

Concordância interobservador na DISE: Vemos todos o mesmo?

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

RESUMO

Objetivos: Avaliar a concordância interobservador na classificação da DISE (*Drug Induced Sleep Endoscopy*) e na decisão terapêutica subsequente.

Desenho do estudo: Transversal.

Material e métodos: Gravaram-se 20 DISEs realizadas em adultos num Hospital. As imagens foram apresentadas a 3 especialistas e 3 internos de formação específica (Otorrinolaringologia). Todos classificaram os achados das DISEs e sugeriram o tratamento. Calculou-se o valor kappa(k) para avaliar a concordância interobservador.

Resultados e Conclusões: A concordância entre os dois grupos em relação à presença de obstrução da via aérea foi 100% ao nível do palato(k=1), 90% na orofaringe(k=0,69), 85% na base da língua(k=0,68) e 85% na epiglote(k=0,69). Quando se avaliou a gravidade da obstrução, a concordância interobservador foi moderada a alta a todos os níveis exceto na base da língua(k=0,38). Verificou-se uma maior frequência de sugestão para cirurgia basilingual no grupo dos internos(7/20) em relação aos especialistas(1/20); k=0,34 (p=0,04).

Desta forma, a concordância entre internos e especialistas parece ser bastante aceitável exceto em termos de classificação da gravidade de obstrução ao nível da base da língua (concordância fraca) e de decisão terapêutica nessa área anatómica (concordância fraca).

Palavras-chave: Endoscopia sob sedação; Apneia do sono; Base da língua; Faringe

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

Correspondência:

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

Artigo recebido a 9 de Setembro 2021. Aceite para publicação a 25 de Setembro de 2021.

INTRODUÇÃO

O síndrome de apneia-hipopneia obstrutiva do sono (SAHOS) é o distúrbio respiratório do sono mais frequente na idade adulta, tendo uma frequência de 2-4% nesta faixa etária⁽¹⁾. O diagnóstico de SAHOS é feito por polissonografia noturna (PSG). O tratamento de primeira linha nos casos de apneia moderada a grave é o CPAP (*continuous positive airway pressure*), no entanto, a sua eficácia é frequentemente limitada por dificuldades de adaptação e má adesão por parte dos doentes^{(2);(3);(4);(5)}. O CPAP pode também ser utilizado em casos de SAHOS ligeiro que se associe a comorbilidades como hipertensão arterial, doença coronária, doença cerebrovascular ou sintomas que perturbem o dia-a-dia do doente (sonolência diurna, diminuição da capacidade cognitiva, alterações do humor, etc)^{(4);(5)}.

Nos casos em que se considera outro tipo de tratamento (exemplos: cirurgia, dispositivo de avanço mandibular, estimulador do nervo hipoglosso), a apreciação metódica de eventuais colapsos da via aérea superior é fundamental. A DISE, exame com popularidade crescente desde o seu aparecimento em 1991⁽⁶⁾, é uma ferramenta especialmente útil neste campo uma vez que permite o diagnóstico topográfico da obstrução através da visualização da via aérea superior (VAS) de forma dinâmica durante o sono farmacologicamente induzido⁽¹⁾. Apesar de ser um exame bastante difundido, existem algumas contra-indicações à sua realização. A DISE não deve ser realizada em indivíduos com *score* ASA (*American Society of Anesthesiology*)⁽⁴⁾, grávidas, casos de alergia a algum dos fármacos utilizados (contra-indicações absolutas) nem em indivíduos com obesidade mórbida (contra-indicação relativa). O exame pode ser realizado no bloco operatório ou numa sala de exames equipada com material de ressuscitação em caso de emergência⁽¹⁾. Além da necessidade de nasofaringolaringoscópio flexível, equipa médica (anestesta; otorrinolaringologista) e monitorização anestésica genérica (SpO₂, ECG, pressão arterial), o local de realização da DISE deve estar dotado de monitor de BIS (*bispectral index*) e bomba infusora TCI (*targeted controlled infusion*)⁽⁷⁾.

Alguns estudos referem que a subjetividade da DISE não deve ser ignorada e, a esse respeito, ressaltam que, apesar das várias classificações validadas, a experiência do observador poderá ter um papel importante na fiabilidade dos resultados deste exame^{(8);(9);(10)}.

