International Journal of Speech-Language Pathology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/iasl20

Understanding user experience and normative data in pharyngeal residue rating scales used in flexible endoscopic evaluation of swallowing (FEES): A scoping review

Thomas Wilson, Martin Checklin, Nadine Lawson, Alissa J. Burnett, Theresa Lombardo & Amy Freeman-Sanderson

To cite this article: Thomas Wilson, Martin Checklin, Nadine Lawson, Alissa J. Burnett, Theresa Lombardo & Amy Freeman-Sanderson (11 Nov 2024): Understanding user experience and normative data in pharyngeal residue rating scales used in flexible endoscopic evaluation of swallowing (FEES): A scoping review, International Journal of Speech-Language Pathology, DOI: 10.1080/17549507.2024.2420606

To link to this article: https://doi.org/10.1080/17549507.2024.2420606

Published online: 11 Nov 2024.

International Journal of Speech-Language Pathology, 2024; Early Online: 1–14

Understanding user experience and normative data in pharyngeal residue rating scales used in flexible endoscopic evaluation of swallowing (FEES): A scoping review

THOMAS WILSON[1,2,3], MARTIN CHECKLIN[4,5], NADINE LAWSON[6], ALISSA J. BURNETT[2], THERESA LOMBARDO[0] & AMY FREEMAN-SANDERSON[7,8,9,0]

[1]Footscray Hospital, Western Health, Footscray, Australia, [2]Deakin University, Geelong, Australia, Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences, Geelong, Australia, [3]Austin Hospital, Austin Health, Heidelberg, Australia, [4]Epworth Healthcare, Richmond, Australia, [5]Discipline of Speech Pathology, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Australia, [6]Speech Pathology & Dietetics, Cabrini Health, Malvern, Australia, [7]Graduate School of Health, University of Technology Sydney, Sydney, Australia, [8]Royal Prince Alfred Hospital, Sydney, Australia, [9]Critical Care Division, Faculty of Medicine, The George Institute for Global Health, UNSW Sydney, Sydney, Australia, and, [10]Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

Abstract

Purpose: Pharyngeal residue rating scales are often used to rate pharyngeal residue observed during flexible endoscopic evaluation of swallowing. Despite the widespread use of pharyngeal residue rating scales, there is no data that has systematically explored user experience. The aim of this scoping review was to investigate specific reporting of user experience, user centred design principles, and normative data in the development of pharyngeal residue rating scales.

Method: A scoping review was conducted across four electronic databases inclusive of all dates until June 2024. Grey literature searching occurred in March–April 2023 and was repeated in June 2024. This review followed the Preferred Reporting Items for Systematic Review and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) protocol. Titles/abstracts, full texts, and data extraction were reviewed by two independent reviewers.

Result: A total of 22 sources were included, with 18 unique pharyngeal residue rating scales identified. Two studies referred to user experience, seven included at least one user centred design principle, and four studies reported on normative data.

Conclusion: The findings of this review highlight few pharyngeal residue rating scales include the experience of the intended user and establish normative data in the initial development phase. User experience, user centred design principles, and normative data may be useful considerations to optimise functionality.

Keywords: *swallowing; deglutition disorders; endoscopy; Pharyngeal residue; user experience and normative data*

Introduction

Pharyngeal residue refers to secretions, food, or fluid boli accumulated in the pharyngeal recesses and is correlated with an increased risk of aspiration and worsened severity of dysphagia (Langmore, 2017; Leonard et al., 2011; Sabry & Abou-Elsaad, 2023; Shapira-Galitz et al., 2019a, 2019b; Steele, Peladeau-Pigeon, Barrett, et al., 2020; Stokely et al., 2015; Yoon et al., 2019). An international consensus study recommended the use of rating tools in dysphagia assessment, including the evaluation and systematic rating of pharyngeal residue, as this is paramount to the comprehensive assessment, characterisation, and treatment of oropharyngeal dysphagia (Espitalier et al., 2018). For pharyngeal residue observed in

Correspondence: Amy Freeman-Sanderson, Graduate School of Health, University of Technology Sydney, Sydney, NSW, Australia. E-mail: amy.freeman-sanderson@uts.edu.au

ISSN 1754-9507 print/ISSN 1754-9515 online © 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
Published by Taylor & Francis
DOI: 10.1080/17549507.2024.2420606

flexible endoscopic evaluation of swallowing (FEES), multiple scoring systems exist, with variance in tool content, measurement methods, and overall purpose (Miles et al., 2021; Neubauer et al., 2016; Starmer, 2022; Swan et al., 2019).

A systematic review by Neubauer et al. (2016) compared the psychometric integrity of seven pharyngeal residue rating scales across criteria including severity definitions, scale type, number of raters, experience of raters, randomisation of images, intra- and inter-rater reliability, and construct validity. The authors concluded the Yale pharyngeal residue severity rating scale was the only tool that provided sufficient detail across all the specified criteria and was therefore recommended for clinical use (Neubauer et al., 2016). Subsequent research has confirmed that the Yale pharyngeal residue rating scale demonstrates satisfactory reliability. However, this is influenced by whether residue is rated from videos or still images, bolus consistency, and training of raters (Rocca et al., 2022, 2024). In addition to this, Swan et al. (2019) evaluated the psychometric properties of published visuo-perceptual measures for instrumental swallowing assessments including nine FEES scales using the consensus-based standards for the selection of health status measurement Instruments checklist. This checklist was used to determine internal consistency, test-retest, inter-rater and intra-rater reliability, measurement error, content validity, structural validity, and hypothesis testing. The authors concluded that psychometric properties of all scales were weak and were unable to recommend the use of any individual scale (Swan et al., 2019). Whilst providing important comparative data across pharyngeal residue rating scales, previous research has not explored the concepts of user experience, user centred design principles, and normative data, which are useful to evaluate and understand the clinical utility of scales. As pharyngeal residue rating scales are widely available for clinical use, investigation into user experience and normative data is warranted to better understand how instrument design considerations influence clinical workflow, decision-making, diagnosis, and broader dysphagia management.

The effectiveness and practical application of a clinical tool, instrument, or service may be inextricably linked to positive user experience (Hartson & Pyla, 2012). In the design sciences, user experience design is used to improve the experience of software interfaces, systems, and spaces through systematic research which involves the intended user, often in the desired context or environment (Ahram et al., 2022). User experience design is conceptually associated with the broader concept of 'design thinking' and is gaining momentum in healthcare as a framework to enhance innovation, promote sustainable resource management, and respond to the increasing demands on healthcare systems globally (Altman et al., 2018; Roberts et al., 2016). User experience research involves multiple cycles of ideation, prototyping, and testing with the aim of streamlining the product, software, or system through observation of prospective end-users (Ahram et al., 2022). This results in an outcome in which the initial development is led by end-users and occurs before formal rollout of the system (Altman et al., 2018; Bate & Robert, 2007).

User experience design does not have one formal definition nor standardised elements, however, one model is Morville's 'user experience honeycomb' (Morville, 2004). The model encompasses elements of usefulness, usability, desirability, value, findability, accessibility, and credibility (Morville, 2004). This model was initially designed for computer software, but the concept is applicable to healthcare (Bate & Robert, 2007). User Experience is context dependent, meaning that each of these elements may not have an equal weight for a single product or tool. For example, when considering pharyngeal residue rating scales, desirability (i.e. the emotional and aesthetically appealing aspect of a tool) is unlikely to influence user experience as much as usability, or the effectiveness, efficiency, and satisfaction of using the scale (Morville, 2004). When considering the design of clinical instruments, such as pharyngeal residue rating scales, positive user experience of a scale is contingent on user centred design principles. While it is acknowledged that design principles are emerging concepts in healthcare, it is believed that they are important areas for consideration to improve reporting of dysphagia in clinical settings.

Involving end-users in healthcare research through co-design approaches is considered best practice, as it increases the relevance and quality of healthcare research (Bird et al., 2021; Evans et al., 2021; Sanders & Stappers, 2008). While healthcare has not explored usability to the same extent as the design sciences, consideration of user centred design in the development of health instruments and personal health tools has been reported (Vaisson et al., 2021). User centred design also seeks to involve the people who are likely to use or interact with a tool, service, or instrument in the design and development process to optimise the overall user experience (Witteman et al., 2021). While the concept of 'end users" is broad and may include researchers, clinicians, and patients alike, for the purposes of this review 'end-users' are defined as clinicians based in patient settings, to enhance the clinical relevance of this review. Bate and Robert (2007) asserted that if something is more 'usable', it is likely to lead to fewer errors and better performance. However, the concept of user centred design in dysphagia has not been widely reported, providing new opportunities to apply this to the development of assessment tools.

