

# Interobserver agreement on DISE: Do we all see the same?

Nuno O'Neill Mendes • Joselina Antunes • João Rito • Gustavo Pedrosa • Cristina Adónis • Ana Guimarães • Filipe Freire

## ABSTRACT

**Objectives:** Evaluate interobserver agreement on DISE classification and consequent therapeutic decision-making.

**Study design:** Cross-sectional.

**Material & methods:** 20 DISEs performed in adults at one Hospital were recorded. Videos were reviewed by six observers (3 senior Otolaryngologists and 3 residents). The observers rated the exam findings and suggested the best treatment for each case. Kappa coefficient(K) was used to evaluate interobserver agreement.

**Results and Conclusions:** Airway obstruction was assessed by seniors and residents and, in this parameter, there was 100% agreement between groups at the level of soft palate(K=1), 90% at oropharynx(K=0,69), 85% at tongue base(K=0,68) and 85% at epiglottis(K=0,69). Concerning to the severity of the obstruction, interobserver agreement was moderate to high at every level except at tongue base(K=0,38). With regard to the treatment, we found a higher rate of indication for tongue base surgery from residents(7/20) comparing to specialists(1/20);  $k=0,34$  ( $p=0,04$ ).

Consequently, agreement between residents and seniors seems to be quite acceptable except on grading severity of obstruction at the level of tongue base (weak agreement) and therapeutic decision on the same anatomical area (weak agreement).

**Keywords:** Drug Induced Sleep Endoscopy; Sleep Apnea; Tongue base; Pharynx

### Nuno O'Neill Mendes

Hospital Professor Doutor Fernando Fonseca, Portugal

### Joselina Antunes

Hospital Professor Doutor Fernando Fonseca, Portugal

### João Rito

Hospital Professor Doutor Fernando Fonseca, Portugal

### Gustavo Pedrosa

Hospital Professor Doutor Fernando Fonseca, Portugal

### Cristina Adónis

Hospital Professor Doutor Fernando Fonseca, Portugal

### Ana Guimarães

Hospital Professor Doutor Fernando Fonseca, Portugal

### Filipe Freire

Hospital Professor Doutor Fernando Fonseca, Portugal

### Correspondence:

Nuno O'Neill Mendes  
nunomendes@campus.ul.pt

Article received on September 9, 2021. Accepted for publication on September 25, 2021.

## INTRODUCTION

Obstructive sleep apnea-hypopnea syndrome (OSAHS) is a sleep-related breathing disorder that is more common among adults, occurring at a frequency of 2–4%<sup>(1)</sup>. It is diagnosed using nocturnal polysomnography (PSG). The first-line treatment for moderate-to-severe apnea is continuous positive airway pressure (CPAP); however, its efficacy is often limited by difficulties with adaptation and poor patient compliance<sup>(2);(3);(4);(5)</sup>. CPAP can also be applied to treat mild OSAHS associated with comorbidities such as high blood pressure, coronary disease, cerebrovascular disease, or symptoms that affect a patient's daily life, such as daytime sleepiness, reduced cognitive ability, and mood changes<sup>(4);(5)</sup>.

Other types of treatment such as surgery, mandibular advancement devices (MADs), and hypoglossal nerve stimulators require meticulous evaluation for potential upper airway (UA) collapse. Drug-induced sleep endoscopy (DISE), has become an established method of assessment since it emerged in 1991<sup>(6)</sup>. It is particularly useful because it allows the topographic diagnosis of the obstruction through the dynamic visualization of the UA<sup>(1)</sup>. However it is contraindicated for pregnant women, individuals with an American Society of Anesthesiology score of 4, drug allergies (absolute contraindications) or morbid obesity (relative contraindication). The examination can proceed in an operating room or an examination room with resuscitation equipment in case of emergency<sup>(1)</sup>. In addition to the need for a flexible nasopharyngoscope, a medical team (anesthetist and otorhinolaryngologist), and general anesthesia monitoring (SpO<sub>2</sub>, ECG, blood pressure), sites where patients undergo DISE should have a bispectral index (BIS) monitor and an infusion pump for targeted controlled infusion (TCI)<sup>(7)</sup>.

