PHYSIOLOGICAL REVIEW

The upper airway in sleep: physiology of the pharynx

Indu Ayappa and David M. Rapoport

Division of Pulmonary and Critical Care Medicine, New York University School of Medicine

Summary
The upper airway is the primary conduit for passage of air into the lungs. Its physiology has been the subject of intensive study: both passive mechanical and active neural influences contribute to its patency and collapsibility. Different models can be used to explain behavior of the upper airway, including the “balance of forces” (airway suction pressure during inspiration versus upper airway dilator tone) and the Starling resistor mechanical model.

As sleep is the primary state change responsible for sleep disordered breathing (SDB) and the obstructive apnea/hypopnea syndrome (OSAHS), understanding its effects on the upper airway is critical. These include changes in upper airway muscle dilator activity and associated changes in mechanics and reflex activity of the muscles. Currently SDB is thought to result from a combination of anatomical upper airway predisposition and changes in neural activation mechanisms intrinsic to sleep.

Detection of SDB is based on identifying abnormal (high resistance) breaths and events, but the clinical tools used to detect these events and an understanding of their impact on symptoms is still evolving. Outcomes research to define which events are most important, and a better understanding of how events lead to physiologic consequences of the syndrome, including excessive daytime somnolence (EDS), will allow physiologic testing to objectively differentiate between “normal” subjects and those with disease. © 2002 Elsevier Science Ltd. All rights reserved.

INTRODUCTION

Over the past 20 years, there has been growing interest in the role of the human upper airway during breathing, especially during sleep. In large part, this interest has come from the increased recognition of the entity of Obstructive Sleep Disordered Breathing. While technically, sleep disordered breathing is a general term for all disorders of breathing during sleep, for simplicity, we will use the term sleep disordered breathing (SDB) in this review to mean only obstructive syndromes of abnormal upper airway resistance, i.e. a generalization of the Obstructive Sleep Apnea/Hypopnea Syndrome. This disorder is both frequent in the general population [1] and has important clinical implications, ranging from disruption of sleep with daytime sequellae of excessive sleepiness (2) to suspected long term cardiovascular consequences [3–5].

Because of this greatly increased interest in the diagnosis of sleep disordered breathing, it has become clear that a better understanding of normal physiology of the upper airway, as well as re-examination of the current definitions and measurement techniques for identifying abnormal respiratory “events” during sleep is necessary. The present review deals with these issues.

Correspondence should be addressed to: David M. Rapoport, MD, Department of Medicine, New York University Medical Center, 550 First Ave., New York, NY 10016, USA. Tel: +1-212-263–6407; Fax: +1-212-263–7445; E-mail: david.rapoport@med.nyu.edu

1087–0792/02/$ – see front matter © 2002 Elsevier Science Ltd. All rights reserved.
NORMAL AIRWAY

Upper airway anatomy

The upper airway is a complicated structure that is usually divided into 4 anatomical subsegments (see Fig. 1):

- Nasopharynx – between the nares and hard palate;
- Velopharynx or retropalatal oropharynx – between the hard palate and soft palate;
- Oropharynx – from the soft palate to the epiglottis;
- Hypopharynx – from the base of the tongue to the larynx.

This total structure forms a passage for movement of air from the nose to the lungs and also participates in other physiological functions such as phonation and deglutition [6]. The properties of the upper airway are a compromise between these different functions, which variably require maintenance of patency (during breathing) or closure of the airway (as in swallowing).

There are 20 or more upper airway muscles surrounding the airway that actively constrict and dilate the upper airway lumen [7, 8]. They can be classified into four groups – muscles regulating the position of the soft palate (alai nasi, tensor palatini, levator palatini), tongue (genioglossus, geniohyoid, hyoglossus, styloglossus), hyoid apparatus (hyoglossus, genioglossus, digastric, geniohyoid, sternohyoid) and the posterolateral pharyngeal walls (palatoglossus, pharyngeal constrictors). These groups of muscles interact in a complex fashion to determine the patency of the airway. Soft tissue structures form the walls of the upper airway, and include the tonsils, soft palate, uvula, tongue and lateral pharyngeal walls [9]. The main craniofacial bony structures that determine the airway size are the mandible [10] and the hyoid bone [11]; these presumably act by providing the anchoring structures to which muscles and soft tissue attach. However, it is clear that complex relationships occur even in these “fixed” structures, as some of these like the hyoid bone “float” without any attachment to other bony or cartilaginous structure. They act as complex levers where muscle action, instead of moving the structure, may tense some of the adjacent soft tissues (e.g. tracheal pull).

In normal non-obese subjects, the mean minimum cross-sectional area across multiple segments of the upper airway has been measured using several techniques: estimates vary from 320–450 mm$^2$ (acoustic reflection) [12–14], 59 mm$^2$ (fast CT at FRC) [15], 64 mm$^2$ (MRI) [9], 144 mm$^2$ [16], 188 mm$^2$ [17] and 138 mm$^2$ [18] (conventional CT). This wide range in size reflects the differences due to individual variability but is also due to differing locations of measurement, positional change (sitting/supine), and differences imposed by the choice of imaging modality (e.g. mouth open is required for acoustic reflection). There is substantial overlap in measurements between normal subjects and those with OSAS. However, it is important to remember that most of the reported measurements were made during wakefulness – thus they combine truly anatomical properties (such as bone structure, fat deposition) and activation of the upper airway dilator muscles.

The minimum caliber of the upper airway in the wake state is primarily in the retropalatal oropharynx [9, 19], which makes it a site of interest as the potential location of collapse during sleep. The anterior wall of the oropharynx is composed primarily of the soft palate, tongue and lingual tonsils; and the posterior wall is bounded by a muscular wall made up of the superior, middle and inferior constrictor muscles that lie in front of the cervical spine. The lateral pharyngeal walls are a complex structure made up of muscles (hyoglossus, styloglossus, stylohyoid, stylop-haryngeus, palatoglossus, palatopharyngeus, the pharyngeal constrictors), lymphoid tissue and pharyngeal mucosa. This complexity of the interactions
between these different muscles makes the oropharynx an extremely difficult structure to evaluate.

**Gravity/position**

Since the upper airway lacks a fixed rigid structural support, shape and size of the airway are dependent on the position of soft tissue structures like the soft palate, tongue and the walls of the oropharynx. These can be influenced by gravity. In the supine position the tongue and soft palate have been shown to move posteriorly, reducing the oropharyngeal area \([20, 21]\), thereby increasing the supraglottic airway impedance \([22]\) and collapsibility \([23]\). Clinical correlation for this anatomic and functional change is demonstrated by the observation that snoring is much more prominent in the supine position.

**Landmarks: bony, fat, muscle, airway**

Upper airway imaging techniques have also been used to visualize the airway lumen and to define the surrounding structures. In awake subjects Schwab et al. \([19]\) have shown that the normal upper airway has a longer lateral than AP dimension using MRI techniques. In addition, by using fast cine CT, they also showed that airway size stays fairly constant during inspiration and reaches a minimum during end expiration, suggesting that muscular stabilization of the airway lumen during inspiration against the negative intraluminal pressure is more important than actual dilatation, as had previously been believed. According to these authors, most respiratory related changes (i.e. end-expiratory loss of diameter) are predominantly in the lateral dimension \([19]\).

