating and validating a training set of breaths for visually scoring IFL. Further discussions on strategies for developing a standardized IFL algorithm took place through teleconferences after the workshop, and online surveys were used in areas of disagreement. In October 2015, a consensus-based scoring algorithm was applied during a second online challenge for manually scoring IFL, which also addressed the role of various filter settings on scoring agreement. All workshop participants had the opportunity to review and revise the manuscript in its final form.

**Workshop Findings and Recommendations**

The following key research objectives were identified by the workshop.

**Signal Analysis and Flow Measurement Techniques in Assessment of Flow Limitation**

Indirect evidence of elevated upper airway resistance can be inferred from the shape of the inspiratory airflow tracing; however, acquisition, processing, and conditioning affect the signal and trade-offs occur (8). The goal is to capture maximum information about instantaneous flow while eliminating unwanted aspects of the signal (noise, drift, and distortion due to the transducers and amplifiers).

True flow-measuring devices (e.g., pneumotachographs) are cumbersome and rarely used in diagnostic clinical sleep studies. The surrogate measure of flow derived from a nasal pressure signal is often from a nasal cannula/pressure transducer system. In this context, the nasal aperture provides a resistance that generates pressure related to the patient’s airflow via a linear or quadratic relationship (9). There may also be a component of the pressure signal that is generated by the Pitot-like behavior of the nasal cannula if it is kept aligned with the direction of nasal airflow (9). Although the actual relationship of the pressure signal to flow is close to quadratic (i.e., flow is equal to the square root of pressure measured), the unmodified pressure signal has been shown to be a useful first approximation, and extensive experience has shown it gives sensitive and reproducible detection of hypopnea. However, complete cessation of airflow (apnea) may be, at times, overread (10, 11).

When analog pressure signals are converted at a finite sampling frequency for digital representation, the sampling rate must be at least two times the highest frequency to be analyzed from the signal, according to Nyquist’s Sampling Theorem. This requirement implies that, to capture frequencies of 30 Hz, the sampling rate needs to be set to at least 60 Hz. For breathing rate, low sampling frequencies (10 Hz) may be sufficient, but higher sampling frequencies are necessary for analysis of the shape of inspiratory flow rate if sharp contours (i.e., “flattening”) are to be captured. Empiric trials suggest that frequencies of 25–50 Hz are sufficient to capture most of the shape characteristics of IFL, whereas frequencies of greater than 100 Hz are needed to define snoring. Using lower sampling frequencies may “cost” information and distort shape by smoothing. Before signals are acquired and digitized, signals are amplified using either AC- or DC-coupled amplifiers, and may be filtered at the time of collection (hardware filtering) or afterwards with software filters. Appropriately chosen AC-coupled amplification effectively produces high-pass filtering. By eliminating low frequencies (such as baseline drift), this can isolate high frequencies (e.g., snoring). However, AC-coupled amplification will significantly distort the signal if applied inappropriately. DC-coupled amplification is usually coupled with a low-pass filter to remove high frequencies that are deemed irrelevant to the signal’s information (e.g., “noise”). At the collection stage, it is often desirable to use minimal hardware filtering, thus avoiding inadvertent and irreversible distortion of the signal; digital filtering can then be applied after acquisition through software. Commercial pressure transducers are usually packaged with AC-coupled amplifiers. DC-coupled amplifiers come at added cost, and can drift over time.

