The Egyptian Journal
of Otolaryngology

ORIGINAL ARTICLE

Inter-observer and intra-observer agreement in drug-induced sedation endoscopy — a systematic approach

Kasra Zainali-Gill[1*] , Jannik Buus Bertelsen[1,2], Christian von Buchwald[3], Therese Ovesen[1,2] and Eva Kierkegaard Kiaer[3]

Abstract

Objective: To evaluate the feasibility of systematic observer training in drug-induced sedation endoscopy (DISE) interpretation.

Methods: Fifty DISE videos were randomly selected from a group of 200 videos of cohort of patients with OSA and/or snoring. The videos were assessed blindly and independently using a modified VOTE classification by an expert observer and by two novice observers starting their training. A systematic approach was initiated. Clusters of 10 videos were scored individually by each observer and then re-evaluated as the expert observer clarified the decision-making in plenum. Kappa coefficient (κ) was calculated as a measure of agreement.

Results: The intra-observer variation for the total agreement and kappa values for the expert observer ranged between moderate to substantial agreement in VOTE classification, whereas the novices varied between "less than chance agreement" to "moderate agreement." The inter-observer variation showed increased agreement and kappa values from day 1 to day 2 for both novice observers except at the velum level for observer C. The total agreement and kappa values for each site also improved compared to results of day 1, except at velum for observer C. The velum site seemed to be more difficult to evaluate. The learning curve varied during the study course for each site of the upper airways.

Conclusion: This study shows that systematic intensive training is feasible, although expert level is not acquired after 100 evaluations. The learning curve for the expert observer showed "moderate to substantial agreement" but differed between the trainees.

Level of evidence: 2

Keywords: Sleep apnea, OSA, OSAS, DISE, Systematic education, Surgical treatment of OSA, ENT

Background

Obstructive sleep apnea (OSA) is a disorder associated with excessive daytime sleepiness and cognitive disturbances as lack of concentration and impaired short-term memory seen in both children and adults. The prevalence of OSA is reported as 10–20% for females and males, respectively [1]. The etiology is the repeated collapse of the upper airway during sleep followed by desaturation leading to oxidative stress and subsequent sleep arousal and a reduced quality of sleep. Adults with untreated OSA are more prone to develop hypertension [2], cardiovascular [3] and cerebrovascular [4] incidents, and type-2 diabetes [5].

First-line treatment of moderate to severe OSA is still considered to be continuous positive airway pressure

*Correspondence: kasrzain@rm.dk
[1] Department of Otorhinolaryngology, University Clinic for Flavour, Balance and Sleep, Hospitalsenheden Vest, 7500 Holstebro, Denmark
Full list of author information is available at the end of the article

(CPAP) treatment, this being defined as the golden standard of treatment. However, adherence to CPAP is reported as low as 30–60% [6]. Subsequently, a substantial proportion of these patients seek different and alternative treatments including upper airway surgery.

Surgical treatment in OSA patients remains controversial with inconsistent outcomes, especially when it comes to long-term results [7]. Nonetheless, the current surgical procedures vary from simple tonsillectomy to multilevel surgery of the upper airway and implantation of upper airway stimulating devices which further emphasizes the need to thoroughly evaluate and plan the surgical procedure to improve the sleep quality.

Drug-induced sedation endoscopy (DISE) allows an overview of the upper airway during a pharmacologically induced simulated sleep. DISE visualizes the specific anatomical site of airway collapse. The technique originally described in 1991 by Croft and Pringle is considered a guiding tool in making treatment decisions [8]. It has been validated by multiple authors and is considered a simple and safe technique [9–11]. Although a classification of upper airway collapse is standardized in validated classifications such as the VOTE classification [12], the decision is also based upon subjectivity and upon observer experience.

This study was designed as a preliminary step towards the establishment of a surgical unit within our ENT department, which in the future should manage OSA patients in a secure and evidence-based manner. As DISE is considered a guiding tool in treatment decisions, this study was the first step in the establishment of sleep surgery. The use of educational sessions, hands-on sessions, and video reviews are widely used in different health workers education, such as emergency doctors and the use of emergency ultrasound [13].

The present study was conducted to prepare two non-experienced ENT doctors for DISE evaluations. The aim was to evaluate the feasibility of systematic training (theoretical education, hands-on video evaluation, and in-plenum discussion) in DISE interpretation assessed by learning curve and inter-rater variability between an experienced surgeon and a novice trainee. The secondary aim was to assess the intra-observer variability.

