

Oxygen Desaturation Index calculation: impact of different methodologies

Diego Munduruca Domingues^{1,2}, Filipe Vilela Soares², Geraldo Lorenzi-Filho^{2,3},
Henrique Takachi Moriya¹

¹Laboratório de Engenharia Biomédica,
Escola Politécnica da Universidade de São Paulo (EPUSP)
São Paulo (SP) – Brasil

²Biologix Sistemas Ltda.
São Paulo (SP) – Brasil

³Laboratório do Sono, Divisão de Pneumologia, Instituto do Coração (InCor)
Faculdade de Medicina da Universidade de São Paulo (FMUSP)
São Paulo (SP) – Brasil

{diego.domingues, htmoriya}@usp.br
filipe.soares@biologix.com.br, geraldo.lorenzi@gmail.com

Abstract. *Obstructive Sleep Apnea (OSA) is a syndrome characterized by episodes of airway obstruction, which causes oxygen desaturation events. These events can be identified by oximetry analysis and are used as one of the parameters to diagnose OSA. However, desaturation events have an inaccurate definition in manuals and in most of the literature. Thus, this work aims to evaluate whether different methodologies for the calculation of desaturation events impact the Oxygen Desaturation Index (ODI) and the diagnosis of OSA. The results indicated that the ODI values are significantly different from each other ($p < 0.001$) and the methodologies provided variable performance for the diagnosis of OSA.*

Resumo. *A Apneia Obstrutiva do Sono (AOS) é uma síndrome caracterizada por episódios de obstrução das vias aéreas, que provoca eventos de dessaturação de oxigênio. Estes eventos podem ser identificados pela análise da oximetria e são utilizados como um dos parâmetros para diagnosticar a AOS. Porém, sua definição é pouco precisa em manuais e na maior parte da literatura. Desta forma, este trabalho busca avaliar se diferentes metodologias para o cálculo de eventos de dessaturação causam impacto no Índice de Dessaturação de Oxigênio (IDO) e no diagnóstico de AOS. Os resultados indicaram que os valores de IDO são significativamente diferentes entre si ($p < 0,001$) e que as metodologias apresentaram performance variável para diagnóstico de AOS.*

1. Introdução

A Apneia Obstrutiva do Sono (AOS) é uma doença caracterizada por repetidos episódios de interrupção total (apneia) ou parcial (hipopneia) das vias aéreas superiores, o que leva a intermitentes dessaturações de oxigênio e sono fragmentado [Nigro et al. 2011]. Esta doença possui uma prevalência de 5% a 15% [Romem et al. 2014] na população

em geral. Na população de meia idade, essa prevalência aumenta para 20% a 30% [Chiang et al. 2017].

A AOS está associada a vários sintomas, como ronco, sonolência diurna, redução de atenção, fadiga e depressão, o que acaba tendo um impacto negativo na qualidade de vida [Flemons et al. 1999]. A AOS não tratada está associada a doenças cardiovasculares, como hipertensão, arritmias, doença arterial coronariana, acidente vascular cerebral e morte cardiovascular [Jones 2009].

O diagnóstico de AOS é usualmente feito com base no Índice de Apneia e Hipopneia (IAH), o qual é obtido com a polissonografia (PSG). Neste exame, o paciente dorme com diversos sensores, incluindo eletrodos de eletrocardiografia (ECG) e eletroencefalografia (EEG), sensor de fluxo nasal, entre outros. Porém, a PSG tem limitações, pois é um exame caro, tecnicamente complexo e inconveniente para o paciente devido à quantidade de fios e sensores que devem se manter conectados a ele durante o sono [Ling et al. 2012, Romem et al. 2014].

Por isso, outras modalidades de diagnóstico de AOS buscam simplificar e tornar o exame mais acessível à população, utilizando um número limitado de canais respiratórios [Erman et al. 2007, Corral et al. 2017]. Com isso, o paciente consegue fazer o exame na própria casa, o que é mais confortável e econômico. Um dos canais respiratórios frequentemente utilizado nestas soluções é a oximetria [Yamashiro and Kryger 1995, Rashid et al. 2021], que pode ser obtida com uso de um oxímetro de pulso, de forma simples e não invasiva.