The subjectivity of DISE should not be ignored; although several classifications have been validated, observer experience may play a critical role in the reliability of the results<sup>(8);(9);(10)</sup>.

Thus, the present study aimed to determine interobserver agreement regarding DISE classifications and treatment choices based on the findings.

## MATERIAL AND METHODS

Twenty adults with OSAHS at the Prof. Doutor Fernando Fonseca Hospital underwent DISE during the first semester of 2019. They were diagnosed with OSAHS using level 2 PSG at home. Patients who refused and/



or did not adapt to CPAP were included. The exclusion criteria comprised mean apnea-hypopnea index (AHI) > 30 (severe OSAHS is usually managed by pulmonology and occasionally surgery to improve adaptation to CPAP) and previous surgery to treat sleep apnea, which can alter the anatomical configuration of the UA and skew the DISE classification. Sociodemographic, biometric, anatomic, and polysomnographic data were retrieved from medical records.

All DISEs proceeded in the OR on an outpatient basis. The patients were sedated using a target-controlled infusion (TCI) of propofol. Sedation depth was monitored using the bispectral index (BIS) and the reference interval was 50–70<sup>(4)</sup>. The patients were initially placed in the supine position then in the position described in their PSG results as the position that generated the most respiratory events (other sleeping positions were not evaluated). All patients underwent the Esmarch and the chin lift maneuvers. Local/nasal anesthetics or vasoconstrictors were not applied. At least two respiratory cycles were observed at each airway level in all examinations (one respiratory cycle was defined according to the European consensus on DISE<sup>(1)</sup> as snoring, desaturation with obstructive hypopnea/apnea, then breathing.

Three specialists as well as one and two residents in their third and fourth years of specific otorhinolaryngology training, respectively assessed the images. The only clinical information given to the observers before the classification of DISE was the AHI value for each patient. The severity of apnea can aid physicians in the choice of treatment in terms of surgical extension and numbers of anatomical levels subjected to intervention. No other clinical information was provided. The observers blindly and independently categorized the DISE findings according to the modified VOTE classification (Table 1)<sup>(9)</sup> in which the configuration of the obstruction can be classified at the level of the base of the tongue as lateral or concentric (in addition to the anterior-posterior configuration in the classic VOTE classification)<sup>(9)</sup>. After classifying the DISE findings, the observers suggested a specific treatment for each patient from the following options at different anatomical levels: barbed repositioning pharyngoplasty

(RP), expansion pharyngoplasty (EP)<sup>(11)</sup>, Pang anterior palatoplasty (AP), and radiofrequency (RFP) for the palate, bilateral extracapsular tonsillectomy (A) for the oropharynx, radiofrequency (RFBT), coblation (CBL), mandibular advancement device (MAD) for base of the tongue, and partial epiglottectomy (PE) for the epiglottis. The following parameters were compared between the observer groups: presence or absence of an obstruction at different levels of the UA, and the degree and configuration of the obstruction. Agreement regarding treatment options between the groups was classified evaluated using the Cohen kappa coefficient ( $\kappa$ )<sup>(12)</sup> as null (0), low ( $\leq 0.01$ –0.2), weak (0.21–0.4), moderate (0.41–0.6), substantial (0.61–0.8), and high/strong (0.81–1.0). Data were statistically analyzed using SPSS 25 (IBM Corp., Armonk, NY, USA). Values with  $p < 0.05$  were considered statistically significant.

## RESULTS

### Characterization of the sample

Twenty consecutive patients (, 65% men; mean age, 52.1  $\pm$  13.4 [18–74] years) were assessed by endoscopy. The AHI of the population determined by PSG before DISE was 16.4  $\pm$  8.8 (5.5–29.5). The mean body mass index (BMI) was 25.7  $\pm$  2.7 (21–29) kg/m<sup>2</sup>.

