

## INTRODUÇÃO

A videoendoscopia da deglutição (VED) é uma avaliação instrumental amplamente conhecida. Apesar de frequentemente utilizada na clínica, a análise dos parâmetros da deglutição através da VED é perceptivo-visual, portanto, subjetiva, pois depende da inspeção e interpretação da imagem. Esse procedimento exige habilidades de percepção visual e cognitiva, que podem ser aprimoradas através da experiência.

A variabilidade na forma como a VED é interpretada e é descrita na literatura<sup>(1)</sup>, sendo a classificação de resíduos faríngeos e episódios de penetração laríngea e aspiração laringotraqueal dificuldades relatadas. O treinamento perceptivo-visual destes parâmetros diagnósticos tem sido realizado para melhorar as classificações, normalmente durante o processo de posição de escalas.

O treinamento para a análise da VED tem melhorado a confiabilidade inter e intraexaminador<sup>(2-5)</sup>. Um estudo propôs uma estrutura de treinamento para classificar parâmetros por meio de uma ferramenta perceptivo-visual, onde avaliadores inexperientes obtiveram melhora significativa na precisão das medidas<sup>(2)</sup>. Estes resultados sugerem que currículos de treinamento estruturados podem ser eficazes para desenvolver as competências necessárias para interpretar achados da VED de forma mais precisa<sup>(6)</sup>.

Apesar das evidências relatadas, a subjetividade na avaliação dos parâmetros da VED continua sendo um desafio. O treinamento das habilidade perceptivo-visuais é emergente, pois aumentará a confiabilidade das classificações do exame. Métodos com alta concordância e reprodutíveis poderão ser implementados e resultarão em melhor manejo clínico da disfagia.

## OBJETIVO

Identificar e mapear as evidências disponíveis sobre treinamento para a análise de parâmetros da videoendoscopia da deglutição (VED).

## MÉTODO

Esta revisão de escopo foi conduzida de acordo com a metodologia do *Joanna Briggs Institute* (JBI) e as diretrizes PRISMA-ScR. O protocolo desta revisão foi registrado no Open Science Framework em 10 de novembro de 2021 (<https://osf.io/4xst5/>).

Estratégias de busca abrangentes para identificar estudos na literatura publicada e não publicada foram elaboradas. As seguintes bases de dados foram utilizadas: MEDLINE/PubMed, Web of Science, Scopus, Cochrane Library, Embase e CINAHL. Para a pesquisa de estudos não publicados e literatura cinzenta foram pesquisadas Google Scholar, ProQuest e MedNar. Todas as buscas

foram feitas em modo avançado, sem aplicação de filtros ou limitação de idiomas.

A busca em três etapas<sup>(7)</sup> foi adotada. Uma estratégia de busca para a base de dados MEDLINE/PubMed foi relatada em estudo a priori publicado pelos autores desta revisão<sup>(8)</sup>. Para a seleção dos estudos, uma série de etapas foi seguida: 1. Os artigos identificados foram importados para o Rayyan, *software* gratuito e online que permite o aprimoramento da triagem de dados; 2. As duplicatas foram identificadas e removidas; 3. Dois revisores cegos para os julgamentos um do outro classificaram cada artigo pelo resumo para inclusão ou exclusão; 4. Foi mantido um registro das decisões na plataforma; 5. Os textos completos dos resumos incluídos foram recuperados e considerados para a revisão. Estas etapas foram conduzidas por dois revisores de forma independente. Quando houve divergências, tanto para a etapa de leitura de resumos quanto do texto completo, estes foram discutidos e resolvidos por consenso. Se isso não fosse viável, havia a possibilidade de convocar um terceiro revisor.

O acrônimo PCC (população, conceito e contexto) foi utilizado para formular os critérios de elegibilidade do estudo. Os critérios de inclusão foram: estudos relacionados ao treinamento para análise da VED para estudantes de graduação ou pós-graduação ou profissionais de saúde, em qualquer idioma, publicados a partir do ano de 1988 (ano da primeira publicação sobre VED), utilizando exames de VED realizados na população acima dos 18 anos. Consideramos dados coletados de fontes múltiplas, como preconiza o desenho da revisão de escopo, incluindo: periódicos revisados por pares, livros didáticos, editoriais, anais de congressos e dissertações/teses. Os critérios de exclusão foram: estudos que não descreviam o treinamento, apresentando apenas resultados e os estudos que avaliassem a disfagia esofágica.