A oximetria permite medir a saturação de oxigênio no sangue (SpO_2). A SpO_2 tem uma alta relação com os eventos de apneia, pois o bloqueio da passagem de ar, característica desses eventos, causa uma interrupção no fornecimento de oxigênio [Alvarez-Estevez and Moret-Bonillo 2015]. Para realizar o diagnóstico, a principal informação derivada da oximetria é o Índice de Dessaturação de Oxigênio (IDO), que é o número médio de eventos de dessaturação por hora de sono.

Tanto na polissonografia quanto nas soluções que utilizam somente a oximetria, os eventos de dessaturação de oxigênio têm um papel chave. Para o cálculo do IAH, apesar das regras para identificar hipopneias terem sido alteradas ao longo do tempo [Berry et al. 2012, Hirotsu et al. 2019], todas utilizam como critério a redução de fluxo aéreo associada a eventos de dessaturações com 3% ou 4% de profundidade e/ou um microdespertar.

Porém, eventos de dessaturação de oxigênio não possuem uma definição única e precisa. O próprio manual da *American Academy of Sleep Medicine* (AASM) se limita a definir a profundidade do evento, em porcentagem, e a linha de base precedente ao evento, ao qual a queda deve ser comparada, sem nenhuma indicação de como deve ser feito o seu cálculo [Berry et al. 2012]. Dessa forma, é possível encontrar na literatura diversos trabalhos que fazem suas próprias definições de eventos de dessaturação de oxigênio — alguns utilizam quedas de 3%, outros utilizam 4%; outra parte define critério de ressaturação, duração mínima e máxima do evento etc. [Alvarez-Estevez and Moret-Bonillo 2015].

Em relação à definição da linha de base para identificação dos eventos de dessaturação, há dois problemas principais: alguns trabalhos citam que utilizaram a linha de base para detecção da queda na SpO_2 , porém não mencionam como esta foi

calculada. O segundo problema está relacionado com a diversidade de metodologias para realizar o cálculo da linha de base. Na revisão sistemática feita por Rashid et al. [Rashid et al. 2021], 8 trabalhos foram analisados e a definição da linha de base foi diferente em cada um deles. Essas discrepâncias podem fazer com que os valores de IDO e IAH, calculados a partir de diferentes metodologias, apresentem uma significativa variabilidade entre si, o que impactaria inclusive no diagnóstico de AOS.

Portanto, esse trabalho tem como objetivo analisar a variabilidade no valor do IDO ao utilizar diferentes definições para cálculo da linha de base.

2. Métodos

2.1. Pacientes

Foram analisados dados de 91 pacientes adultos com suspeita de AOS (tabela 1), que realizaram o exame de polissonografia no Laboratório do Sono do Instituto do Coração do Hospital das Clínicas da Universidade de São Paulo, no período de julho de 2017 a julho de 2018. Foram excluídos pacientes com diagnóstico de insuficiência cardíaca, insuficiência renal, doença hepática, com necessidade de oxigênio suplementar ou fazendo tratamento com pressão positiva contínua na via aérea (CPAP, do inglês *continuous positive airway pressure*). Também foram excluídos os pacientes que realizaram menos de 4 horas de exame. O comitê de ética local aprovou o estudo sob o protocolo SDC 4515/17/015 e um termo de consentimento livre e esclarecido foi obtido de cada participante.

Tabela 1. Característica da população

	Valor	Intervalo
Gênero (homem/mulher)	49/42	
Idade (anos)	55,8 ± 13,4	29 - 86
IMC (kg/m ²)	30,3 ± 5,5	21,6 - 44,4
Escala de sonolência de Epworth	9,8 ± 4,8	0 - 21
IAH com 3% (eventos/h)	36,7 ± 29,4	0,0 - 116,2

Os valores são apresentados com média ± desvio padrão. Os intervalos são apresentados utilizando os valores máximos e mínimos de cada parâmetro. Abreviações: IMC = Índice de Massa Corporal; IAH = Índices de Apneia e Hipopneia;

2.2. Sistema de monitorização

Todos os pacientes foram submetidos a um exame de polissonografia padrão, que incluiu EEG (C3/A2, C4/A1, O1/A2, O2/A1), eletro-oculografia (EOG), eletromiografia (EMG) tibial anterior esquerda e direita, ECG, esforço toracoabdominal, fluxo de ar oronasal (medição de fluxo baseada em termistor e pressão nasal), SpO₂ com oximetria de pulso e posição corporal (S7000, Embla Systems, EUA). Os dados de SpO₂, o qual utilizamos

nesse trabalho para calcular as dessaturações a partir de diferentes metodologias, foram obtidos com frequência de amostragem de 10Hz.

