

# **Videofluoroscopy of Swallow: A Study of Intra and Inter-Rater Reliability of SLT Ratings of Aspiration, Vallecular Residue and Pyriform Sinus Residue**

**Julie Coats**

**Speech and Language Therapist  
NHS Greater Glasgow and Clyde**

# Are we reliable in our interpretation of videofluoroscopy studies?

# Research Questions

- 1) What are the current levels of intra and inter-RR for SLT ratings of a) aspiration b) vallecular residue and c) pyriform sinus residue?
- 2) Does level of training and experience of SLT raters impact on intra-RR and inter-RR?
- 3) What are the implications for clinical practice associated with the findings from questions 1 and 2?

# Method

## Participants

9 out of 11 VF practitioners participated.

## Process

- 40 single swallows of normal fluids
- Aspiration, vallecular residue and pyriform sinus residue rated on specific scales
- Participants blinded to patient details
- Same swallows rated 1 month later in randomised order

# Statistics

- Inter-rater reliability -> Light's Kappa and Intraclass correlation
- Intra-rater reliability -> Weighted Kappa
- Data inputted and statistics applied using 'R'
  
- Statistical values linked to Landis and Koch (1977) terms:
  - **0.41 – 0.60 = moderate**
  - **0.61 - 0.80 = substantial**
  - **0.81 - 1 = almost perfect**

# Aspiration Scale (Kuhlemeier et al., 1998)

None	0
Penetration Only	1
Mild aspiration	2
Moderate aspiration	3
• Severe aspiration	4

# Reliability of Aspiration Ratings

## Inter-RR (analysed as 2 groups)

Initial swallow = 'moderate' (kappa 0.488)

Whole video = 'almost perfect' (kappa 0.810)

## Intra-RR

- Highest degree of variability amongst the raters

"Moderate" intra-RR	2
"Substantial" intra-RR	3
"Almost perfect" intra-RR	3

# Pharyngeal Residue Scale (Kelly et al., 2006)

<b>No pharyngeal residue or coating</b>	<b>0</b>
<b>Coating of the pharyngeal mucosa; no pooling</b>	<b>1</b>
<b>Mild pooling/ residue</b>	<b>2</b>
<b>• Moderate pooling/ residue</b>	<b>3</b>
<b>Severe pooling/ residue</b>	<b>4</b>



# Reliability of Residue Ratings

## Inter-RR

Inter-RR was 'substantial' for ratings of both:

- Vallecular residue ( $\kappa= 0.644$ )
- Pyriform sinus residue ( $\kappa= 0.715$ ).

## Intra-RR

- Vallecular residue ( $\kappa$  0.620-0.841)
- Pyriform sinus residue ( $\kappa$  0.699-0.892)

# Intra-Rater Reliability per Rater (Weighted Kappa)

Level of Training	Aspiration	Vallecular Residue	Pyriform Sinus Residue
Level 3	0.422	0.806	0.822
Level 2	0.619	0.662	0.699
Level 2	0.867	0.695	0.729
Level 3	0.947	0.620	0.811
Level 3	0.962	0.823	0.868
Level 2	0.597	0.791	0.892
Level 3	0.665	0.841	0.747
Level 3	0.767	0.728	0.712

**moderate**; 0.61- 0.80 = **substantial**; 0.81- 1 = **almost perfect**; Landis and Koch, 1977)

*Delivering better health*

# Rater Experience

- No relationship found between level of experience and intra-RR
- Inter-RR was higher when ratings made by Level 3 practitioners were compared

# Summary of Findings

- High levels of inter and intra-rater reliability were not consistently achieved.
- Clinically should be achieving at least 80% agreement (Martin-Harris et al., 2008)

# Implications for Practice

- Establishing shared terminology of swallow structures
- Use of agreed descriptions/ rating scales and operationalisation of terms
- Restructuring of peer-review to focus on improving reliability with 80% level of agreement deemed 'reliable'
- Introducing need for reliability assurance before trainee VF practitioners are signed off.

# References

- Belafsky, P.C. and Kuhn, M.A. (2014) *The Clinician's Guide to Swallowing Fluoroscopy*. New York: Springer Science and Business Media .
- Eisenhuber, E., Schima, S., Schober, E., Pokieser, P., Stadler, A., Scharitzer, M. and Oschatz, E. (2002) 'Videofluoroscopic Assessment of Patients with Dysphagia: Pharyngeal Residue is a Predictive Factor for Aspiration', *Gastrointestinal Imaging*, 178 (2), pg. 393-398.
- Kelly, A.M., Leslie, P., Beale, T., Payten, C. and Drinnan, M.J. (2006) 'Fibreoptic endoscopic evaluation of swallowing and videofluoroscopy: does examination type influence perception of pharyngeal residue severity?', *Clinical Otolaryngology*, 31, pp. 425-432.
- Kim, D.H., Choi, K.,H., Kim, H.M., Koo, J.H., Kim, B.R., Kim, T.W., Ryu, J.S., Im. S., Choi, I.S., Pyun, S.B., Park, J.W., Kang, J.Y. and Yang, H.S. (2012) 'Inter-rater Reliability of the Videofluoroscopic Dysphagia Scale', *Annals of Rehabilitation Medicine*, 36 (6), pp. 791-796.
- Kuhlemeier, K.V., Yates, P. and Palmer, M.D. (1998) 'Intra- and Interrater Variation in the Evaluation of Videofluorographic Swallowing Studies', *Dysphagia*, 13, pp. 142-47.

# References

Landis, R.J. and Koch, G.G. (1977) 'The Measurement of Observer Agreement for Categorical Data', *Biometrics*, 33 (1), pp. 159-174.

Martin-Harris, B., Brodsky, M.B., Michel, Y., O'Castell, D.O., Schleicher, M., Sandidge, J., Maxwell, R. and Blair, J. (2008) 'MBS Measurement Tool for Swallowing Impairment- MBSImp: Establishing a Standard', *Dysphagia*, 23, pp. 392-405.

Pearson Jr, W.G., Hindson, D., Langmore, S.E. and Zumwalt A.C. (2012) 'Evaluating Swallowing Muscles Essential for Hyolaryngeal Elevation by Using Muscle Functional Magnetic Resonance Imaging', *International Journal of Radiation, Oncology, Biology and Physics*, 85 (3), pp. 735-740.

- Rommell, N., Borgers, C., Beckevoort, D.V., Goeleven, A., Dejaeger, E. and Omari, T.I. (2015) 'Bolus Residue Scale: An Easy-to-Use and Reliable Videofluoroscopic Analysis Tool to Score Bolus Residue in Patients with Dysphagia', *International Journal of Otolaryngology*, <http://dx.doi.org/10.1155/2015/780197>

Stoeckli, S.J., Huisman, T.A.G.M., Seifert, B. and Martin-Harris, B.J.W. (2003) 'Interrater Reliability of Videofluoroscopic Swallow Evaluation', *Dysphagia*, 18, pp. 53-57.