**Models of behavior of the airway**

In describing the dynamic behavior of the airway during cyclic respiration, it is useful to employ various physical and/or mathematical models to reduce this complex structure to a more simple and understandable one. The most basic approach is to treat the airway as a rigid tube and analyze its resistance (i.e. assume a fixed or “average” relationship between driving pressure and flow). An extension of this model is to consider “resistance” as varying during the inspiratory cycle composed of a dynamic interaction between flow and pressure. The need for this more complex model becomes evident when one observes that negative intrathoracic pressure transmitted to the passive upper airway during inspiration promotes a reduction in pharyngeal cross-sectional area \([24]\). According to the “balance of pressures” concept \([25, 26]\) the size (and thus resistance) of the airway depends on the balance between collapsing intraluminal pressures generated during inspiration by subatmospheric pressures in the thorax and outward contracting forces of the upper airway dilator muscles. In this analysis, patency of the airway depends on transmural pressure \((\text{Ptm})\), which is the difference between the negative intraluminal pressure caused by inspiratory efforts and the positive dilating pressure from the upper airway musculature.

To help in understanding this “dynamic collapse” an alternative complementary approach has been to describe the upper airway as a collapsible tube. This model deals better with the dynamic collapse described above, and also helps to explain why airway caliber may increase in association with lung inflation \([13, 15]\).

In this collapsible tube model, one can examine multiple components that influence collapse of the susceptible airway. These factors include at least the following:

- Respiratory driving pressure across the region susceptible to collapse – determined by the negative intra-thoracic inspiratory pressure and any fixed resistances of anatomic structures such as the nose;
- Intrinsic properties of the airway wall – this is called the “tube law” and is determined by the size, collapsibility and longitudinal tension on the tube;
- Neural input to the upper airway – which dictates the behavior of the dilating/stabilizing musculature.

A more detailed discussion of these approaches follows.

**Passive static properties of the airway**

Static properties of a tube can be inferred from its behavior during steady state flow. Thus resistance has been used to describe the upper airway, but this is primarily because it is the simplest measure to describe a pressure/flow relationship through a conduit. However because of collapsibility and the influence of dilator muscles, the upper airway does not have a fixed cross-sectional area, which is a minimum requirement for having the linear pressure/flow relationship necessary to define “a resistance”. To overcome this limitation of the “resistance” concept, pharyngeal behavior has been described using either
the resistance calculated at a single fixed flow rate or resistance at the peak flow rate during a maneuver. This approach has been used to describe the relative properties of different segments of the upper airway under specific conditions.

Static properties of the upper airway have also been evaluated by direct measurement of the compliance of the pharyngeal wall (slope of the volume to pressure or cross-sectional area to pressure relationship). However, to evaluate true compliance of this structure one must suppress the neuromuscular influences on the upper airway, as these can fluctuate during the measurement process. In order to remove these neuromuscular influences studies have been performed in anesthetized and paralyzed subjects or during sleep with the application of CPAP. The genioglossus and other pharyngeal muscles are presumably hypotonic on the first breath following the removal of CPAP [27]. Using these techniques Isono et al. evaluated the properties of the pharyngeal wall plus the surrounding structures in normal subjects and patients with obstructive sleep apnea [24, 28, 29]. The relationship between pressure and cross-sectional area of the velopharynx and oropharynx of the passive airway was described as exponential; thus, the airway is more collapsible (i.e. the derivative of area to pressure becomes greater) as the pharyngeal airway becomes smaller (see Fig. 2). There was a fundamental difference between the airway of the normal subjects and those with OSAS in that the airway was closed at atmospheric pressure in subjects with OSAS, while active negative pressure was required to close the normal airway, even in the absence of upper airway muscle activity.

In the presence of upper airway muscle activity, the relationship between area and pressure should shift to the left indicating a decrease in compliance (see Fig. 3). However, in the only data which addresses this directly (in a dog model), Fouke et al. (1986), reported an unchanged pressure/volume relationship during active and passive force generation, which is not compatible with the role of the active muscular component in affecting airway stiffness [7].

**Passive dynamic (independent of tone)**
Whereas the upper airway in normal adult man generally remains patent in the absence of all muscle tone when intraluminal pressure is zero [24, 30], application of intraluminal negative pressure (as during inspiration) or flexion [31] may result in complete collapse. The pressure at which this occurs has been referred to as Pclose [28]. This same collapse can also occur due to increased extra-luminal pressure: an example of this is loss of airway patency in the supine position from passive collapse due to gravitational forces generated by craniofacial structures or adipose tissue surrounding the upper airway. Mass loading of the anterior neck may also increase the collapse of

![Figure 2](image-url)

Figure 2  Static pressure–area curves of the passive velopharynx. The curves can be fitted to an exponential function and curves of patients with OSAS lie below and to the right of the normal subjects indicating greater collapsibility. (From Isono S, Remmers JE, Tanaka A, Sho Y, Sato J, Nishino T. Anatomy of pharynx in patients with obstructive sleep apnea and in normal subjects. *J Appl Physiol* 1997; **82**(4): 1319–1326; with permission.)
the passive airway. This has been shown in anesthetized rabbits with mass loading to simulate excessive adipose tissue resulted in narrowing of the pharynx and increase in upper airway resistance [32]. Conversely, negative pressure applied around the neck of anesthetized dogs decreased upper airway resistance, indicating passive dilation [33].

Active (i.e. muscle dilator) contribution

The upper airway is rich in neural receptors, which play a part in controlling baseline tonic genioglossal EMG. Any loss of this EMG tone, as occurs at sleep onset, probably contributes to raising pharyngeal resistance [34]. The existence of topical receptor mechanisms in the nasopharynx that may influence dilator muscle activity has been investigated following the application of upper airway anesthesia [35]. In addition, a decrease in phasic and tonic genioglossal (GG) activity has been measured during stomal breathing when compared to nasal breathing in tracheostomized subjects, suggesting the influence of local upper airway stimuli [36]. During inspiration there is also phasic inspiratory activity of many dilatory upper airway muscles, including the genioglossus [37, 38] and geniohyoid that has been demonstrated in both human and animal models [39, 40]. This phasic activation of the muscles of the nose, pharynx and larynx has been shown to occur before the diaphragm and intercostals muscle activity suggesting pre-activation of the upper airway muscles in preparation for the development of negative pressure [41, 42]. In experimental situations, it does appear that the upper airway dilator muscles may enlarge the airway by shortening of the muscle fibers (shortening is often assumed to have occurred when increased EMG activity is seen). The presence of a dilating force concurrent with upper airway activation during early inspiration has been shown in the isolated sealed upper airway in a dog model [8]. A more recent study [43] examined the relationship between EMG of the genioglossus and pharyngeal dimensions in laryngectomized patients breathing through a tracheostoma. In this study, inspiratory-related muscle activation was associated with enlargement of the glossopharyngeal airway, all in the absence of flow through the upper airway because of the tracheostomy. Despite the above data, it remains unclear whether the activity of “dilator” muscles is in fact to dilate the airway, or whether these increases in muscle tone act to stabilize the airway and maintain patency against the collapsing forces present during inspiratory airflow.