2.3. Rotinas computacionais

Foi desenvolvida uma rotina computacional utilizando a linguagem de programação Python, em que 3 principais tarefas são executadas (figura 1): decodificação dos arquivos *European Data Format* (EDF) gerados pelo programa do fabricante do equipamento de polissonografia; identificação e contagem das dessaturações de acordo com a metodologia de definição de linha de base escolhida; geração de 2 arquivos: um no formato EDF, com inclusão dos tempos onde foram encontradas as dessaturações, e outro no formato *comma-separated values* (CSV), com os valores de IDO referentes a cada uma das metodologias.

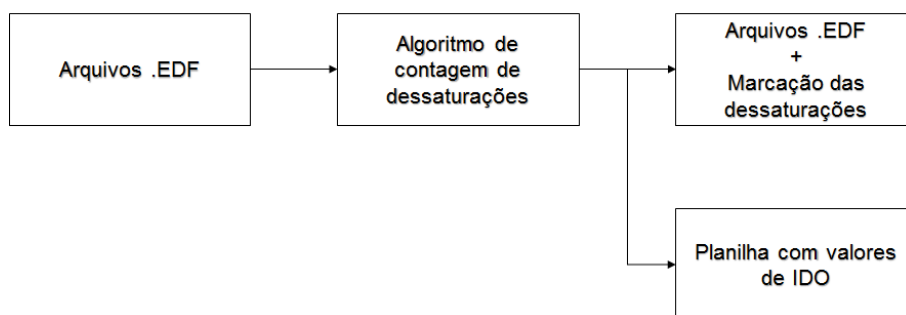
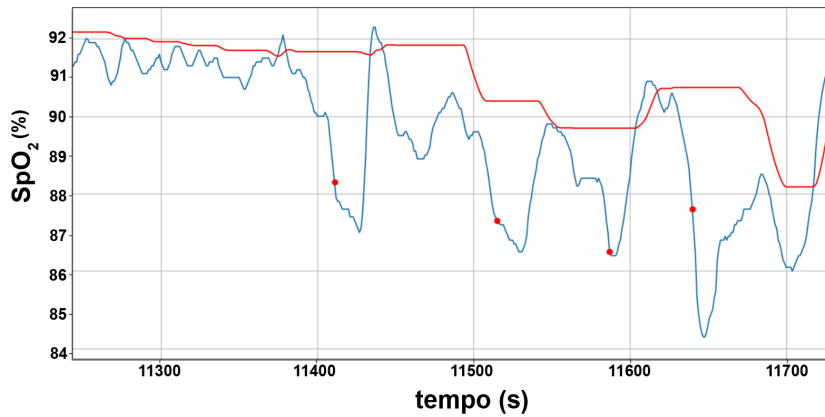


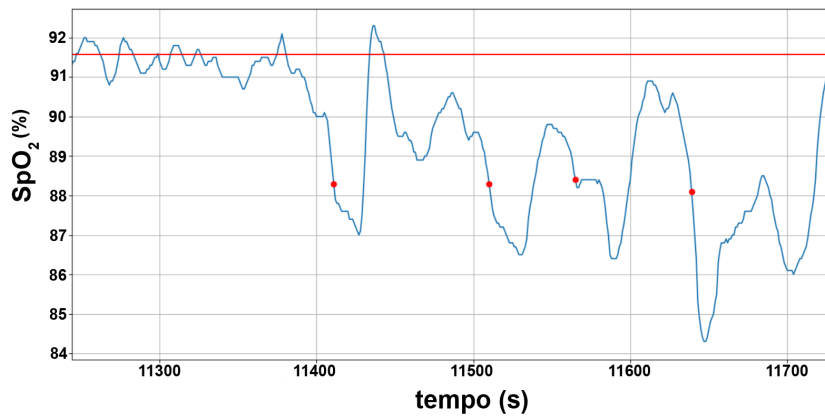
Figura 1. Fluxograma do processamento para contagem de dessaturações