### Findings of obstruction of any degree in the upper airway

Agreement between the two groups regarding the presence of airway obstruction (any degree) was 100% ( $\kappa = 1$ ) at the level of the palate and substantial at those of the oropharynx (90%,  $\kappa = 0.69$ ), base of the tongue (85%,  $\kappa = 0.68$ ), and epiglottis (85%,  $\kappa = 0.69$ ).

### Severity of obstruction according to anatomical level

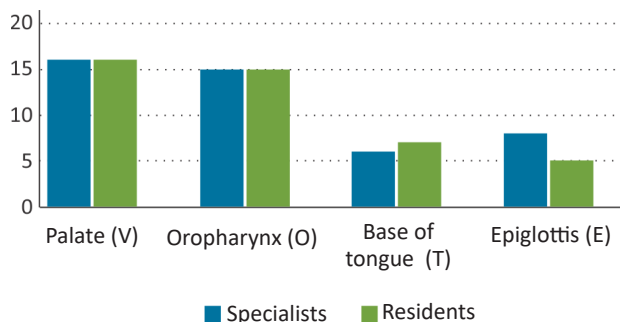
Agreement regarding the severity of the obstruction was substantial between the two groups at all anatomical levels except for that at the base of the tongue (weak agreement;  $\kappa = 0.38$ ; Table 2). In this context, the less-experienced observers (residents) classified the obstruction at the level of the tongue base as complete more frequently than did the specialists (6:1).

**TABLE 1**  
Modified VOTE classification

Structure	Degree of Obstruction	Configuration		
		Anterior-posterior	Lateral	Concentric
Palate (V)	0: None 1: Partial (vibration) 2: Complete (collapse) X: Not visualized			
Oropharynx (O)				
Base of tongue (T)				
Epiglottis (E)				

**GRAPH 1**

Presence of obstruction of any degree according to anatomical level (xx,: anatomical areas; yy, absolute number of individuals with obstruction of any degree)



### Configuration of the obstruction according to the anatomical level

Agreement between the groups regarding the configuration of the obstruction varied between moderate and strong (Table 3). According to the modified VOTE classification, a potential obstruction at the oropharyngeal level should always be classified as lateral. Therefore, the agreement regarding the classification of configuration at this level cannot be evaluated.

### Treatment

The observers proposed treatments that they deemed adequate for each patient, as described in the Material and Methods section. At least one type of treatment was proposed for all patients (Table 4).

## DISCUSSION

Interobserver variability should be considered whenever the reliability of the results of observer-dependent examinations such as DISE are assessed. Because the results of DISE can be applied to the choice of treatment targeted at the type of an obstruction affecting individuals, agreement between such choices should also be assessed<sup>(8);(9)</sup>.

Overall, the agreement obtained between the groups in the present study was adequate, except for that related to the base of the tongue. The weak agreement regarding the classification of the obstruction severity at this level might have led the residents to suggest basilingual surgery more often than it was suggested by the group of specialists.

A previous study of interobserver variability of DISE among 97 observers<sup>(10)</sup> used six videos of DISE and Cohen  $\kappa$ . We emphasize that although seven of these 97 observers had little experience with DISE, the entire group was not divided according to their differing experience levels. Nevertheless, the authors of that study identified the base of the tongue and the palate as the anatomical areas of substantial agreement. The treatment options of the observers were not evaluated in that study. A comparison of their results with those of the current study reveals differences at the level of the base of the tongue (weak vs. substantial agreement). The fact that the observers were not grouped according to experience in the previous study prevented a reliable comparison with the present study's findings. Therefore, we considered the observed differences as unreliable.