Whereas phasic firing of the upper airway neural pathways occurs with inspiration, it may also be increased by reflex activation, independently of chemoreceptor stimulation. In several studies, the experimental application of large negative pressure to the airway produced such a reflex and also suggested changes in the timing between upper airway and diaphragm muscle activation [42, 44–46], and in one study [47] inspiratory activity of the genioglossus closely tracked epiglottic negative pressure under a variety of more physiologic conditions. In the case of the response to high negative pressures one can question whether response may be more relevant to the reflexes induced by abnormal upper airway obstruction than normal upper airway function.

Starling resistor model

As pointed out above, the upper airway has been shown to collapse variably under conditions of negative intraluminal pressure, as occurs during inspiration. This type of behavior has been empirically described by several approaches described above. However, a pattern of dependence of flow on the driving pressure occurs in “Starling resistors”, which are a specific model of “collapsible tube” behavior. This behavior is characterized by a pattern of flow which initially increases as driving pressure increases; however,
above a critical driving pressure, there is a progressive plateauing of flow at some maximal level despite continued increase in driving pressure (flow limitation) [48, 49] (Fig. 4).

One can study this behavior mathematically, or demonstrate it with a physical model consisting of a thin walled elastic (collapsible) tube enclosed in a chamber. The pressure in the chamber (Pcrit) can be varied to be less than or greater than the pressure inside the tube. The analysis of such a system predicts that maximal flow through the tube segment is determined by two separate factors: the resistance of the upstream segment (e.g. at nose) and the transmural pressure surrounding the collapsible segment. This corresponds to a “tissue pressure” in the collapsible parts of the upper airway. When Pcrit is greater than 0, there is airway collapse at rest and obstruction of the tube occurs. When Pcrit is strongly negative, the airway will remain patent even under large inspiratory efforts. There is strong evidence that the hypotonic pharyngeal airway behaves like a Starling resistor. In normal subjects a negative Pcrit has been measured, whereas in snorers Pcrit approaches 0 [48, 50] and in obstructive apnea patients, Pcrit may be positive [51, 52]. However, it must be emphasized that this description of the airway refers to the passive airway, i.e. without added muscle activity that cannot easily be measured during normal breathing.

To further model the airway in the intact human beyond its passive properties, one must add the effects of upstream resistance (anatomy) and changes in collapsibility mediated by changes in baseline and phasic muscle tone.

At submaximal flows, increasing upper airway muscle tone has been shown to result in decreasing resistance [53]. However, at maximal flow, increasing upper airway tone (activation) changes the value of Pcrit for a Starling resistor. This causes an increase in the maximum flow through the Starling resistor by

---

**Figure 4** Inspiratory pressure/flow contour of a collapsible tube – Starling resistor behavior. When the tube is patent a linear relationship is seen between flow and pressure. With an increase in resistance (partial obstruction), flow increases only up to a point and then remains constant with any further increase in driving pressure. (From Condos R, Norman RG, Krishnasamy I, Peduzzi N, Goldring RM, Rapoport DM. Flow limitation as a noninvasive assessment of residual upper-airway resistance during continuous positive airway pressure therapy of obstructive sleep apnea. *Am J Respir Crit Care Med* 1994; 150(2): 475–480; with permission.)
decreasing airway collapsibility of the flow-limiting segment, [54] and can occur even without a change in the resistance (i.e. both flow and pressure can increase). Thus the Starling model helps better characterize the changes in flow that occur with changes in muscle tone.

Factors that influence resistance/collapsibility of the upper airway in humans

**Upstream resistance within the nasal airway**

As pointed out above, the upper airway can be modeled by the more complex Starling resistor. This leads to the understanding that not only collapsibility of the critical segment of the airway needs to be examined, but that upstream (e.g. nasal and nasopharyngeal) resistance will affect the behavior of the entire system. Even under normal circumstances the nose has a relatively high resistance, which is increased by airway narrowing that occurs with mucosal congestion.

**Gender effects**

The data on the effect of gender on pharyngeal cross-sectional area and resistance is conflicting. While some studies have indicated that differences in pharyngeal size can be explained by differences in body surface area [13] other studies have reported greater pharyngeal size in men after normalizing for body weight [12]. Although larger airways should imply lower resistance in males, pharyngeal resistance has been measured to be twice as high in males than females [55], suggesting that there may be gender specific differences in the airway mechanics. Some of these discrepancies may be due to the intrinsic limitations of using the “resistance” concept – see discussion above, and may be better addressed by concepts of collapsibility, i.e. Pcrit. Additionally, it has been shown that changes in pharyngeal cross-sectional area are more dependent on lung volume [13], and increases in pharyngeal resistance in response to load are greater in men than women [12, 56]. This suggests that the differences in mechanics (greater stability and less dependence on lung volume) in women may play a more important role than size.

**Hormonal status**

In addition to gender there is some evidence that hormonal status might influence the upper airway dilator activity [57] with premenopausal women having significantly greater genioglossal muscle activity compared with post memenopausal women, and age-matched men [58, 59], during inspiratory resistive loading.

**Age effects**

Modifications in pharyngeal characteristics with age have been studied but there is no consensus on the effect. Increases [60] and decreases [12, 23] in pharyngeal cross sectional area with age have been reported in subjects when studied awake. Age-related increases in pharyngeal resistance during sleep [61] and wakefulness [34] have been reported along with decreases in the activity of genioglossus and tensor palatini muscles [62]. However, other studies during wakefulness [60] and sleep have reported that upper airway resistance and genioglossus activity during quiet breathing are similar in elderly and young people.

**Extrinsic anatomic and static factors affecting the upper airway**

**Neck and Jaw posture**

Closure of the upper airway has been observed with neck flexion in both the supine and prone positions in anesthetized spontaneously breathing patients [63]. In addition, jaw posture has also been documented to influence the size of the upper airway.

**Surface adhesive forces**

The adhesive forces between the airway luminal surfaces may promote closure of the airway or impede subsequent opening of the airway [30]. It has been shown that the positive pressure needed to open an already closed airway is greater than the closing pressure in both adults and infants [30, 50].

**Obesity**

Obesity can influence the airway properties by increasing the mass loading on the upper airway that could result in airway compression. Since it is one of the main risk factors for obstructive sleep apnea, and weight loss results in improvement of the apnea in some patients [64], it is a factor to consider. However, the exact mechanism whereby obesity predisposes to sleep apnea is still unclear.

**Tracheal tug**

The cross-sectional area of the upper airway appears to increase with lung volume, along with a reduction in the closing pressures and compliance of the upper airway [65, 66]. No data exist to suggest this is an
active neural reflex, and in fact lung inflation is generally inhibitory to respiration (Hering Breur reflex) [67]. The increase in cross-sectional area is probably a passive mechanical effect that results from the axial forces transmitted through the trachea. As the lung volume increases, caudal displacement of the intrathoracic trachea appears to exert longitudinal forces on the pharynx which stabilize it and prevent passive collapse. Support for this interpretation of the mechanical effect of lung inflation comes from direct measurements of cross-sectional area during the entire respiratory cycle measured using cine CT. These data showed that the cross-sectional area stays large at end-inspiration (when neural tone is falling and lung volume is elevated), and only falls to a minimum at end-expiration (when lung volume is small) [19].

