Oropharyngeal Examination to Predict Sleep Apnea Severity

Xavier Barceló, MD; Rosa M. Mirapeix, MD, PhD; Jordi Bugés, MD; Albert Cobos, MD, PhD; Christian Domingo, MD, PhD

Objective: To evaluate the usefulness of the examination of the upper airway, paying special attention to the Friedman tongue position (FTP), to confirm obstructive sleep apnea syndrome (OSAS) and its severity.

Design: Prospective, single-center, cross-sectional study.

Setting: Sleep disorders unit of a community hospital.

Patients: A total of 301 consecutive patients admitted to the sleep disorders unit due to suspicion of OSAS. Assessments included body mass index calculated as weight in kilograms divided by height in meters squared (BMI); neck perimeter measurement; oropharyngeal examination; fiberendoscopy; rhinomanometry; and a sleep study.

Main Outcome Measures: Apnea-hypopnea index (AHI), FTP, the uvula size, and certain complementary examinations (sex, age, BMI, cervical perimeter, nasal flow) whose importance has not been clearly established, and to explore their potential value as predictors of the AHI.

Results: Findings included the following: the mean (SD) age of the patients was 51 (12) years; 71.1% were male; the mean (SD) BMI was 29.8 (4.6); and the mean (SD) cervical perimeter, 40.5 (3.7) cm. In 94.0% of the patients the AHI value was at least 5.0/hour. Patients with FTP scores of 2 and 3 accounted for 74.1% of the whole cohort: 14.3% had an FTP score of 1, and only 11.6% had a score of 4. Of the 6.0% of cases with a normal AHI, 16 patients were classified as having FTP scores of 1, and 2 as having a score of 2. Tonsil size score (P = .005), uvula score (P = .003), BMI (P < .001), cervical perimeter (P < .001), nasal flow at 150 Pa (P = .02), and age (P = .007) were related to OSAS severity. Curiously, AHI in patients who had undergone tonsillectomy was higher than in the TS1 group (tonsils inside the tonsillar fossa) and quite similar to the TS 2 group (tonsils that extend beyond the tonsillar pillars). In the multiple regression model, only the FTP score showed a relevant relationship to OSAS severity.

Conclusions: First, since the FTP score is almost the only parameter related to OSAS severity, a simple oropharyngeal examination can provide key information on this issue. Second, tonsillectomy does not seem to protect against development of OSAS.
The purpose of this study was to validate the clinical usefulness of easy-to-assess clinical signs such as the Friedman tongue position (FTP), uvula size, and certain complementary examinations whose importance has not been clearly established, and to explore their potential value as predictors of the AHI.

## Methods

### Patients

Patients who were naive for sleep studies were referred to the sleep laboratory for diagnosis and treatment of OSAS. Clinical suspicion of OSAS was established on the basis of snoring and daytime sleepiness recorded during the clinical interview with the Spanish version of the Epworth Sleepiness Scale. Patients with snoring history and clinical suspicion of OSAS entered the study protocol.

### Inclusion and Exclusion Criteria

For inclusion, an individual had to be an adult patient (>18 years) referred to our Institution to diagnose or rule out OSAS. Exclusion criteria included (1) having other respiratory diseases, (2) having undergone a tracheostomy, (3) having a clinical suspicion of other sleep disorders, (4) having other comorbidities (chronic renal, cardiac, or hepatic failure), (5) having active neoplasms, or (6) refusal to follow the study protocol of our sleep disorders unit.

### Sample Size

We divided patients into 4 groups: group 1 comprised individuals without OSAS; group 2, patients with mild OSAS; group 3, patients with moderate OSAS; and group 4, patients with severe OSAS. We planned to have a minimum of 30 patients in the smallest FTP group. We considered that many of the patients referred to our sleep laboratory would present some degree of sleep apnea; only a minority were expected to be snorers without sleep disorder (<10%). We also estimated a loss to follow-up of 10%. In view of our past experience, we expected the frequency of sleep apnea to be 3 times higher in groups 2 and 3 than in groups 1 and 4; we therefore aimed to include a minimum of 268 patients.