In healthcare, normative data research involves comparisons between people with a shared condition and a reference group of healthy participants with the aim of establishing what is normal' to then identify and investigate patterns of abnormality (O'Connor,

1990). In dysphagia care, this may involve the assessment and reporting of swallowing function in healthy people without dysphagia under the same conditions (i.e. determining swallowing performance across bolus consistency or bolus volume). Previous research has shown that the assessment of pharyngeal residue, even with validated scales, is subject to human measurement error and influenced by factors, such as perceptual bias, bolus consistency, and the rater's experience and training (Pisegna, 2022; Pisegna, Borders, et al., 2018; Pisegna, Kaneoka, et al., 2018; Pisegna et al., 2020; Rocca et al., 2022, 2024). This suggests that evaluating the severity or amount of pharyngeal residue may not effectively distinguish between normal and impaired swallowing functions without comparative normative values for the same scale. Therefore a scale which assigns a quantity or severity level, but which does not define a normal range with a healthy, non-dysphagic reference group, may have inherent limitations to its interpretation and application in practice, and may negatively impact the end user's experience with this scale (Humbert et al., 2018; Molfenter & Steele, 2013). Clinically, this may result in diagnostic inaccuracies or inappropriate therapeutic recommendations (Vose et al., 2018). The inclusion of normative data within pharyngeal residue rating scales may also facilitate increased accuracy in quantifying pharyngeal residue across bolus variables and contribute to aspiration risk stratification (Molfenter & Steele, 2013; Steele, Peladeau-Pigeon, Nagy, et al., 2020).

Whilst previous literature has described critical issues of psychometric properties of pharyngeal residue rating scales used in FEES reporting, to date, there has been no evaluation of user experience, user-centred design principles, and normative data (Neubauer et al., 2016; Pisegna, 2022; Swan et al., 2019). Given this is an unexplored area that has the potential to shape future research approaches and scale design, further investigation is warranted. As user experience is considered a novel concept for dysphagia research, a scoping review was selected to broadly yet systematically ascertain the breadth of available evidence, analyse knowledge gaps, and provide a preliminary summary of research related to a contemporary concept (Levac et al., 2010; Munn et al., 2018).

The aims of this scoping review were: a) to identify reporting of user experience in pharyngeal residue rating scales development; b) to identify and assess user centred design principles in pharyngeal residue rating scales development, and c) to investigate reporting of normative swallowing data within the initial stages of pharyngeal residue rating scales development.

Method

Protocol and registration

This review followed the Preferred Reporting Items for Systematic Review and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) (Tricco et al., 2018). The protocol was prospectively registered with Open Science Framework on 19th June 2022 (Registration link https://osf.io/nf7yh).

Eligibility criteria

Study eligibility included: research articles that reported on the initial development of published pharyngeal residue and/or secretion management rating scales used in FEES; involved humans over 18 years of age; were published in English; and reported on the initial design processes of pharyngeal residue rating scales. As user experience design typically involves the intended end-user as early as initial concept development, our review focused on initial design processes only (i.e. creation of scoring items, initial validation, pilot testing, or other early design initiatives used to establish a novel rating tool). Studies were excluded if they reported on pharyngeal residue interpretation on instrumental assessments other than FEES (e.g. videofluroscopic swallowing studies [VFSS]), reported on user experience outside of pharyngeal residue rating scales (i.e. reported on the user experience of documentation software, not a rating scale), and only described the use of a pharyngeal residue rating scales (e.g. as an outcome measure, reference standard, or in the validation of an existing scale with a new population), rather than the initial development of the scale. As we sought to investigate specific details of pharyngeal residue rating scales design practices, secondary evidence that summarised multiple studies, such as systematic reviews were excluded. Case studies/series, letters to the editor, commentary/opinion pieces, theses, conference abstracts, and supplemental material were also excluded to enhance the clinical relevance of the review.

Information sources and search strategy

A systematic literature search was conducted using four electronic databases covering multiple professions (CINAHL, Medline, Embase, speechBITE) to meet reported recommendations on using >2 databases for a 95% recall (Ewald et al., 2022). Medical subject headings (MeSH) and free text mapping of keywords in the title and abstracts adapted from Neubauer et al. (2016) were used on each database from inception to June 2022 and repeated in June 2024 (see Appendix A for Medline search terms). To complement this systematic search, a further exploratory search was conducted using the first 200 results on Google Scholar (Bramer et al., 2017; Haddaway et al., 2015). Search criteria were reviewed by five of the authors, one academic, and a medical librarian. A grey literature search was conducted following the guidelines outlined by Arksey and O'Malley (2005) and involved the review of references cited in position statements and/or clinical practice guidelines pertaining to FEES from professional bodies. Professional bodies included Speech

Pathology Australia, The Royal College of Speech and Language Therapists, the American Speech-Language-Hearing Association, the New Zealand Speech-language Therapists' Association, The European Society for Swallowing Disorders, and the Union of the European Phoniatricians. Citation and reference checking of previous systematic reviews were conducted, as well as citation and reference checking of all articles in the final yield (Neubauer et al., 2016; Swan et al., 2019). The search occurred during March–April 2023 and was inclusive of all records from inception up to April 2023 and was repeated in June 2024. Additionally, authors of included studies were contacted for access to any unpublished or in-press material related to the inclusion criteria.

Selection of sources of evidence

The yield of all searches was exported to Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia) for review, with duplicates removed via the automatic function. At both title and abstract and full text phases, all records were screened independently by two reviewers against the inclusion/exclusion criteria. All disagreements were resolved either through consensus discussion or a third reviewer if required. There was discussion at the beginning, middle, and end of each phase of the review to reduce ambiguity in source selection as recommended by Levac et al. (2010).

Data charting process

Data were extracted from the final set of included studies by two authors independently. Data were extracted into a purpose built excel database developed by the first author and pilot tested and refined by five of the authors. All authors completed an initial calibration task which involved data extraction of the same study followed by consensus discussion to affirm the data extraction method, table layout, data points, and operational guidelines.

Data items

Data items extracted from sources included article title, name of pharyngeal residue rating scale, authors, country of origin, year of publication, residue measurement outcome (e.g. oral trials or secretions), measurement method (e.g. type of scale used), user experience, principles of user centred design, and normative data. For studies that included user experience, data extracted included methods of acquiring user experience (e.g. focus groups, surveys, and metrics), characteristics of users (e.g. clinicians or researchers), and user experience elements considered as relevant to the sources obtained. No restrictions were applied regarding the method/methodology of user experience data collection (e.g. questionnaires, contextual inquiry, interviews, and experience co-design). For sources that included user centred design principles, six items from the User-Centred Design 11-item Measure (UCD-11) were adopted and modified to capture the extent to which user experience was considered or explored during the development of pharyngeal residue rating scales (Witteman et al., 2021). For sources that included normative data, additional information was extracted including sample size, age range of participants, male to female distribution, race/ethnicity, food/fluid and bolus size of consistencies trialled, and swallowing conditions/instructions.

Synthesis of result

Descriptive statistics were used to summarise the data collected. Information pertaining to the three review aims were synthesised in tabular format and were presented in relation to the individual pharyngeal residue rating scales identified, rather than across included studies. When a pharyngeal residue rating scale was described across multiple studies, all were included, and data were extracted from all papers.

Result

Selection of sources of evidence

The search results are outlined in the PRISMA flowchart (Figure 1). The full search yielded a total of 7592 sources. Following the removal of duplicates, 4681 sources were screened at the title and abstract level. Full text review was completed for 209 sources, resulting in the inclusion of 22 studies for data extraction. For all included studies, the corresponding author was contacted to establish if any additional unpublished or in-press studies were available. While six of the corresponding authors replied, no material from this was included in the final review following screening against the inclusion criteria. A further three professional bodies/societies were contacted for position statements/guidelines, however no further position documents on FEES were obtained.

