



The Diagnostic Test Accuracy of Clinical Swallow Assessment for Oropharyngeal Aspiration: A Systematic Review

Marissa Romano BSpPath

Student number: a1619892

The Joanna Briggs Institute and School of Translational Health Science,
Faculty of Health Sciences, the University of Adelaide.

marissa.romano@adelaide.edu.au

Table of Contents

The Diagnostic Test Accuracy of Clinical Swallow Assessment for Oropharyngeal Aspiration: A Systematic Review.....	i
Table of Contents	ii
Table of Figures.....	v
Table of Tables	vi
Glossary of terms.....	vii
Acronyms.....	ix
Acknowledgments.....	x
Student Declaration	xi
Abstract	xii
1 Background	16
1.1 Systematic reviews	16
1.2 The systematic review protocol	17
1.3 Systematic review versus traditional literature review	18
1.4 Studies of diagnostic test accuracy	19
2 Introduction.....	22
2.1 Introduction to oropharyngeal aspiration.....	22
2.1.1 Definition of oropharyngeal aspiration.....	22
2.1.2 Neurology of swallowing and airway protection	22
2.1.3 Pathophysiology of chronic oropharyngeal aspiration	23
2.1.4 Causes and prevalence of oropharyngeal aspiration	24
2.1.5 Response to oropharyngeal aspiration in paediatrics compared with adults.....	27

2.1.6	Diagnostic testing for oropharyngeal aspiration	27
3	Systematic Review Protocol	31
3.1	The systematic review protocol	31
3.1.1	Background.....	31
3.1.2	Inclusion criteria	35
3.1.3	Focus of the review	37
3.1.4	Search strategy	38
3.1.5	Methods of the review	39
4	Results.....	41
4.1	Results.....	41
4.1.1	Results of the search	41
4.1.2	Methodological quality of the included papers	43
4.1.3	Results from individual QUADAS checklist items.....	46
4.1.4	Assessing methodological quality using the STARD checklist ..	52
4.1.5	Data from individual studies	58
4.1.6	Data analysis and synthesis.....	74
4.1.7	Meta-analysis.....	76
4.1.8	Sources of heterogeneity	80
4.1.9	Sub-group analysis	81
4.1.10	Summary of findings	83
5	Discussion	86
5.1	Discussion	86
5.1.1	Summary of main results	86
5.1.2	Findings of meta-analysis	86
5.1.3	Population and settings.....	87
5.1.4	Reference test methods.....	87

5.1.5	Index test methods.....	88
5.1.6	Effect of fatigue and aspiration.....	89
5.1.7	Heterogeneity of study results.....	89
5.1.8	The diagnostic test accuracy of elements of the CSA	94
5.1.9	Prevalence and positive and negative predictive values	96
6	Conclusions.....	99
6.1.1	Implications for practice	99
6.1.2	Concluding Summary.....	102
6.1.3	Implications for research	103
6.1.4	Limitations of the review.....	103
7	References	104
8	Appendices.....	114
8.1	Appendix I: Initial Search Terms.....	114
8.2	Appendix II: Electronic databases	115
8.3	Appendix III: PubMed Search Terms.....	116
8.4	Appendix IV: Critical appraisal tool – the QUADAS check list ⁸³	117
8.5	Appendix V: Data extraction tool – the STARD check list ⁸⁴	118
8.6	Appendix VI: Data extracted from individual studies and combined in meta-analysis	121
8.7	Appendix VII: Table of included studies.....	122
8.8	Appendix VIII: Table of excluded studies.....	132

Table of Figures

Figure 1 Flow chart outlining the study identification process.....	42
Figure 2 Methodological quality graph.....	44
Figure 3 Methodological quality item for each study included in meta-analysis.....	45
Figure 4 Forest plot of the included studies.	76
Figure 5 Summary receiver operator characteristic (sROC) plots of the sensitivity and specificity for all included studies.	78
Figure 6 Flowchart demonstrating the likelihood of a test result being accurate differs for a positive or negative test result.....	85