### Study Description and Aims

This was a single-center, observational, cross-sectional study performed at the sleep disorders unit of a community hospital (Esperit Sant de Santa Coloma, Barcelona, Spain) with a reference area of 150,000 inhabitants. The study’s primary aims were (1) to determine the relationship between AHI values and potential explanatory variables (sex, age, body mass index [BMI]; calculated as weight in kilograms divided by height in meters squared), (2) to investigate the role of these variables as predictors of the AHI. The secondary aim was to determine a predictive model to establish the diagnosis and severity of OSAS.

### Study Protocol and Instrumentation

The following examinations were performed in all patients during the first visit: (1) BMI determination, (2) neck perimeter measurements, (3) oropharyngeal examination, (4) fiberendoscopy, (5) phinomanometry, and (6) a sleep study.

At entry, after external physical examination, the otorhinolaryngologist (X.B.) conducted the oropharyngeal examination and the rhinopharyngolaryngoscopy or nasal fibroscopy. The oropharyngeal examination was performed by the same physician in all cases using a frontal light to establish the FTP (Figure 1) and tonsil size (TS) score (Figure 2) using Friedman’s criteria with the patient seated. We added a uvula score predesigned for this study according to the following criteria:

**Uvula Score**

1. Uvula clearly visible without hypertrophy
2. Uvula slightly hypertrophied without touching the tongue
3. Hypertrophic uvula in contact with the tongue base
4. Markedly hypertrophic uvula lying on the tongue base

Rhinomanometry was then performed the day before the sleep study. The fiberendoscopy was performed with a Rhino-Laryngo fiberscope (model ENF-GP; Olympus Co, Tokyo, Japan) and the rhinomanometry with a Sibelmed Rhinospir pro (Sibel S.A., Barcelona, Spain) in accordance with international guidelines. The sleep study was performed using a previously validated portable respiratory recording device (PRRD) (Sibel home Plus; Sibel S.A.). The system records oronasal airflow by thermistor; nasal airflow via a nasal cannula, which acts as a pressure transducer; chest and abdominal respiratory movements; tracheal sounds; heart rate; oxygen saturation; and body position. The PRRD was programmed to start 10 minutes after patients went to sleep. Sleep studies were performed at the patient’s home using the unattended mode.

The AHI was calculated as the sum of the number of episodes of apnea and hypopnea per hour of recording. The analysis was performed manually using the criteria recommended by the American Academy of Sleep Medicine. Respiratory events were characterized as apnea when there was a cessation or reduction of more than 90% of airflow lasting 10 seconds and as hypopnea when any clearly discernible reduction in airflow (>30% and <90%) lasting 10 seconds was observed, associated with a dip of at least 3% in pulse oximeter oxygen saturation.

---

**Figure 1.** The different Friedman tongue position scores (reproduced with permission). A, Grade 1: the examination allows a complete view of the uvula, the soft palate and the tonsils; B, grade 2: the uvula is visible, but not the tonsils; C, grade 3: the soft palate can be seen, but not the uvula; D, grade 4: only the hard palate can be seen.
To assess OSAS severity, patients' AHI was classified as follows: an AHI value below 5.0/hour was considered normal; an AHI value of 5.0/hour to 14.0/hour, mild; an AHI value of 15.0/hour to 30.0/hour, moderate; and an AHI value higher than 30.0/hour, severe; in accordance with the guidelines of the American Academy of Sleep Medicine.11