Os dados foram extraídos dos estudos incluídos na revisão por um revisor e corroborados por outros dois revisores utilizando uma matriz de extração dos dados descrita no protocolo da revisão<sup>(8)</sup>, que foi refinada durante o processo da revisão.

## RESULTADOS

Após consulta às bases de dados, 3.111 trabalhos foram identificados. Destes, foram excluídos 783 por duplicação nas bases. Dos 2.328 trabalhos, foram excluídos 2.307 após leitura dos títulos e resumos. Vinte e um artigos foram selecionados para a leitura completa do texto e análise considerando os critérios de elegibilidade, remanescendo seis artigos que compuseram a amostra final (Tabela I).

**Tabela I.** Informações gerais dos estudos elegíveis.

Autor	Ano/ País	Local	Objetivo
Warnecke et al. <sup>(5)</sup>	2009/ DE	Hospital	Avaliar se um protocolo de VED previamente estabelecido com base na identificação de achados indicativos de disfagia relacionada ao AVC pode ser aprendido e adotado por clínicos inexperientes.
Kaneoka et al. <sup>(3)</sup>	2013/ USA	Universidade	Examinar a confiabilidade e validade da <i>Boston Residue and Clearance Scale</i> (BRACS).
Neubauer et al. <sup>(4)</sup>	2015/ USA	Universidade	Desenvolver, padronizar e obter evidências de validade da <i>Yale Pharyngeal Residue Severity Rating Scale</i> .
Gerschke et al. <sup>(9)</sup>	2019/ DE	Plataforma online	Validar a versão alemã da <i>Yale Pharyngeal Residue Severity Rating Scale</i> e investigar o impacto da experiência e treinamento do avaliador.
Borders et al. <sup>(10)</sup>	2020/ USA	Centro médico	Examinar a confiabilidade inter e intra-avaliador das classificações clínicas do reflexo adutor laríngeo durante o teste de sensação laríngea com o método de toque.
Curtis et al. <sup>(2)</sup>	2022/ USA & NZ	Universidade	Descrever o desenvolvimento da <i>Visual Analysis of Swallowing Efficiency and Safety</i> (VASES) e explorar a viabilidade de treinar novatos para interpretar a VED usando a VASES.

Legenda: AVC = Acidente Vascular Cerebral; VED = Videoesoscopia da deglutição; ORL = Otorrinolaringologia; VFD = videofluoroscopia da deglutição.

- O ano de publicação dos estudos variou entre 2009 e 2022, com intervalos de 2 a 4 anos entre os anos de 2009 e 2019, e intervalos menores, de 1 a 2 anos, entre os anos de 2019 e 2022;
- Três estudos foram realizados nos Estados Unidos<sup>(3,4,10)</sup>, dois na Alemanha<sup>(5,9)</sup> e um em parceria entre pesquisadores dos Estados Unidos e Nova Zelândia<sup>(2)</sup>;
- De maneira geral, os estudos tinham como objetivo apresentar escalas desenvolvidas para a avaliação dos parâmetros da deglutição, principalmente resíduos faríngeos e penetração/aspiração;
- Os locais onde os estudos foram realizados variaram entre hospital, centro médico, uma plataforma online e em universidades, sendo este último predominante.
- Os estudos consideraram um ou mais especialistas para realizarem classificações de referência para acerto dos participantes (Tabela II);
- Houve heterogeneidade significativa em relação aos anos de experiência em VED tanto dos especialistas (5 a 27 anos) quanto dos participantes do treinamento (sem experiência a aproximadamente 8 anos) (Tabela II).

**Tabela II.** Características dos membros do treinamento.

Estudo	Nº especialistas	Experiência especialista(s)	Nº de participantes	Experiência participantes
Warnecke et al. <sup>(5)</sup>	2	Não informado	17	Sem experiência
Kaneoka et al. <sup>(3)</sup>	1	≥ 5 anos	4	≥ 5 anos
Neubauer et al. <sup>(4)</sup>	2	26 anos	20	≈ 8 anos
Gerschke et al. <sup>(9)</sup>	2	27 anos	28	≥ 1 ano
Borders et al. <sup>(10)</sup>	-	-	4	-
Curtis et al. <sup>(2)</sup>	6	-	26	Sem experiência

(-) não informado.