2.4. Metodologias para definição de linha de base

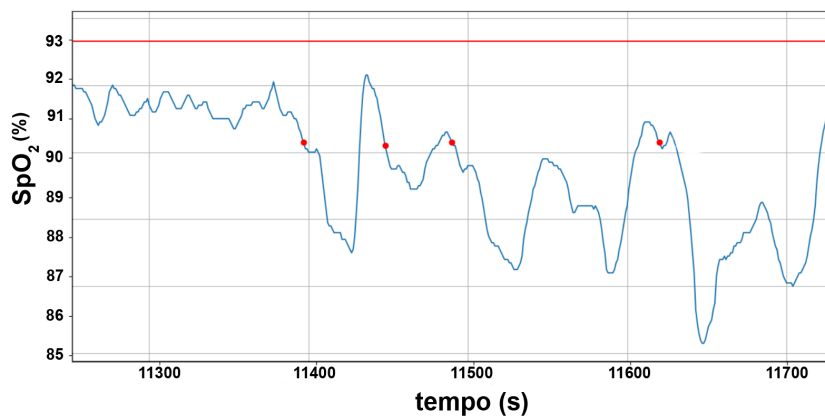
Foram escolhidas três metodologias para cálculo da linha de base. A primeira delas, nomeada neste trabalho como *Linha de Base do Minuto Precedente* (LBMP), utilizou a média dos 20% maiores valores de SpO₂ dentro do minuto que precedeu o ponto da curva de SpO₂ analisado (figura 2a). A segunda metodologia, nomeada como *Linha de Base de Todo o Exame* (LBTE), definiu a linha de base como a SpO₂ média ao longo de todo o exame (figura 2b). Por fim, uma terceira metodologia definiu a linha de base como a média dos 3 minutos iniciais do exame (figura 2c). Esta última foi nomeada como *Linha de Base dos Minutos Iniciais* (LBMI).



(a)



(b)



(c)

Figura 2. Exemplo da detecção de eventos de dessaturação utilizando as metodologias: (a) Linha de Base do Minuto Precedente (LBMP), (b) Linha de Base de Todo o Exame (LBTE) e (c) Linha de Base dos Minutos Iniciais (LBMI). Curva em azul: SpO₂. Curva em vermelho: linha de base calculada. Pontos em vermelho: eventos de dessaturação de oxigênio.

Nas três metodologias, foram detectadas dessaturações com 3% de profundidade em relação à linha de base. Assim, quando um evento de dessaturação era detectado, o algoritmo poderia selecionar um novo evento somente a partir do momento em que a distância da curva em relação à linha de base fosse menor do que 3%. Não foi definida a duração máxima ou mínima para cada evento e não foi utilizado nenhum critério de ressaturação.

2.5. Análise estatística

O teste de Shapiro-Wilk mostrou que os valores de IDO e IAH não apresentam distribuição normal e, portanto, foram utilizados testes estatísticos não paramétricos. Os valores de IDO foram exibidos em um *boxplot*. As diferenças entre as metodologias foram avaliadas por meio do teste de Friedman, com teste *post hoc* de Wilcoxon comparando par a par. Foi determinada a prevalência de AOS para cada uma das metodologias, utilizando as faixas de IDO: menor que 5, entre 5 e 15, entre 15 e 30 e maior que 30. Esses valores foram utilizados pois são os limiares para diagnóstico de apneia normal, leve, moderada e acentuada, respectivamente. [Epstein et al. 2009].

Para analisar a concordância entre os valores de IDO e o IAH, foi utilizada a correlação de Pearson. A Curva Característica de Operação do Receptor (ROC, do inglês *Receiver Operating Characteristic Curve*) foi adotada para o cálculo dos limiares que maximizam a sensibilidade e a especificidade de cada metodologia para diagnóstico de apneia alta ou severa (IAH maior que 15), utilizando o IAH como padrão ouro. Por fim, foram construídos gráficos Bland-Altman entre os valores de IDO e o IAH.

2.6. Resultados

Os valores de IDO para cada uma das metodologias propostas são apresentados no *boxplot* da figura 3. Todos os IDOs foram significativamente diferentes uns dos outros ($p < 0,001$) pelo teste de Friedman e comparações *post hoc*.