**STATISTICAL ANALYSIS**

Data are given as mean (SD). The bivariate relationships between AHI values and potential explanatory variables (sex; age; BMI; cervical perimeter; FL150; and FTP, TS, and uvula scores) were assessed using the Spearman correlation coefficient or the Wilcoxon rank sum test as appropriate. A multivariate regression analysis was conducted to investigate the predictive role of potential explanatory variables in more depth. To achieve normality and homoscedasticity, a logarithm transformation was used, so that the logarithm of the AHI (log AHI) was modeled rather than the AHI itself. Two different strategies were used to select the best-fitting model from a total of 511 alternatives. First, we looked for the model with lowest Mallows’ \( C_p \) value. Second, we used a stepwise selection procedure with a 0.10 probability level of the \( F \) statistic entering or staying in the model. Once a model was selected, 2-way and 3-way interaction terms were also explored. The goodness of fit of the final model was assessed by inspecting residual plots. All statistical analyses were performed with SAS software (version 9.1; SAS Institute Inc, Cary, North Carolina).

The study protocol was approved by the hospital's institutional review board. Only patients’ oral informed consent for the anonymous treatment of their data was required since the study protocol was the same as that used for the standard medical care of these patients in our institution.

**RESULTS**

We prospectively studied 301 consecutive patients from January 2006 to December 2009. Their ages ranged from 18 to 82 years (mean [SD], 51 [12] years),12 71.1% were male, and 28.9% were female, the mean (SD) BMI was 29.8 (4.6), the cervical perimeter was 40.6 (3.7) cm, and the FL150 was 0.80 (0.26) cm³/s.

The distribution of FTP, TS, and uvula scores, along with other patients' characteristics, are displayed in Table 1. Note that for these 3 scores, the 2 middle categories accounted for about 80% of the sample. The AHI values ranged from 0.7/hour to 89.5/hour (mean [SD], 27.6/hour [19.4/hour]). Ninety-four percent had an AHI value higher than 5.0/hour.

**BIVARIATE ANALYSIS**

The AHI values showed a statistically significant correlation with FTP scores, TS scores, uvula scores, BMI, cervical perimeter, and age. In all cases, the correlation was positive, but its intensity varied markedly: while a high correlation was found in the case of the FTP score (Spearman \( r=0.88 \) (Table 2)), the correlation was only moderate for the cervical perimeter \( (r=0.40) \) and very low for the TS scores, uvula scores, and age (Table 2). We
did not detect a significant correlation between AHI values and the FL150.

The AHI values were significantly higher in males than in females (Wilcoxon rank sum test; \(P = .02\)): the median AHI value was 24.3/hour for males and 18.0/hour for females.

**CRUDE DESCRIPTION OF AHI INDEX AND FTP CLASSIFICATION**

Almost three-quarters of patients (74.1%) had FTP scores of 2 and 3 (Table 1). Findings from 18 of 301 patient studies (6%) were considered normal (AHI < 5.0/hour): 16 were classified as FTP 1, and only 2 as FTP 2.

All patients with an FTP score of 1 had an AHI value lower than 16.0/hour; in most of them (64.0%), the AHI value was 5.0/hour to 10.0/hour. Most patients with an FTP score of 2 (58.7%) had an AHI value of 10.0/hour to 20.0/hour. Two patients had normal values, and only 5 had an AHI value of 30.0/hour (all of them with a uvula index of 3). The FTP scores and AHI categories are shown in Table 3.

One patient with an FTP score of 3 had an AHI value of 13.4/hour, but all the others had an AHI value of 16.0/hour. The AHI value in most patients with an FTP score of 3 ranged from 32 to 48, with 3 patients showing an AHI value higher than 60.0/hour. All patients with an FTP score of 4 had an AHI value higher than 30.0/hour, corresponding to severe OSAS. The relation between FTP score and AHI is plotted in Figure 3.

Just over three-quarters of patients (77.4%) had TS scores of 1 and 2 and only 15 had a TS score of 3 (5.0%). None had a TS score of 4. Tonsil size scores of 3 and 4 are common in the pediatric population but are rare in adult OSAS. Fifty-two of the 301 patients (17.3%) had previously undergone tonsillectomy (Table 1).