Methods

The study was conducted in accordance with the ethical standards in the 1964 Declaration of Helsinki. Since only anonymous videos were evaluated, ethic approval for this specific learning study was not necessary. Videos were extracted from a patient database for a previous study [10]. All patients had given orally and written consent to participate in the database. The database is approved by the local scientific board and licensed by the Danish Data Protection Agency. The 50 videos were all anonymous and stripped for patient data. Fifty DISE videos from the archive of the Danish Center of Sleep Surgery, Department of Head and Neck Surgery and Audiology, Copenhagen University Hospital (Rigshospitalet), Copenhagen, Denmark, were randomly selected by an expert observer. The archive consists of patients diagnosed with obstructive sleep apnea or with severe snoring, who were seeking alternatives to CPAP. The 50 videos were randomly selected from a group of 200 videos. The patients were 20–76 years old; mean age of 44 (SD 11.8) and 82% was male. Maximum BMI was 35 kg/m^2, and median BMI was 27 kg/m^2. The distribution by OSA class was 34% with mild, 27% with moderate, and 31% with severe OSA. Only 8% was severe snorers without OSA (AHI < 5).

All DISE examinations were performed by the expert observer. Expert status was attributed due to the investigator's yearlong DISE experience and status as sleep surgeon. All data were recorded in a database. The patient database consists of all patients undergoing DISE in Copenhagen. Since only anonymous videos were evaluated, ethic approval was not needed. DISE was performed in an outpatient setting at the hospital, and all patients went home within an hour after the examination. Sedation was performed with propofol as the single sedative agent in accordance with NAPS guidelines (nurse-administered propofol sedation) and the methodology described by Kiaer et al. [10].

The DISE videos were assessed blindly and independently using a modified VOTE classification [12]. The modified VOTE classification (Table 1) evaluates the different structures of the upper airway that can contribute to a collapse (velum, oropharynx lateral wall, tongue base, and epiglottis). The classification proposes three degrees of severity: 0 = no collapse (0–25%), 1 = partial obstruction (25–75%) or vibration, and 2 = complete obstruction or collapse (75–100%). The classification also classifies the configuration of the collapse as either anteroposterior, lateral, or concentric. In this

Table 1 Modified VOTE classification

Structure	Degree of obstruction (0–2)	Configuration		
		Anterior-posterior	Lateral	Concentric
Velum (soft palate)				
Oropharynx lateral walls				
Tongue base				
Epiglottis				

0, no obstruction; 1, partial obstruction (vibration); 2, complete obstruction

modified version of the VOTE classification, we consider all configurations possible at all levels.

Design

The videos were evaluated by an expert observer (A) and by two novice observers (B and C) starting DISE training (one ear-nose and throat (ENT) specialist and one ENT resident). At first, the two inexperienced observers were given 60 min of theoretical introduction on how to assess the DISE videos using the modified VOTE classification [12]. Then, three videos were assessed and discussed in community, and hence, the basic introduction was completed.

The DISE videos were evaluated blindly and independently in clusters of 10 videos by all three observers. After each block, the 10 videos were re-evaluated in the group, and the expert clarified the decision-making. However, it was not possible to change the initial evaluations. The configuration of the collapse was also described by each investigator; however, these data proved difficult to interpret statistically. Hence, they are not reported.

The fifty DISE videos were completed in five clusters. The next day, the same 50 DISE videos were assessed again in clusters of 10 videos but in a completely different order determined by computer list randomization (www.random.org, IP: 128.0.73.13, timestamp: 2019-05-20 11:12:34 UTC). The total course was completed in two consecutive days.

Statistical analysis

The statistical software STATA was used for the data analysis. The inter-observer and intra-observer agreement were expressed in percentage. Kappa coefficient (κ) and confidence intervals of 95% as a measure of agreement were calculated to define the level of agreement on the scale proposed by Landis and Koch [14]: < 0 less than chance agreement, 0.01−0.20 slight agreement, 0.21−0.40 fair agreement, 0.41−0.60 moderate agreement, 0.61−0.80 substantial agreement, and 0.81−0.99 almost perfect agreement. The level of significance was 0.05 (Table 2).

Table 2 Kappa interpretation

Kappa	Agreement
< 0	Less than chance agreement
0.01−0.20	Slight agreement
0.21−0.40	Fair agreement
0.41−0.60	Moderate agreement
0.61−0.80	Substantial agreement
0.81−0.99	Almost perfect agreement

The inter-observer agreement was measured using a proportion of the agreement, named observed agreement and Cohen's k for each of the variables V, O, T, and E. Bootstrap method was used to calculate the percentile-based confidence interval for the k.

The agreements were calculated for the day 1 and day 2 observations separately and for the combined data. Learning curve of the agreement was also calculated as the observers measured the outcome in a batch of 10 patients, both at day 1 and day 2. The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Results

A total of 98 videos were evaluated (2 videos were excluded on day 2 because of technical difficulties). Hence, data collection consists of 48 unique DISE examinations classified twice using the modified VOTE score system in two consecutive days. In this study, only the degree of collapse (0, 1, or 2) was reported.