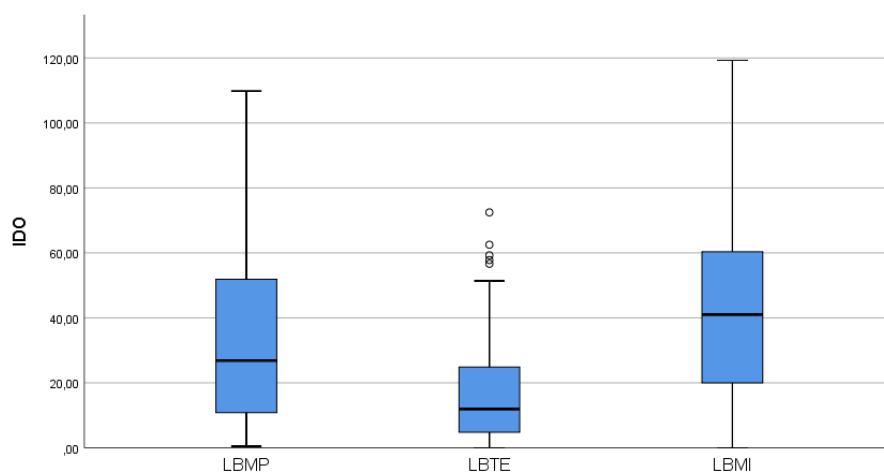


Figura 3. *Boxplot* dos valores de IDO obtidos a partir das metodologias: Linha de Base do Minuto Precedente (LBMP), Linha de Base de Todo o Exame (LBTE) e Linha de Base dos Minutos Iniciais (LBMI).

A figura 4 apresenta a prevalência de AOS ao utilizar as metodologias LBMP, LBTE e LBMI. As faixas de valores utilizadas foram: IDO menor que 5, entre 5 e 15, entre

15 e 30 e maior que 30. Para fins de comparação, também foi apresentada a prevalência utilizando o IAH, para as mesmas faixas de valores.

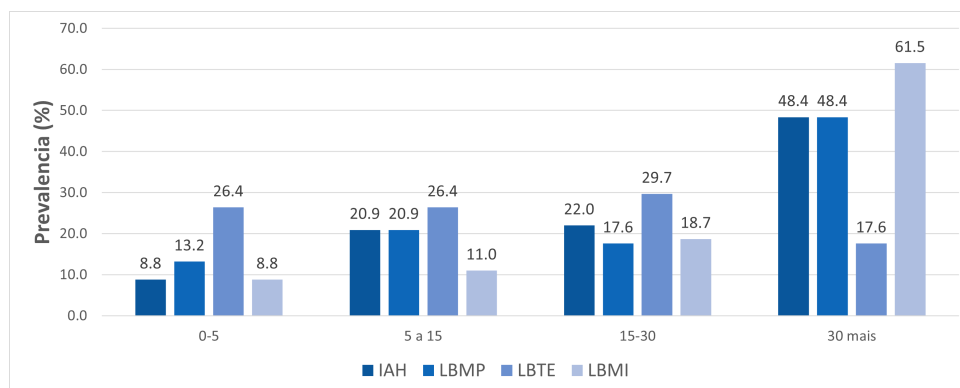
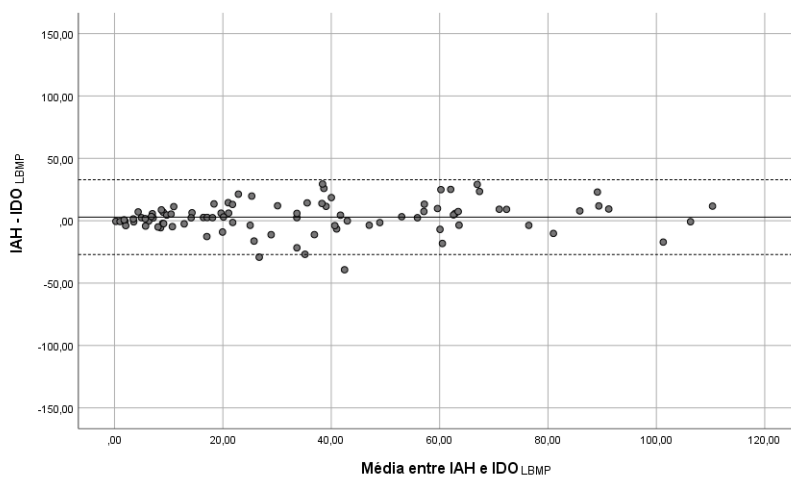