A question that was extensively discussed by the workshop participants was whether the flow signal needs to be linearized (i.e., by taking the square root of the nasal pressure signal to derive flow) for visual scoring. Linearization of the nasal pressure signal by taking the square root is theoretically indicated to accurately describe flow from turbulent systems; it may be the most accurate linearization of the nasal pressure signal. However, this transformation comes at a significant cost if the baseline of the flow cannot be defined absolutely. Existing pressure transducers currently on the market are predominantly AC coupled. Accordingly, the flow signal information contained within most clinical and research sleep studies has been collected using this configuration. The impact of using a square root transformation is generation of data with a “floating” baseline. The assumption underlying this is that the volume of inspiration equals the volume of expiration, and thus the average over time will define zero volume and flow. This assumption can fail when there is significant unidirectional mouth expiration or with any drift of a bias flow (as during CPAP). In these situations, there is potentially an incorrect “zero” to the signal that arises from applying the “linearization,” and this error cannot be detected from the hardware output or reversed in software. Furthermore, amplitude is necessarily altered by the square root transform, and can affect scoring that is amplitude dependent (e.g., the definition of hypopnea). Thus, despite the theoretical advantage of linearization of the pressure-to-flow conversion, the workshop participants concluded that it is better to recommend the recording of “raw” nasal pressure signals. As noted here, there is also good empirical evidence that visual assessment of IFL can be accomplished similarly with and without the square root transformation applied to the nasal pressure signal in existing clinical and research tracings, provided that the scorer has been trained on the same signal. There are also practical advantages associated with the availability of a predominance of data sets that have tracings where nasal pressure has been recorded without linearization or square root transform. Automated algorithms for analysis of IFL can add linearization or other
transformation in software, and some current algorithms for detecting IFL probably are already are tuned to the square root–transformed pressure signal. Thus, the workshop recommended continued recording of flow signals without linearization.

IFL is unlikely to be a binary phenomenon. Nonetheless, to date, most analyses have used “presence” or “absence” of IFL to make decisions about diagnosis (e.g., UARS) or therapy (e.g., CPAP titration) without explicitly addressing the issue of IFL severity. Unfortunately, validated visual and/or automated algorithms to quantitate severity of IFL are not yet available. Therefore, the workshop was unable to recommend approaches for quantifying IFL. This important topic should be addressed in the future. By following the proposed data collection procedures, workshop participants anticipate that data will be available in the future to support this research.

Recommendations.

- Use transducers for nasal pressure that have good resolution in the range of nasal cannula signal during quiet breathing (similar to transducers used in human physiological research with pneumotachograph heads for resting quiet breathing). This pressure range is typically ±2 cm H2O.
- Pressure transducers that are intended for use with CPAP flows typically have a range of 0–20 cm H2O and, therefore, should not be used with nasal cannula pressure signals unless they are of high resolution in the low range and adequately digitized by the hardware of the CPAP unit.
- Where possible, use sampling rates of over 100 Hz to obtain information on snoring.
- The square root transform should not be applied at collection to signals obtained with AC coupling or filtering that suppresses baseline shift in the pressure signal. Visualization of flow signals should be conducted without the square root transform. Any algorithm to recognize IFL should be tested for specificity to a signal that is, or is not, transformed, as it may not generalize to the other condition.
- To avoid confusion about polarity, it is recommended that inspiration corresponds to an upward deflection when displaying flow signals. Biocalibration of the nasal flow signal should be performed to ensure this polarity, or at least to identify the orientation of inspiration versus expiration when analyzing the breaths for flow limitation.

Physiologic Basis for Flow Limitation with Respect to Flow Shape Characteristics Described in the Literature

There is a limited body of literature validating analytic approaches to the manual identification of IFL. Nine articles demonstrated six features used to identify the presence or classify severity of IFL, as follows: (1) flattening of airflow—supraglottic (or esophageal) pressure relationship (12–17); (2) presence of negative effort dependence (NED) (12, 14–16); (3) the shape or function of the airflow–time curve (1, 14, 15, 18, 19); (4) increased inspiratory duty cycle (17); (5) high-frequency fluctuations/oscillations; and (6) airflow–effort asynchrony (16). All studies were performed during non–rapid eye movement (REM) sleep (one study included REM, one study included a whole night) and included a sample size varying from 5 to 26 subjects (median = 11 subjects), with all but 1 using pneumotachograph airflow as the reference standard. Six studies used a nasal pressure signal from either nasal cannula or mask, and supraglottic or Pes was recorded in all but one study. NED (falling flow rate for at least >1 cm H2O drop in pharyngeal pressure signal) (12) was used to classify IFL in five out of nine reports (12, 14–16). When evaluating criteria using only the nasal cannula pressure signal, the workshop participants concluded that, for developing a validated set of criteria for identification of IFL, the following four features, for which the physiologic basis is described subsequently here, should be included:

Flattening. Flattening of the inspiratory flow signal as an indicator of upper airway obstruction is based on the Starling resistor model. In some cases, the upper airway during sleep is believed to behave like a Starling resistor (20, 21), such that when the downstream pressure decreases below the patient’s “critical pressure,” a choke point develops at the point of pharyngeal collapse. Once the choke point develops, further reductions in downstream pressure cannot propagate upstream past the choke point. Consequently, the pressure gradient in the upstream segment of the airway (from the nose to the site of collapse) remains fixed; hence, the volume flow rate through the upstream segment (and thus the downstream segment as well) remains fixed. The result is that flow remains flat during the progressively increasing effort of inspiration.