**PREDICTIVE MODEL FOR AHI VALUES**

Table 4 shows the correlation of the AHI with the explanatory variables. The model with the lowest \(C_p\) value included the FTP score, FL150, the cervical perimeter, age, and the TS score (Table 4). Despite the high correlation coefficient, the TS score did not enter into the model when using the stepwise selection procedure because it did not reach the specified significance level of 0.10. Both

---

**Table 1. Characteristics for the 301 Patients in the Study Group**

<table>
<thead>
<tr>
<th>Study Variable</th>
<th>No. (%) in Each Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTP score</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>43 (14.3)</td>
</tr>
<tr>
<td>2</td>
<td>114 (37.9)</td>
</tr>
<tr>
<td>3</td>
<td>109 (36.2)</td>
</tr>
<tr>
<td>4</td>
<td>35 (11.6)</td>
</tr>
<tr>
<td>Tonsil size score(^b)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>52 (17.3)</td>
</tr>
<tr>
<td>1</td>
<td>124 (41.2)</td>
</tr>
<tr>
<td>2</td>
<td>109 (36.2)</td>
</tr>
<tr>
<td>3</td>
<td>15 (5.0)</td>
</tr>
<tr>
<td>Uvula score</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14 (4.6)</td>
</tr>
<tr>
<td>2</td>
<td>147 (48.8)</td>
</tr>
<tr>
<td>3</td>
<td>126 (41.9)</td>
</tr>
<tr>
<td>4</td>
<td>14 (4.6)</td>
</tr>
</tbody>
</table>

\(^a\)Friedman classified tonsil size in 5 grades from 0 to 4 (grade 0, patients who have had their tonsils removed; grade 1, when the tonsils are inside the tonsillar fossa; grade 2, when they extend beyond the tonsillar pillars; grade 3, when they extend beyond the tonsillar pillars but do not reach the midline; and grade 4, when they extend as far as the midline.

\(^b\)We did not include group 5 because there were no cases of grade 4 (group 5) in our series.

**Table 2. Spearman Correlation of AHI With Explanatory Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Spearman Correlation</th>
<th>(P) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.16</td>
<td>.007</td>
</tr>
<tr>
<td>FL150</td>
<td>0.02</td>
<td>.71</td>
</tr>
<tr>
<td>Cervical perimeter</td>
<td>0.40</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>BMI</td>
<td>0.25</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>FTP score</td>
<td>0.88</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Tonsil size score</td>
<td>0.16</td>
<td>.005</td>
</tr>
<tr>
<td>Uvula score</td>
<td>0.17</td>
<td>.003</td>
</tr>
</tbody>
</table>

Abbreviations: AHI, apnea-hypopnea index; BMI, body mass index; FL150, nasal flow, measured in decimeters cubed per second at 150 Pa; FTP, Friedman tongue position.

**Table 3. FTP Scores and AHI Categories (AASM)**

<table>
<thead>
<tr>
<th>FTP Score</th>
<th>Patients per AHI Category, No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal (37.2)</td>
</tr>
<tr>
<td>1</td>
<td>16 (37.2)</td>
</tr>
<tr>
<td>2</td>
<td>2 (1.7)</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Abbreviations: AASM, American Academy of Sleep Medicine; AHI, apnea-hypopnea index, measured as number of episodes of apnea and hypopnea per hour of recording (see “Methods” section); FTP, Friedman tongue position.
Two-way and three-way interactions between the variables selected by the stepwise procedure were not significant (F tests) (Table 4).

The model finally selected included the FTP score, FL150, cervical perimeter, and age. The root mean square error for this model was 0.454, and the R² statistic was 0.748. The parameter estimates and F tests are shown in Table 4 along with partial R² values. Note that although F tests were significant, the partial contribution of FL150, age, and cervical perimeter to the fit (as measured by the partial R² statistic) was negligible. The final predictive equation for the expected log AHI was −0.320 FTP (score) − 0.409 FL150 (dm³/s) − 0.033 Cervical Perimeter (cm) − 0.006 Age (years).