Characteristics of sources of evidence

A summary of the 18 pharyngeal residue rating scales identified is outlined in Tables I and II. Of the 22 included studies, 18 unique scales were identified. Three were specific to interpretation and severity of accumulated pharyngo-laryngeal secretions, 14 were specific for assessment of pharyngeal residue with oral trials, and one which included both secretions and oral trials. The included studies originated from nine countries, published between 1996–2023. Two studies included the use of concurrent FEES and VFSS (Kim & Jung, 2013; Park et al., 2015) and one pharyngeal residue rating scales was specific to the laryngectomy population (Coffey et al., 2018). Four pharyngeal residue rating scales included normative data (Curtis et al., 2023; Donzelli et al., 2003; Kelly et al., 2008; Murray et al., 1996), two reported on user experience

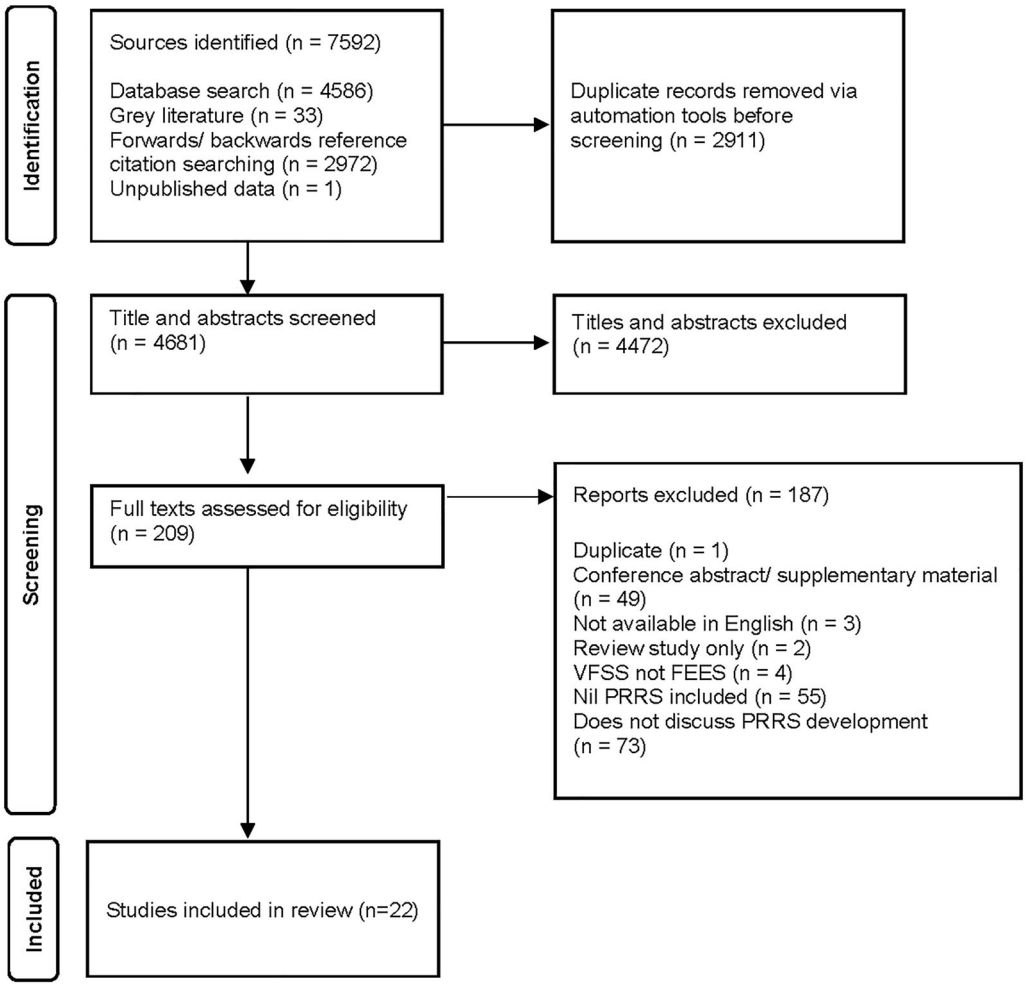

Figure 1. PRISMA flowchart.

Table I. Pharyngeal residue rating scales for secretions.

Scale name	Article publication	Country of origin	Measurement method	Normative data	User experience	User centred design principles
Marianjoy Secretion Scale (Donzelli et al., 2003)	2003	United States	Ordinal scale	Yes	No	No
New Zealand Secretion Scale (NZSS) (Miles & Hunting, 2019)	2019	New Zealand	Ordinal scale	No	No	Yes
Secretion Severity Rating Scale (Murray et al., 1996)	1996	United States	Ordinal scale	Yes	No	Yes

(Curtis, Borders, et al., 2022; Kaneoka et al., 2013), and seven reported on some degree of user centred design principles (Coffey et al., 2018; Curtis, Borders, et al., 2022; Kaneoka et al., 2013; Miles & Hunting, 2019; Murray et al., 1996; Starmer et al., 2021; Tohara et al., 2010), though the level of detail varied across studies and scales. Three pharyngeal residue rating scales were reported over multiple papers (Curtis, Borders, et al., 2022; Curtis, Borders, Perry, et al. 2022; Curtis et al., 2023; Farneti, 2008; Farneti et al., 2014; Kelly et al., 2006, 2008).

User experience

A summary of the user experience data and user centred design principles reported in the identified scales is outlined in Table III. No studies explicitly reported assessment of user experience or incorporation of user experience findings into initial scale development. Two scales reported on initial feasibility during the process of creating a scale with the use of measures obtained by observing potential end users, displaying a consideration of user experience (Curtis, Borders, et al., 2022; Kaneoka et al., 2013). There were no reported changes made to the design or function of the scale based on the measures collected in either study. The authors of the Boston residue and clearance scale (BRACS) measured the average time to complete ratings using the scale compared to using a traditional, non-standardised ordinal rating scale commonly referred to in the literature (i.e. 'none, coating, mild, moderate, severe') (Kaneoka et al.,

Table II. Pharyngeal residue rating scales for oral trials.

Scale name	Article publication	Country of origin	Measurement method	Normative data	User experience	User centred design principles
Boston Residue and Clearance Scale (BRACS) (Kaneoka et al., 2013)	2013	United States	Ordinal scale	No	Yes	Yes
Dynamic Imaging Grade of Swallowing Toxicity for Flexible Endoscopic Evaluation of Swallowing: DIGEST-FEES (Starmer et al., 2021)	2021	United States	Ordinal scale	No	No	Yes
Mansoura Fiberoptic Endoscopic Evaluation of Swallowing Residue Rating Scale (MFRRS) (Sabry et al., 2021)	2021	Egypt	Ordinal scale	No	No	No
Pooling score (P-score) (Farneti, 2008; Farneti et al., 2014)	2008, 2014	Italy	Ordinal scale	No	No	No
The New Scale (Kim & Jung, 2013)	2013	South Korea	Ordinal scale	No	No	No
Unnamed (Kelly et al., 2006, 2008)	2006, 2008	United Kingdom	Ordinal scale	Yes	No	No
Unnamed (Tohara et al., 2010)*	2010	Japan	Ordinal scale	No	No	Yes
Unnamed (Park et al., 2015)	2015	South Korea	Ordinal scale	No	No	No
Unnamed (Baijens et al., 2015)	2015	Netherlands	Ordinal scale	No	No	No
Unnamed (Pilz et al., 2016)	2016	Netherlands	Ordinal scale	No	No	No
Unnamed (Coffey et al., 2018)	2018	United Kingdom	Mixed binary and visual analogue scale	No	No	Yes
Unnamed (Manor et al., 2019)	2019	Israel	Ordinal scale	No	No	No
Unnamed (Simon et al., 2020)	2020	United States	Ordinal scale	No	No	No
Visual Analysis of Swallow Efficiency and Safety (VASES) (Curtis, Borders, et al., 2022; Curtis, Borders, Perry, et al. 2022)	2021, 2022, 2023	United States	Visual analogue scale or verbal numerical ratings	Yes	Yes	Yes
Yale Pharyngeal Residue Severity Rating Scale (Neubauer et al., 2015)	2015	United States	Ordinal scale	No	No	No

*This pharyngeal residue rating scale encompassed an assessment of both secretions and oral trials however only features in this table.

2013). The mean time to complete individual ratings was calculated and reported as ~2 min 37 s. On the initial measure, the BRACS ratings took twice as long as traditional rating methods but on a repeat measure, the total time to complete all BRACS ratings reduced (Kaneoka et al., 2013). User-experience measures reported in the Visual Analysis of Swallow Efficiency and Safety scale (VASES) included accuracy of ratings pre- and post-training, and time to complete ratings for each FEES clip (Curtis, Borders, et al., 2022). There were statistically significant improvements in accuracy post-training, though variability was seen across different anatomical landmarks (Curtis, Borders, et al., 2022). The average time to complete ratings post-training was 1 min 30 sec, which was nearly half the time it took pre-training. All sources had high accessibility, a crucial element of user experience, as they were located within accessible published articles and did not require commercial training or learning materials.

