

Amanda Feliciano¹ 
Mara Behlau¹ 
Jordana Balbinot¹ 
Samuel Ribeiro de Abreu² 
Vanessa Veis Ribeiro³ 
Leonardo Wanderley Lopes⁴ 

Descritores

Voz
Acústica
Distúrbios da Voz
Disfonia
Diagnóstico

Keywords

Voice
Acoustic
Voice Disorders
Dysphonia
Diagnostic

Endereço para correspondência:

Leonardo Wanderley Lopes
Departamento de Fonoaudiologia,
Centro de Ciências da Saúde,
Universidade Federal da Paraíba –
UFPB
Cidade Universitária - Campus 1,
Bairro Castelo Branco, João Pessoa
(PB), Brasil, CEP: 58051-900.
E-mail: lwlopes@hotmail.com

Recebido em: Agosto 04, 2025

Aceito em: Dezembro 15, 2025

Editora: Ana Carolina Constantini.

Acurácia e ponto de corte de diferentes medidas acústicas na tarefa de vogal sustentada em falantes do português brasileiro

Accuracy and cut-off values of different acoustic measures in the sustained vowel task of Brazilian Portuguese speakers

RESUMO

Objetivo: Definir os valores de corte das medidas acústicas extraídas da tarefa de vogal sustentada [a] para discriminação entre presença e ausência de distúrbio vocal em falantes do português brasileiro. **Método:** A amostra foi composta por 376 participantes, 288 do gênero feminino e 88 do gênero masculino, com idade média de $41,20 \pm 14,04$ anos. Os indivíduos foram submetidos ao exame de avaliação laringológica, e a uma gravação vocal, sendo coletada uma amostra de fala com a vogal [a] sustentada e posteriormente realizado o julgamento perceptivo auditivo e acústico da voz, dos quais 277 foram classificados com distúrbio vocal e 99 sem distúrbio vocal. Foram extraídas 44 medidas acústicas e analisadas por meio da curva ROC para identificação do ponto de corte de cada medida, bem como suas métricas de desempenho. **Resultados:** Foram definidos pontos de corte para 23 das 44 medidas acústicas inicialmente extraídas. Dentre elas, seis medidas demonstraram desempenho robusto e clinicamente relevante para discriminar vozes com e sem distúrbio vocal em falantes do português brasileiro. As medidas com valores de AUC superiores a 0,70 e sensibilidade e especificidade equilibradas acima de 70% foram: Amplitude Variability Index (AVI) ($\leq 0,91$), Glottal-to-noise excitation (GNE)2000Hz ($> 0,92$), Spectral Noise Level (SNL)100-2600Hz ($\leq 7,71$), SNL100-3000Hz ($\leq 3,38$), SNL100-8000Hz ($\leq 10,95$) e SNL5100-8000Hz ($\leq -1,84$). **Conclusão:** Seis medidas apresentaram desempenho satisfatório para identificar a presença de distúrbio vocal em falantes do português brasileiro. Dez medidas apresentaram bom desempenho para triagem vocal e sete para confirmação diagnóstica de distúrbio vocal.

ABSTRACT

Purpose: To define cut-off values for acoustic measures extracted from the sustained vowel [a] task to discriminate between the presence and absence of vocal disorders in Brazilian Portuguese speakers. **Methods:** The sample consisted of 376 participants, 288 female and 88 male, with a mean age of 41.20 ± 14.04 years. All individuals underwent a laryngological evaluation and a voice recording session, during which a sustained vowel [a] sample was collected. Subsequently both auditory-perceptual and acoustic analyses of the voice were performed. Among the participants 277 were classified as having a vocal disorder and 99 as having no vocal disorder. A total of 44 acoustic measures were extracted and analyzed using ROC curves to determine the cut-off value for each measure, along with their performance metrics (sensitivity, specificity and AUC). **Results:** Cut-off values were established for 23 of the 44 extracted acoustic measures. Six measures showed robust and clinically relevant performance in discriminating between voices with and without vocal disorders in Brazilian Portuguese speakers. The measures with AUC values greater than 0.70 and balanced sensitivity and specificity above 70% were: Amplitude Variability Index (AVI) (≤ 0.91), Glottal-to-noise excitation (GNE)2000Hz (> 0.92), Spectral Noise Level (SNL)100-2600Hz (≤ 7.71), SNL100-3000Hz (≤ 3.38), SNL100-8000Hz (≤ 10.95) and SNL5100-8000Hz (≤ -1.84). **Conclusion:** Six acoustic measures demonstrated satisfactory performance in identifying the presence of vocal disorders in Brazilian Portuguese speakers. Ten measures showed good performance for voice screening purposes and seven for diagnostic confirmation of vocal disorders.

Trabalho realizado no Centro de Estudos da Voz – CEV - São Paulo (SP), Brasil.

¹ Centro de Estudos da Voz – CEV - São Paulo (SP), Brasil.

² Universidade Federal de Rondonópolis – UFR - Rondonópolis (MS), Brasil.

³ Departamento de Fonoaudiologia, Faculdade de Ciências e Tecnologias em Saúde, Universidade de Brasília – UnB - Brasília (DF), Brasil.

⁴ Departamento de Fonoaudiologia, Centro de Ciências da Saúde, Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.

Fonte de financiamento: nada a declarar.

Conflito de interesses: nada a declarar.

Disponibilidade de Dados: Os dados de pesquisa estão disponíveis no corpo do artigo.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/>), que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

INTRODUÇÃO

A análise acústica faz parte dos procedimentos de avaliação multidimensional da voz em conjunto com o julgamento perceptivo auditivo (JPA), a análise aerodinâmica, o exame estrutural e vibratório da laringe, e a autoavaliação vocal⁽¹⁾. A partir da análise acústica é possível obter medidas para quantificar as características do sinal vocal e realizar a avaliação descritiva dos padrões visuais desses sinais⁽²⁾. Isso contribui para triagem, confirmação diagnóstica e para o monitoramento da eficácia das intervenções terapêuticas em sujeitos com distúrbio vocal^(3,4).

Há uma grande diversidade de medidas acústicas para avaliação do sinal vocal. Dentre as medidas convencionais e mais amplamente utilizadas estão a frequência fundamental ou de oscilação (f_0), medidas de curto termo de perturbação de frequência (*jitter*), de intensidade vocal (*shimmer*) e Proporção Harmônico Ruído (*Harmonic-Noise Ratio* - HNR)⁽¹⁾. As medidas convencionais são eficazes, sobretudo, em casos de desvios vocais leves. Já as medidas espectrais e cepstrais, que vem sendo mais amplamente utilizadas e estudadas nos últimos anos, resultam da dupla transformação do sinal acústico da voz, que revelam o grau de organização dos harmônicos no espectro, sendo efetivas para análise de vozes desviadas, visto que são fortes preditoras da presença de desvio vocal⁽⁵⁾.