Intra-observer variation

The intra-observer variation between day 1 and day 2 is depicted in Table 3. The total agreement and *k*-values for each upper airway collapse (UAC) site were compared between the 2 days.

The expert observer A had an observed agreement (OA) with *k* = values from 0.49 to 0.62 equal to "moderate agreement" and "substantial agreement" as shown in Table 2. The novice observer B had *k*-values between 0.46 and 0.60 equal to "moderate agreement." The k coefficient

Table 3 Intra-observer variation day 1 versus day 2

UAC site	Observed agreement	Kappa	95% confidence interval	
			Lower limit	Upper limit
Observer A				
V	70.83%	0.49	0.27	0.70
O	77.08%	0.53	0.35	0.71
T	81.25%	0.49	0.49	0.86
E	75.00%	0.62	0.43	0.79
Observer B				
V	72.92%	0.47	0.22	0.69
O	79.17%	0.60	0.38	0.79
T	77.08%	0.60	0.39	0.80
E	64.58%	0.46	0.26	0.66
Observer C				
V	27.08%	0.00	−0.16	0.16
O	64.58%	0.33	0.11	0.55
T	60.42%	0.40	0.19	0.59
E	62.50%	0.37	0.16	0.60

Zainali-Gill *et al. The Egyptian Journal of Otolaryngology*   (2022) 38:52

Page 4 of 9

of the novice observer C was 0.00 at the V site and 0.33 to 0.40 in the other sites corresponding to "fair agreement."

Inter-observer variation

The inter-observer agreement is depicted in Table 4. Observer A was compared to novice observer B and to novice observer C. The OA and *k*-values increased from day 1 to day 2 for both novice observers except the velum for observer C. The total agreement and *k*-values for each UAC site also improved compared to results of day 1, except at velum for observer C. Overall, the velum site seemed to be more difficult to evaluate than the oropharynx lateral wall, tongue base, and epiglottis. The highest OA was for oropharynx lateral wall collapse with $k = 0.76$ (A vs B) and $k = 0.42$ (A vs C). The OA for the tongue

Table 4 Inter-observer variation day 1 versus day 2

UAC site	Observed agreement	Kappa	95% confidence interval	
			Lower limit	Upper limit
Observer A versus observer B day 1				
V	60.00%	0.27	0.06	0.50
O	80.00%	0.63	0.44	0.81
T	74.00%	0.55	0.33	0.73
E	66.00%	0.46	0.25	0.66
Observer A versus observer B day 2				
V	77.08%	0.57	0.35	0.78
O	95.83%	0.91	0.77	1.00
T	79.17%	0.64	0.39	0.82
E	87.50%	0.81	0.65	0.94
Observer A versus observer B total				
V	68.38%	0.41	0.22	0.58
O	87.76%	0.76	0.63	0.86
T	76.53%	0.60	0.42	0.74
E	76.53%	0.64	0.49	0.77
Observer A versus observer C day 1				
V	50.00%	0.20	0.01	0.43
O	64.00%	0.34	0.10	0.55
T	58.00%	0.35	0.14	0.53
E	64.00%	0.43	0.22	0.61
Observer A versus observer C day 2				
V	14.58%	−0.10	−0.23	0.03
O	77.08%	0.52	0.29	0.72
T	72.92%	0.55	0.33	0.76
E	64.58%	0.46	0.26	0.66
Observer A versus observer C total				
V	32.66%	0.02	0.21	0.58
O	70.41%	0.42	0.64	0.86
T	65.31%	0.45	0.41	0.74
E	64.29%	0.44	0.49	0.77

base and the epiglottis UAC were similar, but for velum, *k* was 0.41 (A vs B) compared to 0.02 (A vs C). Nonetheless, all parameters except velum score for novice C improved in day 2 suggesting an improvement in DISE interpretation.

Learning curve

Each UAC parameter with *k*-values for the variation between novice observer and the experienced observer after each cluster of ten examinations is shown as Figs. 1, 2, 3, and 4. During the 2 days of observation, the learning curves changed for each structure of the upper airway. For observer B, an increase in *k*-values was seen for the velum and epiglottis, whereas observer C was challenged in the interpretation of velum. A stable increase in k coefficients is depicted for oropharynx for both novices and for tongue base on day 2. However, a large variation between the clusters was observed.

