Critical Review:
Is the endoscopic swallowing assessment more sensitive than the videofluoroscopic swallowing assessment at identifying penetration or aspiration in adults with dysphagia?

Beth Jeronimo
M.Cl.Sc Speech-Language Pathology Candidate
University of Western Ontario: School of Communication Sciences and Disorders

This critical review examines whether the endoscopic swallowing assessment is more sensitive than the videofluoroscopic swallowing assessment at identifying penetration or aspiration in adults with dysphagia. Six studies, all within-subjects designs, are reviewed. Overall, research suggests that endoscopic assessment is a sensitive, reliable method for identifying aspiration or penetration, and evaluating swallowing safety in patients with dysphagia. However, evidence supporting the use of endoscopic over videofluoroscopic assessment is inconclusive, and it is recommended that these methods be used as complimentary, rather than exclusive, tools.

Introduction

Videofluoroscopic assessment of swallowing, or the Modified Barium Swallow (MBS), has long been viewed as the 'gold standard' of instrumental swallowing assessments. Recently however, evidence has emerged in support of a new assessment technique using nasendoscopy, or Fiberoptic Endoscopic Evaluation of Swallowing (FEES). The MBS captures views of the oral, pharyngeal and esophageal stages of swallowing through radiographic imaging taken while patients trial foods of different consistencies mixed with barium (Madden, 2000). The MBS is conducted in a radiological suite, not at bedside, and relies on the availability of the radiologist and the SLP (Madden, 2000). FEES, on the other hand, enables clinicians to assess the function of the palate, pharynx and larynx through use of a nasolaryngoscope while patients trial foods of different consistencies mixed with food dye (Bastian, 1993).

FEES has gained popularity due to its advantages over the MBS which include conducting this assessment at bedside, the ability to repeat the assessment multiple times due to no exposure to radiation, and its use as a biofeedback tool to help patients develop a safe swallow (Leder, 1998). However, an advantage of the MBS, that FEES does not allow for, is visualization of the oral and esophageal phases of the swallow (Madden, 2000). With these factors in mind, the following studies were conducted with the goal of determining which assessment is most sensitive at identifying penetration or aspiration in adults with dysphagia.

Objectives

The primary objective of this paper is to provide a summary and critical evaluation of existing literature on the sensitivity of the endoscopic and videofluoroscopic swallowing assessments in identifying aspiration or penetration. The secondary objective is to provide recommendations for evidence-based clinical practice. This information will allow clinicians to make informed decisions regarding the most appropriate assessment given the aspect of the swallow that is being evaluated, the medical status of the patient, and the impact it will have on dysphagia management.

Methods

Search Strategy

Articles were found by searching computerized databases, including PubMed, GLOBUS and Google Scholar using the following terms: (dysphagia) AND (assessment) AND (Modified Barium Swallow) OR (Videofluoroscopy) AND (Fiberoptic Endoscopic Evaluation of Swallowing) OR (nasendoscopy). The search was limited to English journals and limitations were not set on date of publication.

Selection Criteria

Studies selected for inclusion in this review were required to evaluate the use of videofluoroscopy and nasendoscopy on adults. All studies were required to investigate the identification of aspiration and/or penetration. All subjects were referred for a swallowing assessment due to pre-existing conditions that are known to affect swallowing.

Data Collection

Results of the literature search yielded six articles which met the above selection criteria. The articles include six within-subjects designs.

Results

Kelly, Drinnan, & Leslie (2007) investigated whether the type of assessment, FEES or videofluoroscopy, influenced the scoring of penetration and aspiration using a within-subjects design. The purpose of this study was to determine the best way of assessing for risk of aspiration pneumonia, and for making decisions

about oral intake recommendations. In order to compare assessments on the same swallow, 15 participants underwent assessments simultaneously. Scoring was based on the Penetration Aspiration Scale, an 8-point scale that rates penetration and aspiration based on depth of entry of food into the airway, and whether or not the material is cleared.

Raters in this study were carefully selected and were blinded to participant information and the pairing of FEES and videofluoroscopic recordings. Intra- and interrater reliability were calculated using weighted Kappa. A five-way ANOVA was appropriately used to assess differences in ratings, and patient and examination type were found to have the most significant effect on scores.