Desta forma, o presente estudo teve como objetivos



avaliar a concordância interobservador na classificação da DISE e na decisão terapêutica baseada nos achados deste exame.

MATERIAL E MÉTODOS

Foram gravadas, no Hospital Prof. Doutor Fernando Fonseca, ao longo do primeiro semestre do ano de 2019, 20 DISE's realizadas em adultos com Síndrome de Apneia Hipopneia Obstrutiva do Sono (SAHOS). O diagnóstico de SAHOS foi feito por polissonografia nível 2 no domicílio. As DISEs foram realizadas em doentes que recusaram e/ou não se adaptaram ao tratamento com CPAP. Foram excluídos do estudo doentes com IAH>30 (doentes com SAHOS grave são habitualmente geridos pela Pneumologia e o tratamento cirúrgico é realizado apenas de forma pontual para melhorar adaptação ao CPAP) ou doentes previamente submetidos a cirurgia para tratamento de apneia do sono (cirurgia pregressa altera a configuração anatômica da VAS e pode enviesar a classificação da DISE). Os dados sociodemográficos, biométricos, anatômicos e polissonográficos foram registados após consulta informática nos respetivos processos clínicos.

Todas as DISEs foram realizadas no bloco operatório em regime de ambulatório. No que à técnica anestésica diz respeito, foi utilizado propofol em infusão por TCI. A profundidade da sedação foi monitorizada por BIS, tendo como intervalo de referência valores entre 50-70⁽¹⁾. Em termos de posicionamento, os doentes foram colocados inicialmente em decúbito dorsal e posteriormente na posição descrita na respetiva PSG como sendo a mais geradora de eventos respiratórios (não se avaliaram outras posições de sono). Foram realizadas em todos os indivíduos as manobras de *Esmarch* e *chin lift*. Não foram utilizados anestésicos ou vasoconstritores locais/nasais. Em todos os exames foram observados 2 ou mais ciclos respiratórios para cada nível da via aérea (um ciclo respiratório foi definido como a seguinte sequência de eventos conforme descrito no consenso europeu de DISE(1): ronco – dessaturação com hipopneia/apneia obstrutiva – respiração). As imagens foram apresentadas

individualmente a 3 especialistas e a 3 internos de formação específica em Otorrinolaringologia (1 no 3º ano e 2 no 4º ano de formação). A única informação clínica que foi dada aos 6 observadores previamente à classificação da DISE foi o valor de IAH de cada doente (a severidade da apneia pode auxiliar o classificador na decisão terapêutica, nomeadamente, em termos de extensão do tratamento cirúrgico e número de níveis anatômicos abordados). Não foi fornecida informação clínica adicional. Cada um dos observadores classificou, de forma independente e cega, os achados da DISE segundo a classificação VOTE modificada⁽⁹⁾. A classificação VOTE modificada (Tabela 1) tem a particularidade de, ao nível da base da língua, permitir classificar a configuração da obstrução como lateral ou concêntrica (além da configuração anteroposterior já contemplada na classificação VOTE clássica)⁽⁹⁾. Depois de classificarem os achados da DISE, os observadores sugeriram o tratamento indicado para cada caso. Para cada nível anatômico, foram fornecidas aos observadores diferentes opções de tratamento. Ao nível do palato: Faringoplastia de reposicionamento com sutura farpada (BRP); Faringoplastia de expansão (FE)(11); Palatoplastia anterior de Pang (PA); Radiofrequência (RFP). Ao nível da orofaringe: Amigdalectomia extracapsular bilateral (A). Ao nível da base da língua: Radiofrequência (RFBL); Coblation (CBL); Dispositivo de avanço mandibular (DAM). Ao nível da epiglote: Epiglotectomia parcial (EP). Foram comparados, entre os dois grupos de observadores, os seguintes parâmetros: presença/ausência de obstrução nos vários níveis da VAS; grau de obstrução; configuração da obstrução. Foi ainda avaliada a concordância entre os dois grupos no que diz respeito à decisão terapêutica. Utilizou-se o coeficiente kappa(k) de Cohen para avaliar a concordância entre os dois grupos. A concordância foi classificada(12) como nula ($K=0$), baixa ($0,01 \leq K \leq 0,2$), fraca ($0,21 \leq K \leq 0,4$), moderada ($0,41 \leq K \leq 0,6$), substancial ($0,61 \leq K \leq 0,8$) e forte/alta ($0,81 \leq K \leq 1$).