Discussion

This study showed that the systematic approach with an initial theoretical insight in the evaluation of DISE examinations and the continuous debriefing of the previous 10 videos led to an overall increase in agreement between the expert observer and the novices in 2 days. This study also highlights the variations in the learning curve for the novice observers on day 1 (batch ID 1–5) and day 2 (batch ID 6–10). Furthermore, the learning curve suggests that observer B and observer C have focused on different UAC sites at different times, and a large amount of evaluations are needed before one is familiar with DISE interpretation. The intra-observer agreement was highest for the expert observer, and the inter-observer agreement also showed an increase in agreement except on the level of velum.

Intra-observer variation

Vroegop et al. investigated the intra-observer variation and the inter-observer variation in a large-scale study [15]. Due to their study design, 52 non-experienced surgeons and five experienced surgeons completed the intra-observer test, whereas inter-observer agreement was investigated in a group of seven experienced observers and 90 inexperienced ENT surgeons [15]. Their intra-observer data suggested fair to almost perfect agreement in the experienced group at the different UAC sites when comparing day 1 and day 2, whereas the intra-observer k coefficients for the non-experienced group differed from no agreement to moderate agreement at the different UAC sites. In the non-experienced group, they reported palatal OA = 0.92 ($k = −0.01$), oropharynx OA = 0.62 ($k = 0.07$), tongue base OA = 0.74 ($k = 0.38$), hypopharynx OA = 0.56 ($k = 0.38$), and epiglottis OA = 0.76 ($k = 0.52$). We found better intra-observer
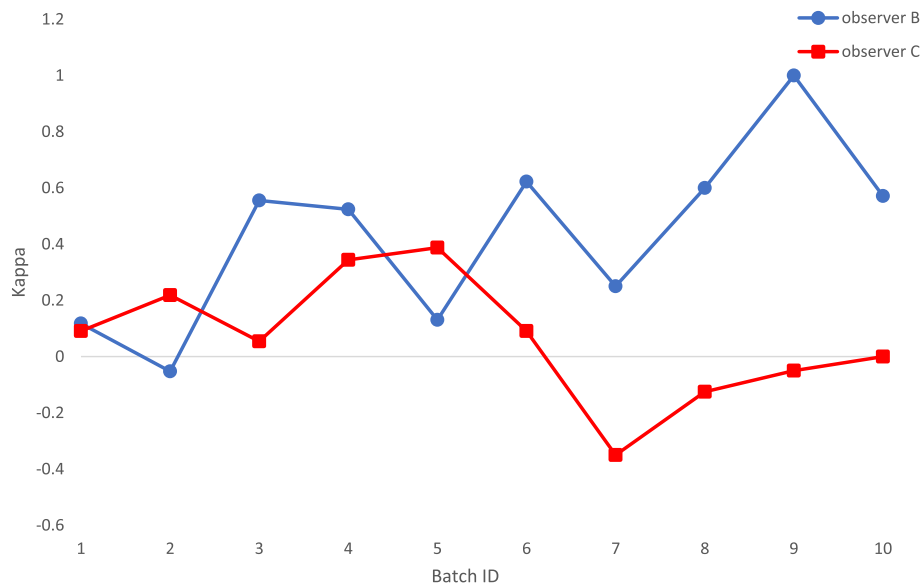
Fig. 1 Kappa values for the variation between the novice observers and the experienced observer on the velum level after each cluster of 10 examinations
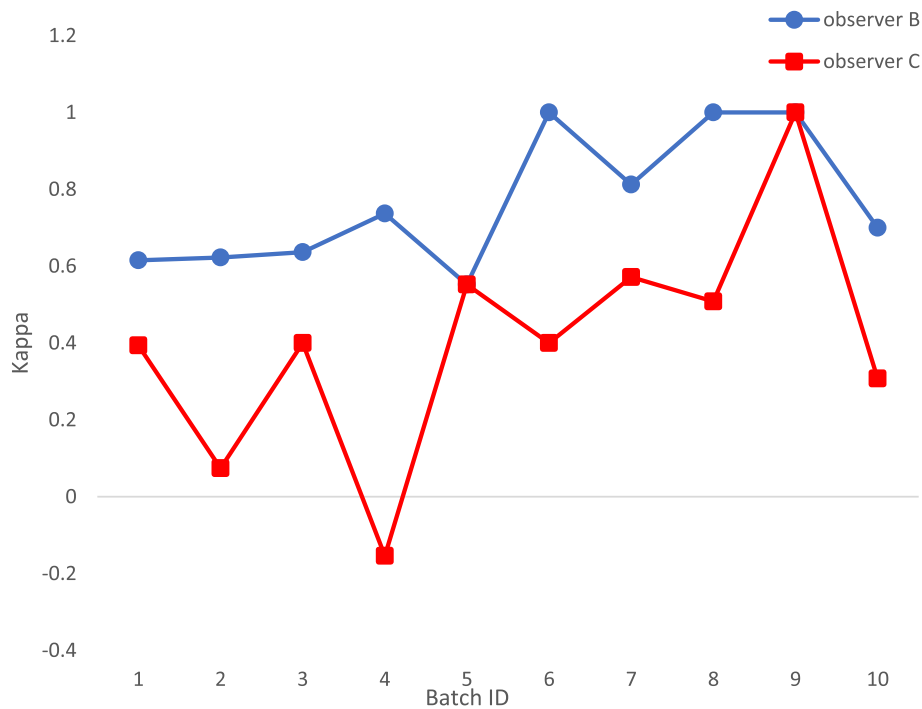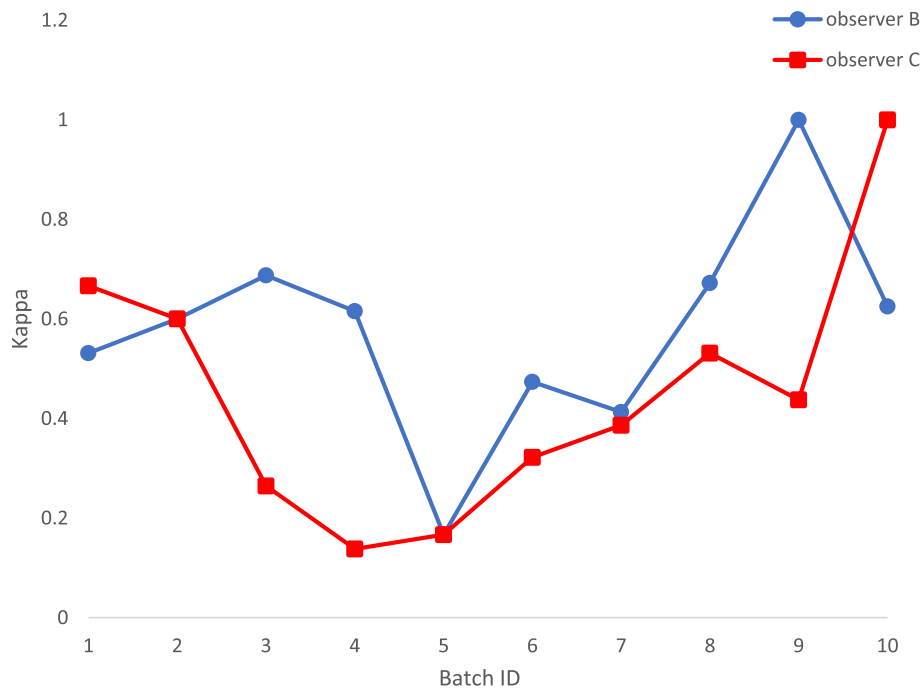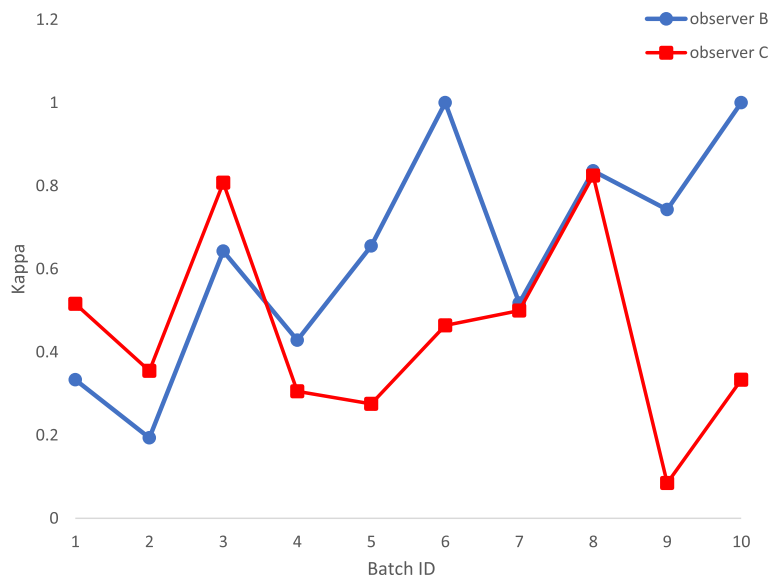

