

# Inter- and intra-rater reliability for assessment of swallowing using a hand-held portable ultrasound device

R. Hammond<sup>1,2</sup>, A. Dimmock<sup>2</sup>; K. Winiker<sup>1</sup>; M.L. Huckabee<sup>1</sup>



<sup>1</sup>The University of Canterbury Rose Centre for Stroke Recovery and Research, School of Psychology Speech and Hearing, Christchurch, New Zealand

<sup>2</sup>Waitematā District Health Board, Auckland, New Zealand



## Introduction

Ultrasound (US) assessment of swallowing has been documented as reliable in both healthy<sup>1</sup> and dysphagic participants<sup>2,3</sup>, primarily based on offline image analysis. Despite this, US has not translated into standard clinical practice. This may be due to cost as well as the time required to analyse imaging offline. Recently developed inexpensive, portable wireless US technology has the potential for increased access and immediate results. Reliability of image acquisition and measurement of US images obtained with Clarius™ in a pressured clinical environment is explored.

## Materials & Methods

- + **Participants:** Eight patients, aged 33-96 with mixed aetiologies, were recruited following referral for a videofluoroscopic swallowing study (VFSS).
- + **Instrumentation:** A curvilinear Clarius™ US device (frequency range: 2-6 MHz, depth: 3-30 cm), wirelessly connected to an iPad, was used to acquire images from dynamic swallowing gestures - hyoid excursion and thyrohyoid approximation. A linear Clarius™ transducer (frequency range: 4-13 MHz, depth: 1-7 cm) was used to collect measures of tongue thickness and cross sectional area of submental muscles at rest.
- + **Data Acquisition and Measurement:** Data were independently collected by two investigators within the same day. Online measurements of images were completed during the exam, using Clarius™ software on an iPad. Offline measurements were completed using ImageJ software on a large screen.
- + **Reliability assessment:** Inter-rater reliability was calculated with intraclass correlation coefficient (ICC).
  - + Effect of data acquisition on reliability was explored by calculating online inter-rater ICC and comparing with offline inter-rater ICC.
  - + Effect of environmental, equipment and time constraints on online measurement was explored by calculating ICC of online and offline measurement of the same acquired images.