**Effect of sleep on upper airway resistance/collapsibility**

Radiographic measurements have shown that during wakefulness patency of the upper airway is well maintained in different postures, although the reflexes which control this seem to be critical [68]. However, with the onset of sleep there are several modifications that may occur in the factors affecting patency of the upper airway including changes in neuromuscular activation, ventilation, chemical and mechanical load responses.

Significant increases in upper airway resistance associated with sleep have been shown in animals [69] and humans [70–72]. Supraglottic resistance has been shown to increase from low values (1–2 cm H₂O/L/s) [22, 72] to values as high as 5–10 cm H₂O/L/s and to 50 cm H₂O/L/s in heavy snorers [73]. Less is known about changes in airway caliber, but most studies suggest that it decreases during sleep, with the lateral pharyngeal walls playing an important role in this narrowing [74] (Fig. 5). These changes in mechanics induced by sleep could result in either hypoventilation (loss of the reflex response to increased airway load), or a reflex induced increase in ventilatory output with maintained ventilation and blood gases. Some degree of hypoventilation occurs at sleep onset in normals. That this is a consequence of the mechanics rather than a change in set-point for CO₂ has been shown in a study where unloading by CPAP or breathing He/O₂ mixtures (characterized by reduced density) returned mildly elevated sleeping PCO₂ to awake levels [75].

![Figure 5](https://via.placeholder.com/150)

**Figure 5** Axial magnetic resonance images of a normal subject during wakefulness and sleep. The airway is reduced in both the AP and lateral dimensions during sleep compared with wakefulness. There is also thickening of the lateral pharyngeal walls during sleep. (From Trudo FJ, Gefter WB, Welch KC, Gupta KB, Maislin G, Schwab RJ. State-related changes in upper airway caliber and surrounding soft-tissue structures in normal subjects. *Am J Respir Crit Care Med* 1998; 158(4): 1259–1270; with permission.)

Although it is likely that sleep in normals induces increased collapsibility in the airway (decreased tone to the upper airway dilators), the effect on calculations of resistance is confounded by reflex responses (or their absence). Thus it has been shown that sleep has an effect on multiple aspects of upper airway behavior.

**Effect of sleep on muscle tone in the upper airway**

The change in muscle activity of the upper airway with onset of sleep has been investigated directly by measuring the EMG, by using measured changes in pharyngeal wall compliance or by using derived values of ventilation and airway resistance. Many studies have shown that the phasic inspiratory activity of the genioglossus [26, 37, 76, 77] and geniohyoid are maintained during sleep in normals. Conversely, decreases in both tonic and phasic tone have been shown in the genioglossus, geniohyoid [78], tensor palatini [72], levator palatini, palatoglossus [77, 79] and other respiratory muscles at the onset of sleep. These have been shown to be associated with transient decreases in ventilation and increased upper airway
resistance [62, 72]. However, in normal subjects these decreases are short lived and parallel those seen in the diaphragm [76] and intercostal muscles, which rise as a response to induced obstructive hypoventilation and mild hypercapnea (1–2 mmHg) seen during sleep. The role that both tonic and phasic tone play in maintaining airway patency is shown by denervation studies, which in an animal model resulted in collapse of the airway during sleep [26]. In healthy human subjects, dense upper airway anesthesia increases upper airway resistance during sleep and can cause prolongation of apnea [80]. Despite this, the marked decrease in EMG seen in all airway muscles during REM sleep [81] does not result in uniform obstruction. Some of this last paradox may be due to the protective effect of the reduced inspiratory airway pressures generated during REM sleep, again illustrating the difficulty in using flow and resistance of the collapsible airway as indices of function during conditions of changing effort.

Effect of sleep on load response
Whereas during wakefulness application of a resistive load to the airway results in increased respiratory drive, this response may be lost or greatly attenuated during sleep [71]. There is debate over whether a similar response to airway loading exists by which the upper airway muscle tone directly increases in response to resistive loading of the system. However, it is difficult to separate this possible direct effect from a non-specific increase in ventilatory drive. Despite numerous studies that have measured upper airway resistance and genioglossal EMG during wakefulness, sleep and with resistive loading, the direct relationship between GG EMG activity and upper airway resistance during sleep is unclear, with conflicting results in multiple studies [71, 78, 82]. Possible reasons for conflicting observations include the assumption that muscle activity (EMG) is a surrogate of muscle fiber shortening, and the fact that in most experiments muscle activity can only be measured in a few locations/muscles, which do not fully reflect the total airway muscle response.

Effect of CO₂ on muscle activity during sleep
In the awake state, elevation of CO₂ is a powerful respiratory stimulant. There is a large literature on the effect of sleep on this response, and the consensus is that this may be only minimally affected by sleep, at least in non-REM stages. Less is known about the effect of sleep on CO₂ responses of the upper airway muscles [75], independent of general respiratory stimulation. The results from studies examining changes in GG activation with hypercapnia are variable [77, 83], and hypercapnia has been shown to decrease collapsibility of the airway similar to tracheal displacement [84].

Special human considerations
Gender
The data on gender differences in airway resistance measured during sleep are conflicting. While Trinder et al. [85] measured a greater increase in upper airway resistance during sleep in healthy men compared with women Thurnheer et al. [86] concluded that there were no major age or gender differences in the changes in airway resistance during sleep.

SNORING
One of the most common consequences of the changes in the properties of the upper airway that occur during sleep is the occurrence of snoring. Whether to call this pathology, if it occurs in the absence of frank airway obstruction (apnea, hypopnea) or repetitive arousal, is still debated. The prevalence of habitual snoring is extremely high and reported to be roughly 40% in men and 20% in women [1, 87, 88]. The snoring sound is the result of vibrations of the soft tissues of the pharynx, soft palate and uvula having specific acoustic characteristics with frequencies ranging from 5 to 136 Hz [89]. Early cineradiographic recordings made by Lugaresi et al. [90] showed that snoring was associated with increased esophageal pressure swings during inspiration with partial pharyngeal obstruction. In studies comparing snorers and non-snorers during sleep it has been shown that snorers have more negative inspiratory pressure [91], greater pulmonary resistance during snoring breaths [73] and flow limitation [51, 73, 92]. These events are all thought to be normal consequences of sleep, and not necessarily pathologic. However, many suggest snoring can be considered as part of the spectrum of sleep disordered breathing because the magnitude of these changes lies between that of the non-snoring (normal) subject and obstructive sleep apneic.

Anatomical measurements of the upper airway using multiple techniques like MRI [93], acoustic reflections [14], CT scanning [94] and x-ray cephalometry [95]
have demonstrated differences between snorers and non-snoring asymptomatic subjects. As discussed previously, since the airway area depends on body posture, gender, lung volume, state of consciousness and body size it is difficult to attribute the differences in airway dimensions to snoring alone. Nevertheless, it has been reported that snorers have reduced airway area at the level of velopharynx [96], tongue base, hyoid bone, have reduced sagittal dimensions, longer soft palates and longer and wider uvulas than healthy volunteers [97]. There is also however the alternate hypothesis that the reduction in airway cross section is a result of snoring, due to inflammation of the mucosa and edema [98].