User-centred design

None of the identified studies and respective pharyngeal residue rating scales met all modified UCD-11 criteria, however, seven tools included at least one UCD-11 element. No studies reported on assessing the needs of potential end-users either before, or during the development of the pharyngeal residue rating scales. Three studies reported on involving potential end-users in designing the prototype of the scale (Coffey et al., 2018; Kaneoka et al., 2013; Miles & Hunting, 2019) and three studies reported on involving potential end users in evaluating a prototype or final version (Kaneoka et al., 2013; Miles & Hunting, 2019; Starmer et al., 2021). Two studies reported observing potential end-users (i.e. working clinicians external to the research team) using the pharyngeal residue rating scales (Curtis, Borders, et al., 2022; Murray et al., 1996). Two studies explicitly reported the changes made through the iterative cycles of the design process (Miles & Hunting, 2019; Tohara et al., 2010), and two studies involved an expert panel (i.e. external to the research team) in the development of the pharyngeal residue rating scales (Coffey et al., 2018; Miles & Hunting, 2019).

Normative data

A summary of normative data reported in the identified pharyngeal residue rating scales is outlined in Table IV. Four studies reported on normative data across four different scales (Curtis et al., 2023; Donzelli et al., 2003; Kelly et al., 2008; Murray et al., 1996). Two studies included the examination of pharyngo-laryngeal secretions (Donzelli et al., 2003; Murray et al., 1996) and the other two reported on residue following oral bolus trials (Curtis et al., 2023; Kelly et al., 2008). The two secretion management scales compared secretion accumulation with aspiration of diet/fluids. The consistency of diet/fluids under examination was not reported in either study

Table III. User experience and modified user-centered design 11-item measure.

Scale name	User experience data	Were potential end users involved in any steps to help understand users and their needs?	Were potential end users involved in any steps of designing, developing, or refining a prototype	Were potential end users involved in any steps intended to evaluate prototypes or a final version of the tool?	Were potential end users observed using the tool in any way?	Were changes made between iterative cycles explicitly reported in any way?	Was an expert panel involved?
Boston Residue and Clearance Scale (BRACS) (Kaneoka et al., 2013)	Yes	No	Yes	Yes	No	No	No
Dynamic Imaging Grade of Swallowing Toxicity for Flexible Endoscopic Evaluation of Swallowing: DIGEST-FEES (Starmer et al., 2021)	No	No	No	Yes	No	No	No
Mansoura Fiberoptic Endoscopic Evaluation of Swallowing Residue Rating Scale (MFRRS) (Sabry et al., 2021)	No	No	No	No	No	No	No
Marianjoy Secretion Scale (Donzelli et al., 2003)	No	No	No	No	No	No	No
New Zealand Secretion Scale (NZSS) (Miles & Hunting, 2019)	No	No	Yes	Yes	No	Yes	Yes
Pooling score (P-score) (Farneti, 2008; Farneti et al., 2014)	No	No	No	No	No	No	No
Secretion Severity Rating Scale (Murray et al., 1996)	No	No	No	No	Yes	No	No
The New Scale (Kim & Jung, 2013)	No	No	No	No	No	No	No
Unnamed (Kelly et al., 2006, 2008)	No	No	No	No	No	No	No
Unnamed (Tohara et al., 2010)	No	No	No	No	No	Yes	No
Unnamed (Park et al., 2015)	No	No	No	No	No	No	No
Unnamed (Baijens et al., 2015)	No	No	No	No	No	No	No
Unnamed (Pilz et al., 2016)	No	No	No	No	No	No	No
Unnamed (Coffey et al., 2018)	No	No	Yes	No	No	No	Yes
Unnamed (Manor et al., 2019)	No	No	No	No	No	No	No
Unnamed (Simon et al., 2020)	No	No	No	No	No	No	No
Visual Analysis of Swallow Efficiency and Safety (VASES) (Curtis, Borders, et al., 2022; Curtis, Borders, Perry, et al. 2022)	Yes	No	No	No	Yes	No	No
Yale Pharyngeal Residue Severity Rating Scale (Neubauer et al., 2015)	No	No	No	No	No	No	No

(Donzelli et al., 2003; Murray et al., 1996). In both studies, the healthy participants were reported to have normal or minimal pooled secretions, and normal swallow function, suggesting no aspiration of diet/fluids was observed. Further data on pharyngeal residue of the diet/fluids trialled were not reported in either study. For the pharyngeal residue rating scales which examined pharyngeal residue following oral diet/fluids trials, a range of bolus consistencies and volumes were reported (Curtis et al., 2023; Kelly

Table IV. Normative data reported in pharyngeal residue rating scales.

Scale name	Study sample size* (*n*)	Age range—year mean (range)	Sex (M, F)	Race/ethnicity	Diet/fluid consistencies and bolus volume trialled	Swallowing conditions/ instructions
Marianjoy Secretion Scale (Donzelli et al., 2003)	4	46 (NR–NR)	M: 3 F: 1	NR	• Secretions • Diet/fluid consistencies not reported	NR
Unnamed (Kelly et al., 2008)	51	Young group: 30.5 (23–38) Older group: 75 (65–88)	Young group M: 10 F:11 Elderly group M: 18 F: 12	NR	• Liquid (5 ml, 10 ml, large mouthful) • Smooth vanilla yoghurt (10 ml) • Chopped banana (10 ml) • Cheese sandwich (3 × 3 cm)	Natural swallow
Secretion Severity Rating Scale (Murray et al., 1996)	22	Young Group: NR (24–40) Older group: NR (60–83)	M: 22 F: 0	NR	• Secretions • "Food" • "Liquid"	NR
Visual Analysis of Swallow Safety and Efficiency (VASES) (Curtis et al., 2023)	39	50.2 (27–83)	M: 19 (48%) F: 20 (51%)	Asian: 1 (2%) Black or African American: 10 (25%) Hispanic or Latino: 5 (12%) Multi-racial: 3 (7%) White: 20 (51%)	• Thin fluids/IDDSI 0 (self-selected volumes including natural and single swallow, 5, 10, 20, 90 ml) • Pudding/IDDSI 4 (5 ml) • Cracker/IDDSI 7 (prompted normal bite size)	Natural swallow, single swallow, consecutive sips

NR: not reported; M: male; F: female; IDDSI: Internal Dysphagia Diet Standardisation Initiative (Steele et al., 2018).
Note. *Sample size of healthy participants only.

et al., 2008). Kelly et al. (2008) proposed an anatomically defined, 6-point ordinal scale to assess pharyngeal residue in healthy participants across a range of bolus consistencies. Healthy older and younger adults had equally efficient pharyngeal clearance, with pharyngeal residue only seen in a small number of participants, typically in small amounts (Kelly et al., 2008). They reported that younger healthy adults had marginally more residue than healthy older adults, with coating or mild residue in the laryngeal inlet present in 21 younger adults and 11 older adults (Kelly et al., 2008). Similarly, Curtis et al. (2023) used the VASES to assess pharyngeal residue in healthy adults. While pharyngeal residue was present in healthy adults across most oral trials, it was typically estimated to cover 2–3% of an anatomical surface across various bolus consistencies (Curtis et al., 2023). Nil studies reported normative data of thickened fluid bolusi. Three of the four studies reported data categorised by young and older healthy adults as well as gender (Donzelli et al., 2003; Kelly et al., 2008; Murray et al., 1996). One study reported on additional information including gender, race/ethnicity, height, weight, and body mass index (Curtis et al., 2023). Nil studies reported embedding normative reference values into the scoring system of the pharyngeal residue rating scales in the initial development.

Discussion

This scoping review is the first to describe user experience, user-centred design principles, and normative data in the initial development of pharyngeal residue rating scales used in FEES. Through a systematic literature search, 18 unique pharyngeal residue rating scales were identified from 22 studies. There was heterogeneity across scales regarding content, measurement methods, anatomic landmarks, and overall purpose (e.g. assessment of secretions only or assessment of oral trials). Pharyngeal residue rating scales commonly encompassed additional and interconnected parameters within a single tool, including assessment of physiological response to residue, effectiveness of clearance attempts, and airway protection (penetration/aspiration). Despite pharyngeal residue rating scales gaining greater attention, with seven of the 18 scales published within the last five years (Coffey et al., 2018; Curtis, Borders, et al., 2022; Curtis, Borders, Perry, et al. 2022; Curtis et al., 2023; Manor et al., 2019; Miles & Hunting, 2019; Sabry et al., 2021; Simon et al., 2020; Starmer et al., 2021), there is limited data pertaining to the inclusion of user experience, user centred design principles and normative data within the initial stages of scale development.

