The Relationship between the Mallampati Scoring System, the Berlin Questionnaire, and Epworth Sleepiness Scale

Joshua F. Gonzales, MA, RRT-NPS
S. Gregory Marshall, PhD, RRT, RPSGT
Chris J. Russian, MEd, RRT-NPS, RPSGT

1. Assistant Professor of Respiratory Care, Texas State University, San Marcos, Texas
2. Chair, Associate Professor of Respiratory Care, Texas State University, San Marcos, Texas
3. Associate Professor of Respiratory Care, Director of Clinical Education, Texas State University, San Marcos, Texas


ABSTRACT
The purpose of this study is to determine if there is a relationship between a person’s Mallampati score and their scores on the Berlin Questionnaire and Epworth Sleepiness Scale. The Berlin Questionnaire and Epworth Sleepiness Scale are currently used as prescreening tools for persons who may suffer from sleep disorders. The Berlin Questionnaire is used to identify persons at risk for sleep apnea syndrome and the Epworth Sleepiness Scale is used to determine the level of a person’s daytime sleepiness. The Mallampati Scoring System is an assessment of the anatomy of the oral cavity to predict the ease of intubation. The relationship between the Mallampati Score System, Berlin Questionnaire, and Epworth Sleepiness Scale were analyzed. Seventy-seven students participated in this pilot study. The correlation and regression analysis demonstrated the following results: The researcher found no significant correlations between the Mallampati Scoring System and the Epworth Sleepiness Scale and Berlin Questionnaire. The regression analysis revealed that a Mallampati Score could not be used as a predictor of Berlin Questionnaire and Epworth Sleepiness Scale results. Results of this pilot study reveal that the Mallampati Scoring System has no correlation and may not be used as a predictor to Berlin Questionnaire and Epworth Sleepiness Scale results.

INTRODUCTION
In 1975, Guilleminault et al. published an extensive case series of 250 patients complaining of problem sleepiness. The authors diagnosed 35 patients with “abnormal sleep apnea syndrome” experiencing “a minimum of 30 apneic episodes during seven hours of nocturnal sleep.” An apneic period was defined as “a cessation of air flow at nose and mouth enduring for a minimum of 10 seconds.” Within this case series, Guilleminault et al. observed a possible link between sleep apnea syndrome and hypertension, excessive daytime sleepiness, snoring, and a fragmented sleeping pattern. The authors concluded that men with a history of sleep problems should be evaluated for episodic abnormal respiratory events during sleep. Current medical literature substantiates many of the observations by Guilleminault et al., and most physicians would agree that a sleep assessment is warranted when the signs and symptoms of a sleep disorder are present.

Obstructive sleep apnea (OSA) can be described as repetitive obstruction of the upper airway often resulting in arousal from sleep and oxygen desaturation. Risk factors for OSA include increased neck circumference, obesity, craniofacial abnormalities, hypothyroidism, and acromegaly. In recent decades, recognition of this problem and the number of patients diagnosed with OSA has grown in the US. Young et al. report that approximately 1 in 5 adults has at least mild OSA and 1 in 15 adults has OSA of...
The Relationship Between the Mallampati Scoring System, the Berlin Questionnaire, and the Epworth Sleepiness Scale.

The BSQ assesses the presence of risk factors for OSA, such as snoring, wake-time sleepiness, the presence of obesity, and hypertension. The ESS is a self-administered eight-item questionnaire designed to assess daytime sleepiness in adults. Both instruments are self-reporting tools requiring honest and accurate answers in order to be effective. A third instrument, which is controversial in its relationship to OSA, is the Mallampati scoring system (MSS). The MSS is a medical scoring system that utilizes a visual oropharynx airway classification and was originally developed to identify patients at risk for difficult endotracheal intubation before surgery.

The BSQ has demonstrated an impressive ability to survey and screen sleep disorder subjects. Sharma et al. showed that the administration of the BSQ may help identify high-risk subjects and potentially avoid unnecessary polysomnography studies. Sharma’s study comparing Berlin survey scores with full polysomnography sleep studies demonstrated a sensitivity of 86% and a specificity of 96% toward predicting OSA. In another study involving primary care patients, Netzer found a predictive sensitivity of 86% and specificity of 77% (at a cutoff apnea hypopnea index (AHI) greater than 5), a predictive sensitivity of 54%, and a specificity of 97% (at a cutoff AHI greater than 15) for predicting the presence of OSA as confirmed by overnight polysomnography sleep study.