The changes in upper airway physiology that occur during sleep have all been implicated in the development of snoring. In addition to anatomy, the functional properties of the upper airway appear to be different in snorers and normal subjects. During wakefulness snorers (non-apneic) have more collapsible upper airways than non-snoring normals [50]. Also, the critical pressures needed to collapse the upper airway lies between that of normal subjects and patients with OSAS [51]. The reason for the increased collapsibility in snorers is still unclear. Possible mechanisms include neuromuscular deficiencies of the upper airway dilator muscles, abnormal tissue properties and abnormal linkage between the pharyngeal dilator muscles and pharyngeal tissue due to fatty infiltration [99].

The results from studies comparing the upper airway properties of apneic snorers and non-apneic snorers have been conflicting because their classification is dependent on the technique used to measure the respiratory events. A subject classified as a non-apneic snorer by the respiratory index based on nasal airflow measured by thermistry could be classified as an apneic snorer by a more sensitive measure of flow/effort like nasal cannula or esophageal manometry. Therefore, while some studies showed no difference in collapsibility [100], snoring frequency, intensity and nasal resistance [101] or pharyngeal cross-sectional areas at different lung volumes [14], others have reported higher pharyngeal distensibility [102, 103] and sound intensity [100, 104] in apneic snorers versus non-apneic snorers. The greater sound intensity appears to be the result of greater negative pressures on resumption of breathing in apneic snorers, resulting in high flow rates, turbulent flow and greater forces on the vibrating structures. More recently, Series et al. have reported differences in the properties of the musculus uvulae (an upper airway dilator located in the uvula) [105] and the genioglossus [106] in snorers and subjects with OSAHS. The force generated by these muscles is greater in apneics than non-apneic snorers during wakefulness and sleep, possibly as an adaptive response to high-intensity resistive loading. However, the sleep-induced decrease of this upper airway dilator activity results in greater alterations in mechanical efficiency of apneic subjects [107].

Several mathematical models of the upper airway as a collapsible tube have been developed to study snoring. A tube of given geometry, elastic constant, resistance, density of gas and flow parameters will become unstable at a certain location leading to vibration of the walls and dynamic collapse with the appearance of sound structures similar to snoring [108]. Since these parameters are different in individual subjects the site of snoring sound generation will vary depending on where the instability occurs. Direct observations during sleep have confirmed that the location of the vibrations along with the narrowing of the upper airway is variable [92, 109]. Other models coupling mechanics to neuromuscular physiology through reflex wall stiffening predict that snoring develops with reduced strength and increased latency of the reflex [110].

Despite the reported greater prevalence of snoring in men the gender effects on snoring intensity are unclear. Wilson et al. [104] have reported that women do not snore as loudly as men, independent of severity of SDB or BMI, in contrast to Metes et al. [101] who showed no difference in the maximum snoring intensity between genders. Both studies however did not show any age effect on the intensity of snoring [101, 104]. A recent epidemiological study has reported chronic nocturnal nasal congestion as a risk factor for habitual snoring [111].

SLEEP DISORDERED BREATHING

Since the 1960s, there has been an increasing recognition of the syndrome of Sleep Disordered Breathing. This term includes all forms of abnormal breathing patterns associated with sleep. By far the most common are the variants in which there is transient complete (apnea) or partial (hypopnea, snoring) obstruction of the upper airway. Early hypotheses to explain the hypersomnolence seen in the most severe form of sleep disordered breathing, then
called “the Pickwickian Syndrome”, suggested this symptom was caused by CO₂ narcosis. However, later studies reported that obstructive apneas were caused by the “obstruction of the upper outlet, by the backward movement of the tongue”, which then resulted in frequent awakenings and disruption of sleep, explaining the hypersomnolence [112]. In 1972, Walsh et al. [113] directly observed the upper airway during sleep using cinefluoroscopy in obese, hypersomnolent patients with OSA, and observed that upper airway obstruction was produced by the “tongue retracting into apposition with the posterior pharyngeal wall”. Establishment of an airway with a tracheostomy or nasopharyngeal tube abolished the apneas and usually cures the symptoms of the syndrome. A landmark study by Remmers et al. in 1978 [25] established the location of upper airway collapse during sleep as the oropharynx in patients with obstructive sleep apnea syndrome, and proposed that this resulted from an imbalance between negative pharyngeal pressure and the opposing force of the upper airway musculature (genioglossal). These events (apnea) were shown to resolve with a burst of GG EMG activation. The location of this collapse has since been shown to be at variable locations in the upper airway and can occur simultaneously at multiple sites [27, 114–117]. Thus, presently, abnormal function of the upper airway is thought to be the main source of obstruction and symptoms in the syndrome of sleep disordered breathing.

Anatomy

Much work has been done to address the hypothesis that abnormal anatomy underlies sleep disordered breathing. Several imaging techniques have been used to study the upper airway size and changes in the airway size and soft tissue structures that surround the airway in patients with SDB, including acoustic reflection [10, 14, 65, 102], fluoroscopy [16], nasopharyngoscopy [118], cephalometry [119], CT and MRI imaging [16, 19, 120–122]. Reduced awake supine upper airway caliber by CT [123] has been shown in male patients with obstructive sleep apnea syndrome at the naso, oro and hypopharyngeal levels when compared to normal subjects. These images showed that the narrowest region in both patients with OSA and normal non-obese controls, while awake, is the region posterior to the soft palate, and that the cross sectional area of this retropalatal region is smaller in patients than controls. No abnormal collections of fat density were found to explain these differences. However, a recent series of observations have pointed out that, independent of airway caliber, airway configuration may differ in this area in OSA patients. Thus, in contrast to normals who have the major axis of the pharyngeal airway oriented in the lateral dimension, patients show an axis oriented anteroposteriorly, which corresponds to lateral narrowing at the critical level [93] (see Fig. 6). Two structures lateral to the upper airway that could cause this narrowing, the lateral pharyngeal walls and the lateral pharyngeal fat pads, have been examined in these studies. The lateral pharyngeal walls have been shown to be thicker in apneics [124]; however, the pharyngeal fat pads were not closer together, and the area and width of the fat pads were not larger in apneics at the level of the minimum airway [19]. These unexpected findings have led the authors to conclude that the thickness of the lateral pharyngeal walls, and not the actual size of the soft palate, tongue or fat pads, was the predominant anatomic factor causing airway narrowing in apneic patients. The reason for the thicker walls is not known; in particular, spectroscopic analysis of tissue in this area shows no differences in fat or water content between normal subjects and apneics [19]. It has been speculated that an increase in muscle mass due to weight gain or the “exercise” of overcoming apnea might explain the increase in the size of the lateral soft-tissues without increase in the direct fat deposition. However, application of CPAP results in increase of the airway volume and area in the oropharynx [125], with most of the changes occurring again in the lateral rather than anteroposterior axis. This has led some to invoke “folding” of the mucosa as the explanation of the lateral wall thickening, with “unfolding” during CPAP. In any case, these consistent observations emphasize the importance of the lateral structures in contrast to other (tongue, soft palate) airway structures in the pathogenesis of obstruction.