The contour of the inspiratory flow tracing has been demonstrated to be related to upper airway resistance during titration of CPAP (8). Speciﬁcally, a plateau in the breath contour correlates well with increased airway resistance. Similarly, in non–CPAP studies, it has been shown that a plateau on the inspiratory flow tracing from the nasal cannula (6) also signiﬁes increased upper airway resistance and the presence of flow limitation (1). Figure 1 visually demonstrates flattening during an individual breath.

NED. Instead of flattening, some breaths can exhibit NED, in which flow decreases as inspiratory effort increases (Figure 2). NED has been observed in patients with obstructive sleep apnea (OSA), but the extent to which it is present varies between individuals. It has been suggested that NED may be impacted by variability in upper airway pharyngeal dilator activity (22).

The physiologic basis for NED is as follows. Unlike the Starling resistor described previously here, in patients with NED, reductions in downstream pressure propagate upstream, past the choke point, and narrow the upstream segment—the greater the reduction in downstream pressure, the more the narrowing of the upstream segment. The more the narrowing of the upstream segment, the more airflow is reduced. The result is that, as someone breathes in, the flow rate drops across the breath, producing a “scooped out” inspiratory flow trace, as shown in Figure 2. In many patients, NED may alternate or coexist with inspiratory flattening (23).

Increased inspiratory time. An increase in upper airway resistance will also increase the inspiratory time (Ti), which could therefore be used as an indicator of IFL. In an effort to preserve tidal volume, patients tend to increase their Ti when breathing through a high resistance (i.e., they increase their “neural Ti”), meaning the diaphragm activates longer. Because respiratory frequency will affect Ti, no absolute Ti is
likely to characterize elevated resistance across patients, and Ti/Ttot (where Ttot is the total duration of the respiratory cycle) has a very small dynamic range. Changes in absolute Ti within a patient have been shown to have more value in detecting periods of high resistance (24, 25). See the online supplement for further details.

Snoring. In small samples of research subjects, snoring, which refers to a vibratory sound of the pharyngeal structures, has been shown to be related to increased total

Figure 1. Flattening is shown pictographically here in a nasal cannula flow tracing. Flow-limited breaths, characterized by flattening, demonstrate a nonlinear relationship between driving pressure and flow in contrast to normal-contour breaths. Reprinted by permission from Reference 1.

Figure 2. Negative effort dependence (NED) within individual breaths is demonstrated here. After peak inspiration is reached, flow decreases throughout the remaining part of the breath cycle despite increases in epiglottic pressure or effort. An increase in flow is sometimes observed at the end of a breath cycle (terminal peak) as driving pressure reaches its baseline. Adapted by permission from Reference 22.
pulmonary resistance and flow limitation (26). Moreover, snorers tend to have more collapsible airways (27). Snoring has also been shown to be associated with increased effort during inspiration as measured by Pes (25, 28, 29).

**Recommendations.** To detect IFL in individual breaths, visual inspection of the nasal cannula pressure signal should rely on the following:

- As determined by expert consensus, evidence of flattening and/or scooping (to indicate NED) for 75% or greater of the duration of the inspiratory cycle on the nasal pressure signal as a representation of flow limitation. When the presence of flattening and/or scooping occurs for less than 75% of the duration of the inspiratory cycle, expert consensus is less clear as to whether this is still indicative of flow limitation.
- When these criteria are ambiguous or only partially met, consider additional criteria to identify IFL, including the presence of a prolonged Ti (by approximately 10%) (24) relative to the breath cycle or visualized snoring identified by fast vibrations on the nasal pressure signal. These ancillary criteria should not be used to score IFL in the unequivocal absence of inspiratory flattening or NED unless both of these criteria are present simultaneously.