The Epworth Sleepiness Scale

The ESS is a self-reporting, eight-question survey developed through the observation of the frequency and repetition of symptoms associated with daytime sleep and sleepiness in adults. The questionnaire asks subjects to rate, on a scale of 0-3, the chances when dozing is likely in eight specific situations common in daily activity (i.e. 0 = would never doze; 3 = high chances of dozing). Participants in this study were asked to distinguish dozing behaviors from feelings of general tiredness. The ESS total score is the sum of the eight item scores and has a range of 0 to 24 points with scores greater than 9 suggestive of excessive daytime sleepiness.

The ESS has also been shown to be a reliable method for prescreening OSA patients by measuring persistent daytime sleepiness in adults. Johns found the ESS to be reasonably reliable in the test-retest trials. In Johns’ study, ESS scores did not significantly change and were highly correlated when tested and retested 5 months later in students whose daytime sleepiness was expected to stay the same. More importantly, Johns also found that high ESS scores of patients with confirmed OSA were reduced to more normal scores after 3 months of treatment with nasal continuous positive airway pressure.
The Mallampati Scoring System

The MSS was developed as an oral airway classification system and has been used for more than 20 years to identify patients presenting with a potentially difficult endotracheal intubation. Subjects are instructed to open their mouths to allow for direct visualization of the oropharyngeal region by the assessor. Assessors using the MSS must be able to identify the four anatomical landmarks found in the oropharyngeal region. Classifications of MSS are categorized as Class I, Class II, Class III, and Class IV based upon how many of the four normally visualized anatomical structures observed. Class I is considered “normal” with all four anatomical structures seen: the soft palate, hard palate, uvula, and tonsillar pillars. In MSS Class II, only the soft palate, hard palate, and upper portions of the uvula and tonsils are visible. Class III is defined by visualization of only the soft and hard palate and base of the uvula, while in the MSS Class IV, only the hard palate is visualized. In short, the MSS classifies the amount of “breathing room” an individual has around anatomical structures. Recent studies have shown that the MSS can also be useful as an assessment tool for subjects at high risk for OSA due to anatomical crowding of the oropharynx. The MSS instrument requires the assessor to identify four key anatomical structures that are easily and quickly observed when the subject opens the mouth widely, as shown in Figure 1.

Figure 1. Mallampati Scoring System

Although the ESS and BSQ are both self-reported questionnaires, the physician or healthcare provider can easily administer them. Administration of the MSS requires a qualified healthcare provider familiar with oral anatomy, such as a physician, physician-assistant, respiratory therapist, or nurse. However, the MSS is non-invasive and easily conducted in less than 15 seconds by trained healthcare providers. The ESS, BSQ, and MSS instruments have independently demonstrated clinical usefulness in identifying symptomology associated with high probability OSA. Since the ESS, the BSQ, and the MSS are primarily used in patient assessment to identify excessive sleepiness, snoring, and anatomical room for mouth-breathing respectively, the researchers developed the following hypotheses: 1) the BSQ and the MSS may predict ESS and 2) the MSS may predict BSQ. The null hypotheses states: 1) the BSQ and the MSS are not predictive of the ESS and 2) the MSS is not predictive of the BSQ.

METHODS

Participants
Prior to selecting participants, this study was approved by the Institutional Review Board at Texas State University-San Marcos. A convenience sample of college students enrolled in a health professions program was chosen. Eligibility criteria included freshman or sophomore students enrolled in the respiratory care (RC) program. Subjects self-reporting recent illness were excluded from participating in the study. Participation was strictly voluntary and participants were recruited by self-selection. Neither scholastic extra credit nor financial incentive was provided for participating in the study. All participants completing the consent form also completed all study interventions. The study details were presented to all participants at the conclusion of a freshman or sophomore RC course. Interested students were directed to an office in the RC department to receive further instructions and sign the consent form. Gender or racial data were not collected for this study. All participants appeared to be in...
good health at the time of the study. Although student health was not assessed specifically for this study, students enrolled in the RC program completed a health screen to include a physical exam and immunizations.

Interventions
Once consent was provided, participants were escorted to a department laboratory to complete the three portions of the research. Scores on all portions were anonymous to all participants. Study participants were asked to complete both the BSQ and the ESS as truthfully as possible. Upon completion of the questionnaires, each subject was assessed by the same research investigator using the MSS through direct visual examination of the oral cavity by asking the subject to open his/her mouth. Total time to complete all three portions ranged from 5 to 10 minutes. Subject responses to the ESS and BSQ questionnaires were carefully transferred to an Excel® spreadsheet (Microsoft Corp, Redmond, WA) along with the MSS classification score. All three portions were completed in consecutive order on the same day that consent was provided. The ESS and BSQ have user instructions printed at the top of each questionnaire. An investigator was present at all times to answer questions in the event of confusion.