Results indicate that the type of assessment does influence judgment of the severity of penetration or aspiration, and therefore these assessments cannot be used interchangeably. When the same swallow was assessed using both tools, Penetration Aspiration Scale scores were significantly higher with FEES. This suggests that penetration and aspiration are rated as more severe when using FEES. However, it is recommended by Kelly et al. (2007) that more research is needed to determine whether one assessment has a more clinically significant impact in terms of predicting the likelihood of aspiration pneumonia.

Strengths of this study include use of FEES and videofluoroscopy simultaneously to assess swallowing function. This was also the only study reviewed that used a standardized scoring method. Selection criteria and reliability measures used for raters of these assessments, and appropriate statistical analysis of the data are also strengths of this study. A limitation of this study is the small sample size and the fact that the sample selected is not representative of the general population of those with dysphagia. There is also no participant selection criteria identified, aside from the fact that they were referred for a swallowing assessment. Based on these limitations and strengths, this study provides a suggestive level of evidence.

Rao, Brady, Chaudhuri, Donselli, & Wesling (2003) conducted a prospective pilot study on 11 patients to determine sensitivity and specificity values for laryngeal penetration and tracheal aspiration for the videofluoroscopic swallow study (VFSS) and FEES. All participants had suspected laryngeal or pharyngeal abnormality or dysphonia, and underwent VFSS and FEES simultaneously. This study differs from the others presented because sensitivity and specificity were calculated both when VFSS was the gold standard, and when FEES was the gold standard. This was done by the researchers in order to determine the validity of each assessment. Interrater reliability was calculated, as well as 2x2 contingency tables, Kappa correlation and Fisher's exact test, all appropriate for this study.

Data is clearly presented to support the findings that sensitivity, in identifying aspiration and penetration, was higher when FEES was used as the gold standard. Conversely, specificity values were higher when VFSS was used as the gold standard. Furthermore, most similar agreement between VFSS and FEES was found when identifying aspiration. Overall, researchers conclude that both assessments should be considered the gold standard and use should be based on clinical decision-making and equipment availability.

Strengths of this study include both assessments being completed simultaneously, both assessments being used as the 'gold standard', and appropriate data analysis. However, a stated limitation of this study is that patients were taught safe swallow strategies before being assessed, to reduce the risk of penetration and aspiration. Although this was done in order to ensure safety of patients, decreasing the likelihood that patients will aspirate or penetrate makes it difficult to assess these swallowing issues. Researchers recommend that in order to enhance the validity of this research this teaching should be avoided in the future. Another limitation is the small, heterogeneous sample. Researchers recognize that this study should be replicated with more participants. While data analysis and methods of this study are clear strengths, factors related to participant teaching and sample size lead to an equivocal level of evidence.

Leder, Sasaki, & Burrell (1998) used a within-subjects design to compare the reliability of FEES and MBS in identifying silent aspiration in dysphagic patients. The basis for determining the reliability of these methods was that the clinical bedside swallowing evaluation does not assess the pharyngeal phase of the swallow and therefore is not able to identify silent aspiration.

This study included 400 subjects, reflective of the 'general hospital population' of those with dysphagia. One subgroup (343 subjects) were evaluated using only FEES. Of relevance to the present review is the second subgroup (57 subjects), which was assessed using both FEES and MBS. Agreement between assessments regarding presence of aspiration in subgroup 2 was found for 96% of cases.

Based on this data it was concluded that FEES is a reliable method of identifying silent aspiration, compared to the MBS. Support in favor of FEES is provided based on its advantages (avoids radiation

exposure, can be repeated, can be videotaped, can be transported).

A strength of this study includes its methods. First of all, it is stated by the researchers that protocol for MBS and FEES examinations were followed. Additionally, both assessments were reviewed by an otolaryngologist and a radiologist, and raters were blinded to each other's results. Reviewers also reached 100% agreement regarding identification of silent aspiration. A significant limitation however is that assessments were not conducted simultaneously. Another limitation is that although the participant group of relevance to this review contains a moderate sample size, it is not indicated whether this subgroup is representative of those with dysphagia. Based on these strengths and limitations this study provides a suggestive level of evidence.

Wu, Hsiao, Chen, Chang, & Lee (1997) investigated whether more support could be given to FEES or videofluoroscopy in the assessment of swallowing safety. The purpose of identifying a new method of assessment was that videofluoroscopy is not always accessible, and subjective bedside assessments have many limitations.