**COMMENT**

Obstructive sleep apnea syndrome is very common and affects a considerable proportion of the population. If untreated, it causes considerable increases in cardiac and cerebrovascular morbidity and mortality. Many patients with obstructive OSAS remain undiagnosed owing to inadequate resources for case detection and investigation, a situation that has clear repercussions for their quality of life. Moreover, OSAS represents a considerable socioeconomic burden, owing to comorbidity, increased health care utilization in the primary and secondary health care sectors, increased use of medication, impact on employment, and loss of income. Treatment of OSAS reduces morbidity, mortality, and hospitalization rates, and its cost-effectiveness has been demonstrated. So, there is a clear need to establish optimal procedures to identify patients at risk.

Traditionally, OSAS diagnosis has been performed using in-hospital, overnight polysomnography. This procedure is time consuming and expensive, and, since its capacity is limited, waiting lists tend to lengthen.

In recent decades, several attempts have been made to find predictive models based on clinical examination of the upper airway, combined or not with clinical data such as the Epworth Sleepiness Score and/or specific radiological examinations. These models have proved difficult to implement, at least to some extent because the description of the anatomical regions involved has not been uniform. Several studies have shown a statistical relationship between certain morphological abnormalities and OSAS severity. Tonsil size and cervical perimeter seem to be related to OSAS severity, although their specific weight in the pathophysiological mechanisms of the disease is not clear. The Mallampati score seems to be the only parameter that is strongly related to OSAS severity.
endotracheal intubation. Friedman et al\(^9\) definitively re-
termined that the position of the palate with respect to
the tongue was an indicator of the ease or difficulty of
OSAS (Figure 4).

Friedman tongue position\(^9\) suggests that the degree of
nasal obstruction has a neg-
ligible influence on the pathophysiological mechanisms
of the disease. These results question the usefulness
of grades A and B in Fujita’s classification.\(^{34}\) Interestingly,
the cervical perimeter shows the second largest \(R^2\) value,
which may point to obesity as a risk factor (BMI was also
significantly correlated with AHI but did not enter the
model, possibly because cervical perimeter is more spe-
cific as a measure of local obesity).

Similarly, TS was irrelevant. Curiously, tonsillectomy
performed during childhood did not seem to be pre-
ventive but in fact seemed to favor the development
of OSAS (Figure 4).

Previous studies have shown a statistical relationship be-
tween the FTP (frequently called the modified Mallampati
score) and sleep disturbances.\(^{9,27,29-33}\) To our knowledge,
our study is the first to quantify the specific weight of these
maneuvers in a multivariate regression model. Our results
suggest that a patient with an FTP score of 3 or 4 is un-
likely to have an AHI value within the normal range. In other
words, findings from a correct oropharyngeal physical ex-
amination may contradict those from a sleep study con-
sidered to be normal. The strength of the association be-
tween the FTP scores and the OSAS severity is further
illustrated in Figure 3, in which we show the joint distri-
bution of FTP scores and AHI values categorized accord-
ing to the AASM criteria.\(^{12}\) Note that all patients with an
FTP score of 4 had severe OSAS; most of the patients with
an FTP score of 3 (71.6%) also had severe OSAS, but
all the remaining patients but 1 had moderate OSAS. Almost
all patients with an FTP score of 2 were evenly split between
mild and moderate categories, with only 2 cases (1.7%) being
normal and 4 cases (3.5%) being severe. Finally, patients
with an FTP score of 1 had either normal or mild OSAS but
very rarely moderate and hardly ever severe OSAS.