In the VASES and BRACS, usability metrics frequently used in user experience design are reported, however, they are described in a way that seeks to ascertain initial feasibility of using the scale (i.e. if the measurement methods/tool content is fit for purpose) rather than understanding the user to improve the experience of using the scale (Curtis, Borders, et al., 2022; Kaneoka et al., 2013). The metrics explored provide preliminary data to ascertain usability, which is only one element of user experience and did not result in changes to improve the usability of either scale. To enhance pharyngeal residue rating scales implementation, usability metrics, such as time to

complete ratings, time to complete documentation, and/or number of viewings required to score, may be beneficial. Adjustments from these tests may increase usability and contribute to enhanced user experience. Post pharyngeal residue rating scales development, Messina et al. (2024) explored usability metrics of the Yale pharyngeal residue severity rating scale, the BRACS, the Pooling score, and the residue ordinal rating scale. Intra-rater and inter-rater reliability, time to complete ratings and perceived difficulty/ease of use were compared across these scales. The authors concluded the Yale pharyngeal residue severity rating scale yielded the highest scores for reliability and the Pooling score and BRACS had the lowest for inter-rater and intra-rater reliability, respectively. The residue ordinal rating scale required the least amount of time to rate and was perceived as the easiest to use, while the BRACS took the longest and was also perceived as the most difficult (Messina et al., 2024). It is also important to note that while one tool may be perceived as easier to use, it may be less detailed in the information it provides. This data suggests that enhanced usability may increase the reliability of scales. These findings provide clinicians with useful data to drive decisions for tool use in clinical practice.

In the absence of user experience knowledge for pharyngeal residue rating scales, evaluation of the user-centeredness of their design is one way to gauge how the authors considered the experience of their intended end-user in the scale's development (Witteman et al., 2021). While involving end-users in instrument design is highly desired to increase clinical utility and maximise adoption in clinical practice, user centred design principles were present in fewer than half of evaluated scales. There was no individual scale which comprehensively integrated all elements of the modified UCD-11 framework in the development phase. Notably, none of the identified pharyngeal residue rating scales included an assessment of the needs or preferences of clinicians either before or during the development phase. Qualitative methods in user experience research, such as focus groups, surveys, and think-aloud testing may bolster quantitative usability measures and add data across multiple user experience elements (Hartson & Pyla, 2012; Wolcott & Lobczowski, 2021). We acknowledge the rationale for pharyngeal residue rating scales development varies, initial aims need to align with measurement properties, and this is perhaps a key reason for the absence of user-experience, user-centred design, and normative data in the initial tool development. This notion is supported by subsequent accumulative research on these aspects following scale development and this should be acknowledged (Rocca et al., 2022, 2024; Sabry & Abou-Elsaad, 2023; Sutton et al., 2024). However, given it is not uncommon for scales to be subsequently implemented into comparable clinical settings, further consideration of these concepts in the initial tool design phase may leverage important insights to streamline the translation of research evidence into clinical practice. Considering this, it would be valuable for future research to examine the usability of tools and assess the practice patterns, needs, and preferences for pharyngeal residue rating scales from a clinician's point of view. Data of this nature could then be used to inform further development or modification of pharyngeal residue rating scales to better align scale functionality with current or desired practice patterns.

One of the primary functions of pharyngeal residue rating scales is to measure and understand the severity of swallowing impairment or abnormality. Reference to normative data in scoring parameters is imperative to accurately capture the extent to which swallowing function deviates from the expected pattern and to be confident that a score assigned actually reflects disordered swallowing (Steele, Peladeau-Pigeon, Barrett, et al., 2020; Steele, Peladeau-Pigeon, Nagy, et al., 2020). Despite this, only four of the 18 pharyngeal residue rating scales included normative data or reported on comparisons between healthy participants and participants with dysphagia in the initial development phase (Curtis et al., 2023; Donzelli et al., 2003; Kelly et al., 2008; Murray et al., 1996). In three of the papers, data were limited by insufficient detail regarding diet/fluid consistencies trialled, minimal report of swallowing cues provided, small sample sizes, unequal male to female distribution, and limited representation of multiple races/ethnicities (Donzelli et al., 2003; Kelly et al., 2008; Murray et al., 1996). Data of this nature provides a general summary of healthy pharyngeal residue and may have limited application to clinical practice (Vose et al., 2018). The VASES was the only tool that included broad normative data (diet/fluid consistencies, bolus size, contrast agent, swallowing instructions, gender, age, height, and body mass index) (Curtis et al., 2023). Furthermore, the VASES provides a normal range (<3% coverage), with normative reference values aligning closely with previous research on bolus size and pharyngeal residue in VFSS (Garand et al., 2023). Using the VASES to determine a normal volume of residue, rather than categorising severity levels, may also contribute to advancing FEES pharyngeal residue rating scales to match the sophistication of quantitative, norm-referenced scales used in VFSS. It may be beneficial to ascertain if the end-users of the VASES would require a central table or schematic which includes normal values to further enhance the user experience of the VASES. Lastly, the current study researchers acknowledge existing normative data for pharyngeal residue viewed endoscopically, however, this was not included as the purpose of this review, rather it was to look at pharyngeal residue rating scales development (Butler et al., 2009; Sutton et al., 2024; Veiga et al., 2014). There were no pharyngeal residue rating scales identified that employed or referenced this

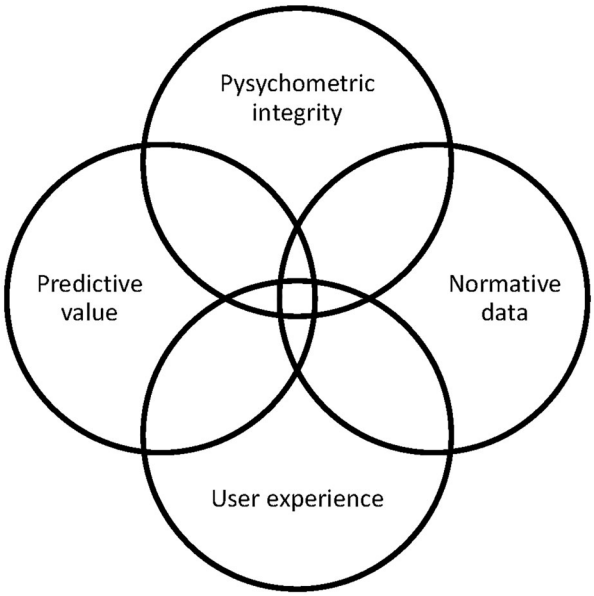

Figure 2. Four-part model for key components of dysphagia residue rating scales tools.

existing data. While consideration of normative data in some of the pharyngeal residue rating scales identified is commended, normative data pertaining to pharyngeal residue observed on FEES remains underreported and lacks clinical applicability when compared to the literature for VFSS (Steele et al., 2023; Street et al., 2024).

It is acknowledged that designing a clinical interpretation rating scale has many pertinent concepts to consider. Each concept may have varying importance depending on the environmental context and the unique needs of diverse end-users, including researchers, clinicians, and patients. For example, while one scale may be faster to use, it may be lacking in psychometric integrity or detail. Additionally, one scale may be more attractive over another due to the predictive value it possesses, which may be particularly useful with a specific patient group. While the focus of this review is to explore the initial design phase of pharyngeal residue rating scales as this is customary in user centred design, it is acknowledged that there is an accumulation of evidence for a pharyngeal residue rating scale over time that demonstrates continued development and improvement. For example, while normative data was not established in the initial design phase of the Yale pharyngeal residue severity rating scale, recent research has established preliminary normative data across various bolus volumes and consistencies for this scale long after its initial development (Sutton et al., 2024). To conceptualise key elements for a pharyngeal residue rating scale, a four-part model is proposed (see Figure 2) to support clinicians with appraising literature and adopting pharyngeal residue rating scales into clinical practice, or conversely, to guide future research of pharyngeal residue rating scales by conceptualising components which may enhance functionality.