Explanations accounting for role of sleep in causing upper airway obstruction/increased resistance

While the actual location of collapse in sleep disordered breathing events is often the oro or hypopharynx where an underlying anatomic change may be present, it is clear that patients with severe obstruction during sleep do not have significant functional...
airway impairment while awake. Thus to date, few if any physiologic tests of function during wakefulness reliably detect SDB [126]. It is clear that the onset of some aspect of sleep must play an essential part in the physiology. As described earlier, upper airway muscle activity is reduced during sleep in normals and sleep apneics. The past twenty years of work have repeatedly focused on whether anatomy alone explains sleep apnea, whether abnormal neural control of the airway is the cause of collapse, or whether abnormal airway structure is an “anatomic substrate” on which normal physiological mechanisms produce the loss of patency at sleep onset. The current thinking is that this last explanation is most likely.

Remmers et al. were the first to suggest that an overall balance of airway pressure and anatomic effects, which included the dilating force generated by upper airway muscles, should be considered as determining airway patency [24]. Thus the forces promoting closure, which include collapsing negative intraluminal pressure, were opposed by anatomic and neural mechanisms and the whole process could be considered as a “seesaw” (see Fig. 7).

In awake subjects balance is maintained in favor of patency. In OSA patients, anatomic airway narrowing may occur, but the opposing dilating forces are sufficient to maintain patency. During sleep in normals reduced airway dilating forces, due to reduced upper airway activity, do not alter the balance sufficiently to cause collapse. In OSA, however, collapse occurs because the forces (now affected by sleep induced changes) are imbalanced. In favor of this “balance” theory is the observation that even normal subjects with no obstructive sleep disordered breathing can develop repetitive obstructive events when they are subjected to negative airway pressure [127, 128], indicating that no intrinsic neural abnormality is needed in the face of sufficient collapsing pressure.

Another important advance in our understanding of the etiology of the upper airway collapse in patients with obstructive sleep apnea came with the work of White and colleagues, who first suggested that the critical event was loss of awake compensation for increased anatomic resistance by the upper airway dilators. According to this line of thinking, OSA patients, in order to compensate for their abnormal anatomy and/or more collapsible pharyngeal airway, have augmented genioglossus muscle activity during wakefulness as compared with healthy subjects [129].

Figure 6  Schematic of airway geometry during respiration. The patients with OSA have a greater anterior posterior configuration compared with normal subjects who have a greater lateral dimension. (From Schwab RJ, Gefter WB, Hoffman EA, Gupta KB, Pack AI. Dynamic upper airway imaging during awake respiration in normal subjects and patients with sleep disordered breathing. Am Rev Respir Dis 1993; 148(5): 1385–1400; with permission.)
This reflex compensatory neuromuscular mechanism is lost at sleep onset in both normals and patients, but is associated with pharyngeal collapse under conditions of chronic airway loading [130].

Depending on the model chosen to understand the behavior of the upper airway, different observations can be generated to elucidate components of the underlying abnormality leading to the tendency to airway collapse in SDB. Thus, complete airway occlusion during induced central apnea has been shown in some patients with obstructive sleep apnea syndrome, and interpreted to suggest a component of passive collapse due to extraluminal structures [131]. A similar observation during anesthesia showed that normal subjects maintain a patent airway, and require negative intraluminal pressures for closure, whereas subjects with OSA showed closure even at positive airway pressures below some critical positive value [24]. These observations again have suggested that changes in the surrounding extraluminal structures may underlie the propensity to obstruction during sleep.

There are several influences of sleep on factors other than the airway structure and the effect of neuromuscular control on the dilating upper airway musculature that may be important to the development of obstruction in SDB. Thus, ventilatory drive (i.e. neural output to the diaphragm and other pump muscles) itself may determine the amount of collapsing inspiratory pressure seen by the upper airway for any given degree of upstream resistance. It is well

**Figure 7** Schematic model of the balance of pressures concept. The pharyngeal patency depends on the balance of upper airway muscle activities (UA) and airway pressure on two sides of the fulcrum that represents the intrinsic properties of the pharynx. In panels B and D (subject with OSAS) the fulcrum is to the right of the normal subject (A and C) (From Isono S, Remmers JE, Tanaka A, Sho Y, Sato J, Nishino T. Anatomy of pharynx in patients with obstructive sleep apnea and in normal subjects. *J Appl Physiol* 1997; 82(4): 1319–1326; with permission.)
established that patterns of breathing in both delta sleep (slow deep breaths with low inspiratory flow rates are common) and REM (rapid, but small tidal volumes, again with potentially low flow rates) differ from the typical respiratory patterns seen during wake and stage 1 and 2 non-REM sleep (where obstruction is most commonly seen). These changes may be involved in the observation that apnea is least common in delta sleep (although low grade stable elevation of airway resistance is common), and that occasional patients show less apnea in REM than non-REM sleep. This last observation should not be confused with the more common observation that if obstruction does occur in REM, it is likely to be more severe (due to loss of muscle tone) and last longer (due to impaired arousal mechanisms) [132]. The contrast between the effect of REM on the upper airway directly, and the ventilatory drive to the pump, highlights the benefits of looking at either the balance of forces concept or the Starling resistor models of the upper airway rather than focusing only on the simple mechanics of the tube.

Finally, there are some data which suggest that the upper airway is not the only area of relevance in obstructive apnea. These authors found increased pulmonary resistance with reduced flow in the absence of changes in the driving pressure (Pes) concluding that the upper airway obstruction is only one aspect of OSA, and decrease in central nervous activity diminishes the respiratory drive before upper airway obstruction [133].

**Human considerations**

The prevalence of SDB from epidemiological studies has been reported to be higher in men than in women. Clinic based studies reported a very strong male predominance of ratios of 8:1 or greater for OSAS [134]. Newer community based studies by Young et al. [1] found smaller differences in the prevalence of OSAS with male-to-female ratios of 2:1 or 3:1. The reasons for the lower prevalence in women are still being examined. As described earlier there may be differences in the upper airway size or mechanics that predispose men to SDB. The greater central body fat distribution and larger neck dimensions of men account for some differences in prevalence [135]. Some authors have suggested that female sex hormones may provide a protective effect on the upper airway [57]. There is also a suggestion that SDB may be under diagnosed in women [136]. However, the reason for this is not clear as no differences have been found in the symptoms reported by men and women for the same degree of polysomnographic abnormality [137]. These issues are being examined in large scale multi center trials like the Sleep Heart Health Study [138].

**EVALUATION OF OBSTRUCTIVE SLEEP DISORDERED BREATHING**

**Detection/definition of the critical changes in upper airway patency**

The review thus far has focused on the concepts underlying the physiology of the upper airway that lead to changes in patency – i.e. what determines the resistance, collapsibility and ultimately ability of the upper airway to function as a conduit for respiratory gases. However, breathing is a fundamentally cyclic process. Individual breaths rather than a sustained “flow” define normal ventilation, and sleep disordered breathing consists of episodic chains of abnormal breaths alternating with normal ones, which together determine the pattern of gas exchange. Thus one needs to go beyond describing the physiology of steady flow through a tube to fully understand the derangements of obstructive SDB. Furthermore, the
techniques used to detect and define these abnormalities in breathing are intimately involved in our ability to understand the process.