Foi utilizado o programa SPSS statistics 25 (IBM) para o tratamento estatístico dos dados. Um valor $p < 0,05$ foi utilizado para traduzir significância estatística.

TABELA 1

Classificação de VOTE modificada

Estrutura	Grau de Obstrução	Configuração		
		Anteroposterior	Lateral	Concêntrico
Palato (V)	0: sem obstrução 1: obstrução parcial (vibração) 2: obstrução completa (colapso) X: não visualizado			
Orofaringe (O)				
Base da língua (T)				
Epiglote (E)				

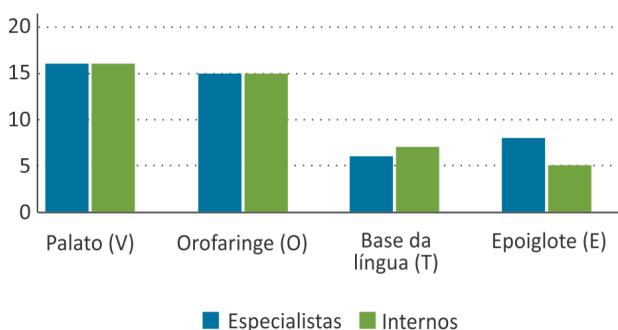
RESULTADOS

Caracterização da amostra

Foram englobados no estudo 20 indivíduos (de forma consecutiva) submetidos a endoscopia sob sedação (65% do sexo masculino; média de idades $52,1 \pm 13,4$ anos; min:18; máx: 74). O índice de apneia-hipopneia (IAH) médio da população na PSG realizada antes da DISE foi de $16,4 \pm 8,8$ (min: 5,5; máx: 29,5). O índice de massa corporal médio (IMC) foi de $25,7 \pm 2,7$ kg/m² (min: 21; máx: 29).

GRÁFICO 1

Presença de obstrução de qualquer grau por nível anatómico (xx: áreas anatómicas; yy: número absoluto de indivíduos com obstrução de qualquer grau)



Presença obstrução de qualquer grau na via aérea superior

A concordância entre os dois grupos em relação à presença de obstrução da via aérea (qualquer grau) foi total ao nível do palato (100%; $k=1$). Documentou-se

concordância substancial nos outros três níveis avaliados: 90% ao nível da orofaringe ($k=0,69$), 85% na base da língua ($k=0,68$) e 85% na epiglote ($k=0,69$).

Gravidade da obstrução por nível anatómico

No que diz respeito à gravidade de obstrução por nível anatómico, a referir concordância substancial entre os dois grupos a todos os níveis exceto ao nível da base da língua (concordância fraca; $k=0,38$) – Tabela 2.

Neste parâmetro, a referir que o grupo de observadores menos experientes (internos) classificou a obstrução ao nível da base da língua como total mais frequentemente que os especialistas (proporção de 6:1).

Conformação da obstrução por nível anatómico

A concordância entre os dois grupos no que concerne à conformação da obstrução variou entre moderada e forte (Tabela 3). Uma vez que segundo a classificação VOTE modificada, uma eventual obstrução ao nível da orofaringe deve ser sempre classificada como lateral, não é possível avaliar a concordância para a classificação da conformação a este nível.

Tratamento

Foi proposto, por parte dos observadores, o tratamento que consideraram adequado para cada caso conforme descrito em “material e métodos”. Todos os casos foram propostos para pelo menos um tipo de tratamento. A decisão terapêutica de ambos os grupos (especialistas e internos) para cada nível anatómico está descrita na tabela 4.