Limitations

The strengths of this review included adherence to the PRISMA-ScR process, double reviewer screening across all review stages, and systematic approach to searching and data extraction (Tricco et al., 2018). However, there are some limitations; firstly, the review only included studies published in English, meaning that pharyngeal residue rating scales in other languages may have been missed and therefore not represented in this data set. Future reviews would benefit from the allocation of funds to enable the use of a professional academic translation service. Secondly, while the systematic approach used to extract data was a strength in our overall strategy, the use of the UCD-11 in its original form was not appropriate for our aims. Therefore, the modified version used in this study was not validated and the data it provided was purely descriptive. Lastly, as the aim was to obtain articles reporting pharyngeal residue rating scales in their initial development phase, it is possible that studies that have evaluated user experience, user centred design principles, and normative data in later stages are not captured in the final yield and results should be interpreted accordingly.

Future research

Future development of pharyngeal residue rating scales may benefit from assessment against a range of user experience measures and methodologies, as well as implementation of user centred design principles. Contextual inquiry, observational studies, discovery interviewing, or focus groups may be methods employed to investigate user experience further in pharyngeal residue rating scales. Ultimately, user experience concepts may be extended beyond clinicians and include the patient's perspective to gain insight into what this group expect and value when others report on their swallowing function. Given the paucity of normative data and its importance in clinical decisions, future work should prioritise inclusion of this data with larger sample sizes, reporting across all bolus consistencies described in the International Dysphagia Diet Standardisation Initiative (Steele et al., 2018) using reliable methods of rating pharyngeal residue. Finally, it would be valuable to determine and understand what frontline clinicians value and need in pharyngeal residue rating scales to optimise future research priorities and clinical implementation.

Conclusion

This review highlights that user experience, user centred design principles, and normative data are seldom systematically reported in the initial development of pharyngeal residue rating scales. These findings add

to prior research and offer a novel perspective on optimising pharyngeal residue rating scales through the integration of key pillars including psychometric integrity, predictive value, user experience, and normative data. There is an opportunity to enhance the functionality of scales and their use in clinical practice by considering these principles in pharyngeal residue rating scales development.

Acknowledgements

The authors acknowledge and thank the Epworth HealthCare and Deakin University Library services for their input in the search, and Laura Vat for her contribution to the screening process.

Ethical approval

Ethics approval was not required for this scoping review.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research project did not receive any grants from funding agencies in the public, commercial, or not-for-profit sectors.

ORCID

Thomas Wilson http://orcid.org/0000-0002-8620-0630
Martin Checklin http://orcid.org/0000-0002-9758-4334
Nadine Lawson http://orcid.org/0000-0001-9153-9922
Alissa J. Burnett http://orcid.org/0000-0002-0156-7039
Amy Freeman-Sanderson http://orcid.org/0000-0003-2161-6421

References

Ahram, T. Z., Rebelo, F., & Soares, M. M. (2022). *Handbook of usability and user-experience*. CRC Press.

Altman, M., Huang, T. T. K., & Breland, J. Y. (2018). Design thinking in health care. *Preventing Chronic Disease*, *15*, E117. https://doi.org/10.5888/pcd15.180128

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Baijens, L. W. J., Pilz, W., Kremer, B., & Passos, V. L. (2015). Identifying patterns of FEES-derived swallowing trajectories using group-based trajectory model. *Dysphagia*, *30*(5), 529–539. https://doi.org/10.1007/s00455-015-9632-3

Bate, P., & Robert, G. (2007). *Bringing user experience to healthcare improvement: The concepts, methods and practices of experience-based design*. Radcliffe Publishing.

Bird, M., McGillion, M., Chambers, E. M., Dix, J., Fajardo, C. J., Gilmour, M., Levesque, K., Lim, A., Mierdel, S., Ouellette, C., Polanski, A. N., Reaume, S. V., Whitmore, C., & Carter, N. (2021). A generative co-design framework for healthcare innovation: Development and application of an end-user engagement framework. *Research Involvement and Engagement*, *7*(1), 12. https://doi.org/10.1186/s40900-021-00252-7

Bramer, W. M., Rethlefsen, M. L., Kleijnen, J., & Franco, O. H. (2017). Optimal database combinations for literature searches in systematic reviews: A prospective exploratory study. *Systematic Reviews*, *6*(1), 245. https://doi.org/10.1186/s13643-017-0644-y

Butler, S. G., Stuart, A., & Kemp, S. (2009). Flexible endoscopic evaluation of swallowing in healthy young and older adults. *The Annals of Otology, Rhinology, and Laryngology*, *118*(2), 99–106. https://doi.org/10.1177/000348940911800204

Coffey, M. M., Tolley, N., Howard, D., Drinnan, M., & Hickson, M. (2018). An investigation of the post-laryngectomy swallow using videofluoroscopy and fiberoptic endoscopic evaluation of swallowing (FEES). *Dysphagia*, *33*(3), 369–379. https://doi.org/10.1007/s00455-017-9862-7

Curtis, J. A., Borders, J. C., & Troche, M. S. (2022). Visual analysis of swallowing efficiency and safety (VASES): Establishing criterion-referenced validity and concurrent validity. *American Journal of Speech-Language Pathology*, *31*(2), 808–818. https://doi.org/10.1044/2021_ajslp-21-00116

Curtis, J. A., Borders, J. C., Dakin, A. E., & Troche, M. S. (2023). Normative reference values for FEES and VASES: Preliminary data from 39 nondysphagic, community-dwelling adults. *Journal of Speech, Language, and Hearing Research*, *66*(7), 2260–2277. https://doi.org/10.1044/2023_jslhr-23-00132

Curtis, J. A., Borders, J. C., Perry, S. E., Dakin, A. E., Seikaly, Z. N., & Troche, M. S. (2022). Visual analysis of swallowing efficiency and safety (VASES): A standardized approach to rating pharyngeal residue, penetration, and aspiration during FEES. *Dysphagia*, *37*(2), 417–435. https://doi.org/10.1007/s00455-021-10293-5

Donzelli, J., Brady, S., Wesling, M., & Craney, M. (2003). Predictive value of accumulated oropharyngeal secretions for aspiration during video nasal endoscopic evaluation of the swallow. *The Annals of Otology, Rhinology, and Laryngology*, *112*(5), 469–475. https://doi.org/10.1177/000348940311200515

Espitalier, F., Fanous, A., Aviv, J., Bassiouny, S., Desuter, G., Nerurkar, N., Postma, G., & Crevier-Buchman, L. (2018). International consensus (ICON) on assessment of oropharyngeal dysphagia. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, *135*(1S), S17–S21. https://doi.org/10.1016/j.anorl.2017.12.009

Evans, J. M., Gilbert, J. E., Bacola, J., Hagens, V., Simanovski, V., Holm, P., Harvey, R., Blake, P. G., & Matheson, G. (2021). What do end-users want to know about managing the performance of healthcare delivery systems? Co-designing a context-specific and practice-relevant research agenda. *Health Research Policy and Systems*, *19*(1), 131. https://doi.org/10.1186/s12961-021-00779-x

Ewald, H., Klerings, I., Wagner, G., Heise, T. L., Stratil, J. M., Lhachimi, S. K., Hemkens, L. G., Gartlehner, G., Armijo-Olivo, S., & Nussbaumer-Streit, B. (2022). Searching two or more databases decreased the risk of missing relevant studies: A metaresearch study. *Journal of Clinical Epidemiology*, *149*, 154–164. https://doi.org/10.1016/j.jclinepi.2022.05.022

Farneti, D. (2008). Pooling score: An endoscopic model for evaluating severity of dysphagia. *Acta Otorhinolaryngologica Italica*, *28*(3), 135–140.

Farneti, D., Fattori, B., Nacci, A., Mancini, V., Simonelli, M., Ruoppolo, G., & Genovese, E. (2014). The Pooling-score (P-score): Inter- and intra-rater reliability in endoscopic assessment of the severity of dysphagia. *Acta Otorhinolaryngologica Italica*, *34*(2), 105–110.

Garand, K. L. F., Grissett, A., Corbett, M. M., Molfenter, S., Herzberg, E. G., Kim, H. J., & Choi, D. (2023). Quantifying pharyngeal residue across the adult life span: Normative values by age, gender, and swallow Task. *Journal of Speech, Language, and Hearing Research*, *66*(3), 820–831. https://doi.org/10.1044/2022_jslhr-22-00413

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of google scholar in evidence reviews and its applicability to grey literature searching. *PLOS One*, *10*(9), e0138237. https://doi.org/10.1371/journal.pone.0138237

Hartson, H. R., & Pyla, P. S. (2012). *The UX book: Process and guidelines for ensuring a quality user experience*. Morgan Kaufmann.