Figura 4. Prevalência de AOS ao utilizar o IAH da polissonografia e o IDO a partir das metodologias LBMP, LBTE e LBMI. As faixas de valores utilizadas foram: normal (0-5), leve (5-15), moderada (15-30) e acentuada (30 ou mais). IAH = Índice de Apneia e Hipopneia; IDO = Índice de Dessaturação de Oxigênio; LBMP = Linha de Base do Minuto Precedente; LBTE = Linha de Base de Todo o Exame; LBMI = Linha de Base dos Minutos Iniciais.

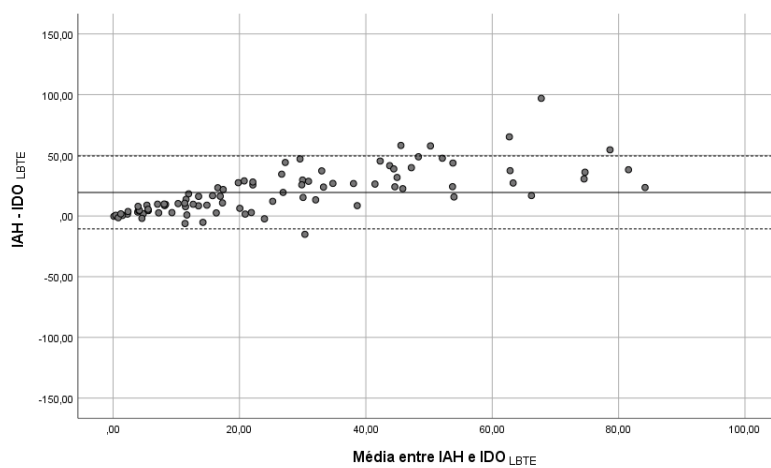
Os limiares que maximizam a sensibilidade e especificidade de cada metodologia para diagnóstico de apneia alta ou severa, assim como os valores destas métricas, são apresentados na tabela 2. Os valores da correlação de Pearson entre IAH e IDO são 0,91, 0,81 e 0,43 para as metodologias LBMP, LBTE e LBMI, respectivamente. Por fim, a figura 5 apresenta os gráficos de Bland-Altman entre o IDO obtido por cada uma das metodologias e o IAH.

Tabela 2. Limiares que maximizam a sensibilidade e especificidade para diagnóstico de apneia alta ou severa utilizando as metodologias: Linha de Base do Minuto Precedente (LBMP), Linha de Base de Todo o Exame (LBTE) e Linha de Base dos Minutos Iniciais (LBMI).

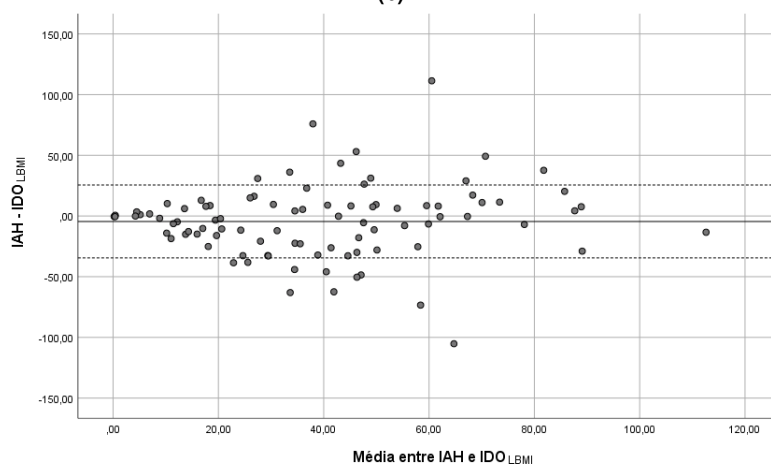
	Limiar equivalente	Sensibilidade (%)	Especificidade (%)
LBMP	15	90,6	92,6
LBTE	12	92,9	92,6
LBMI	31	70,3	66,7