TABELA 2

Gravidade da obstrução por nível anatómico (A gravidade é classificada como parcial/vibração ou total/colapso)

Nível anatómico	Palato		Orofaringe		Base da língua		Epiglote	
	Parcial (nº de doentes)	Total (nº de doentes)						
Especialistas	6	11	6	10	6	1	4	5
Internos	4	13	6	8	2	6	4	2
Concordância (valor Kappa)	Substancial (0,64)		Substancial (0,68)		Fraca (0,38)		Substancial (0,64)	

TABELA 3

Conformação da obstrução por nível anatómico

Nível anatómico	Palato			Orofaringe	Base da língua			Epiglote	
	Lateral (nº de doentes)	AP (nº de doentes)	Concêntrico (nº de doentes)	Lateral (nº de doentes)	Lateral (nº de doentes)	AP (nº de doentes)	Concêntrico (nº de doentes)	Lateral (nº de doentes)	AP (nº de doentes)
Especialistas	-	3	14	16	-	5	-	1	8
Internos	-	4	13	16	2	8	-	1	5
Concordância (valor Kappa)	Forte (0,83)			N/A	Moderada (0,51)			Substancial (0,61)	

N/A: Não aplicável (nos casos em que existiu concordância quanto à presença de obstrução, esta foi obrigatoriamente lateral)

TABELA 4

Decisão terapêutica (A concordância foi calculada em relação à decisão de intervir em cada nível anatómico e não em relação ao tipo de tratamento)

	Palato (tipo de tratamento)	Orofaringe (tipo de tratamento)	Base da língua (tipo de tratamento)	Epiglote (tipo de tratamento)	Cirurgia multinível
Especialistas	17 (16x BRP; 1x PA)	16 (A)	2 (CBL; DAM)	4 (EP)	3/20
Internos	17 (13x BRP; 4x PA)	14 (A)	7 (6x CBL; 1x RFBL)	2 (EP)	6/20
Concordância (valor Kappa)	Total (1,00;p<0,001)	Substancial (0,69;p=0,002)	Fraca (0,34;p=0,04)	Substancial (0,69;p=0,0019)	

BRP: Faringoplastia de reposição; PA: Palatoplastia anterior; CBL: Coblation da base da língua; RFBL: Radiofrequência da base da língua; DAM: Dispositivo de avanço mandibular; EP: Epiglotectomia parcial.

NOTA: Cirurgia multinível foi definida como aquela que visa a correção de pelo menos 2 níveis anatómicos; no caso da cirurgia do palato, se for acompanhada somente de amigdalectomia é contabilizada como um único nível.

DISCUSSÃO

A variabilidade interobservador é um conceito que deve ser estudado sempre que se pretende avaliar a fiabilidade dos resultados de um exame observador-dependente como é exemplo a DISE. Sendo a DISE realizada na perspectiva de oferecer um tratamento dirigido ao tipo de obstrução de cada indivíduo, a concordância entre as decisões terapêuticas decorrentes deste exame também pode e deve ser avaliada^{(8);(9)}.

Neste trabalho, de um modo geral, a concordância entre os dois grupos foi adequada, exceto ao nível da base da língua. A fraca concordância em termos de classificação da gravidade da obstrução a este nível terá levado a que o grupo dos internos tenha sugerido com maior frequência cirurgia basilingual comparativamente com o grupo de especialistas.

Um estudo prévio realizado por um grupo da Antuérpia procurou avaliar a variabilidade interobservador da DISE⁽¹⁰⁾. O referido trabalho foi realizado com um considerável número de observadores (97), no entanto, baseou-se em apenas seis vídeos de DISE's. Tal como no trabalho por nós realizado, a concordância entre os observadores foi calculada pelo coeficiente kappa(k) de Cohen. Importa frisar que dos 97 observadores, 7 deles eram inexperientes em DISE, no entanto, não foram criados grupos de categorização dos observadores por diferentes graus de experiência. Apesar disso, os autores desse artigo documentaram que as áreas anatómicas onde se verificou maior concordância foi ao nível da base da língua e do palato (substancial em ambos). Não foi avaliada a decisão terapêutica dos observadores. Comparando os resultados estudo citado com aqueles obtidos no nosso estudo, verificou-se que os resultados do nosso trabalho são diferentes ao nível da base da língua (concordância fraca vs concordância substancial). Consideramos que o facto de, no trabalho citado, não terem sido criados grupos de observadores com diferentes experiências inviabiliza uma comparação fidedigna entre os dois trabalhos e por isso não valorizamos as diferenças verificadas.