Humbert, I. A., Sunday, K. L., Karagiorgos, E., Vose, A. K., Gould, F., Greene, L., Azola, A., Tolar, A., & Rivet, A. (2018). Swallowing kinematic differences across frozen, mixed, and ultrathin liquid boluses in healthy adults: Age, sex, and normal Variability. *Journal of Speech, Language, and Hearing Research*, *61*(7), 1544–1559. https://doi.org/10.1044/2018_JSLHR-S-17-0417

Kaneoka, A. S., Langmore, S. E., Krisciunas, G. P., Field, K., Scheel, R., McNally, E., Walsh, M. J., O'Dea, M. B., & Cabral, H. (2013). The Boston residue and clearance scale: Preliminary reliability and validity testing. *Folia Phoniatrica et Logopaedica*, *65*(6), 312–317. https://doi.org/10.1159/000365006

Kelly, A. M., Leslie, P., Beale, T., Payten, C., & Drinnan, M. J. (2006). Fibreoptic endoscopic evaluation of swallowing and videofluoroscopy: Does examination type influence perception of pharyngeal residue severity? *Clinical Otolaryngology*, *31*(5), 425–432. https://doi.org/10.1111/j.1749-4486.2006.01292.x

Kelly, A. M., Macfarlane, K., Ghufoor, K., Drinnan, M. J., & Lew-Gor, S. (2008). Pharyngeal residue across the lifespan: A first look at what's normal. *Clinical Otolaryngology*, *33*(4), 348–351. https://doi.org/10.1111/j.1749-4486.2008.01755.x

Kim, H.-R., & Jung, S.-H. (2013). Development of a new scale for pharyngeal residue using videofluoroscopic swallowing study and fiberoptic endoscopy. *Journal of the Korean Dysphagia Society*, *3*(2), 70–77. https://doi.org/10.34160/jkds.2013.3.2.005

Langmore, S. E. (2017). History of fiberoptic endoscopic evaluation of swallowing for evaluation and management of pharyngeal dysphagia: Changes over the years. *Dysphagia*, *32*(1), 27–38. https://doi.org/10.1007/s00455-016-9775-x

Leonard, R., Rees, C. J., Belafsky, P., & Allen, J. (2011). Fluoroscopic surrogate for pharyngeal strength: The pharyngeal constriction ratio (PCR). *Dysphagia*, *26*(1), 13–17. https://doi.org/10.1007/s00455-009-9258-4

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*(1), 69. https://doi.org/10.1186/1748-5908-5-69

Manor, Y., Oestreicher-Kedem, Y., Gad, A., Zitser, J., Faust-Socher, A., Shpunt, D., Naor, S., Inbar, N., Kestenbaum, M., Giladi, N., & Gurevich, T. (2019). Dysphagia characteristics in Huntington's disease patients: Insights from the fiberoptic endoscopic evaluation of swallowing and the swallowing disturbances questionnaire. *CNS Spectrums*, *24*(4), 413–418. https://doi.org/10.1017/s1092852918001037

Messina, F., Rocca, S., Manca, B., Scarponi, L., Ninfa, A., Schindler, A., & Pizzorni, N. (2024). Pharyngeal residue scoring in fiberoptic endoscopic evaluation of swallowing: Reliability comparison and applicability among different scales. *Dysphagia*, *39*(5), 872–880. https://doi.org/10.1007/s00455-024-10669-3

Miles, A., & Hunting, A. (2019). Development, intra- and inter-rater reliability of the New Zealand Secretion Scale (NZSS). *International Journal of Speech-Language Pathology*, *21*(4), 377–384. https://doi.org/10.1080/17549507.2018.1458901

Miles, A., Hunting, A., Fletcher, H. V., & Allen, J. (2021). Current approaches to reporting pharyngo-laryngeal secretions. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *29*(6), 479–486. https://doi.org/10.1097/MOO.0000000000000740

Molfenter, S. M., & Steele, C. M. (2013). The relationship between residue and aspiration on the subsequent swallow: An application of the normalized residue ratio scale. *Dysphagia*, *28*(4), 494–500. https://doi.org/10.1007/s00455-013-9459-8

Morville, P. (2004). *User experience design*. Semantic Studios. Retrieved October 4, 2022, from https://semanticstudios.com/user_experience_design/

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, *18*(1), 143. https://doi.org/10.1186/s12874-018-0611-x

Murray, J., Langmore, S. E., Ginsberg, S., & Dostie, A. (1996). The significance of accumulated oropharyngeal secretions and swallowing frequency in predicting aspiration. *Dysphagia*, *11*(2), 99–103. https://doi.org/10.1007/bf00417898

Neubauer, P. D., Hersey, D. P., & Leder, S. B. (2016). Pharyngeal residue severity rating scales based on fiberoptic endoscopic evaluation of swallowing: A systematic review. *Dysphagia*, *31*(3), 352–359. https://doi.org/10.1007/s00455-015-9682-6

Neubauer, P. D., Rademaker, A. W., & Leder, S. B. (2015). The Yale pharyngeal residue severity rating scale: An anatomically defined and image-based tool. *Dysphagia*, *30*(5), 521–528. https://doi.org/10.1007/s00455-015-9631-4

O'Connor, P. J. (1990). Normative data: Their definition, interpretation, and importance for primary care physicians. *Family Medicine*, *22*(4), 307–311.

Park, W. Y., Lee, T. H., Ham, N. S., Park, J. W., Lee, Y. G., Cho, S. J., Lee, J. S., Hong, S. J., Jeon, S. R., Kim, H. G., Cho, J. Y., Kim, J. O., Cho, J. H., & Lee, J. S. (2015). Adding endoscopist-directed flexible endoscopic evaluation of swallowing to the videofluoroscopic swallowing study increased the detection rates of penetration, aspiration, and pharyngeal residue. *Gut and Liver*, *9*(5), 623–628. https://doi.org/10.5009/gnl14147

Pilz, W., Vanbelle, S., Kremer, B., van Hooren, M. R., van Becelaere, T., Roodenburg, N., & Baijens, L. W. J. (2016). Observers' agreement on measurements in fiberoptic endoscopic evaluation of swallowing. *Dysphagia*, *31*(2), 180–187. https://doi.org/10.1007/s00455-015-9673-7

Pisegna, J. M. (2022). What's the evidence? A commentary on FEES research. *Perspectives of the ASHA Special Interest Groups*, *7*(6), 1941–1959. https://doi.org/10.1044/2022_PERSP-22-00128

Pisegna, J. M., Borders, J. C., Kaneoka, A., Coster, W. J., Leonard, R., & Langmore, S. E. (2018). Reliability of untrained and experienced raters on FEES: Rating overall residue is a simple task. *Dysphagia*, *33*(5), 645–654. https://doi.org/10.1007/s00455-018-9883-x

Pisegna, J. M., Kaneoka, A., Coster, W. J., Leonard, R., & Langmore, S. E. (2020). Residue ratings on FEES: Trends for clinical application of residue measurement. *Dysphagia*, *35*(5), 834–842. https://doi.org/10.1007/s00455-019-10089-8

Pisegna, J. M., Kaneoka, A., Leonard, R., & Langmore, S. E. (2018). Rethinking residue: Determining the perceptual continuum of residue on FEES to enable better measurement. *Dysphagia*, *33*(1), 100–108. https://doi.org/10.1007/s00455-017-9838-7

Roberts, J. P., Fisher, T. R., Trowbridge, M. J., & Bent, C. (2016). A design thinking framework for healthcare management and innovation. *Healthcare*, *4*(1), 11–14. https://doi.org/10.1016/j.hjdsi.2015.12.002

Rocca, S., Negri, L., Valenza, N., Schindler, A., & Pizzorni, N. (2024). Pharyngeal residues scoring through the Yale pharyngeal residue severity rating scale (YPRSRS): Efficacy of training. *Dysphagia*. https://doi.org/10.1007/s00455-024-10725-y

Rocca, S., Pizzorni, N., Valenza, N., Negri, L., & Schindler, A. (2022). Reliability and construct validity of the Yale pharyngeal residue severity rating scale: Performance on videos and effect of bolus consistency. *Diagnostics*, *12*(8), 1897. https://doi.org/10.3390/diagnostics12081897

Sabry, A., & Abou-Elsaad, T. (2023). Pharyngeal residue severity and aspiration risk in stroke patient using fiber-optic

endoscopic evaluation of swallowing. *Folia Phoniatrica et Logopaedica*, *75*(3), 158–163. https://doi.org/10.1159/000528204

Sabry, A., Coyle, J. L., & Abou-Elsaad, T. (2021). Mansoura fiberoptic endoscopic evaluation of swallowing residue rating scale (MFRRS): An anatomically based tool – A preliminary study. *Folia Phoniatrica et Logopaedica*, *73*(6), 478–490. https://doi.org/10.1159/000512158