**Current Automated Approaches for Detecting Flow Limitation in the Literature**

During CPAP titration, manual classification of breaths into normal, intermediate, and definite IFL from the airflow signal shape alone has been reported. Two studies inferred validity of classification by demonstrating higher peak resistance/ Pes in IFL breaths compared with normal breaths (8, 30). Algorithms for automatic classification based on shape have been developed and are the basis of auto-titrating CPAP devices (31). However, these algorithms are proprietary, and little validation is available. Their validity for IFL detection has been inferred from empiric experience in successfully titrating patients with OSA. However, this cannot be generalized to the analysis of nasal pressure signals on diagnostic recordings without validation.

In the analysis of diagnostic studies using the nasal cannula airflow signal, automated classification of breaths into three to four distinct shapes has been reported using custom software (1), artificial neural networks (32), and finite automata analysis (33). Validity of classification has been inferred from resistance (measured using Pes/flow (1, 32)) and by showing good agreement with manual classification (32, 33). Other approaches used an automated clustering algorithm to classify breaths into 1 non-IFL and 23 IFL shapes (33), or discriminant feature analysis using machine learning methods (34–36).

The intent was to identify clinical subtypes by flow pattern and compare them with the detection accuracy of an expert, but the utility and clinical relevance of the larger number of distinct shapes remains unproven.

All automated analyses to date have required significant preprocessing and manual review for acceptable data quality, particularly for accurate breath detection. Preliminary unpublished experience indicates that algorithms developed for analyzing flow during CPAP are not applicable for nasal cannula data without further tuning.

**Recommendations.**

- Manual visual scoring of IFL is time consuming and cumbersome for large datasets of sleep studies, and the development of automated techniques is desirable. However, before using automated algorithms, these must be tested against either physiologic parameters or a proposed consensus recognition. In particular, validation should be specifically performed on the nasal pressure transducer signal during diagnostic studies rather than from CPAP devices used during therapy.

**The Utility of Resistance Measurements and Transcutaneous CO₂ in Validating Manual Scoring Methods**

The gold standard method for assessing IFL consists of: (1) measuring V at the airway opening with a pneumotachograph; (2) estimating driving Pes with an esophageal catheter; and (3) documenting that an increase in driving pressure results in no increase, or even decrease, in flow (1, 8, 30). Accordingly, the condition of flow limitation is associated with an increase in the effective resistance (R) of the upper airway computed as Pes/V. The resulting value for R summarizes the information contained in Pes and V over the entire breath into a unique variable representing apparent airway obstruction (although R changes dynamically during IFL). The question of whether R could be a signal useful to detect flow limitation was raised in early (8) and subsequent (24, 37) studies. The potential interest in R as a marker of obstructive sleep events was further enhanced by using the forced oscillation technique (FOT), which is based on applying a high-frequency (e.g., 5 Hz) pressure oscillation at the airway opening (P_{FOT}). Given that the muscle pump generates virtually no pressure at FOT frequency, driving pressure can be estimated solely from P_{FOT}, and thus R can be noninvasively estimated as P_{FOT}/V (38, 39). However, all the available data on R measured in sleep studies, either invasively with a catheter or noninvasively by FOT, indicate that, whereas this variable is a good marker of airway obstruction, it does not specifically distinguish between hypopneas with or without flow limitation (24, 38). Indeed, in a rigid wall airway, V is exclusively determined by the driving pressure. By contrast, in a collapsible airway (e.g., a Starling resistor), V depends on both driving pressure and transmural pressure, as this gradient determines the luminal section and the degree of flow limitation, and hence effective R (40). Therefore, a high value of R indicates increased airflow obstruction, but does not inform on whether flow limitation is present.

**Transcutaneous CO₂.** The partial pressure of transcutaneous CO₂ (P_{TCO₂}) starts to increase at sleep onset until it levels off with stable, slow-wave sleep. This high, steady level of P_{TCO₂} (plateau) represents the set-point, providing a target ventilation during non-REM sleep. Prolonged periods of progressive IFL during sleep are associated with an increase in P_{TCO₂}, which drives breathing only when P_{TCO₂} increases above the set-point, the optimal P_{TCO₂} level during sleep. Whether this increase in P_{TCO₂} and respiratory effort lead to adverse outcomes is still unknown, but this elevation has been suggested to be associated with sympathetic overactivity and higher variability in blood pressure (43).