Table of Tables

Table 1 A typical 2x2 table used to classify patient test results and the presence or absence of disease or target condition.	20
Table 2 Differences between Bours review ⁸⁷ and this systematic review. ...	43
Table 3. Meta-analysis of diagnostic accuracy.	77
Table 4 A 2x2 classification table constructed using data extracted from the included studies.	80
Table 5 Summary of study findings.....	84
Table 6 Between study differences and similarities used to rate the presence or absence of aspiration during clinical swallow assessment and the cut-off values used to determine the score.	93
Table 7 Diagnostic accuracy of various features of CSA as presented in 5 of the 13 included studies..	95
Table 8 Comparison of PPV and NPV using varying prevalence of aspiration.	97
Table 9 A 2x2 classification used to determine PPV and NPV for prevalence of 60%	97
Table 10 A 2x2 classification used to determine PPV and NPV for prevalence of 10%.	98

Glossary of terms

Clinical Swallow Assessment: non-instrumental, non-radiologic assessment of swallow function.

Diagnostic test accuracy: the ability of a test to distinguish between patients with a target disease or condition from those without the disease or condition.

Dysphagia: difficulty in any of the four phases of swallowing.

False negative: index test result is negative, reference test result is positive.

False positive: index test result is positive, reference test result is negative.

Index test: the 'new' test or test in question, in a study of diagnostic test accuracy.

Oropharyngeal aspiration: the entry of food and/or fluids below the level of the vocal cords.

Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist: 14 point checklist used to assess the methodological quality of studies of diagnostic test accuracy.

Reference test: the 'gold standard' test against which the index test is compared in a study of diagnostic test accuracy.

Sensitivity: the ability of a test to correctly identify those with the disease or target condition.

Specificity: the ability of a test to correctly identify those without the disease or target condition.

Standards for Reporting of Diagnostic Accuracy (STARD) checklist: 25 item checklist used to extract data of studies of diagnostic test accuracy.

Summary receiver operating characteristic plot: graphical representation used to describe the performance of a diagnostic test based on data from meta-analysis.

True negative: index test result and reference test result are negative.

True positive: index test result and reference test result are positive.

Video Fluoroscopic Swallow Study: radiographic assessment of swallow function.

Acronyms

CSA: clinical swallow assessment

CVA: cerebrovascular accident

DTA: diagnostic test accuracy

FEES: fiberoptic endoscopic evaluation of swallowing

FN: false negative

FP: false positive

QUADAS: Quality Assessment of Diagnostic Accuracy Studies

sROC plot: summary receiver operating characteristic plot

STARD: Standards for Reporting of Diagnostic Accuracy

TN: true negative

TP: true positive

VFSS: video fluoroscopic swallow study

Acknowledgments

I would like to thank the following people for the generosity of their time and expertise during the writing of this thesis: supervisors Dr Tim Schultz, Dr Andrew Tai, Dr Sarahlouise White, research librarian Maureen Bell, statistician Thomas Sullivan, second reviewer May Thwin, the Speech Pathology Department, Women's and Children's Hospital and Haruka Tohara, author of one of the included studies for providing the raw data clearly describing the number of participants in the study with aspiration present and aspiration absent.

Student Declaration

This work contains no material that has been accepted for the award of any other degree or diploma in any University of any other tertiary institution, and, to the best of my knowledge and belief, contains no material previously published or written by any other person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Catalogue, and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed:

Dated:29/4/2014

Abstract

Background

Oropharyngeal aspiration, the recurrent entry of food and/or fluids below the level of the vocal cords, can result in a range of complications including: chronic lung diseases, aspiration pneumonia, malnutrition and/or dehydration. Video fluoroscopic swallow study is the Gold Standard assessment of oropharyngeal aspiration but is resource intense, exposes the patient to radiation and is not available in all hospitals and centres. The Clinical Swallow Assessment is a bedside swallow assessment widely used to screen and/or assess for oropharyngeal aspiration. The evidence base behind the diagnostic test accuracy of the Clinical Swallow Assessment has not previously been synthesised.