Abnormal breath detection

Having identified that one wishes to detect that a change has occurred in the overall behavior of the upper airway associated with "sleep disordered breathing" the next step is to operationalize the detection of the abnormality that will be the basis of identifying an "abnormal" breath. The earliest practical approach was to identify breaths that did not occur despite clear evidence of respiratory effort (i.e. obstructive apnea). This is easily done with any detector of respiratory gas at the nose and mouth; in the usual form, these detectors rely on the different characteristics of inspired and expired air as a marker. Thus temperature oscillation (used by thermistors or thermocouples) between cooler inhaled room air temperature and warmer exhaled air (near body temperature) and oscillations in PCO₂ (only present in exhaled gas) were, and remain, good markers of the presence of a breath. Alternatively, respiratory efforts, measured as movement of the chest and abdomen, can also be calibrated to represent a "tidal volume". When these movements are fully out of phase, this may indicate total obstruction to airflow because there is no net change in the volume of the chest cavity. However, all of the above are best at detecting "all or nothing" air movement. In the case of incomplete airway obstruction the relationship between temperature change and true flow is not straightforward [139, 140].

By definition, measurement of airway resistance must rely on actual measurement of a true index of both airflow and a separate measure of intrathoracic effort or pressure. It has been suggested that the reference standard for measurement of respiratory effort should be esophageal pressure measurement [141]. This technique consists of inserting a trans-nasal flexible pressure catheter to directly measure intrathoracic (and presumably pleural) pressure swings during inspiration as a continuous measure of effort. However, the determination of airway resistance actually requires more than effort and should be based on its ratio to airflow (directly measured at mouth and nose with a snug fitting mask).

The concept of "hypopnea" was introduced by Block et al. as periods of shallow breathing with oxygen desaturation [142]. The assumption was made that reductions in ventilation were due to increases in airway resistance. Ventilation (tidal volume) was thus the critical variable to be measured, and this was done by derivation from non-invasive Inductance Plethysmography (which measures chest/abdominal movement). Transient reductions in the signal were confirmed as being reductions in ventilation (presumed due to increases in resistance) by looking for consequences such as a drop in O₂ saturation. While originally successful at defining a characteristic pattern of abnormal breaths, the rapid extrapolation of these definitions to use with other more qualitative monitors (thermistors, thermocouples and end-tidal PCO₂ devices) led to great non-uniformity in what was detected and called "hypopnea" [143, 144]. More recently, it has been suggested a return is necessary to measurements which truly measure volume or flow and correctly quantitate the underlying ventilation. Use of nasal/full facemasks with pneumotachographs, although theoretically necessary, have tended to be reserved for research. In the past several years, we and others have suggested an attractive surrogate for true pneumotachography; the use of a combination of a nasal cannula and pressure transducer, which has been validated for non-quantitative (relative) detection of ventilation [139, 140, 145, 146].

The combination of this return to measurement of actual airflow or its closest analogues, and the recognition that breaths may occur that have high resistance/abnormal airway collapse without actual reduction in airflow has led to evaluation of techniques to detect airway resistance directly on a breath by breath basis [147]. In the current approaches to detecting SDB, the shift has been away from detecting only actual ventilation, to trying to detect unexpected increases in respiratory drive (inspiratory EMG or esophageal pressure [2]) during "normal" appearing breaths (high resistance), or to measuring indirect evidence of a change in airway patency by detecting consequences (arousal, desaturation, elevation in markers of sympathetic activity [148], or finally to measuring surrogates of flow/pressure relationships which may indicate abnormal behavior. The last of these has been the subject of much recent work by our laboratory and is based on the "flow limitation" concept.

Flow limitation

"Flow limitation", as used in the context of SDB, refers to the characteristic flattened shape of the
pressure-flow relationship which has been observed when increasing effort does not result in increasing airflow in the upper airway. In addition to measuring amplitude of the flow (useful to detect conventional apnea and hypopnea), the identification of a plateau on the inspiratory flow/time waveform correlates with an elevated upper airway resistance (measured as pressure divided by flow) during CPAP titration [49] and during spontaneous breathing [140, 149]. Snoring generally occurs with flow limitation, but flow limitation appears to be a more sensitive marker of elevated resistance [150, 151], and treatment aimed at normalizing flow limitation has resulted in improvement of EDS [152].

Early studies to identify flow limitation were performed with a conventional pneumotachograph flow signal but required a tight-fitting face mask, which was excessively intrusive for routine sleep monitoring. The nasal cannula/pressure transducer circumvents this limitation, and provides a simple alternative to the pneumotachograph that provides a semi-calibrated signal proportional to inspiratory airflow. It detects the small pressure fluctuations caused by inspiration and expiration inside the nose. We and others have shown that the signal obtained from this device is comparable in both shape and amplitude to that of a conventional pneumotachograph [49, 145, 153, 154].

The relationship between this signal and a simultaneous pneumotachographic flow signal is essentially linear over the relevant range of flow measured, at least over periods of interest (20–100 breaths) (Fig. 9). This system has recently been accepted by the general sleep community in a position paper as one of the two acceptable methods of monitoring respiration during sleep [141]. The simplicity, low cost and non-obtrusiveness of the device make it ideal for quantitative monitoring of respiration during sleep [155].

Events

Sleep Disordered Breathing is characterized by intermittent sequences of abnormal breaths. Thus, apart from identifying individual breaths that are abnormal, it is necessary to define which “sequences” of breaths define pathological “events”. Breaths are an intermittent “sampling” of the airway behavior and it is transient abnormalities (i.e. elevations in resistance) which need to be detected during sleep. Apnea is easily defined and identified as an abnormal respiratory event that has a specific duration. The logical extension of the concept of apnea is to describe an abnormal period lasting more than a given fixed duration, now including breaths that have only partial airway collapse. Thus one must detect (and define) events composed of a series of breaths with a characteristic (low volume, low flow, high effort, a waveform indicating flow-limitation) indicating a period of sustained but intermittently elevated instantaneous resistance (Fig. 10). One can then study the consequences (desaturation, arousal etc.) of such events. Alternatively, one can specify a sequence of breaths that are different (e.g. smaller) and end in an arousal or desaturation and assume these are caused by high resistance.

Identification of these sequences of abnormal breaths (“events”) in clinical practice has traditionally been performed in two logical steps. Thus one first uses a breathing detector (e.g. thermistor/thermocouple) to find events of reduced airflow; these are then validated as true “events” or hypopneas by checking for the coincidence with desaturation or arousal on EEG. Typical definitions in use until recently required a fixed decrease in the thermistor signal (50–75%) associated with oxyhemoglobin desaturation of 2–4% or EEG microarousal. This two step algorithm was necessary because the detector is not as sensitive as desired, but also not very specific if its “gain” to detect potential events is not maximal. Unfortunately, even with this approach, a substantial number of arousals are detected where no “event” is suspected. This number may be as high as 20% in typical OSAS, but may be much higher (up to 70%
in milder SDB, e.g. the upper airway resistance syndrome (UARS). In fact the UARS is defined by the presence of Respiratory Effort-Related Arousals – which are arousals associated with increased respiratory effort in the absence of any detectable events (at least by thermistry).