A aplicabilidade clínica das medidas acústicas depende da definição de valores de corte conhecidos e válidos, com desempenho diagnóstico consistente e adequado à população-alvo. No contexto do português brasileiro (PB), poucas medidas acústicas têm valores de corte definidos, o que limita a comparação intersujeitos e a comparação da voz do paciente com valores de referência⁽⁶⁾, não contribuindo para uma prática clínica embasada e consciente.

Recentemente, foram estabelecidos valores de corte para dois índices multiparamétricos, o *Acoustic Voice Quality Index* (AVQI) e o *Acoustic Breathiness Index* (ABI) para auxiliar na identificação de alterações vocais em falantes do PB⁽⁷⁾. A acurácia diagnóstica do AVQI foi excelente e o ponto de corte foi de 1,33, apresentando sensibilidade de 78,8% e especificidade de 90,6%. Já o ABI apresentou acurácia diagnóstica boa, com ponto de corte de 2,94, sensibilidade de 75,3% e especificidade de 93,4%⁽⁷⁾. Esses dados evidenciam a robustez do AVQI e do ABI como ferramentas objetivas para a avaliação e monitoramento de distúrbios vocais^(7,8).

Outras duas medidas têm ganhado destaque como parâmetros objetivos na análise acústica da voz, o *Cepstral Peak Prominence* (CPP) e o *Smoothed Cepstral Peak Prominence* (CPPS), visto que apresentam boa correlação com o julgamento perceptivo-auditivo e alta sensibilidade na detecção de desvios vocais e não dependem da extração da frequência fundamental. Há valores de corte para discriminar vozes com e sem distúrbio vocal no PB⁽⁶⁾. O estudo identificou que valores abaixo de 28,15 decibéis (dB) (vogal [ɛ]), 28,77 dB (vogal [a]) e 28,58 dB (frase *Consensus Auditory-Perceptual Evaluation of Voice* - CAPEV) e para o CPPS valores abaixo de 16,42 dB (vogal [ɛ]), 17,02 dB (vogal [a]) e 11,30 dB (todas as sentenças do CAPEV) são indicativos de presença de distúrbio vocal⁽⁶⁾.

O ponto de corte é extraído a partir da melhor relação entre sensibilidade e especificidade. Nessa análise, é essencial a

identificação de medidas com desempenho equilibrado, que apresentem simultaneamente valores satisfatórios de sensibilidade e especificidade. Essas medidas têm a capacidade de identificar corretamente tanto indivíduos com distúrbios vocais (verdadeiros positivos) quanto indivíduos sem a condição (verdadeiros negativos), reduzindo a ocorrência de falsos diagnósticos. Esse equilíbrio é geralmente representado por valores de área sob a curva (AUC - *Area Under the Curve*) superiores a 0,70 em análises ROC, indicando bom desempenho discriminativo⁽¹⁾.

A classificação quanto à presença ou ausência de condições específicas de saúde depende da precisão do procedimento de medição ou teste. Tal precisão reflete a concordância entre os achados de um novo teste, chamado teste índice, e os obtidos por meio do método mais confiável existente para detectar o desfecho de saúde em questão, o padrão de referência⁽⁹⁾. Portanto, a seleção de um padrão de referência apropriado é imperativa para validar as métricas de acurácia de um teste diagnóstico emergente. Em relação à escolha do teste de referência, os clínicos enfrentam obstáculos ao estabelecer critérios para normalidade vocal devido ao seu caráter intrinsecamente multidimensional. Atualmente, não existe um único método de referência capaz de diferenciar com precisão indivíduos vocalmente saudáveis daqueles com distúrbio vocal⁽¹⁰⁾. Na ausência de evidências científicas robustas e consenso entre especialistas sobre a detecção de certas condições de saúde, padrões de referência alternativos devem ser desenvolvidos⁽⁹⁾.

Embora o JPA e a autoavaliação sejam tradicionalmente utilizadas como métodos preliminares para identificar distúrbios vocais⁽¹¹⁾, a inspeção visual da laringe (incluindo laringoscopia e estroboscopia) é considerada o padrão-ouro para confirmar a categorização diagnóstica dos distúrbios da voz⁽¹⁰⁾. A síntese de informações de procedimentos distintos para construir um critério de referência integrativo, derivado da correlação entre os resultados de uma avaliação visual da laringe (realizada por um médico) e da JPA (realizada por um fonoaudiólogo), pode servir para avaliar a capacidade discriminativa de medidas acústicas na distinção entre indivíduos com e sem distúrbios vocais.

Na prática clínica da área de voz, pode-se utilizar diferentes tipos de tarefas para a identificação de distúrbios vocais. O tipo de tarefa selecionado para a análise acústica, influencia diretamente no seu resultado. A vogal [a] é amplamente utilizada na análise acústica e espectrográfica vocal devido à sua produção estável e representatividade da ressonância vocal. Estudos recentes indicam que a escolha da vogal pode influenciar significativamente a precisão diagnóstica, sendo [a] a mais adotada internacionalmente por facilitar a padronização dos procedimentos e a comparação entre resultados⁽¹²⁾. A predominância da vogal [a] em protocolos acústicos está também associada à facilidade de obtenção de amostras de fala sustentada com mínima interferência articulatória, o que favorece a análise espectrográfica detalhada dos formantes e das características temporais da voz⁽¹³⁾. Porém, observa-se na literatura a ausência de alguns valores de corte e a necessidade de identificar e validar medidas acústicas com desempenho diagnóstico robusto para a tarefa da vogal sustentada [a] em falantes do PB.

Nesse contexto, o objetivo do presente estudo foi definir os valores de corte das medidas acústicas extraídas na tarefa

de vogal sustentada [a] para discriminação entre presença e ausência de distúrbio vocal em falantes do PB.

MÉTODO

Trata-se de um estudo transversal, retrospectivo e descritivo. O estudo seguiu as diretrizes do *Standards for Reporting Diagnostic Accuracy* (STARD)⁽¹⁴⁾. O estudo foi aprovado pelo Comitê de Ética em Pesquisa da instituição responsável sob o número 02954612.9.0000.5188. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido (TCLE).

No presente estudo o teste índice foi a análise acústica da vogal sustentada [a] e o teste padrão de referência (*gold standard*) foi a presença ou ausência de distúrbio vocal, definida com base na avaliação clínica composta por exame laringológico e JPA. Assim, apesar do teste índice avaliar a voz (medidas acústicas), a condição-alvo que buscou-se detectar/discriminar foi a presença de distúrbio vocal. O presente estudo utilizou um resultado derivado da concordância entre o exame visual da laringe e os resultados da JPA como padrão de referência para determinar o desfecho clínico, ou seja, a presença ou ausência de distúrbio vocal.