Fig. 2 Kappa values for the variation between the novice observers and the experienced observer on the oropharynx level after each cluster of 10 examinations

agreements. This again proposes benefits from the systematic theoretic teaching of DISE video interpretation. Vroegop et al. also concluded that experience in performing DISE is necessary to obtain reliable observations, and training under guidance of an experienced surgeon might be helpful for those inexperienced with DISE.

Fig. 3 Kappa values for the variation between the novice observers and the experienced observer on the tongue base level after each cluster of 10 examinations

Fig. 4 Kappa values for the variation between the novice observers and the experienced observer on the epiglottis level after each cluster of 10 examinations

This allows for an increase in the learning curve for the non-experienced reviewer. The intra-observer agreement between day 1 and day 2 was highest for the expert observer with agreement *k*-values stating moderate to substantial agreement. With the intentional evolution and improvement of evaluation of DISE examinations in novice observers, the intra-observer variation should imply that some of the scores given at day 2 subsequently

Zainali-Gill *et al. The Egyptian Journal of Otolaryngology*   (2022) 38:52

Page 7 of 9

are not the same as day 1, since the observers have increased their evaluation skills. This is depicted in the learning curves for each UAC site. However, the increase was not stable, and the increase was not the same for novice observer B and novice observer C.

Inter-observer variation

Carrasco-Llatas et al., in a similar study with 31 videos, investigated the inter-observer variation between an expert observer and an ENT resident in training [14]. The degree of collapse using VOTE classification was evaluated, and they found OA for velum 80% ($k = 0.17$), OA for oropharynx 78.57% ($k = 0.67$), OA for tongue base 54.84% ($k = 0.35$), and OA for epiglottis 61.29% ($k = 0.43$). The low k-value for velum was explained by the high prevalence of velum collapse in their material. Nonetheless, they had problems agreeing on velum as we had, suggesting that this is a difficult evaluation. This study compared the expert observer A to novice B and to novice C, respectively. The agreement between A and B increased substantially from day 1 to day 2 in all UAC sites. When observer A was compared to novice C, the agreement in all UAC sites except the velum also increased. Our results suggest that a systematic theoretical teaching of understanding the dynamics in the collapse of upper airways contributes fairly to improving the evaluation of DISE examination.

In the study by Vroegop et al., the inter-observer agreement in a group of 7 experienced observers and 90 inexperienced ENT surgeons was evaluated. Although they did not use the VOTE classification, they found that inter-observer agreement in general was higher in the experienced group than in the non-experienced group. Even the inter-observer agreement in the palatal site was high in the experienced group with $OA = 0.88$ ($k = 0.51$) suggesting moderate agreement, whereas the non-experienced group with $OA = 0.88$ ($k = -0.03$) suggested no agreement. Again, the k coefficient for palatal examination in the non-experienced group was low.

The non-experienced group presented at the oropharyngeal level an $OA = 0.45$ ($k = 0.09$), the tongue base OA = 0.63 ($k = 0.33$), and the epiglottis OA = 0.57 ($k = 0.23$). Although these results are not directly comparable to our study, since we compared the results of the expert to the results of the novices, the total OA values and the k coefficients for A vs B and A vs C in our study are higher. This could imply a benefit from the systematic and structural introduction and in-plenum discussion of DISE videos.

Learning curve

The approach of evaluating the DISE procedures in clusters of ten followed by debriefing in our opinion

contributed to an increase in inter-observer agreement and hence an increased learning curve for the novice observers. The agreement was poor for velum and epiglottis but better for oropharynx and tongue base. Vroegop et al. found that experience in performing DISE was necessary to obtain reliable observations, since both inter-observer and intra-observer agreement was higher in experienced versus inexperienced ENT surgeons [16].

Even among the experienced operators, Ong et al. concluded that DISE remained a subjective examination, and more studies investigating the improvement in inter-rater reliability after implantation of training videos were needed [17]. A French study from the ENT sleep experts group found DISE to be a technique with a limited inter-observer agreement in the detection of obstructive sites, and no learning curve effect was observed [18]. In order to improve the reliability of the DISE classification, Veer et al. recently introduced the PTLTbE (palate, tonsils, lateral pharyngeal wall, tongue base, epiglottis) where tonsillar obstruction is separated from lateral pharyngeal wall collapse [19]. This classification also uses images of the different degrees of obstruction and configuration as a guidance tool. They found the inter-observer reliability to be improved and the learning curve of the system to be short and steep even among junior doctors.

Due to the present results and since the VOTE classification is universally used all over the world, the VOTE classification is still used at our ENT department. Our results suggest that each DISE observer needs a large amount of DISE interpretations and continuous education, before expert level can be achieved. We have established a comprehensive sleep database on OSA patients referred to the department. DISE videos are reviewed independently by the same two ENT specialists, and treatment is suggested. Finally, a therapeutic intervention for the specific patient is planned in plenum, and follow-up is also registered. The VOTE classification is stored in the database, and this allows for a more comprehensive comparison of in-future inter-observer variation.