Objectives

To synthesise the best available evidence on the diagnostic test accuracy (sensitivity and specificity) of clinical swallow assessment compared with Video Fluoroscopic Swallow Study in diagnosing oropharyngeal aspiration in children and adults with dysphagia.

Inclusion criteria

Types of participants

Any patients referred for swallowing assessment, specifically assessed for oropharyngeal aspiration were included and there was no exclusion based on age or gender. Study results were excluded for head and neck cancer patients, patients with a tracheostomy *in situ* and patients with craniofacial anomalies.

Focus of the review

The focus of the review was to examine the diagnostic test accuracy of clinical swallow assessment, as compared with Video Fluoroscopic Swallow Study.

Types of studies

This systematic review considered any relevant cross sectional study that measured diagnostic test accuracy.

Types of outcomes

Outcomes of interest were the sensitivity and specificity of the clinical swallow, as compared with the video fluoroscopic study and the positive and negative predictive values. Where this data was not reported in the studies, these measures were calculated from the reported raw data.

Search strategy

Thirteen major databases were searched from their inception until April 31st 2012. There were no limits during the search stage as relevant studies were omitted if search filters such as 'English' and 'Human' were applied.

Methodological quality

Methodological quality was assessed using the QUADAS checklist. Data was collected using the STARD checklist. Sensitivity and specificity measures were combined in meta-analysis to generate a summary receiver operator characteristic plot.

Results

There were 1787 titles initially identified. Following duplicate removal and screening against inclusion criteria, 37 papers were retrieved for detailed examination and 24 papers were excluded as they did not meet the inclusion criteria. The most common reason for exclusion was that the paper was not a study of diagnostic test accuracy. There were 13 studies included in the systematic review and found to have high methodological quality. Data extracted from individual studies was statistically combined in meta-analysis to produce a forest plot and summary receiver operating characteristic (sROC) plot. Heterogeneity was evident in the forest plot, particularly for sensitivity as evidenced by the wider confidence intervals for sensitivity compared with specificity. The test sensitivity varied from 21% to 93%, the specificity from 46% to 93%. The summary mean sensitivity and specificity was calculated as 71% and 76% respectively. Positive predictive value was

calculated as 60% and negative predictive value was 81%. The scatter of points around the curve on the sROC plot also indicated heterogeneity. Sources of heterogeneity were identified and explored. The shape of the sROC curve strongly supported the finding of a threshold effect, which is expected for studies in which there is a strong interpretative component such as the clinical swallow assessment. This occurs as clinicians may vary in their criteria for what constitutes a positive or negative test result. The overall prevalence of aspiration in the included studies was calculated as 35%. Results are based predominantly on adult, acute post stroke patients.

Conclusion

This thesis provides good evidence for an overall estimate of the sensitivity and specificity of clinical swallow assessment compared with video fluoroscopic swallow study for the assessment of oropharyngeal aspiration. In this population, a clinician can be much more confident in a negative test result than a positive test result. A false positive test result may lead to unnecessary patient care and costs, including with-holding oral medications and prescription of modified diets and/or fluids. A false negative test result may lead to compromised lung health and/or pneumonia.

Implications for practice

Using calculations of the positive predictive values and negative predictive values, 60% of patients who test positive for aspiration are truly aspirating and 81% of patients who test negative for aspiration are truly not aspirating. Positive and negative test results are affected by the prevalence of the condition in the population. To summarise, the PPV increases and the NPV decreases as prevalence increases and the PPV decreases and NPV increases as the prevalence decreases. For example if the prevalence is much lower (e.g. 10%) the NPV rises to 96% and the PPV decreases to 24%. This thesis provides data for centres where VFSS is not available regarding the diagnostic test accuracy of clinical swallow assessment for oropharyngeal aspiration.

Implications for Research

Only one of the included studies provided data for infants and children. None of the included studies addressed infants, children or adults without a neurological aetiology. Further research is needed for infants and children with dysphagia as well as neurologically intact and normally developing infants, children and adults.