A amostra do estudo foi constituída a partir de critérios de inclusão e exclusão aplicados por meio dos procedimentos que garantiram a análise do padrão de referência. Foram incluídos no estudo adultos entre 18 e 65 anos que haviam realizado avaliação otorrinolaringológica nas duas semanas anteriores à coleta dos dados, com diagnóstico confirmado de distúrbio vocal, e que não haviam iniciado tratamento vocal (terapia ou cirurgia) antes da coleta. Foram excluídos participantes com distúrbios cognitivos ou neurológicos que impedissem o uso dos procedimentos de gravação, usuários profissionais da voz, e indivíduos que já haviam feito terapia vocal formal ou cirurgia na região da cabeça ou pescoço. Além disso, também foram aplicados critérios de exclusão relacionados à qualidade dos sinais vocais, excluindo-se amostras de voz que possuíam tempo máximo de fonação menor do que 5 segundos, presença de pico de corte no sinal acústico e relação sinal-ruído (RSR) abaixo de 30 dB de nível de pressão sonora (NPS)⁽¹¹⁾.

Como o objetivo da pesquisa foi definir os valores de corte das medidas acústicas extraídas na tarefa de vogal sustentada [a] para discriminação entre vozes com e sem distúrbio vocal em falantes do PB, os critérios de elegibilidade permitiram a inclusão de indivíduos em ambos os grupos. O grupo sem distúrbio vocal foi composto por indivíduos sem queixas vocais, distúrbios laríngeos estruturais ou funcionais e sem desvio na qualidade vocal conforme demonstrado pelo JPA. O grupo com distúrbio vocal foi composto por indivíduos com queixas vocais, distúrbio estrutural ou funcional da laringe e desvio na qualidade vocal detectado pelo JPA.

Foi realizado um cálculo do tamanho da amostra pelo tamanho da população (Equação 1) para estimar quantos participantes eram necessários para cada grupo de pesquisa. Os parâmetros utilizados foram: tamanho da população com distúrbio vocal (N) no Brasil de 15.232.500 (203.100.000*0,075)^(15,16), margem de erro (e) de 10%, nível de confiança de 95% (z=1,96) e proporção populacional de indivíduos com distúrbio vocal não

professores (p) de 0,5. Calculou-se um tamanho amostral (n) de 97 participantes, por grupo.

$$\text{Tamanho da amostra}(n) = \frac{z^2 \cdot p(1-p)}{e^2} \div \left(1 + \frac{z^2 \cdot p(1-p)}{e^2 N} \right) \quad (1)$$

Sendo:

z = 1,96

p = 0,5

e = 0,1

N = 15.232.000 (7,5% de disfônicos não professores em 203.100.000 habitantes no Brasil)

Inicialmente foram incluídos 391 indivíduos (304 mulheres e 87 homens). Desses, 15 foram excluídos com base nos seguintes critérios: foram excluídos três indivíduos com diagnóstico inconclusivo de distúrbio de voz para evitar inconsistência no conjunto de dados e manter a relevância clínica dos resultados do estudo; 12 indivíduos eram profissionais da voz. Assim, a amostra do estudo foi composta por 376 participantes: 277 pacientes com distúrbio vocal (217 mulheres e 60 homens) e 99 indivíduos sem distúrbio vocal (71 mulheres e 28 homens), com média de idade de 41,20 ± 14,04 anos.

O recrutamento dos participantes com distúrbio vocal foi realizado a partir da lista de espera de um laboratório de voz de uma universidade. Todos os pacientes buscaram o serviço por demanda espontânea ou foram encaminhados por otorrinolaringologistas. Os participantes sem distúrbio vocal foram recrutados entre estudantes e funcionários da universidade onde a pesquisa foi realizada.

Os procedimentos de coleta de amostras vocais seguiram o protocolo rotineiro de avaliação vocal inicial do laboratório. Todos os dados foram registrados digitalmente (sinais de voz) ou em prontuário (anamnese e laudo do exame laríngeo). As amostras vocais foram coletadas durante a primeira avaliação clínica vocal, antes do início da terapia vocal. Todos os participantes responderam a uma breve entrevista de histórico médico e foram encaminhados para exame laríngeo visual por um otorrinolaringologista, com solicitação de retorno do laudo em até 15 dias.

Os seguintes equipamentos foram utilizados para gravação: *software Fonoview* (versão 4.5, CTS Informática), computador Dell *all-in-one*, microfone unidirecional cardioide Sennheiser modelo E-835 (em pedestal), acoplado a um pré-amplificador Behringer U-Phoria UMC 204. As gravações foram feitas em cabine acústica com ruído de fundo inferior a 20 dB SPL, taxa de amostragem de 44.000 Hz e 16 bits. O microfone foi posicionado a 10 cm da boca do falante em um ângulo de 45°. A distância foi medida com régua antes de cada tarefa, e o participante era posicionado no limite traseiro da cabine para manter a distância fixa. Orientou-se que o participante não movimentasse a cabeça.

Todos os participantes registraram a tarefa de fala da vogal sustentada [a], em frequência e intensidade habitual auto selecionada. Durante a gravação, os sinais foram monitorados visualmente no *software Fonoview*, observando a duração das vogais (mínimo de 5 s) e a presença de *peak clipping*.

Quando necessário, a gravação foi repetida para garantir a qualidade dos registros. Após a coleta, os participantes foram encaminhados para exame médico otorrinolaringológico com videolaringostroboscopia. Tais exames foram realizados no mesmo laboratório e resultaram em laudo escrito.

Para o JPA as amostras vocais foram apresentadas a três fonoaudiólogos especialistas em voz, com mais de 20 anos de experiência em voz. A sessão de JPA ocorreu em um ambiente silencioso, com fones de ouvido acoplados a um *laptop*, em volume confortável, autorreferido pelos avaliadores. Para o JPA, os fonoaudiólogos escutaram a amostra vocal e na sequência faziam a marcação em uma Escala Visual Analógica (EVA) com intervalo entre 0 e 100 mm. Os valores de corte da EVA⁽¹⁷⁾ foram utilizados para classificar as vozes quanto à presença de desvio vocal e o grau geral (GG). Uma pontuação mais próxima de 0 representa menor desvio na qualidade vocal, e quanto mais próxima de 100, mais desviante é a qualidade vocal⁽¹⁷⁾. Sinais avaliados com valores $\leq 35,5$ mm foram considerados sem distúrbio vocal, enquanto vozes com valores $> 35,5$ mm foram classificadas com distúrbio vocal. Ao final da sessão de JPA, 20% (n=76) das amostras vocais foram repetidas aleatoriamente para analisar a confiabilidade intra-avaliador usando o coeficiente kappa de Cohen. Selecionamos os resultados com a maior confiabilidade intra-avaliador, que obtiveram um coeficiente kappa de 0,89, indicando boa concordância.