- Em relação à apresentação da estrutura do treinamento, houve alta variabilidade (Tabela III);

**Tabela III.** Treinamento nos estudos elegíveis.

Autor	Parâmetro	Estrutura do treinamento	Características dos exames
Warnecke et al. <sup>(5)</sup>	Penetração/Aspiração	1) Palestra instrucional de 30 minutos; Demonstrações em vídeo; 2) Classificação com revisão de vídeos; Discussão; 3) Classificação.	Classificação de 16 exames (parte 2) e 8 exames (parte 3).
Kaneoka et al. <sup>(3)</sup>	Resíduos faríngeos	Sessão de três horas. Classificação independente comparada à do especialista, até que atingisse uma faixa de 3 pontos da escala nas consistências consideradas.	Não relata quantidade de vídeos usados no treinamento. Após treinamento, 63 exames de VED foram usados para classificação.
Neubauer et al. <sup>(4)</sup>	Resíduos faríngeos	Não referiu tempo de duração. Definições escritas, visuais, verbais, e discussões sobre a escala.	Não relata características das imagens usadas no treinamento. Após treinamento, foram usadas 25 imagens para classificação.
Gerschke et al. <sup>(9)</sup>	Resíduos faríngeos	Sessão de oito minutos. Apresentação audiovisual da escala, descrições, práticas de classificação de imagens.	Não relata características das imagens usadas no treinamento. Após treinamento, foram usadas 30 imagens para classificação.
Borders et al. <sup>(10)</sup>	Sensibilidade laríngea	Duas sessões de duas horas. Os participantes desenvolveram diretrizes de classificação; julgaram individualmente 50 vídeos; resolveram discordâncias por consenso após assistir novamente os vídeos.	Classificação de 50 vídeos no treinamento. Após treinamento, 125 vídeos diferentes foram avaliados.
Curtis et al. <sup>(2)</sup>	Resíduos faríngeos e Penetração/Aspiração	1) Apresentação de regras da escala; 2) Prática de classificação; 3) Apresentação de vídeo de 60 minutos; 4) Prática de classificação de 5 vídeos; 5) Sessão de 60 minutos em grupo ao vivo e discussão com um especialista.	Após o treinamento, 35 vídeos foram usados para classificação, sendo 10 repetidos para análise de confiabilidade.

- A estrutura nos estudos variou entre palestras instrucionais, apresentações em vídeo, discussões em grupo com ou sem especialistas e prática de classificação de exames; ou combinações entre uma ou duas destas categorias;
- O tempo de duração do treinamento variou entre um estudo que não referiu tempo de duração do treinamento<sup>(4)</sup> até um que apresentou média de tempo de 6h entre os participantes<sup>(2)</sup>;
- Sobre as características dos exames apresentados para classificação no treinamento, os números variaram entre 24 e 125 vídeos, sendo que alguns estudos utilizaram imagens ao invés de vídeos<sup>(4,9)</sup>;
- Os estudos consideraram a análise de confiabilidade como principal medida de resultado do treinamento (Tabela IV);

**Tabela IV.** Medidas de resultado dos estudos elegíveis.

Estudo	Critérios para aferir competência	Medidas relacionadas ao treinamento		
		Análise de confiabilidade		Precisão das classificações
		intra	inter	
Warnecke et al. <sup>(5)</sup>	-	-	x	x
Kaneoka et al. <sup>(3)</sup>	x	-	x	-
Neubauer et al. <sup>(4)</sup>	x	x	x	-
Gerschke et al. <sup>(9)</sup>	x	-	x	-
Borders et al. <sup>(10)</sup>	x	x	x	-
Curtis et al. <sup>(2)</sup>	x	x	x	x

- Dois estudos<sup>(2,5)</sup> consideraram a precisão das classificações dos participantes em relação à dos especialistas;
- Apenas um estudo estabeleceu critério para aferir, durante o treinamento, se o participante estava treinado o suficiente para realizar a análise<sup>(3)</sup>.

## DISCUSSÃO

Nesta revisão de escopo observou-se que o número de estudos sobre treinamento de profissionais para a avaliação dos parâmetros da VED ainda é pouco expressivo, com número limitado de publicações na literatura publicada e não publicada.

Três artigos foram publicados nos últimos três anos<sup>(2,9,10)</sup>, e os demais, entre sete e doze anos atrás<sup>(3-5)</sup>, o que reflete a falta de consistência nas publicações neste sentido, e ao mesmo tempo, uma ascensão desta temática, visto que este intervalo de tempo entre as publicações tem diminuído. Nota-se que os Estados Unidos lideram as publicações sobre o tema, havendo maior reconhecimento por parte dos pesquisadores americanos sobre a importância do treinamento para a análise da VED.

Os treinamentos foram realizados para oferecer maior confiabilidade à avaliação do profissional utilizando um novo método de classificação apresentado. Estes estudos foram realizados em lugares onde comumente é realizada a VED, como universidades, hospital e centro médico. Porém, um estudo realizado por meio de uma plataforma *online* demonstra a possibilidade do treinamento utilizando recursos computacionais<sup>(9)</sup>.