Sample Size
A priori power calculation demonstrated a sample size of 95 would be required to achieve a power of greater than .8. Therefore, we decided to approach our two largest RC student groups, i.e. the freshmen and sophomore students, in an attempt to meet this sample size.

Objectives
The following hypotheses guided this project.
1. The BSQ score and the MSS score will predict an ESS score.
2. The MSS score will predict a BSQ score.

Statistical Methods
Multiple regression analysis was utilized to determine if the MSS and the BSQ would predict an ESS score. In addition, linear regression was utilized to determine if the MSS, now serving as the independent variable, would predict the BSQ score. MSS scores were entered directly into Excel® without any conversion from the actual participant scores. As stated previously, a MSS score range of 1-4 is possible. The BSQ provides an assessment of sleep apnea risk as either high or low, based on the participant responses. Participant BSQ scores indicating low sleep apnea risk received a score of 1, and BSQ scores indicating high sleep apnea risk received a score of 2. All intervention scores were copied from Excel® into Statistical Package for the Social Sciences (SPSS; SPSS Inc., Chicago, IL) for data analysis.

RESULTS
A total of 77 RC students volunteered to participate in this pilot project. No participants dropped out during the study. Subject ages of this convenient sample ranged between 18 to 46 years of age. Normality of the data was verified by assessing the histogram of the standardized regression residuals following data analysis. All subjects tolerated the Mallampati assessment without incident. No subjects requested to view their ESS or BSQ scores once collected.

Linear regression was used to assess the predictability of the BSQ and the MSS with regards to ESS. Results of the model summary suggest the BSQ score and the MSS score are weak predictors of the ESS (R²=.047, R²adj=.021, F (2,74)=1.825, p>0.05). The BSQ and MSS account for 4.7% of the ESS score. These results indicate that the MSS and BSQ are not strong predictors of the levels of excessive sleepiness experienced by a healthy college student. The standardized beta weights for MSS and BSQ (.114 and .167, respectively) are both positive, indicating that as the score for either predictor increased, so did the ESS.

In addition, linear regression was used to assess the predictability of the MSS with regards to the BSQ. Results suggest the MSS was a non-significant and weak predictor of the BSQ score with regards to predicting the risk of OSA (R²=.026, R²adj=.013, F (1,75)= 2.022, p>0.05). The MSS accounts for 2.6% of the BSQ score when completed by healthy college students. The standardized beta weight for the MSS was .162, once again indicating a positive relationship with the BSQ.

DISCUSSION
Although previous research has shown that all three instruments may independently predict sleep disordered breathing associated with OSA, our study suggests there is no predictability between the instruments. Our research findings suggest that in healthy, young adults without a prior objective diagnosis of sleep disturbances, the use of the MSS is not useful in predicting scores on sleep questionnaires. Although the BSQ and ESS have been shown to be effective tools in screening for sleep...
disorders and patients at risk for OSA, the MSS usefulness has been debatable. Similar research determined that an MSS class of IV was not useful in ruling in patients with severe OSA, and an MSS class I was not useful in ruling out OSA.\textsuperscript{24} Although we did not directly test for OSA, our results show that an MSS class I through IV had no relationship with scores on these effective screening tools. The basic nature of all three evaluations is different, i.e. two are self-reporting and one is based on a visual examination by a clinician. Therefore, it appears that there is very little consistency between all three of these evaluations.

Results of our pilot project may challenge the results provided by Nuckton et al. who showed the MSS to be an independent predictor of the presence and the severity of OSA.\textsuperscript{12} Nuckton found that for every 1 point increase in the Mallampati score, the odds of having OSA increased more than two-fold with a predicted increase of the AHI by more than 5 events per hour.\textsuperscript{12} A similar study confirms that a high Mallampati score represents a predisposing factor for OSA, especially if it is associated with nasal obstruction.\textsuperscript{25} Perhaps a combination of assessment tools, other than the MSS, ESS, and BSQ would be more effective in detecting OSA in healthy, young adults.

While symptoms alone do not directly detect OSA, perhaps an instrument capable of ruling out or confirming OSA based on clinical features and characteristics would be of benefit.\textsuperscript{26} Rowley et al. discovered that the use of clinical prediction formulas in conjunction with a combination of clinical features was effective in detecting an AHI of \( \geq 10 \).\textsuperscript{27} Although the ESS, BSQ, and MSS have been shown to have predictive value independently, a study comparing all three assessment instruments could not be found in the literature. A single instrument that accounts for excessive daytime sleepiness, as identified by the ESS, snoring and high-risk characteristics, as recognized by the BSQ, and anatomical features of the oropharynx, as delineated by the MSS, does not yet exist. The development of such an instrument has the potential for providing broad base screening of individuals with a complaint of sleep disturbance and excessive tiredness prior to, or in lieu of, a complete polysomnography sleep study.