Antes de realizar a extração das medidas acústicas analisadas neste estudo, utilizamos o software de acesso aberto Praat⁽¹⁸⁾ (Paul Boersma e David Weenink, Universidade de Amsterdã, Holanda), versão 6.2.10, para obter a RSR dos sinais vocais. A RSR média das amostras incluídas nesta pesquisa foi de $43,08 \pm 8,02$ dB NPS. Este valor é aceitável (RSR > 30 dB NPS) e garante a qualidade dos sinais para a extração das medidas acústicas⁽¹¹⁾. Nenhum dos sinais vocais utilizados neste estudo apresentou RSR < 30 dB NPS. As medidas foram extraídas automaticamente por meio de um script no software gratuito Praat versão 6.2.10, o *VoxMore*⁽¹⁹⁾. Foram extraídas 44 medidas acústicas, incluindo:

- Estatísticas da frequência fundamental (f_0) – oito medidas: média (f_0 media), mediana (f_0 mediana), desvio-padrão (DPf $_0$) primeiro (f_0 q1) e terceiro quartil (f_0 q3), mínima (f_0 min), máxima (f_0 max) e coeficiente de variação (f_0 CV);
- Estatísticas do período – três medidas: média do período (PER), desvio-padrão do período (PSD) e logaritmo natural do desvio-padrão do período (LNPSD);
- Perturbação de curto termo de f_0 – cinco medidas: *jitter* local, *jitter* absoluto (*jitter*_{ABS}), *jitter* relative average perturbation (*jitter*_{RAP}), *jitter* period perturbation quotient (*jitter*_{PPQ5}), *jitter* difference of differences in pitch (*jitter*_{ddp});
- Perturbação de curto termo da amplitude – sete medidas: *shimmer*_{local}, *shimmer*_{dB}, *shimmer*_{APQ1}, *shimmer*_{APQ3}, *shimmer*_{APQ5}, *shimmer*_{APQ11} e *shimmer*_{DDA}, amplitude variability index (AVI);

- Medidas espectrais e cepstrais – 16 medidas: declínio espectral, queda espectral, seis variações de *spectral noise level* (SNL_{100-2600 Hz}, SNL_{100-3000 Hz}, SNL_{100-5100 Hz}, SNL_{100-8000 Hz}, SNL_{2600-5100 Hz} e SNL_{5100-8000 Hz}), *high-frequency noise* (Hfno), diferença entre a amplitude do primeiro e do segundo harmônico (H1H2), diferença entre a amplitude do primeiro formante e a amplitude do primeiro harmônico (H1A1), diferença entre a amplitude do terceiro formante e a amplitude do primeiro harmônico (H1A3), *spectral flatness of residue signal* (SFR), desvio-padrão do *harmonic-to-noise ratio* (HNRDP), *harmonic-to-noise ratio* Dejonckere (HNRD) e *harmonic-to-noise-ratio* (HNR).

- Filtragem inversa – uma medida: *Pitch amplitud* (PA).

- Perturbações na forma da onda – quatro medidas: autocorrelação e três variações do *Glottal to Noise Excitation*, incluindo GNE_{1000Hz}, GNE_{2000Hz} e GNE_{3000Hz}.

Para gerar a curva ROC foi calculada a área sob a curva ROC a partir do melhor ponto de corte para as medidas acústicas analisadas da vogal sustentada e obteve-se suas métricas de desempenho. A AUC trata-se de uma medida de desempenho diagnóstico de um teste e pode variar de 0,5 a 1, quanto mais alta, melhor o desempenho do teste, sendo classificada como: excelente ($> 0,90$), boa (0,80-0,90), aceitável (0,70-0,80), ruim (0,60-0,70) e nenhuma capacidade de discriminação aceitável ($< 0,60$)⁽²⁰⁾.

Além da AUC, o desempenho das medidas analisadas levou em consideração o valor preditivo positivo (+PV), o valor preditivo negativo (-PV), a razão de verossimilhança positiva (+LR) e a razão de verossimilhança negativa (-LR). No presente estudo, serão apresentados apenas os resultados das medidas acústicas com valores de acurácia, sensibilidade e/ou especificidade $\geq 0,70$, considerados na presente pesquisa como satisfatórios. Valores menores que 0,70 são considerados sem capacidade discriminatória aceitável para detectar vozes com e sem distúrbio vocal⁽²⁰⁾. Assim, das 44 medidas extraídas e analisadas, foi possível definir ponto de corte e índices de acurácia diagnóstica - AUC, sensibilidade, especificidade, +LR, -LR, +PV e -PV de apenas 23 medidas acústicas.

Todas as análises foram realizadas utilizando o software *Statistical Package for the Social Sciences* (SPSS), versão 2.0. O nível de significância adotado foi de 5%.

RESULTADOS

A Tabela 1 foi subdividida em três sessões onde podem ser encontrados os valores de corte e as métricas de desempenho das 23 medidas acústicas que apresentaram AUC acima de 0,7, dentre as 44 analisadas. Foram definidos os seguintes pontos de corte: seis medidas acústicas apresentaram desempenho satisfatório com AUC $> 0,70$, sensibilidade e especificidade equilibradas e superiores a 70%: AVI ($\leq 0,91$), GNE_{2000Hz} ($> 0,92$), SNL_{100-2600Hz} ($\leq 7,71$), SNL_{100-3000Hz} ($\leq 3,38$), SNL_{100-8000Hz} ($\leq 10,95$) e SNL_{5100-8000Hz} ($\leq -1,84$) (Figura 1). Dez medidas apresentaram

Tabela 1. Ponto de corte das medidas acústicas extraídas a partir da vogal [a] sustentada

