Videofluoroscopy of Swallow: A Study of Intra an 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

<table>
<thead>
<tr>
<th>Description</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Moderate” intra-RR</td>
<td>2</td>
</tr>
<tr>
<td>“Substantial” intra-RR</td>
<td>3</td>
</tr>
<tr>
<td>“Almost perfect” intra-RR</td>
<td>3</td>
</tr>
<tr>
<td>Description</td>
<td>Score</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>No pharyngeal residue or coating</td>
<td>0</td>
</tr>
<tr>
<td>Coating of the pharyngeal mucosa; no pooling</td>
<td>1</td>
</tr>
<tr>
<td>Mild pooling/ residue</td>
<td>2</td>
</tr>
<tr>
<td>Moderate pooling/ residue</td>
<td>3</td>
</tr>
<tr>
<td>Severe pooling/ residue</td>
<td>4</td>
</tr>
</tbody>
</table>
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)

<table>
<thead>
<tr>
<th>Level of Training</th>
<th>Aspiration</th>
<th>Valleeular Residue</th>
<th>Pyriform Sinus Residue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 3</td>
<td>0.422</td>
<td>0.806</td>
<td>0.822</td>
</tr>
<tr>
<td>Level 2</td>
<td>0.619</td>
<td>0.662</td>
<td>0.699</td>
</tr>
<tr>
<td>Level 2</td>
<td>0.867</td>
<td>0.695</td>
<td>0.729</td>
</tr>
<tr>
<td>Level 3</td>
<td>0.947</td>
<td>0.620</td>
<td>0.811</td>
</tr>
<tr>
<td>Level 3</td>
<td>0.962</td>
<td>0.823</td>
<td>0.868</td>
</tr>
<tr>
<td>Level 2</td>
<td>0.597</td>
<td>0.791</td>
<td>0.892</td>
</tr>
<tr>
<td>Level 3</td>
<td>0.665</td>
<td>0.841</td>
<td>0.747</td>
</tr>
<tr>
<td>Level 3</td>
<td>0.767</td>
<td>0.728</td>
<td>0.712</td>
</tr>
</tbody>
</table>

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

**Delivering better health**

www.nhsggc.org.uk
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


References