There are three primary limitations associated with this study. First, the use of college students may have caused the ESS scores to be higher compared to individuals of the same age not attending college. Our results cannot include the same age population of people outside of college due to the sleep habits and lifestyle associated with college life and studies. Second, the sample size did not reach the power calculation and third, confirmed OSA patients were not included in the study for comparison.

CONCLUSION
In the current standard-of-care practice, the BSQ, ESS, and MSS instruments are seldom used in combination when OSA screening is being conducted. Although each instrument is predictive of OSA or sleep disturbance issues, the process of arriving at OSA prediction is different for each. Our findings indicate that self-reporting questionnaire results are not consistent with results performed by a clinical evaluator when evaluating healthy young adults. Future research will need to investigate the consistency of a patient’s evaluation of self with a clinician’s assessment supported with the patient’s positive or negative diagnosis of OSA. Researchers involved in this pilot study realize that there are many factors that may lead to sleep disordered breathing and all factors are not covered by the screening tools examined in this research. This study suggests that the BSQ, ESS, and MSS are not significant predictors of each other in healthy young adults.

REFERENCES

© The Internet Journal of Allied Health Sciences and Practice, 2011


APPENDIX A
Berlin Questionnaire

Level of risk is based on responses in symptom categories. In categories 1 and 2, high risk is characterized by continuing symptoms (more than three or four times a week) in two or more questions on snoring and on sleepiness during wake time or driving, respectively. For category 3, high blood pressure or body mass index >30 are considered high risk. Patients with high-risk features in any two of the three categories warrant referral to sleep clinic.

Height (m)______ Weight_______ Age ______ Male/Female________

CATEGORY 1
1. Do you snore?
   _ Yes _ No _Don’t Know

   If you snore:

   2. Your snoring is:
      _ Slightly louder than breathing _As loud as talking _ Louder than talking
      _ Very loud, can be heard in adjacent rooms

   3. How often do you snore
      _ Nearly every day _ 3-4 times a week _ 1-2 times a week
      _ 1-2 times a month _ Never or nearly never

   4. Has your snoring ever bothered other people?
      _ Yes _ No _ Don’t Know

   5. Has anyone noticed that you stop breathing during your sleep?
      _ Nearly every day _ 3-4 times a week _ 1-2 times a week _ 1-2 times a month
      _ Never or nearly never

   CATEGORY 2
6. How often do you feel tired or fatigued after your sleep?
   _ Nearly every day _ 3-4 times a week _ 1-2 times a week _ 1-2 times a month
   _ Never or nearly never

   7. During your waking time, do you feel tired, fatigued or not up to par?
   _ Nearly every day _ 3-4 times a week _ 1-2 times a week _ 1-2 times a month
   _ Never or nearly never

   8. Have you ever nodded off or fallen asleep while driving a vehicle?
   _ Yes _ No

   If yes:

   9. How often does this occur?
      _ Nearly every day _ 3-4 times a week _ 1-2 times a week _ 1-2 times a month
      _ Never or nearly never

   CATEGORY 3
10. Do you have high blood pressure?
    _ Yes _ No _ Don’t know

Adapted from Table 2 in Netzer et al with permission from the American College of Physicians
APPENDIX B

Epworth Sleepiness Scale

Name: ____________________________ Today’s date: ________________

Your age (Yrs): ____________ Your sex (Male = M, Female = F): ____________

How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired?

This refers to your usual way of life in recent times.

Even if you haven’t done some of these things recently try to work out how they would have affected you.

Use the following scale to choose the most appropriate number for each situation:

0 = would never doze
1 = slight chance of dozing
2 = moderate chance of dozing
3 = high chance of dozing

It is important that you answer each question as best you can.

<table>
<thead>
<tr>
<th>Situation</th>
<th>Chance of Dozing (0-3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting and reading</td>
<td></td>
</tr>
<tr>
<td>Watching TV</td>
<td></td>
</tr>
<tr>
<td>Sitting, inactive in a public place (e.g. a theatre or a meeting)</td>
<td></td>
</tr>
<tr>
<td>As a passenger in a car for an hour without a break</td>
<td></td>
</tr>
<tr>
<td>Lying down to rest in the afternoon when circumstances permit</td>
<td></td>
</tr>
<tr>
<td>Sitting and talking to someone</td>
<td></td>
</tr>
<tr>
<td>Sitting quietly after a lunch without alcohol</td>
<td></td>
</tr>
<tr>
<td>In a car, while stopped for a few minutes in the traffic</td>
<td></td>
</tr>
</tbody>
</table>

THANK YOU FOR YOUR COOPERATION

© M.W. Johns 1990-97