Medidas Acústicas	CUT-OFF	AUC	Sensibilidade (%)	Especificidade (%)	+LR	-LR	+PV	-PV
AVI	$\leq 0,91$	0,779	70,97	76,49	3,02	0,38	25,9	95,8
GNE _{2000 Hz}	$>0,92$	0,785	70,97	73,61	2,69	0,39	23,7	95,7
SNL _{100-2600 Hz}	$\leq 7,71$	0,808	83,87	72,12	3,01	0,22	25,7	97,5
SNL _{100-3000 Hz}	$\leq 3,38$	0,816	77,42	80,67	4	0,28	31,6	96,9
SNL _{100-8000 Hz}	$\leq 10,95$	0,814	90,32	70,63	3,08	0,14	26,2	98,4
SNL _{5100-8000 Hz}	$\leq -1,84$	0,747	70,97	73,61	2,69	0,39	23,7	95,7
GNE _{1000Hz}	$>0,92$	0,76	83,87	54,65	1,85	0,3	17,6	96,7
HNR	$>20,73$	0,837	90,32	64,68	2,56	0,15	22,8	98,3
LNPSD	$\leq -9,58$	0,722	93,55	48,88	1,83	0,13	17,5	98,5
PA	$>0,91$	0,709	70,97	61,34	1,84	0,47	17,5	94,8
PSD	$\leq 0,00$	0,722	93,55	48,88	1,83	0,13	17,5	98,5
SHIMMER _{APQ 5}	$\leq 2,26$	0,779	93,55	51,12	1,91	0,13	18,1	98,6
SHIMMER _{APQ 11}	$\leq 2,78$	0,792	100	51,31	2,05	0	19,3	100
SHIMMER _{DB}	$\leq 0,24$	0,78	77,42	66,04	2,28	0,34	20,9	96,2
SHIMMER LOCAL	$\leq 2,78$	0,776	77,42	65,67	2,26	0,34	20,7	96,2
SNL _{100-5100 HZ}	$\leq 11,44$	0,807	90,32	69,52	2,96	0,14	25,5	98,4
GNE _{3000 Hz}	$>0,9$	0,76	64,52	82,16	3,62	0,43	29,4	95,3
HNR D	$>29,78$	0,72	61,29	76,21	2,58	0,51	22,9	94,5
JITTER LOCAL	$\leq 0,31$	0,712	67,74	70,9	2,33	0,46	21,2	95
JITTER _{PPQ5}	$\leq 0,18$	0,725	67,74	70,9	2,33	0,46	21,2	95
SHIMMER _{APQ 3}	$\leq 1,18$	0,765	64,52	74,63	2,54	0,48	22,7	94,8
SHIMMER _{DDA}	$\leq 3,54$	0,765	64,52	75	2,58	0,47	23	94,8
SNL _{2600-5100 HZ}	$\leq 1,21$	0,743	54,84	83,27	3,28	0,54	27,4	94,1

Legenda: CUT-OFF: Valor de corte para a medida acústica. AUC (Area Under the Curve): Área sob a curva. +LR: Razão de verossimilhança positiva. -LR: Razão de verossimilhança negativa. +PV: Valor preditivo positivo. -PV: Valor preditivo negativo. AVI – Amplitude Variability Index. GNE – Glottal-to-Noise Excitation - GNE_{1000Hz}, GNE_{2000Hz}, GNE_{3000Hz}. SNL (Spectral Noise Level) - SNL_{100-2600 Hz}, SNL_{100-3000 Hz}, SNL_{100-5100 Hz}, SNL_{100-8000 Hz}, SNL_{2600-5100 Hz}, SNL_{5100-8000 Hz}. HNR – Harmonic-to-noise-ratio. LNPSD – Logaritmo natural do desvio-padrão do período. PA – Pitch amplitud. PSD – Desvio-padrão do período. Shimmer_{APQ5}, shimmer_{APQ11}, shimmer_{dB}, shimmer_{local}, shimmer_{APQ3}, shimmer_{DDA}. HNR D – Harmonic-to-noise ratio Dejonckere. Jitter local, Jitter_{PPQ5} - jitter period perturbation quotient.

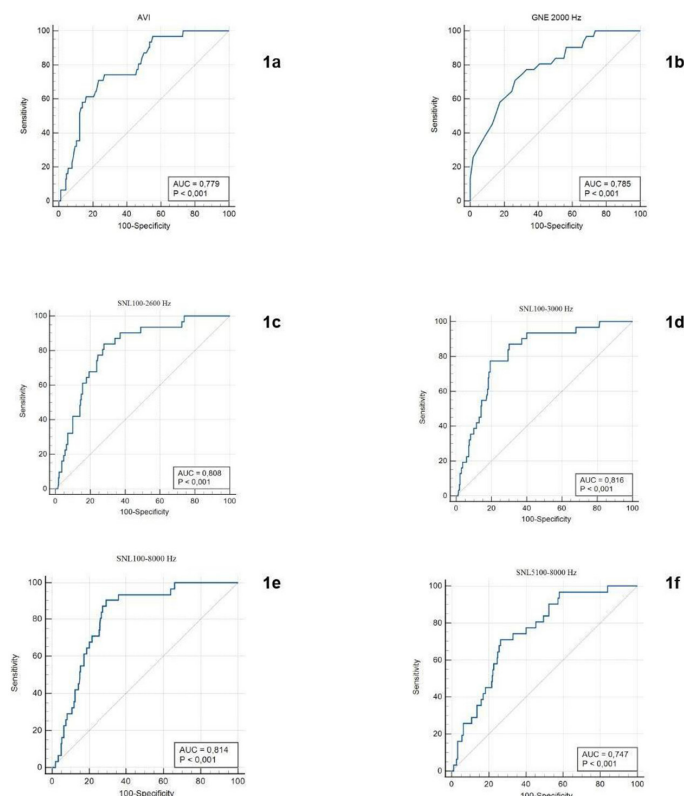


Figura 1. Curvas ROC das medidas acústicas AVI (1a), GNE_{2000Hz} (1b), SNL_{100-2600Hz} (1c), SNL_{100-3000Hz} (1d), SNL_{100-8000Hz} (1e) e SNL_{5100-8000Hz} (1f) na tarefa da vogal [a] sustentada

sensibilidade satisfatória (>70%), foram elas: GNE_{1000Hz} (>0,92), HNR (>0,73), $LNPSD$ ($\leq -9,58$), PA (>0,91), PSD ($\leq 0,00$), $shimmer_{APQ5}$ ($\leq 2,26$), $shimmer_{APQ11}$ ($\leq 2,78$), $shimmer_{dB}$ ($\leq 0,24$), $shimmer_{local}$ ($\leq 2,78$) e $SNL_{100-5100Hz}$ ($\leq 11,44$). Já sete apresentaram especificidade satisfatória (>70%): GNE_{3000Hz} (>0,9), $HNRD$ (>29,78), $Jitter_{local}$ ($\leq 0,31$), $Jitter_{PPQ5}$ ($\leq 0,18$), $Shimmer_{APQ3}$ ($\leq 1,18$), $Shimmer_{DDA}$ ($\leq 3,54$) e $SNL_{2600-5100Hz}$ ($\leq 1,21$).

DISCUSSÃO

O presente estudo se propôs a definir os valores de corte das medidas acústicas extraídas na tarefa de vogal sustentada [a] para discriminação entre vozes com e sem distúrbio vocal em falantes do PB. No presente estudo, seis medidas acústicas apresentaram bom desempenho para discriminar vozes com e sem distúrbio vocal, além de medidas específicas com base em sensibilidade ou especificidade, e seus valores de corte podem ser utilizados na prática clínica vocal.