**TABLE 2**

Severity of obstruction according to anatomical levels in patients (n)

Anatomical level	Palate (n)		Oropharynx (n)		Base of tongue (n)		Epiglottis (n)	
	Partial (nº of patients)	Total (nº of patients)	Partial (nº of patients)	Total (nº of patients)	Partial (nº of patients)	Total (nº of patients)	Partial (nº of patients)	Total (nº of patients)
Specialists	6	11	6	10	6	1	4	5
Residents	4	13	6	8	2	6	4	2
Agreement ( $\kappa$ )	Substantial (0,64)		Substantial (0,68)		Weak (0,38)		Substantial (0,64)	

**TABLE 3**

Configuration of obstruction according to anatomical levels

Anatomical level	Palate (n)			Oropharynx (n)	Base of tongue (n)			Epiglottis (n)	
	Lateral (nº of patients)	AP (nº of patients)	Concentric (nº of patients)		Lateral (nº of patients)	Lateral (nº of patients)	AP (nº of patients)	Concentric (nº of patients)	Lateral (nº of patients)
Specialists	-	3	14	16	-	5	-	1	8
Residents	-	4	13	16	2	8	-	1	5
Agreement ( $\kappa$ )	Forte (0,83)			N/A	Moderada (0,51)			Substantial (0,61)	

Data are shown as numbers (n) of patients. N/A: Not applicable (when all agreed regarding the finding of obstruction, the latter was always lateral).

**TABLE 4**

Choice of treatment

	Palate	Oropharynx	Base of tongue	Epiglottis	Multilevel surgery
<b>Specialists</b>	17 (16x RP; 1x PA)	16 (A)	2 (CBL; MAD)	4 (PE)	3/20
<b>Residents</b>	17 (13x RP; 4x PA)	14 (A)	7 (6x CBL; 1x RFBT)	2 (PE)	6/20
<b>Agreement (<math>\kappa</math>)</b>	Total (1,00;p<0,001)	Substantial (0,69;p=0,002)	Weak (0,34;p=0,04)	Substantial (0,69;p=0,0019)	

Agreement was calculated relative to the decision to intervene at each anatomical level and to the type of treatment.

Multilevel surgery aims to correct at least two anatomical levels. If surgery of the palate is accompanied by tonsillectomy alone then it is regarded as single level surgery. AP, Anterior palatoplasty; CBL, Coblation of the base of the tongue; MAD, mandibular advancement device; PE, Partial epiglottectomy; RFBT, Radiofrequency of the base of the tongue; RP, repositioning pharyngoplasty.

Another study evaluated the agreement regarding the classification of DISE and categorized observers according to their levels of experience<sup>(8)</sup>. That study included one experienced observer (a specialist in otorhinolaryngology) and one inexperienced observer (a resident with specific training in otorhinolaryngology). They classified 31 DISEs and the agreement between them was assessed using Cohen  $\kappa$ . Agreement regarding the severity and configuration of the obstruction was substantial, except for that at the base of the tongue, which was weak because the resident undervalued the basilingual obstruction. Agreement in terms of treatment options was adequate for all levels except for that at the basilingual level, which was weak because the classification of the obstruction at this level differed. A comparison of these and the present results showed that the base of the tongue in both studies was the site of disagreement between residents and specialists in terms of obstruction classification and proposed treatments. Despite the similar results, the above study included only two observers, which hindered a realistic comparison with the present study of agreement among six observers. The most recent literature suggests that DISE is mainly relevant when surgery of the base of the tongue or MAD are considered<sup>(10)</sup>. Thus, because the agreement among observers with different levels of experience was the weakest in the basilingual region, this anatomical level should always be assessed by an experienced specialist. This study had some limitations, including the small sample size (20 individuals). Obstructions were classified based on videos, which did not allow the observers to consider the position of the patients at each moment. All evaluators worked in the same hospital, where the learning methods and style were similar, which could partly explain the agreement. The results might have differed had the study included physicians from different hospitals. Furthermore, our evaluators did not analyze factors other than DISE findings, which should be considered when selecting a treatment. Factors such as age, BMI, and retrognathia/prognathism are important to consider before proposing a specific treatment to a patient. The most important factor is that the otorhinolaryngologist must understand the motivation

and potential compliance of patients towards their suggested therapeutic solutions.

### CONCLUSION

Although DISE is overall a safe and reliable procedure, the experience of the observer is a determinant in the classification of the examination findings and in the choice of treatment, especially at the level of the base of the tongue.