Existe na literatura um outro estudo de um grupo de

Valência que efetivamente avaliou a concordância na classificação da DISE categorizando os observadores por nível de experiência⁽⁸⁾. No referido trabalho participou um observador experiente (especialista em Otorrinolaringologia) e um observador inexperiente (interno de formação específica em Otorrinolaringologia). Classificaram-se 31 DISE's e foi utilizado o coeficiente kappa(k) de Cohen para aferir a concordância entre os dois observadores. Os autores documentaram concordância substancial quando se avaliou a gravidade e conformação da obstrução exceto ao nível da base da língua (concordância fraca à custa de subvalorização da obstrução basilingual pelos internos). Em termos de decisão terapêutica, a concordância foi adequada a todos os níveis exceto ao nível basilingual (concordância fraca à custa das diferenças registadas na classificação da obstrução a este nível). Comparando estes resultados com aqueles verificados no nosso trabalho, verificamos que a área anatómica de discordância entre internos e especialistas em termos de classificação da obstrução e de tratamento proposto é a mesma (base da língua). Apesar da similaridade dos resultados, consideramos que o facto de o estudo citado ter somente dois observadores limita uma comparação apropriada com o presente estudo (seis observadores).

A literatura mais recente sugere que a DISE é sobretudo relevante em situações em que se considere cirurgia de base da língua ou DAM⁽¹⁰⁾. Deste modo, e uma vez que a região basilingual parece ser aquela se associa a menor consenso entre observadores com diferente experiência, pode inferir-se que este nível deve ser sempre observado por um profissional especialmente experiente.

Este estudo apresenta algumas limitações, desde logo em termos de tamanho da amostra (20 indivíduos). Por outro lado, o facto de a classificação ter sido feita baseada em vídeos previamente gravados, não permite ao observador assegurar qual o posicionamento do doente em cada momento ao longo do vídeo. Uma outra limitação do presente trabalho é o facto de todos os classificadores que nele participam exercerem a sua atividade no mesmo Hospital. Este fator pode, por si só,

explicar parte da concordância verificada uma vez que a metodologia e forma de aprendizagem são semelhantes entre os observadores (ao contrário do que aconteceria se o estudo fosse realizado com Médicos de diferentes Hospitais). Além disso, existem outros fatores além dos achados da DISE que devem ser tidos em conta aquando da decisão terapêutica e não foram analisados pelos observadores neste trabalho. A este respeito, é importante ter em linha de conta uma série de fatores antes de propor um indivíduo para determinado tratamento. Desde logo, o Médico Otorrinolaringologista deve preocupar-se em entender qual a motivação e potencial adesão do doente às várias soluções terapêuticas existentes. Um conjunto de outros fatores além daqueles que se observam nas imagens da DISE (ex: idade; índice de massa corporal; retro/prognatismo; entre outros) devem ser devidamente ponderados aquando da escolha da opção terapêutica adequada ao doente.

CONCLUSÃO

A DISE é globalmente fiável e segura, no entanto, o presente estudo sugere que a experiência do observador é um fator preponderante tanto na classificação dos achados deste exame como na decisão terapêutica, sobretudo ao nível da base da língua.

Agradecimentos

Ao Serviço de Otorrinolaringologia do Hospital Professor Doutor Fernando Fonseca, em particular, ao Doutor Filipe Freire e Doutora Ana Guimarães.

Conflito de Interesses

Os autores declaram que não têm qualquer conflito de interesse relativo a este artigo.

Confidencialidade dos dados

Os autores declaram que seguiram os protocolos do seu trabalho na publicação dos dados de pacientes.

Proteção de pessoas e animais

Os autores declaram que os procedimentos seguidos estão de acordo com os regulamentos estabelecidos pelos diretores da Comissão para Investigação Clínica e Ética e de acordo com a Declaração de Helsínquia da Associação Médica Mundial.

Política de privacidade, consentimento informado e Autorização do Comité de Ética

Os autores declaram que têm o consentimento por escrito para o uso de fotografias dos pacientes neste artigo.

Financiamento

Este trabalho não recebeu qualquer contribuição, financiamento ou bolsa de estudos.

Disponibilidade dos Dados científicos

Não existem conjuntos de dados disponíveis publicamente relacionados com este trabalho.

Referências bibliográficas

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.