Uma boa acurácia é uma combinação que reflete o equilíbrio necessário entre a detecção precisa de casos positivos e a minimização de diagnósticos incorretos, sendo fundamental para a aplicabilidade clínica confiável das medidas acústicas avaliadas. Uma medida com alta sensibilidade possui capacidade de identificar corretamente indivíduos com distúrbio vocal, ou seja, os positivos, sendo um instrumento valioso para a triagem de distúrbio vocal, pois minimiza o risco de não detectar casos verdadeiros. Já a alta especificidade indica que as medidas são eficazes na identificação correta dos indivíduos que não apresentam distúrbio vocal, ou seja, os negativos, sendo aplicadas na confirmação diagnóstica para evitar o tratamento desnecessário de indivíduos sem distúrbios vocais⁽⁴⁾.

A medida AVI é descrita na literatura com alta sensibilidade para prever a presença da rugosidade⁽²¹⁾. Em nossa pesquisa, ela apresentou uma AUC de 0,779, com sensibilidade de 70,97% e especificidade de 76,49%, sugerindo um desempenho aceitável (0,70-0,80) na distinção entre grupos com e sem distúrbio vocal, podendo apresentar potencial para ser utilizada como uma ferramenta para triagem e avaliação. Embora o +PV dessa medida seja baixo (25,9%), o -PV de 95,8% indica que um resultado negativo pode ser confiável para excluir o distúrbio vocal⁽²²⁾. No entanto, é necessária uma análise conjunta com medidas complementares para melhor segurança de tomada de decisão no momento da avaliação⁽²³⁾.

A medida acústica GNE tem se mostrado relevante na análise acústica da voz por estimar a quantidade relativa de ruído no sinal sonoro, refletindo a eficiência da fonte glótica em comparação à turbulência do trato vocal⁽³⁾. O GNE pode ser útil na detecção de distúrbios vocais, especialmente quando combinado com outras medidas acústicas⁽²⁴⁾. Dentre suas variantes, o GNE_{2000Hz} destaca-se no presente estudo por apresentar um equilíbrio mais robusto entre sensibilidade e especificidade diagnóstica. O GNE_{2000Hz} alcançou AUC de 0,76, com especificidade de 73,61%, evidenciando melhor desempenho na correta identificação de indivíduos sem distúrbio vocal. Clinicamente, o GNE_{2000Hz} tem aplicação potencial na diferenciação entre produção vocal glótica eficiente e sinais sonoros com maior componente de ruído, sendo particularmente útil na detecção de sopro

e incompetência glótica. Estudos com o GNE_{2000Hz} reforçam sua aplicabilidade como índice complementar na detecção de distúrbios vocais, sobretudo em protocolos multiparamétricos de avaliação⁽²⁵⁾. No que tange às demais variantes dessa medida, o GNE_{1000Hz} , apesar de sua alta sensibilidade (83,87%), demonstra baixa especificidade (54,65%), resultando em maior taxa de falsos positivos. Já o GNE_{3000Hz} , apesar de ter alta especificidade (82,16%), seu desempenho geral não demonstrou vantagem em termos de sensibilidade e estabilidade diagnóstica, o que favorece o uso clínico do GNE_{2000Hz} como uma medida acústica mais eficiente.

A medida acústica SNL tem a finalidade de quantificar a proporção de energia aperiódica (ruído) presente no sinal vocal⁽²¹⁾. Quanto maior o valor de SNL, maior o grau de rugosidade⁽⁸⁾. Dentre as medidas realizadas, a medida $SNL_{100-3000Hz}$ apresentou a maior AUC (0,816), combinando uma sensibilidade de 77,42% e especificidade de 80,67%, fornecendo um bom equilíbrio para detectar indivíduos com distúrbio vocal, tornando-se uma opção viável na prática clínica⁽²⁶⁾. Além disso, apresentou o maior +LR (4,00), sugerindo que um resultado positivo aumenta consideravelmente a probabilidade de distúrbio vocal⁽²⁷⁾. A medida $SNL_{100-8000Hz}$ apresentou uma das maiores sensibilidades (90,32%), um AUC de 0,814 e excelente -PV indicando que essa medida tem um forte potencial para detectar indivíduos com distúrbio vocal. A medida $SNL_{100-2600Hz}$ apresentou um AUC de 0,808, com sensibilidade de 83,87% e especificidade de 72,12%. Seu -LR foi de 0,22, indicando que um resultado negativo reduz significativamente a probabilidade da presença de distúrbio vocal. A medida $SNL_{5100-8000Hz}$ obteve AUC de 0,747, especificidade e sensibilidade equilibradas.

Dentre as medidas analisadas, dez obtiveram bom desempenho para identificar indivíduos com distúrbio vocal, porém falham em identificar indivíduos sem distúrbio vocal. Isso indica que essas medidas possuem maior aplicabilidade em populações de alto risco⁽³⁾, visto que apresentaram alta sensibilidade e AUC maior que 0,7, porém baixa especificidade. São elas: GNE_{1000Hz} , HNR , $LNPSD$, PA , PSD , $shimmer_{APQ5}$, $shimmer_{APQ11}$, $shimmer_{dB}$, $shimmer_{local}$, $SNL_{100-5100Hz}$.

Outras sete medidas obtiveram bom desempenho para identificar indivíduos sem distúrbios vocais, porém falham em identificar indivíduos com distúrbio vocal. Essas medidas são aplicáveis em populações de baixo risco⁽³⁾, visto que apresentaram alta especificidade e AUC maior que 0,7, mas baixa sensibilidade. São elas: GNE_{3000Hz} , $HNRD$, $jitter_{local}$, $jitter_{PPQ5}$, $shimmer_{APQ3}$, $shimmer_{DDA}$, $SNL_{2600-5100Hz}$.

Destaca-se o parâmetro HNR, medida que quantifica a proporção de ruído em relação à proporção de harmônicos em uma amostra vocal, podendo ser de grande valor na diferenciação entre vozes com e sem distúrbio vocal⁽²⁸⁾. No presente estudo, a medida HNR obteve o maior AUC entre as analisadas (0,837), além de uma elevada sensibilidade (90,32%), o que sugere que ela é uma das melhores métricas para detectar a presença de alterações vocais, apesar da sua especificidade moderada (64,68%) a qual indica que pode haver um número de falsos positivos. Esse fenômeno é comum em testes altamente sensíveis e deve ser considerado ao escolher a melhor abordagem diagnóstica⁽²⁷⁾. Esse achado sugere que a HNR pode ser uma medida promissora para triagens

iniciais onde a prioridade é identificar corretamente os casos positivos⁽²⁹⁾. A combinação de medidas com alta sensibilidade (como o HNR) com outras de especificidade elevada (como o SNL 2100-5100Hz) podem otimizar a acurácia diagnóstica. Os resultados evidenciam a importância da análise acústica no diagnóstico e monitoramento da voz, ressaltando a necessidade de abordagem multimodal, combinando diferentes parâmetros para aumentar a precisão diagnóstica^(1,25).