### Acknowledgements

We thank the Otorhinolaryngology Department of the Professor Doutor Fernando Fonseca Hospital, in particular Dr. Filipe Freire and Dr. Ana Guimarães.

### Conflict of Interest

The authors declare no conflict of interest regarding this article.

### Data confidentiality

The authors declare having followed the protocols in use at their workplace regarding the publication of patients' data

### Human and animal protection

The authors declare that the procedures complied with the regulations established by the Ethics and Clinical Research Committee and according to the Declaration of Helsinki (2003; World Medical Association).

### Privacy policy, informed consent and approval by the ethics committee

The authors declare that the patients provided written, informed consent for the use of their photographs in this article.

### Funding

The present study was conducted in the absence of any financial contribution, funding, or grant.

### Availability of scientific data

No datasets related to this study are publicly available.

## References

1. De Vito A, Carrasco Llatas M, Vanni A, Bosi M, Braghiroli A, Campanini A. et al. European position paper on drug-induced sedation endoscopy (DISE). *Sleep Breath*. 2014 Sep;18(3):453-65. doi: 10.1007/s11325-014-0989-6.
2. Pieters T, Collard P, Aubert G, Dury M, Delguste P, Rodenstein DO. Acceptance and long-term compliance with nCPAP in patients with obstructive sleep apnoea syndrome. *Eur Respir J*. 1996 May;9(5):939-44. doi: 10.1183/09031936.96.09050939.
3. Grote L, Hedner J, Grunstein R, Kraiczi H. Therapy with nCPAP: incomplete elimination of sleep related breathing disorder. *Eur Respir J*. 2000 Nov;16(5):921-7. doi: 10.1183/09031936.00.16592100.
4. Epstein L, Kristo D, Strollo P, Friedman N, Malhotra A, Patil S. et al. Clinical guideline for the evaluation, management, and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med*. 2009 Jun 15;5(3):263-76.
5. Coman AC, Todea DA, Popa E, Radu T, Cadar O, Borzan C. Multilateral characterization of masks and tubes surfaces in contact with respiratory system through ventilation. *Journal of Optoelectronics and Advanced Materials*. 2015; 17(9-10): 1563-71.
6. Croft CB, Pringle M. Sleep nasendoscopy: a technique of assessment in snoring and obstructive sleep apnoea. *Clin Otolaryngol Allied Sci*. 1991 Oct;16(5):504-9. doi: 10.1111/j.1365-2273.1991.tb01050.x.
7. Babar-Craig H, Rajani NK, Bailey P, Kotecha BT. Validation of sleep nasendoscopy for assessment of snoring with bispectral index monitoring. *Eur Arch Otorhinolaryngol*. 2012 Apr;269(4):1277-9. doi: 10.1007/s00405-011-1798-1.
8. Carrasco-Llatas M, Zerpa-Zerpa V, Dalmau-Galofre J. Reliability of drug-induced sedation endoscopy: interobserver agreement. *Sleep Breath*. 2017 Mar;21(1):173-179. doi: 10.1007/s11325-016-1426-9.
9. Kezirian EJ, White DP, Malhotra A, Ma W, McCulloch CE, Goldberg AN. Interrater reliability of drug-induced sleep endoscopy. *Arch Otolaryngol Head Neck Surg*. 2010 Apr;136(4):393-7. doi: 10.1001/archoto.2010.26.
10. Vroegop AV, Vanderveken OM, Wouters K, Hamans E, Dieltjens M, Michels NR. et al. Observer variation in drug-induced sleep endoscopy: experienced versus nonexperienced ear, nose, and throat surgeons. *Sleep*. 2013 Jun 1;36(6):947-53. doi: 10.5665/sleep.2732.
11. Pang KP, Woodson BT. Expansion sphincter pharyngoplasty: a new technique for the treatment of obstructive sleep apnea. *Otolaryngol Head Neck Surg*. 2007 Jul;137(1):110-4. doi: 10.1016/j.otohns.2007.03.014.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977 Mar;33(1):159-74.