Participants included 28 adults with a history of dysphagia due to cerebrovascular accident, Parkinson's disease, head injury, or nasopharyngeal carcinoma. Participants underwent videofluoroscopy assessment and FEES during a 2-week period in a within-subjects design. Swallows were evaluated based on 5 features of the swallow including laryngeal penetration and tracheal aspiration.

It was concluded that FEES, overall, is more reliable and sensitive at identifying swallow safety, than videofluoroscopy. Compared to videofluoroscopy, FEES was found to have a higher percentage rate of identification for both laryngeal penetration and tracheal aspiration. In terms of agreement between assessments only 85% agreement was found on these two items, and this is a reported limitation of this study. Researchers caution that this discrepancy between assessments may be due to FEES having a lower false negative rate because it can be repeated, or due to FEES having better viewing of bolus localization and movement.

Other limitations of this study include the small sample size that is not representative of the general population with dysphagia. This is problematic because it restricts the generalization of findings. Additionally, assessments were conducted on separate occasions and therefore evaluated different behaviours, which limits the direct comparability of these results. Researchers report on this limitation and conclude that it prevents a determination of which assessment is 'correct'. Finally, findings from this study are supported by qualitative data only and no statistical analysis was conducted. This study therefore provides an equivocal level of evidence.

Singh et al. (2008) performed a within-subjects design to determine the correlation of milk nasendoscopy and videofluoroscopy in the detection of aspiration in 100 patients with neurologically-based dysphagia. Both assessments were performed on the same day, and aspiration in the pre-swallow, swallow and post-swallow phases was identified. Correlation of results was reviewed by Kappa test and the difference was examined with Chi square test; both appropriate for this study.

Results indicate that there was 58% agreement in the pre-swallow phase, 63% agreement in the swallow phase, and 58% in the post-swallow phase between assessments when identifying aspiration. It was also found that milk nasendoscopy was more sensitive in the identification of aspiration post-swallow, and videofluoroscopy was more sensitive in the pre-swallow phase, with no difference during the swallow.

A strength of this study is that it includes a fairly large homogeneous sample, and therefore a statement of findings could be made about this population. Statistical analysis is also a strength of this study. A limitation is that assessments were not completed simultaneously. Another limitation is that it is not indicated who completed the ratings of these swallowing assessments, and reliability measures are not included, making it difficult to judge the suitability of the raters. Based on stated strengths and limitations this study provides an equivocal level of evidence.

Madden, Fenton, Hughes, & Timon (2000) compared milk-swallow endoscopy and videofluoroscopy in the assessment of swallowing function, due to discussed limitations of videofluoroscopy (exposure to radiation, inability to test bed-ridden patients etc.). A prospective study was carried out on 20 sets of assessments from 17 patients who had dysphagia due to stroke or surgery for head and neck cancer. Four criteria essential for swallowing safety were assessed, including aspiration. Sensitivity, specificity, positive and negative predictive values were analysed, and Fisher exact test was used, all appropriate for this study.

Results indicate that milk-swallow endoscopy has high specificity and sensitivity in detecting the four criteria for swallowing safety, including aspiration. Data also indicates good predictive correlation between assessments, which reflects the clinical usefulness of

endoscopy. Overall, Madden et al. conclude that endoscopy is as sensitive as videofluoroscopy in detecting aspiration, and should be used as a screening tool, for follow-up of patients postoperatively and during therapy. However, the authors do not rule out videofluoroscopy, arguing instead that it should serve as a complementary assessment.

Statistical analysis is a strength of the study. A limitation of this research is that this was a pilot study and was based on a very small sample. Another limitation is that it was not indicated who rated these assessments and whether rater reliability measures were included. Finally, assessments were conducted up to two weeks apart, during which time swallowing status could have changed. Although the researchers mention that time between assessments was minimized to avoid this, it leads to less compelling findings than studies where assessments were performed simultaneously. This study therefore provides an equivocal level of evidence.

Discussion

To date, videofluoroscopy has been considered the 'gold standard' of clinical swallowing assessments. The studies reviewed here demonstrate the effectiveness of a second clinical swallowing assessment: endoscopy. Before drawing conclusions from this research, several limitations listed in this review warrant further discussion.