Limitations

The study has limitations. Considering that the purpose of this study was to evaluate an evolution in video interpretation using systematic training, the large confidence intervals are outweighed as the results overall suggest an improvement in inter-observer agreement on day 2 compared to day 1, plus an increase in learning curve during day 1 and the overall fair agreement in intra-observer agreement.

DISE can be evaluated using different classification protocols described elsewhere as Koo's DISE classification focusing on surgical treatment [20] or Vicini's

Zainali-Gill *et al. The Egyptian Journal of Otolaryngology*   (2022) 38:52

Page 8 of 9

NOHL classification that can be used both during awake and sedation [21]. We choose the VOTE classification due to its broad use in the clinical and academic literature worldwide [22]. However, since the focus of the study was the evaluation of focused training using systematic review and in plenum discussions between expert and novices to DISE, we choose only to focus on whether the airway was, open, partly collapsed or fully collapsed and did not focus on the configuration of the collapse entity (anteroposterior, lateral, or concentric configurations). Nor did we compare notes on further surgical interventions since this was not in the scope of the study even if this comparation would be clinically relevant.

The systematic teaching and educational sessions have intensive benefits but are yet vulnerable for repeated misinterpretations. Even the expert observer could not demonstrate almost perfect agreement within day 1 and day 2. This suggests that some intra-observer variation will continue to exist, and not only DISE experience but also surgical experience seems to be important in order to obtain reliable observations. However, even among this group, surgery planning based on DISE should be done with caution.

Conclusion
This study showed that the systematic approach with an initial theoretical insight in the evaluation of DISE examinations and the continuous debriefing of the previous 10 videos led to an overall increase in agreement between the expert observer and the novices in 2 days. However, a large number of evaluations is needed before one is familiar with DISE interpretation.

Abbreviations
DISE: Drug-induced sedation endoscopy; VOTE: Velum, oropharyngeal, tongue, and epiglottis; CPAP: Continuous positive airway pressure; BMI: Body mass index; OSA: Obstructive sleep apnea; UAC: Upper airway collapse.

Authors' contributions
KZG: study design, data collection, statistical analysis, and main author. JBB: study design, data collection, data collection, statistical analysis, and discussion. CvB: study design, data analysis, and discussion. TO: study design, data analysis, and discussion. EKK: study design, data collection, statistical analysis, and discussion. All authors read and approved the final manuscript.

Availability of data and materials
All VOTE ratings are available at reasonable request. The statistical analyses and data are available for digital upload.

Declarations

Author details
[1]Department of Otorhinolaryngology, University Clinic for Flavour, Balance and Sleep, Hospitalsenheden Vest, 7500 Holstebro, Denmark. [2]Department of Clinical Medicine, Aarhus University, Aarhus N, Aarhus, Denmark. [3]Department of Otorhinolaryngology, Head and Neck Surgery and Audiology, Rigshospitalet, Copenhagen, Denmark.

References
1. Jordan AS, McSharry DG, Malhotra A (2014) Adult obstructive sleep apnoea. Lancet. https://doi.org/10.1016/S0140-6736(13)60734-5
2. Parati G, Lombardi C, Hedner J, Bonsignore MR, Grote L, Tkacova R et al (2012) Position paper on the management of patients with obstructive sleep apnea and hypertension. J Hypertens 30:633–646. https://doi.org/10.1097/hjh.0b013e328350e53b
3. Punjabi NM, Caffo BS, Goodwin JL, Gottlieb DJ, Newman AB, O'Connor GT et al (2009) Sleep-disordered breathing and mortality: a prospective cohort study. PLoS Med. https://doi.org/10.1371/journal.pmed.1000132
4. Redline S, Yenokyan G, Gottlieb DJ, Shahar E, O'Connor GT, Resnick HE et al (2010) Obstructive sleep apnea-hypopnea and incident stroke: the sleep heart health study. Am J Respir Crit Care Med 182:269–277. https://doi.org/10.1164/rccm.200911-1746OC
5. Foster GD, Sanders MH, Millman R, Zammit G, Borradaile KE, Newman AB et al (2009) Obstructive sleep apnea among obese patients with type 2 diabetes. Diabetes Care 32:1017–1019. https://doi.org/10.2337/dc08-1776
6. Abdelghani A, Benzarti W, Ben SH, Gargouri I, Garrouche A, Hayouni A et al (2016) Adherence to treatment with continuous positive airway pressure in the obstructive sleep apnea syndrome. Tunis Med 94:551–562
7. Caples SM, Rowley JA, Prinsell JR, Pallanch JF, Elamin MB, Katz SG et al (2010) Surgical modifications of the upper airway for obstructive sleep apnea in adults: a systematic review and meta-analysis. Sleep 33:1396–1407. https://doi.org/10.1093/sleep/33.10.1396
8. Croft CB, Pringle M (1991) Sleep nasendoscopy: a technique of assessment in snoring and obstructive sleep apnoea. Clin Otolaryngol. https://doi.org/10.1111/j.1365-2273.1991.tb01050.x
9. Berry S, Roblin G, Williams A, Watkins A, Whittet HB (2005) Validity of sleep nasendoscopy in the investigation of sleep related breathing disorders. Laryngoscope 115:538–540. https://doi.org/10.1097/01.mlg.0000157849.16649.6e
10. Kiær EK, Tønnesen P, Sørensen HB, Rubek N, Hammering A, Møller C et al (2019) Propofol sedation in drug induced sedation endoscopy without an anaesthesiologist – a study of safety and feasibility. Rhinology 57:125–131. https://doi.org/10.4193/Rhin18.066
11. de Vito A, Carrasco Llatas M, Ravesloot MJ, Kotecha B, de Vries N, Hamans E et al (2018) European position paper on drug-induced sleep endoscopy: 2017 update. Clin Otolaryngol 43:1541–1552. https://doi.org/10.1111/coa.13213