A definição de pontos de corte para medidas acústicas representa um avanço importante para a prática clínica ao fornecer parâmetros objetivos que podem auxiliar na triagem e no diagnóstico diferencial entre vozes com e sem distúrbio vocal. Contudo, é fundamental destacar que a ausência de desempenho diagnóstico satisfatório para algumas medidas não implica em sua inaplicabilidade clínica. Muitas dessas medidas podem ser importantes em contextos específicos, como na comparação intra sujeito, permitindo monitorar alterações vocais ao longo do tempo, avaliar respostas a intervenções terapêuticas ou técnicas vocais, e identificar variações sutis que não são captadas por avaliações puramente categóricas. Portanto, a integração de múltiplos parâmetros acústicos, associados aos demais procedimentos da avaliação multidimensional da voz, favorece uma abordagem mais abrangente e personalizada no manejo da voz, valorizando tanto os indicadores diagnósticos robustos quanto os dados que contribuem para o acompanhamento e aprimoramento contínuo da saúde vocal⁽³⁰⁾.

Todo estudo científico, por sua natureza, opera sob certas condições e escolhas metodológicas que geram limitações. O presente trabalho não é exceção, e as considerações que se seguem visam oferecer uma análise transparente desses aspectos, delineando o escopo e as fronteiras de aplicabilidade dos resultados. Contudo, embora essas limitações informem a generalização e o contexto dos achados, elas não comprometem a validade interna do estudo nem a robustez de suas principais conclusões, que permanecem relevantes e cientificamente válidas dentro do universo populacional e metodológico investigado.

O desenho transversal e descritivo impede o estabelecimento de relações causais diretas ou a avaliação da estabilidade dos pontos de corte ao longo do tempo. Adicionalmente, apesar do expressivo tamanho da amostra, o desequilíbrio entre o número de participantes com e sem distúrbio vocal, embora reflexo da realidade clínica, dificulta a generalização dos pontos de corte. Portanto, os pontos de corte desse estudo podem ser utilizados como referência na prática clínica, mas faz-se necessário estudos com amostras mais abrangentes e com populações com prevalências distintas de distúrbio vocal. A falta de um padrão-ouro reconhecido para essa finalidade pode ser considerada uma limitação do estudo. A classificação dos participantes em com e sem distúrbio vocal baseou-se na combinação do exame visual laríngeo, realizado por um médico, e do JPA realizado por fonoaudiólogos. Embora a alta confiabilidade do JPA tenha sido assegurada, seu caráter subjetivo pode influenciar a categorização. Em conjunto, essas considerações delimitam o escopo de aplicabilidade dos resultados e ressaltam a complexidade da avaliação vocal, ao mesmo tempo em que expandem e refinam o conhecimento nesta área.

Como recomendação clínica, o conjunto de medidas de eleição frente a uma suspeita de distúrbio vocal extraídas da tarefa de vogal sustentada [a] em falantes do PB são SNL_{100-2600Hz}, SNL_{100-3000Hz}, SNL_{100-8000Hz}, visto que apresentam sensibilidade e especificidade $\geq 80\%$. Em seguida, para fins de triagem, as medidas de eleição com sensibilidade $\geq 80\%$ são GNE_{1000Hz}, HNR, LNPSD, PSD, Shimmer_{APQ5}, Shimmer_{APQ11} e SNL_{100-5100Hz}. Finalmente, as medidas com aplicabilidade para confirmação diagnóstica com especificidade $\geq 80\%$ são GNE_{3000Hz} e SNL_{2600-5100Hz}.

CONCLUSÃO

Conclui-se que 23 de um total de 44 medidas acústicas extraídas da vogal sustentada [a] de falantes do PB apresentaram métricas de desempenho satisfatório. As medidas acústicas AVI, GNE_{2000Hz}, SNL_{100-2600Hz}, SNL_{100-3000Hz}, SNL_{100-8000Hz} e SNL_{5100-8000Hz} demonstraram capacidade discriminatória aceitável para diferenciar entre vozes com e sem distúrbio vocal. As medidas GNE_{1000Hz}, HNR, LNPSD, PA, PSD, shimmer_{APQ5}, shimmer_{APQ11}, shimmer_{dB}, shimmer_{local}, SNL_{100-5100Hz} apresentaram ponto de corte com sensibilidade satisfatória e podem ser utilizadas na prática clínica para fins de triagem. As medidas GNE_{3000Hz}, HNRD, Jitter_{local}, Jitter_{PPQ5}, shimmer_{APQ3}, shimmer_{DDA}, SNL_{2600-5100Hz} apresentaram ponto de corte com especificidade satisfatória e possuem aplicabilidade na confirmação diagnóstica.

REFERÊNCIAS

- Englert M, Lopes L, Vieira V, Behlau M. Accuracy of Acoustic Voice Quality Index and its isolated acoustic measures to discriminate the severity of voice disorders. *J Voice*. 2022;36(4):582.e1-10. <https://doi.org/10.1016/j.jvoice.2020.08.010>. PMID:32873433.
- Jalalinajafabadi F, Gadepalli C, Jarchi D, Cheetham BMG. Acoustic analysis and digital signal processing for the assessment of voice quality. *Biomed Signal Process Control*. 2021;70:103018. <https://doi.org/10.1016/j.bspc.2021.103018>.
- Lopes LW, Alves JN, Evangelista DS, França FP, Vieira VJD, Lima-Silva MFB, et al. Acurácia das medidas acústicas tradicionais e formânticas na avaliação da qualidade vocal. *CoDAS*. 2018;30(5):e20170282. <https://doi.org/10.1590/2317-1782/20182017282>. PMID:30365651.
- Uloza V, Pribušis K, Ulozaite-Stanienė N, Petrauskas T, Damaševičius R, Maskeliūnas R. Accuracy analysis of the multiparametric acoustic voice indices, the VWI, AVQI, ABI, and DSI measures, in differentiating between normal and dysphonic voices. *J Clin Med*. 2023;13(1):99. <https://doi.org/10.3390/jcm13010099>. PMID:38202106.
- Nascimento RCPB, Sampaio MC. Os efeitos do gênero e da idade nas medidas acústicas tradicionais e cepstrais em vozes saudáveis e disfônicas: revisão de literatura sistemática [trabalho de conclusão de curso]. Salvador: Departamento de Fonoaudiologia, Universidade Federal da Bahia; 2022.
- Lopes LW, Abreu SR. Accuracy and cut-off values of cepstral measures in the clinical evaluation of Brazilian Portuguese speakers. *J Voice*. 2024;1997(24):142-5. <https://doi.org/10.1016/j.jvoice.2024.04.021>. PMID:38724311.
- Englert MT, Behlau M, Lucero JC. Validação do Acoustic Voice Quality Index (AVQI v. 03.01) e do Acoustic Breathiness Index (ABI) para o português brasileiro [tese]. São Paulo: Universidade Federal de São Paulo, Escola Paulista de Medicina; 2020.
- Latoszek BBV, Maryn Y, Gerrits E, De Bodt M. A metaanalysis: acoustic measurement of roughness and breathiness. *J Speech Lang Hear Res*. 2018;61(2):298-323. https://doi.org/10.1044/2017_JSLHR-S-16-0188. PMID:29392295.