**Recommendations.**

- Pes (pleural pressure) measurement may contribute to the identification of flow
limitation, whereas measurements of resistance reflect airway obstruction, but not necessarily IFL.

- Transcutaneous CO₂ levels may increase during IFL, although the clinical implications of this remain to be established.
- In that Pes and CO₂ monitoring are not routinely recorded during most partial and complete polysomnography (PSG) studies, workshop participants agreed that development of IFL scoring algorithms should proceed based on analysis of nasal pressure signals alone.

Relationship of IFL to Adverse Outcomes

Although flow limitation during sleep has been quantified using a variety of methods, it remains unknown whether there is a threshold degree of IFL that should be considered pathologic. The focus of a recent study was to determine the extent of IFL that could be considered normal (44). IFL was manually scored through visual analysis in a sample of 163 individuals who underwent overnight PSG without any sleep complaints (Epworth sleepiness score < 10, Chalder fatigue scale score ≤ 4, and no symptoms of OSA) and with an AHI less than 5/h. The mean % IFL was 8.8 (± 10.4%) of breaths, with an upper 95th percentile confidence limit of 31%. Thus, in a sample of normal individuals without sleep complaints, IFL may be present in up to 31% of total breaths during sleep. Workshop participants agreed that establishing the threshold of % IFL (based on standardized scoring) that is clearly associated with adverse clinical outcomes remains an important future research objective.

The available data exploring the relationship between IFL and clinical outcomes are largely reported in studies on UARS. This syndrome has been described based on the hypothesis that snoring and repetitive respiratory events characterized by IFL and RERAs may adversely impact on neurocognition, cardiovascular function (5, 45–48), and depressed mood (49). Currently, the International Classification of Sleep Disorders does not recognize UARS as a distinct diagnostic entity, but rather considers it part of the spectrum of OSA. UARS has typically been defined as the presence of excessive daytime sleepiness in patients with snoring and conventional AHI less than 5/h (5, 45). The American Academy of Sleep Medicine now provides for identification of RERAs based on flattening of the inspiratory nasal pressure signal with the presence of EEG to detect microarousals. Thus, most ambulatory studies will be inadequate for assessing UARS, and in-laboratory studies are required. Studies to date linking UARS to adverse clinical outcomes are largely observational (5, 44–48). The intensity of inspiratory efforts during RERAs has been linked to excessive sleepiness (48). RERAs have been associated with hemodynamic changes and sympathetic activation (5, 50). Observational studies have demonstrated that maternal sleep-disordered breathing during pregnancy (frequently characterized by IFL and mild OSA) is associated with an increased risk of gestational hypertension and gestational diabetes, after adjusting for potential confounders (51–54). In addition, in the pediatric population, milder degrees of prolonged periods of stable upper airway obstruction or airflow limitation have been observed in PSG studies (55). In some studies, even simple snoring without elevated AHI has been associated with neurobehavioral impairment in children (56, 57).

Optimal therapeutic strategies for UARS have not been defined, though oral appliances, as well as CPAP, may be effective. Randomized controlled trials and large prospective cohort studies to further establish the relationship between UARS and health outcomes are currently lacking. In assessing the clinical relevance of IFL, standardizing its measurement is first required.

Recommendations.

- Once standardized (and preferably automated) IFL scoring is available, further research is required to establish normative data for IFL in the general population.
- Further studies are also required to evaluate links between IFL and adverse outcomes in a diversity of patient populations, using both observational and randomized, controlled, interventional studies.

Results of the Crowdsourcing Online Challenges for the Visual Detection of Flow Limitation

Agreement among scoring practices for the manual identification of IFL was evaluated through scoring exercises or “challenges” that were crowdsourced to the workshop participants and broader sleep research community. Two challenges were crowdsourced, and each consisted of 30 breath strips, with each strip representing a 60-second epoch. Only the nasal cannula pressure transducer signal was shown. A broad range of breath types (flat to sinusoidal) was featured from a sampling of sleep studies with low AHI. The challenges were made available online through the NSRR website (sleepdata.org) and sent to all workshop participants, but were also opened to the public with a specific invitation to ATS–Sleep and Respiratory Neurobiology (SRN) members (challenge 1) and an announcement at the 14th International Sleep and Breathing Conference in Brazil (challenge 2) in October 2015 to encourage broader participation.