Zainali-Gill *et al. The Egyptian Journal of Otolaryngology*   (2022) 38:52

Page 9 of 9

12. Kezirian EJ, Hohenhorst W, De Vries N (2011) Drug-induced sleep endoscopy: the VOTE classification. Eur Arch Otorhinolaryngol 268:1233–1236. https://doi.org/10.1007/s00405-011-1633-8

13. Blehar DJ, Barton B, Gaspari RJ (2015) Learning curves in emergency ultrasound education. Acad Emerg Med 22:574–582. https://doi.org/10.1111/acem.12653

14. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159. https://doi.org/10.2307/2529310

15. Vroegop AV, Vanderveken OM, Boudewyns AN, Scholman J, Saldien V, Wouters K et al (2014) Drug-induced sleep endoscopy in sleep-disordered breathing: report on 1,249 cases. Laryngoscope 124:797–802. https://doi.org/10.1002/lary.24479http://wx7cf7zp2h.search.serialssolutions.com?sid=EMBASE&issn=0023852X&id=doi:10.1002%2Flary.24479&atitle=Drug-induced+sleep+endoscopy+in+sleep-disordered+breathing%3A+Report+on+1%2C249+cases&stitle=Laryngoscope&title=Laryngoscope&volume=124&issue=3&spage=797&epage=802&aulast=Vroegop&aufirst=Anneclaire+V.&auinit=A.V.&aufull=Vroegop+A.V.&coden=LARYA&isbn=&pages=797-802&date=2014&auinit1=A&auinitm=V

16. Vroegop AVMT, Vanderveken OM, Wouters K, Hamans E, Dieltjens M, Michels NR et al (2013) Observer variation in drug-induced sleep endoscopy: experienced versus nonexperienced ear, nose, and throat surgeons. Sleep 36:947–953. https://doi.org/10.5665/sleep.2732

17. Ong AA, Ayers CM, Kezirian EJ, Tucker Woodson B, de Vries N, Nguyen SA et al (2017) Application of drug-induced sleep endoscopy in patients treated with upper airway stimulation therapy. World J Otorhinolaryngol Head Neck Surg. https://doi.org/10.1016/j.wjorl.2017.05.014

18. Bartier S, Blumen M, Chabolle F (2020) Is image interpretation in drug-induced sleep endoscopy that reliable? Sleep Breath. https://doi.org/10.1007/s11325-019-01958-5

19. Veer V, Zhang H, Mandavia R, Mehta N (2020) Introducing a new classification for drug-induced sleep endoscopy (DISE): the PTLTbE system. Sleep Breath. https://doi.org/10.1007/s11325-020-02035-y

20. Koo SK, Lee SH, Koh TK, Kim YJ, Moon JS, Lee HB et al (2019) Inter-rater reliability between experienced and inexperienced otolaryngologists using Koo's drug-induced sleep endoscopy classification system. Eur Arch Otorhinolaryngol. https://doi.org/10.1007/s00405-019-05386-9

21. Vicini C, De Vito A, Benazzo M, Frassineti S, Campanini A, Frasconi P et al (2012) The nose oropharynx hypopharynx and larynx (NOHL) classification: a new system of diagnostic standardized examination for OSAHS patients. Eur Arch Otorhinolaryngol 269:1297–1300. https://doi.org/10.1007/s00405-012-1965-z

22. Amos JM, Durr ML, Nardone HC, Baldassari CM, Duggins A, Ishman SL (2018) Systematic review of drug-induced sleep endoscopy scoring systems. Otolaryngol Head Neck Surg 158:240–248. https://doi.org/10.1177/0194599817737966

Publisher's Note