9. Rutjes AWS, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PMM. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ*. 2006;174(4):469-76. <https://doi.org/10.1503/cmaj.050090>. PMID:16477057.
10. Roy N, Barkmeier-Kraemer J, Eadie T, Sivasankar MP, Mehta D, Paul D, et al. Evidence-based clinical voice assessment: a systematic review. *Am J Speech Lang Pathol*. 2013;22(2):212-26. [https://doi.org/10.1044/1058-0360\(2012/12-0014\)](https://doi.org/10.1044/1058-0360(2012/12-0014)). PMID:23184134.
11. Deliyski DD, Shaw HS, Evans MK. Adverse effects of environmental noise on acoustic voice quality measurements. *J Voice*. 2005;19(1):15-28. <https://doi.org/10.1016/j.jvoice.2004.07.003>. PMID:15766847.
12. Maryn Y, Roy N, De Bodt M, Van Cauwenberge P, Corthals P. Acoustic voice analysis: does vowel choice influence the diagnostic value? *J Voice*. 2019;33(5):666.e1-9.
13. Titze IR, Lemke J, Montequin D. Voicing and silence periods in continuous speech. *J Acoust Soc Am*. 2020;107(2):1065-76.
14. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al, and the STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology*. 2015;277(3):826-32. <https://doi.org/10.1148/radiol.2015151516>. PMID:26509226.
15. Brasil. Instituto Brasileiro de Geografia e Estatística. Pesquisa Nacional por Amostra de Domicílios Contínua: resultados 2022. Rio de Janeiro: IBGE; 2022.
16. Behlau M, Zambon F, Guerrieri AC, Roy N. Epidemiology of voice disorders in teachers and nonteachers in Brazil: prevalence and adverse effects. *J Voice*. 2012;26(5):665.e9-18. <https://doi.org/10.1016/j.jvoice.2011.09.010>. PMID:22516316.
17. Yamasaki R, Madazio G, Leão SHS, Padovani M, Azevedo R, Behlau M. Auditory-perceptual evaluation of normal and dysphonic voices using the Voice Deviation Scale. *J Voice*. 2017;31(1):67-71. <https://doi.org/10.1016/j.jvoice.2016.01.004>. PMID:26873420.
18. Boersma P, Weenink D. Praat: doing phonetics by computer. Version 6.0.37 [software]. Amsterdam: University of Amsterdam; 2018.
19. Abreu SR, Moraes RM, Martins PN, Lopes LW. VoxMore: technological artifact to assist voice acoustic evaluation in the teaching-learning process and clinical practice. *CoDAS*. 2023;35(6):e20220166. <https://doi.org/10.1590/2317-1782/20232022166pt>. PMID:37909540.
20. Hosmer DW Jr, Lemeshow S. Applied logistic regression. 2nd ed. New York: Wiley; 2000. <https://doi.org/10.1002/0471722146>.
21. Abreu SR. Índice do grau geral do desvio vocal: desenvolvimento, avaliação e validação de um modelo de suporte à decisão para os falantes do português brasileiro [tese]. João Pessoa: Universidade Federal da Paraíba; 2024.
22. Zhao L, Xu R, Zhu H, Zhang L. Predictive values of acoustic voice parameters for diagnosing vocal fold pathologies. *Clin Otolaryngol*. 2018;43(3):745-51.
23. Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice*. 2009;24(5):540-55. <https://doi.org/10.1016/j.jvoice.2008.12.014>. PMID:19883993.
24. Latoszek BBV, Mayer J, Watts CR, Lehnert B. Advances in clinical voice quality analysis with VOXplot. *J Clin Med*. 2023;12(14):4644. <https://doi.org/10.3390/jcm12144644>. PMID:37510759.
25. Nguyen DD, Novakovic D, Madill C. Voice disorder discrimination using vowel acoustic measures in female speakers. *Int J Lang Commun Disord*. 2024;59(5):2087-102. <https://doi.org/10.1111/1460-6984.13081>. PMID:38884559.
26. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett*. 2006;27(8):861-74. <https://doi.org/10.1016/j.patrec.2005.10.010>.
27. Leite DRA. Desenvolvimento de um modelo de classificação da tipologia dos sinais vocais com base no deep learning [tese]. João Pessoa: Universidade Federal da Paraíba; 2022.
28. Cappellari VM, Cielo CA. Características vocais acústicas de crianças pré-escolares. *Rev Bras Otorrinolaringol*. 2008;74(2):82-8. <https://doi.org/10.1590/S0034-72992008000200018>.
29. Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R. Understanding and using sensitivity, specificity and predictive values. *Indian J Ophthalmol*. 2008;56(1):45-50. <https://doi.org/10.4103/0301-4738.37595>. PMID:18158403.
30. Patel RR, Awan SN, Barkmeier-Kraemer J, Courey M, Deliyski D, Eadie T, et al. Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association Expert Panel to Develop a Protocol for Instrumental Assessment of Vocal Function. *Am J Speech Lang Pathol*. 2018;27(3):887-905. https://doi.org/10.1044/2018_AJSLP-17-0009. PMID:29955816.

Contribuição dos autores

AF: redação – rascunho original; MB: supervisão; JB: redação – rascunho original; SRA: curadoria de dados; VVR: supervisão; LWL: supervisão.

Utilização de tecnologia assistida por inteligência artificial

Os autores declaram que foi utilizada ferramenta de inteligência artificial exclusivamente para revisão gramatical e aprimoramento da linguagem em português durante a preparação do manuscrito. A ferramenta não interferiu no conteúdo científico, nas análises estatísticas ou nas conclusões do estudo. Todo o texto final foi integralmente revisado, validado e aprovado pelos autores, que assumem total responsabilidade pelo conteúdo apresentado.