(a)



(b)



(c)

Figura 5. Bland-Altman comparando (a) IAH e IDO calculado com LBMP (IDO_{LBMP}); (b) IAH e IDO calculado com LBTE (IDO_{LBTE}); e (c) IAH e IDO calculado com LBMI (IDO_{LBMI}). IAH = Índice de Apneia e Hipopneia; IDO = Índice de Dessaturação de Oxigênio; LBMP = Linha de Base do Minuto Precedente; LBTE = Linha de Base de Todo o Exame; LBMI = Linha de Base dos Minutos Iniciais.

3. Discussão

Este trabalho pretendeu demonstrar que diferentes definições para o cálculo da linha de base impactam na identificação de eventos de dessaturação de oxigênio e, consequentemente, no valor do IDO em uma população característica (tabela 1). Apesar deste fato, seria possível que o diagnóstico propriamente não fosse impactado, já que a classificação da apneia é feita por intervalo de valores — 0, 5, 15 e 30. Porém, a partir da análise da distribuição dentro de cada um dos grupos (figura 4), foi observada uma diferença significativa dependendo da metodologia empregada.

O algoritmo LBTE apresentou a menor média de IDO, seguido pelo algoritmo LBMP e LBMI, como pode ser observado no *boxplot* da figura 3. O motivo para isso pode estar ligado ao fato de que, quando acontecem muitas dessaturações ao longo do exame, a saturação de oxigênio média sofre uma queda, ou seja, a linha de base pela metodologia LBTE é mais baixa e, assim, são identificados predominantemente eventos com profundidade elevada. A prevalência de pacientes com AOS (IAH maior que 15) neste trabalho foi de 70,3%, ou seja, a maioria dos pacientes teve muitas dessaturações, o que maximizou este efeito.

Em relação ao algoritmo LBMI, como a saturação média nos minutos iniciais de exame é mais alta, já que o paciente ainda está acordado, sem sofrer eventos de apneia/hipopneia, a linha de base tem um valor mais elevado, o que torna mais fácil que as quedas na saturação sejam identificadas como dessaturações. Porém, nos pacientes com AOS severa ou grave, algumas dessaturações são perdidas ao utilizar esta metodologia, pois nem sempre acontece a ressaturação necessária para identificar um novo evento. Esta característica de superestimar em uma situação e subestimar em outra faz com que esta metodologia apresente a pior correlação com o IAH (0,43), bem abaixo das metodologias LBMP (0,91) e LBTE (0,81).

A metodologia LBMP apresentou o melhor desempenho frente às demais. Ela obteve a melhor correlação com o IAH e, conforme apresentado na tabela 2, foi a metodologia que teve a sensibilidade e especificidade maximizada utilizando o limiar de 15, mesmo valor adotado para o IAH. A partir dos gráficos de Bland-Altman da figura 5, é possível observar que a diferença dos limites de concordância de 95% é menor para esta metodologia — 48, 73 e 119 para LBMP, LBTE e LBMI, respectivamente.

Essas diferenças observadas podem levar a diagnósticos distintos, seja para apneia ou qualquer outra doença respiratória que utilize dessaturação para este fim. No geral, laudos de polissonografia apresentam o critério de limiar de queda utilizado para identificar um evento de dessaturação — geralmente, 3% ou 4% — mas não declaram nenhuma das outras características de funcionamento do algoritmo adotado para detecção dos eventos de dessaturação, como a questão da linha de base, por exemplo.

Dessa forma, o resultado sugere uma necessidade de que eventos de dessaturação de oxigênio sejam melhor definidos em manuais e na literatura. Assim, diferentes pesquisadores, fabricantes de polissonografia e equipamentos relacionados poderão basear seus algoritmos de maneira mais uniforme entre si, de modo a garantir a reprodutibilidade do IDO. Os resultados obtidos neste trabalho trazem evidências de que não é suficiente que o oxímetro seja um dispositivo rigidamente calibrado e certificado pelos órgãos reguladores, se a manipulação feita com o sinal de oximetria não for devidamente padronizada.