Todos os estudos incluídos tinham como objetivo apresentar ou validar escalas desenvolvidas para a avaliação de parâmetros da deglutição, em sua maioria os parâmetros resíduos faríngeos e penetração/aspiração. Estes parâmetros possuem grande relevância clínica e para a pesquisa, pois possuem relação com a segurança da deglutição.

Os estudos consideram uma estrutura de treinamento onde os parâmetros da deglutição são classificados por especialistas, para que estas classificações possam ser utilizadas posteriormente como referência para os participantes. O painel de especialistas considera

geralmente entre 1 a 2 avaliadores, com experiência mínima de um ano. Apenas um estudo<sup>(2)</sup> descreveu um painel de especialistas mais criterioso, com profissionais treinados, certificados e que publicaram pesquisas envolvendo desempenho e interpretação da VED.

O número de participantes foi variável, entre 4 a 28 avaliadores, com tempos diferentes de experiência. Alguns estudos consideraram a experiência dos participantes na interpretação da VED, enquanto outros tiveram foco em participantes sem formação para análise do exame. Essas características indicam que algumas escalas de classificação são delineadas para uso por especialistas, enquanto outras possuem escopo mais abrangente, sugerindo a utilização por profissionais experientes, desde que haja treinamento.

O tipo de treinamento fornecido envolveu o ensino didático, avaliação prática independente e abordagens combinadas. Treinamentos instrucionais com palestras, tutoriais em vídeo, representações escritas, visuais e verbais, com discussões com avaliadores referência são comuns e, de forma geral, aumenta a precisão e confiabilidade das medidas de resultado. Apenas um estudo preocupou-se em estabelecer um critério de competência no treinamento para aferir se o avaliador estava treinado o suficiente durante o treinamento<sup>(3)</sup>.

Sobre a dose adotada para treinar analistas da VED, não existe consenso para a definição do número de horas suficientes. Existem estudos que realizaram treinamentos de 2 a 3 horas, enquanto outro utilizou autorrelato de participantes para contabilizar o número médio de 6 horas<sup>(2)</sup> (mínimo de 4h, máximo de 20h). Mais pesquisas são necessárias para determinar a dose ideal de treinamento para alcançar competência e confiabilidade dos participantes, principalmente quando estes são inexperientes.

A maioria dos estudos considerou as classificações por meio de recortes de vídeos ou exame completo, enquanto outros utilizaram a análise de imagens<sup>(4,9)</sup>. Como a VED é dinâmica, em condições reais esta análise costuma ocorrer por meio de vídeos. Porém, o exame por quadro isolado ainda é comum nas classificações clínicas do exame, principalmente em estudos de validação de escalas, onde busca-se obter controle da validade interna da pesquisa e permitir comparações entre os avaliadores. O número de exames avaliados nos treinamentos possui alta variabilidade, entre 24 e 125 exames.

A maior parte dos estudos limitou-se à análise de confiabilidade intra e interavaliadores como medidas de resultado. Informações que seriam úteis para avaliar a eficácia do treinamento também não foram apontadas, como: conhecimento sobre anatomia e fisiologia, experiência clínica, medidas de autoconfiança, e medidas relacionadas à análise, como precisão e velocidade.

Esta revisão de escopo identificou uma lacuna de conhecimento, visto que a literatura não apresenta resultados específicos sobre o treinamento para a análise de parâmetros da VED, estando limitada ao treinamento como objetivo secundário nos estudos de validação de ferramentas para a classificação destes parâmetros.

Nossos achados sugerem a emergente necessidade do desenvolvimento de métodos de treinamento estruturados e padronizados, considerando variáveis que possam influenciar na certificação de competência de estudantes e profissionais de saúde na análise de parâmetros da VED. Essas metodologias devem considerar aspectos como curva de aprendizagem, relatório de desempenho, níveis de habilidade, progressos e o uso de métodos computacionais para as respostas, oferecendo *feedback* imediato ao participante.

### CONCLUSÃO

O treinamento em análise de parâmetros da VED é mais comum em estudos de validação de escalas, que adotam o treinamento perceptivo-visual para estabelecer melhor confiabilidade na análise. A literatura revisada aponta a necessidade de desenvolver treinamentos para tornar estudantes e profissionais de saúde competentes na análise dos parâmetros da VED, considerando variáveis que possam influenciar no desempenho dos participantes.