Details on challenge 1, including supplemental figures, are summarized in the online supplement. Overall, the agreement was poor among scorers for three categories of degree of IFL: normal; intermediate IFL; and IFL (κ = 0.3). Based on the scoring results from challenge 1, a consensus-based algorithm for scoring (Figure 3) was developed, and included elimination of the “intermediate IFL” category. Furthermore, the algorithm was expanded to include criteria by which to use information from adjacent breaths before assigning a breath score, and for better defining specific shape characteristics (scooping or flattening) when determining if a breath was flow limited.

Challenge 2 occurred several months after the ATS Workshop (October 2015) and was similar in format to challenge 1. The goal was to determine if the use of filter settings altered agreement among scorers and/or changed propensities to score flow limitation. The consensus-based algorithm (Figure 3) was used for scoring. Further details on challenge 2 are contained in the online supplement. Among workshop participants who completed the exercise, the agreement was greater among breath strips that were unfiltered versus the same breath strips that had high-cut filtering set at 30 Hz or low-cut filtering set at 0.05 Hz (κ = 0.6 vs. 0.4).

Based on the results of challenge 2, which demonstrated significant alterations in the shape of the pressure signal, in addition to changing patterns in scoring when specific filters were applied (low-cut [LC] > 0.01), it was decided that, when possible, low-cut filter settings should be avoided and the signal should be analyzed unfiltered. However, due to certain
It is preferable to use unfiltered signals (high and low cut), but if this is not possible due to hardware restrictions, a low-cut filter should be set at 0.01 Hz or lower to preserve the pressure transducer signal for IFL analysis.

**Recommendations.**

- It is preferable to use unfiltered signals (high and low cut), but if this is not possible due to hardware restrictions, a low-cut filter should be set at 0.01 Hz or lower to preserve the pressure transducer signal for IFL analysis.

**Summary of Key Findings and Future Directions**

The ATS workshop on the noninvasive detection of flow limitation in sleep studies was organized to determine optimal strategies to standardize, and eventually automate, the scoring of flow limitation in sleep studies. Ultimately, scoring large datasets for flow limitation, when done consistently, will be key to providing a basis for studies evaluating the clinical significance of flow limitation with respect to adverse outcomes.

The workshop participants unanimously recommended that methodology for scoring IFL be standardized and shared openly with the clinical and scientific communities. Moreover, participants identified shape characteristics of the pressure transducer curves that were considered by all to be more representative of IFL than others. Although the overall agreement between scorers was suboptimal, despite agreement on the definition of IFL, workshop participants believed it was important to proceed and score larger datasets using the consensus-based definition and algorithm to explore the overall probability that a particular breath within a dataset would be scored IFL among experts. This proposed method would establish a training set that would be useful for the development of future automated scoring methods.

Through e-mail correspondence and webinars after the May 2015 Workshop, workshop participants have agreed to extend this work, and immediate future steps include creating longer segments (e.g., 1 h) of breath sequences from several different sleep studies, and asking workshop participants to score these using the current manual algorithm. A probability score of IFL per breath will then be assigned to individual breaths, which could be used in the development and validation of automated scoring algorithms. In the interim, efforts are ongoing by some workshop participants with expertise and resources to develop software programming of such algorithms. These automated algorithms, which reflect the elements of the manual-based algorithms developed by the group, can be compared against one another to obtain a validation of the automated method of scoring. Future research projects to fine tune these areas will be required.

This official Workshop Report was prepared by an ad hoc committee on inspiratory flow limitation, which was part of the Assembly on Sleep and Respiratory Neurobiology.

Members of the committee are as follows:

- **R. John Kimoff, M.D. (Co-Chair)**
- **Sushmita Pami, M.D., M.Sc. (Co-Chair)**
- **Indu Ayappa, Ph.D.**
- **Ramon Farie, Ph.D.**
- **Jason Karkness, Ph.D.**
- **Luciana Palomini, M.D.**
- **Jean-Louis Pepin, M.D., Ph.D.**
References