The FTP score is a modification of the Mallampati score, established by Friedman.\(^9\) Mallampati et al\(^{35}\)
determined that the position of the palate with respect to
the tongue was an indicator of the ease or difficulty of
endotracheal intubation. Friedman et al\(^{35}\) definitively re-
lated the position of the palate and the tongue with the
risk of OSAS. Friedman et al\(^{35}\) made changes to Mallam-
pati’s procedure: whereas Mallampati et al\(^{33}\) performed
the examination with the tongue outside the oropharyn-
geal cavity, Friedman et al\(^9\) maintained the tongue in-
side the oral cavity in a neutral position. Initially, they
called their observations the modified Mallampati grade\(^9\)
but later adopted the term Friedman tongue position\(^{36}\) be-
cause it is, in fact, the position of the tongue in relation to
the palate that is assessed. Friedman et al\(^{35}\) believed that
this modification provides a natural and physiological
language position similar to the one achieved during sleep.

The FTP score is assessed with the patient in a sitting
position (the supine position increases the Mallampati
score, as Tham et al\(^{37}\) observed) by opening the mouth
in a natural, nonforced manner, with the tongue inside
without swallowing or inspiring. We repeated the man-
uever 5 times and took the most frequent result. A cor-
rect performance of the examination is time consuming
but is necessary to avoid significant errors in determin-
ing the FTP score. In our cohort, those with FTP scores
of 2 and 3 accounted for 74.1% of the whole population,
but the severity of OSAS in the 2 groups was quite dif-
ferent. These results emphasize the importance of cor-
rect performance of the maneuver.

Our study has several limitations. First of all, the sleep
studies are performed on an ambulatory basis and might
underestimate the severity of sleep disorders. Nonethe-
less, these procedures are accepted in the literature, and
our experience shows their reliability in clinical prac-
tice.\(^{38,39}\) The second point applies to the target population.
Our study was performed in a white population; because
some studies have reported higher prevalence and severity
levels in Asian and African American populations,\(^{22,40}\)
there is no guarantee that our results apply to these populations.
Furthermore, because the sleep studies were not full poly-
sonomographic studies, the AHI may have been underestimat-
ed. Finally, the oropharyngeal examination was per-
formed by a specialist and may not be so accurate when
implemented in primary care. The reproducibility of the
examination has to be widely implemented.

Despite increasing awareness of the condition and im-
provements in diagnostic procedures, most patients with
OSAS in the community remain undiagnosed and un-
treated. In the primary care sector there is considerable
potential for improved diagnosis. In high-risk groups, such
as patients with acute and chronic cardiovascular or cere-
brovascular disease or those with diabetes mellitus, easy
and inexpensive tests are necessary to help primary care
physicians to detect patients at risk of OSAS. The com-
plexity and cost of diagnosis is high, and many patients who
are candidates for sleep studies may not receive adequate
care owing to geographical or economic limitations. More-
over, the morbidity associated with the disease is dramati-
cally reduced when patients are treated.\(^{22}\) All these char-
acteristics reinforce the importance of our results. We
advocate the use of a protocolized oropharyngeal exami-
nation at primary care level to diagnose and categorize the
severity of the disease and thus to optimize the decision to
send a patient to a sleep unit for further evaluation.

Submitted for Publication: May 20, 2011; final revision
received July 19; accepted August 14, 2011.

Correspondence: Christian Domingo, MD, PhD, Servei
de Pneumologia, Corporació Parc Taulí, Parc Taulí s/n,
08208 Sabadell, Barcelona, Spain (cdomingo@tauli.cat).
Author Contributions: All authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Barceló and Domingo. Acquisition of data: Barceló and Bugé. Analysis and interpretation of data: Barceló, Mirapeix, Cobos, and Domingo. Drafting of the manuscript: Barceló. Critical revision of the manuscript for important intellectual content: Mirapeix, Bugés, Cobos, and Domingo. Statistical analysis: Mirapeix, Cobos, and Domingo. Administrative, technical, and material support: Barceló and Bugés. Study supervision: Mirapeix and Domingo.

Financial Disclosure: Dr Mirapeix has organized courses and received lecture fees from Schering-Plough and Merck Sharp & Dohme. Dr Domingo has received lecture fees from Glaxo Smith-Kline, Astra-Zeneca, Novartis, Boehringer, Menarini, Bayer, and Esteve.

REFERENCES