However, if one uses a more sensitive means for detection of abnormal breaths (e.g. pneumotachograph or nasal cannula surrogate), nearly all abnormal breaths due to increased respiratory effort and/or increased resistance can be detected, and thus the “candidate events” are more easily defined as a sequence of truly abnormal breaths. The need to “filter” these events by looking for the confirmatory desaturation or arousal becomes less compelling. Other techniques that have been proposed to identify sleep disordered breathing with fairly high sensitivity include analysis of pulse transit time (PTT) [156], respiratory effort using thoracoabdominal bands [157], and measurement of upper airway impedance using the forced oscillation technique [158, 159]. To some extent, all of these have been shown to identify high resistance respiratory events with good correlation to esophageal events. There is some difference in the way the various techniques classify an event as an apnea, hypopnea or other high resistance event, but it remains to be shown that this has clinical importance greater than detecting all the events in the first place [141].

Whatever technique one uses to detect “events” – now defined solely by a respiratory change and by some arbitrary minimal duration (usually 10 s in adults) – it remains an open question whether counting the number and duration of these events will be sufficient to explain the spectrum of clinical complaints of patients (e.g. sleepiness and neuropsychiatric dysfunction), as well as to grade the risk of long term cardiovascular sequelae of SDB [160]. To date, most attempts to relate a count of events (AHI or RDI) to severity of patient symptoms or outcomes have resulted in only modest correlations.

One of the interesting questions, about which only a little is known, is what terminates a period of abnormal airway behavior, i.e. what ends a sequence
of abnormal breaths. The general consensus is that some level of brain activation (i.e. arousal) is always seen at the end of an obstructive event, just as loss of the state of arousal is the proximate cause of the onset of airway abnormality/ventilatory instability in SDB [161]. This is not equally true of “central” events, where the predominant abnormality is loss of central neural respiratory “drive” rather than loss of airway patency. Thus at the end of an obstructive event frank EEG arousal (a change in basic frequency and appearance of waves characterizing the “awake” state) is usually seen, but these EEG changes may be absent in up to 20% of obstructive events. Some of these “missed” arousals are clearly due to our inability to define and “rate” EEG changes which are visible to the eye, or even to more sophisticated signal processing techniques. Some may be due to cortical activation having been present at other sites of the brain (frontal) and not detected by the central EEG leads conventionally used in sleep monitoring (C3, C4). In rare (approximately 5%) of cases, no cortical EEG changes are seen at the end of an obstructive event. However, other markers of brain-stem activation (change in autonomic tone, heart rate or BP) are almost invariably detected, indicating some degree of “arousal” was present [162, 163].

All of the above begs the question of what causes the brain to arouse in situations where obstruction of the airway has occurred. Although sufficient chemical (i.e. hypoxic and hypercapnic) stimuli will arouse most individuals (even those with spinal injury and thus unable to respond to express the ventilatory stimulation), it has been suggested that it is the mechanical effort which generates a common signal causing arousal under most physiological conditions. Thus Gleeson et al. [164] showed that for a given individual in a single sleep state, arousal always occurred at the same level of intrathoracic pressure swing with a breath – whether the effort was generated in response to obstruction, hypoxia or hyper-capnia. During clinical sleep studies, many patients tend to show a regular recurrent level of drive at arousal during periods of obstructive apnea, and have regular apnea lengths, minimal saturations and similar large breaths after each apnea, suggesting they have aroused at very similar levels of stimulation for each episode. To date, little has been done to use this physiological “signature” of an individual’s apneas to characterize their clinical disease. In contrast to the mechanoreceptor hypothesis for arousal, Ayas et al. [165] showed that hypercapnia alone could induce arousal from sleep in a study on neurologically complete spinal cord injury patients. Because of the recent improvement in sensitivity and specificity in detecting subtle events (e.g. Respiratory effort related arousals – RERAs) in clinical populations, we recently tested the hypothesis that improved detection (with a nasal cannula to detect respiration) would translate into better correlation with clinical symptoms, and an improved ability to separate normal subjects from those with “disease” based on polysomnography alone. The results of testing several algorithms for event detection [166] suggest that counting all respiratory events (classical apneas, hypopneas and clusters of flow limitation breaths indicative of Respiratory Effort Related Arousals – RERAs) may provide the best approach if symptoms of excessive daytime somnolence are used as the dependent variable. Further refinements of the algorithm by which the magnitude of SDB is described (e.g. a two dimensional approach giving different weights to apnea and non-apnea events to define their impact on EDS) were as good as, but did not improve on the ability to separate normal from diseased subjects (Fig. 11) At this time it seems likely that even with the “best” event detection and analysis, some degree of differential susceptibility of individuals will, at least in part, determine the clinical manifestations at any given level of physiological abnormality of the airway. Despite this, the quest remains to define the respiratory index of abnormal airway behavior which provides the greatest clinical utility and provides at least a partial physiological means to separate patients from normals by an objective test of breathing.

**SUMMARY**

The upper airway is a complex structure with many competing physiological functions, including swallowing, speech and breathing. It is richly innervated and its mechanical behavior is influenced by a multitude of neural controls and reflexes. Attention to its behavior with changes in brain states from arousal to sleep has helped us to define the spectrum of sleep disordered breathing, but our knowledge remains incomplete. Because SDB appears to be very prevalent in both clinic and epidemiologic populations, upper airway dysfunction has assumed an ever-increasing role in clinical medicine. However, much research remains to be done to understand the full range of anatomic, passive mechanical and active neural influences on
upper airway patency and to fully exploit this knowledge in the definition of disease states and the optimization of therapies for disease.

Practice Points

1. The upper airway is a collapsible structure whose patency is dictated by a combination of passive mechanical properties and active neural mechanisms.

2. Sleep produces changes in airway physiology, which, in association with abnormal anatomy, lead to sleep disordered breathing events and the syndrome of obstructive sleep apnea/hypopnea.

Research Agenda

1. The full range of anatomic, passive mechanical and active neural influences on upper airway patency needs to be more fully described.

2. The neural pathways leading to arousal and their implications for EDS need to be better understood.

3. This knowledge needs to be used to improve the definition of disease states, ranging from snoring to obstructive sleep apnea, and to help develop optimal therapies for disease.

Figure 11  Plot of apnea index (AI) against non-apnea event index (RDI\text{Total} - AI) for 103 test subjects. Open circles are asymptomatic subjects, open triangles are snorers without excessive daytime somnolence (EDS) and the closed circles are the subjects with both EDS and snoring. The region bounded by AI < 2/h and RDI\text{Total} - AI < 16/h was defined as identifying the “normal” region. (From Hosselet J J, Ayappa I, Norman R G, Krieger A C, Rapoport D M. Classification of sleep-disordered breathing. *Am J Respir Crit Care Med* 2001; 163(2): 398–405; with permission.)
REFERENCES


UPPER AIRWAY IN SLEEP