One limiting factor is the small sample sizes included in most studies reviewed. Kelly et al. (2007), Rao et al. (2003), Wu et al. (1997) and Madden et al. (2000) all used between 11 and 28 participants. In all studies reviewed there is also little information provided about selection criteria for participants or about participants themselves. In all studies the cause of dysphagia is indicated, however most studies do not include a sample that is representative of all those with dysphagia. This, along with the limited number of participants, makes it difficult to generalize research findings.

Another limiting factor is that assessments are not conducted simultaneously in all studies. If assessments are conducted at different times it means that each assessment is evaluating a different behavior. This is problematic because it then comes into question whether these different behaviours can be compared. Finally, the majority of studies reviewed fail to indicate how raters of the videofluoroscopic and endoscopic assessments were selected, and rater reliability measures are not included. If assessments are not rated consistently by qualified individuals the validity of research findings may need to be questioned.

Conclusion and Clinical Implications

The studies reviewed provide an equivocal level of evidence that the endoscopic assessment of swallowing is a more sensitive tool than videofluoroscopy when identifying aspiration and penetration. Rather, it is concluded in these studies that both methods are equally reliable and should be used as complementary tools. Several of the studies indicate that the evaluation of a patient necessitates the use of both tools in order to obtain an accurate description of swallow function. It is recommended that clinical judgment and equipment availability be used in order to determine the most appropriate assessment.

While these studies do not indicate that one assessment is superior to the other, they do provide clinically useful information. Together these studies indicate which assessment allows for best visualization of each phase or anatomical component of the swallow. It is then left up to the clinician to determine the most appropriate assessment based on each individual client. Suggestions are also made regarding specific populations or instances best suited to each assessment. Rao et al. (2003) suggest that those patients with suspected laryngeal or pharyngeal abnormality are best suited to FEES, whereas those with esophageal abnormality would be best evaluated with VFSS. Madden et al. (2000) also recommend that endoscopy should be used as a screening tool, for follow-up of patients postoperatively and during therapy.

Future Research Recommendations

Future research is still required to determine the sensitivity of both assessment tools in identifying aspiration and penetration. Such research may include larger participant groups with a variety of medical diagnoses. Research may also include videofluroscopic and endoscopic assessments conducted simultaneously, standardized rating methods, and reliability measures obtained on those individuals rating the swallows. Finally, while it is recommended in the studies reviewed that both assessments be conducted on patients to ensure a thorough assessment, there is no consideration of the cost and feasibility of doing so in a hospital setting in terms of time and resources. While assessments in combination offer more information than both in isolation, using both is not cost effective. Future research may consider the sensitivity of these assessment tools in identifying aspiration and penetration, as well as how their use is influenced by these clinical factors.

References

Kelly, A. M., Drinnan, M. J., & Leslie, P. (2007). Assessing penetration and aspiration: How do Videofluoroscopy and Fiberoptic Endoscopic Evaluation of Swallowing compare? *The Laryngoscope, 117*(10), 1723-1727.

Leder, S. B., Sasaki, C. T., & Burrell, M. I. (1998). Fiberoptic Endoscopic Evaluation of Dysphagia to Identify Silent Aspiration. *Dysphagia, 13*(1), 19-21.

Madden, C., Fenton, J., Hughes, J., & Timon, C. (2000). Comparison between videofluoroscopy and milk-swallow endoscopy in the assessment of swallowing function. *Clinical Otolaryngology, 25*(6), 504-506.

Rao N., Brady, S. L., Chaudhuri, G., Donselli J. J., & Wesling, M. W. (2003). Gold-standard? Analysis of the videofluoroscopic and fiberoptic endoscopic swallow examinations. *Journal of Applied Research*, 3(1), 89-96.

Singh, V., Berry, S., Brockbank, M., Frost, R., Tyler, S., & Owens, D. (2009). Investigation of aspiration: Milk nasendoscopy versus videofluoroscopy. *European Archives of Oto-Rhino-Laryngology, 266*(4), 543-545.

Wu, C. H., Hsiao, T. Y., Chen, J. C., Chang, Y. C., & Lee, S. Y. (1997). Evaluation of Swallowing Safety with Fiberoptic Endoscope: Comparison with Videofluoroscopic Technique. *The Laryngoscope, 107*(3), 396-401.