Referências

- Alvarez-Estevez, D. and Moret-Bonillo, V. (2015). Computer-Assisted Diagnosis of the Sleep Apnea-Hypopnea Syndrome: A Review. *Sleep Disorders*, 2015:1–33.
- Berry, R. B., Budhiraja, R., Gottlieb, D. J., Gozal, D., Iber, C., Kapur, V. K., Marcus, C. L., Mehra, R., Parthasarathy, S., Quan, S. F., Redline, S., Strohl, K. P., Ward, S. L. D., and Tangredi, M. M. (2012). Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. *Journal of Clinical Sleep Medicine*, 08(05):597–619.
- Chiang, L. K., Kam, C. W., and Ng, L. V. (2017). Overnight pulse oximetry for screening of obstructive sleep apnea in at-risk adult patients in the primary care setting: Prospective case series. *Family Medicine and Community Health*, 5(3):215–222.
- Corral, J., Sanchez-Quiroga, M. A., Carmona-Bernal, C., Sanchez-Armengol, A., De La Torre, A. S., Duran-Cantolla, J., Egea, C. J., Salord, N., Monasterio, C., Teran, J., Luz Alonso-Alvarez, M., Muñoz-Mendez, J., Arias, E. M., Cabello, M., Montserrat, J. M., De La Peña, M., Serrano, J. C., Barbe, F., and Masa, J. F. (2017). Conventional polysomnography is not necessary for the management of most patients with suspected obstructive sleep apnea noninferiority, randomized controlled trial. *American Journal of Respiratory and Critical Care Medicine*, 196(9):1181–1190.
- Epstein, L. J., Kristo, D., Strollo, P. J., Friedman, N., Malhotra, A., Patil, S. P., Ramar, K., Rogers, R., Schwab, R. J., Weaver, E. M., and Weinstein, M. D. (2009). Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *Journal of Clinical Sleep Medicine*, 5(3):263–276.
- Erman, M. K., Stewart, D., Einhorn, D., Gordon, N., and Casal, E. (2007). Validation of the ApneaLink™ for the screening of sleep apnea: A novel and simple single-channel recording device. *Journal of Clinical Sleep Medicine*, 3(4):387–392.
- Flemons, W. W., Buysse, D., Redline, S., Oack, A., Strohl, K., Wheatley, J., Young, T., Douglas, N., Levy, P., McNicolas, W., Fleetham, J., White, D., Schmidt-Nowarra, W., Carley, D., and Romaniuk, J. (1999). Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research. *Sleep*, 22(5):667–689.
- Hirotsu, C., Haba-Rubio, J., Andries, D., Tobback, N., Marques-Vidal, P., Vollenweider, P., Waeber, G., and Heinzer, R. (2019). Effect of three hypopnea scoring criteria on OSA prevalence and associated comorbidities in the general population. *Journal of Clinical Sleep Medicine*, 15(2):183–194.
- Jones, S. (2009). Sleep Disordered Breathing and Mortality: Eighteen-Year Follow-up of the Wisconsin Sleep Cohort. *Yearbook of Pulmonary Disease*, 31(8):1071–1078.
- Ling, I. T., James, A. L., and Hillman, D. R. (2012). Interrelationships between Body Mass, Oxygen Desaturation, and Apnea-Hypopnea Indices in a Sleep Clinic Population. *Sleep*, 35(1):89–96.
- Nigro, C. A., Dibur, E., and Rhodius, E. (2011). Pulse Oximetry for the Detection of Obstructive Sleep Apnea Syndrome: Can the Memory Capacity of Oxygen Saturation Influence Their Diagnostic Accuracy? *Sleep Disorders*, 2011:1–9.

- Rashid, N. H., Zaghi, S., Scapuccin, M., Camacho, M., Certal, V., and Capasso, R. (2021). The Value of Oxygen Desaturation Index for Diagnosing Obstructive Sleep Apnea: A Systematic Review. *Laryngoscope*, 131(2):440–447.
- Romem, A., Romem, A., Koldobskiy, D., and Scharf, S. M. (2014). Diagnosis of obstructive sleep apnea using pulse oximeter derived photoplethysmographic signals. *Journal of Clinical Sleep Medicine*, 10(3):285–290.
- Yamashiro, Y. and Kryger, M. H. (1995). Nocturnal oximetry: Is it a screening tool for sleep disorders? *Sleep*, 18(3):167–171.