Sanders, E. B. N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign*, *4*(1), 5–18. https://doi.org/10.1080/15710880701875068

Shapira-Galitz, Y., Shoffel-Havakuk, H., Halperin, D., & Lahav, Y. (2019a). Association between laryngeal sensation, preswallow secretions and pharyngeal residue on fiberoptic endoscopic examination of swallowing. *Dysphagia*, *34*(4), 548–555. https://doi.org/10.1007/s00455-019-10001-4

Shapira-Galitz, Y., Shoffel-Havakuk, H., Halperin, D., & Lahav, Y. (2019b). Correlation between pharyngeal residue and aspiration in fiber-optic endoscopic evaluation of swallowing: An observational study. *Archives of Physical Medicine and Rehabilitation*, *100*(3), 488–494. https://doi.org/10.1016/j.apmr.2018.05.028

Simon, S. R., Florie, M., Pilz, W., Winkens, B., Winter, N., Kremer, B., & Baijens, L. W. J. (2020). Association between pharyngeal pooling and aspiration using fiberoptic endoscopic evaluation of swallowing in head and neck cancer patients with dysphagia. *Dysphagia*, *35*(1), 42–51. https://doi.org/10.1007/s00455-019-09992-x

Starmer, H. (2022). Incorporating flexible endoscopic evaluation of swallowing rating scales into clinical practice. *Perspectives of the ASHA Special Interest Groups*, *7*(6), 1984–1990. https://doi.org/10.1044/2022_PERSP-22-00094

Starmer, H. M., Arrese, L., Langmore, S., Ma, Y., Murray, J., Patterson, J., Pisegna, J., Roe, J., Tabor-Gray, L., & Hutcheson, K. (2021). Adaptation and validation of the dynamic imaging grade of swallowing toxicity for flexible endoscopic evaluation of swallowing: DIGEST-FEES. *Journal of Speech, Language, and Hearing Research*, *64*(6), 1802–1810. https://doi.org/10.1044/2021_JSLHR-21-00014

Steele, C. M., Bayley, M. T., Bohn, M. K., Higgins, V., Peladeau-Pigeon, M., & Kulasingam, V. (2023). Reference Values for Videofluoroscopic Measures of Swallowing: An Update. *Journal of Speech, Language, and Hearing Research : JSLHR*, *66*(10), 3804–3824. 10.1044/2023_JSLHR-23-0024637669617

Steele, C. M., Namasivayam-MacDonald, A. M., Guida, B. T., Cichero, J. A., Duivestein, J., Hanson, B., Lam, P., & Riquelme, L. F. (2018). Creation and initial validation of the international dysphagia diet standardisation initiative functional diet scale. *Archives of Physical Medicine and Rehabilitation*, *99*(5), 934–944. https://doi.org/10.1016/j.apmr.2018.01.012

Steele, C. M., Peladeau-Pigeon, M., Barrett, E., & Wolkin, T. S. (2020). The risk of penetration-aspiration related to residue in the pharynx. *American Journal of Speech-Language Pathology*, *29*(3), 1608–1617. https://doi.org/10.1044/2020_AJSLP-20-00042

Steele, C. M., Peladeau-Pigeon, M., Nagy, A., & Waito, A. A. (2020). Measurement of pharyngeal residue from lateral view videofluoroscopic images. *Journal of Speech, Language, and Hearing Research*, *63*(5), 1404–1415. https://doi.org/10.1044/2020_JSLHR-19-00314

Stokely, S. L., Peladeau-Pigeon, M., Leigh, C., Molfenter, S. M., & Steele, C. M. (2015). The relationship between pharyngeal constriction and post-swallow residue. *Dysphagia*, *30*(3), 349–356. https://doi.org/10.1007/s00455-015-9606-5

Street, M., Morrison, M., & Miles, A. (2024). Exploring Reliable and Identifiable Quantitative Anatomical, Timing, and Displacement Measures in the VFS of Speech. *Journal of Craniofacial Surgery*, *35*(1), 91–95. https://doi.org/10.1097/SCS.0000000000009791 .

Sutton, S., Lim, L., Servino, K., To, H., Wang, L., McCoy, Y., Bice, E. M., & Galek, K. E. (2024). Normal values for swallow events during endoscopic evaluation of swallowing: A preliminary study. *European Archives of Oto-Rhino-Laryngology*, *281*(10), 5517–5525. https://doi.org/10.1007/s00405-024-08782-y

Swan, K., Cordier, R., Brown, T., & Speyer, R. (2019). Psychometric properties of visuoperceptual measures of videofluoroscopic and fibre-endoscopic evaluations of swallowing: A systematic review. *Dysphagia*, *34*(1), 2–33. https://doi.org/10.1007/s00455-018-9918-3

Tohara, H., Nakane, A., Murata, S., Mikushi, S., Ouchi, Y., Wakasugi, Y., Takashima, M., Chiba, Y., & Uematsu, H. (2010). Inter- and intra-rater reliability in fibroptic endoscopic evaluation of swallowing. *Journal of Oral Rehabilitation*, *37*(12), 884–891. https://doi.org/10.1111/j.1365-2842.2010.02116.x

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Vaisson, G., Provencher, T., Dugas, M., Trottier, M.-È., Chipenda Dansokho, S., Colquhoun, H., Fagerlin, A., Giguere, A. M. C., Hakim, H., Haslett, L., Hoffman, A. S., Ivers, N. M., Julien, A.-S., Légaré, F., Renaud, J.-S., Stacey, D., Volk, R. J., & Witteman, H. O. (2021). User involvement in the design and development of patient decision aids and other personal health tools: A systematic review. *Medical Decision Making*, *41*(3), 261–274. https://doi.org/10.1177/0272989X20984134

Veiga, H. P., Fonseca, H. V., & Bianchini, E. M. G. (2014). Sequential swallowing of liquid in elderly adults: Cup or straw? *Dysphagia*, *29*(2), 249–255. https://doi.org/10.1007/s00455-013-9503-8

Vose, A. K., Kesneck, S., Sunday, K., Plowman, E., & Humbert, I. (2018). A survey of clinician decision making when identifying swallowing impairments and determining treatment. *Journal of Speech, Language, and Hearing Research*, *61*(11), 2735–2756. https://doi.org/10.1044/2018_jslhr-s-17-0212

Witteman, H. O., Vaisson, G., Provencher, T., Chipenda Dansokho, S., Colquhoun, H., Dugas, M., Fagerlin, A., Giguere, A. M., Haslett, L., Hoffman, A., Ivers, N. M., Légaré, F., Trottier, M.-E., Stacey, D., Volk, R. J., & Renaud, J.-S. (2021). An 11-item measure of user- and human-centered design for personal health tools (UCD-11): Development and validation. *Journal of Medical Internet Research*, *23*(3), e15032. https://doi.org/10.2196/15032

Wolcott, M. D., & Lobczowski, N. G. (2021). Using cognitive interviews and think-aloud protocols to understand thought processes. *Currents in Pharmacy Teaching & Learning*, *13*(2), 181–188. https://doi.org/10.1016/j.cptl.2020.09.005

Yoon, J. A., Kim, S. H., Jang, M. H., Kim, S. D., & Shin, Y. B. (2019). Correlations between aspiration and pharyngeal residue scale scores for fiberoptic endoscopic evaluation and videofluoroscopy. *Yonsei Medical Journal*, *60*(12), 1181–1186. https://doi.org/10.3349/ymj.2019.60.12.1181

Appendix A

Search terms adopted and adapted from Neubauer et al. (2016).

Medline (Ovid)	
1	exp Deglutition/
2	exp Deglutition Disorders/
3	"pharyngeal residue".mp.
4	(swallow* and residue).mp.
5	1 or 2 or 3 or 4
6	Endoscopes/
7	Endoscop*.mp.
8	(fiberoptic and endoscopic).mp.
9	Fiber Optic Technology/
10	exp Fluoroscopy/
11	FEES.mp.
12	6 or 7 or 8 or 9 or 10 or 11
13	test*.mp.
14	evaluat*.mp.
15	scale*.mp.
16	grade*.mp.
17	score*.mp.
18	FEES.mp.
19	13 or 14 or 15 or 16 or 17 or 18
20	aspirat*.mp.
21	Food/
22	food.mp.
23	swallow*.mp.
24	20 or 21 or 22 or 23
25	5 and 12 and 19 and 24
26	limit 75 to (English language and humans)