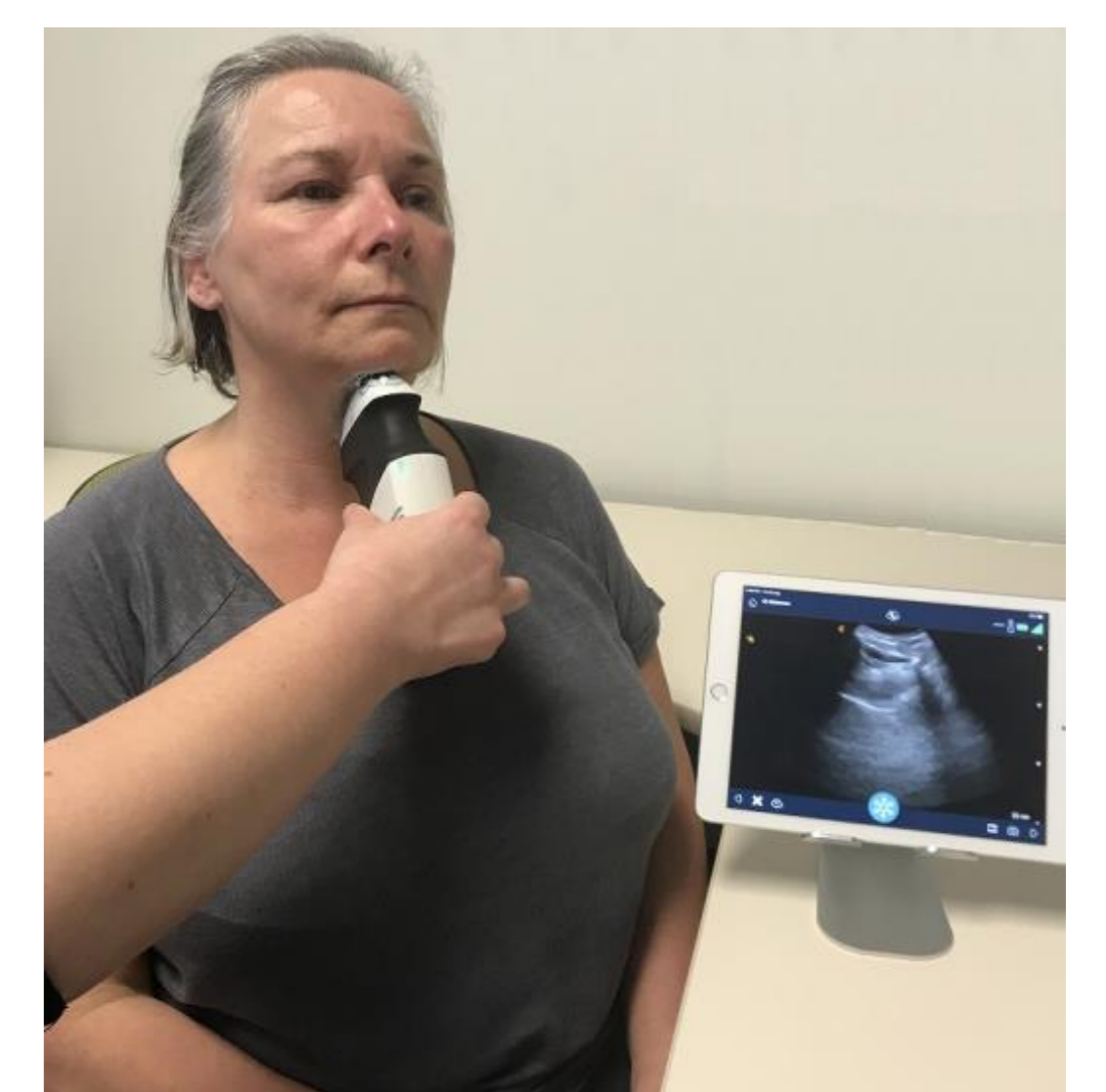


Fig 1. Curvilinear C3 Clarius™ US scanner, hyoid excursion measure

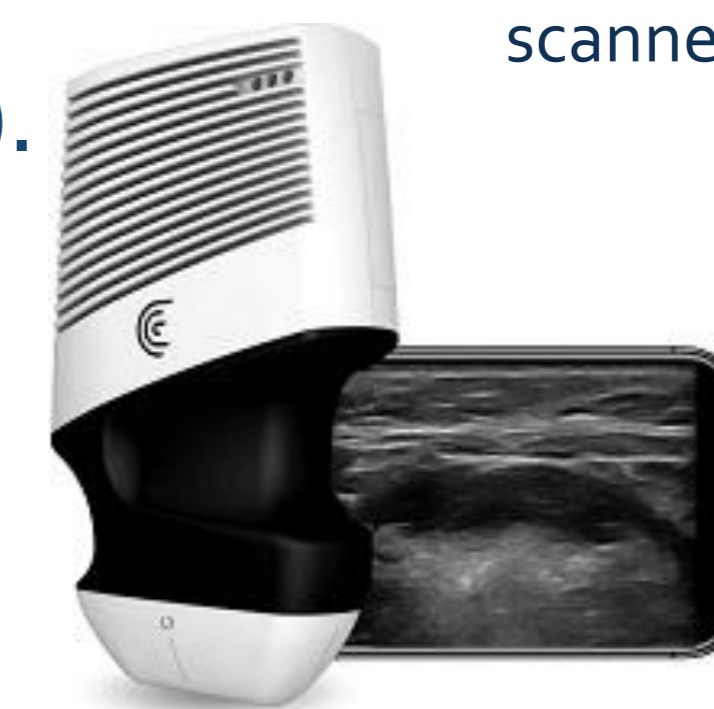


Fig 2. Linear L7 Clarius™ US scanner, cross sectional area of submental muscles

## Results

- + Online acquisition and measurement of inter-rater reliability ranged from poor<sup>4</sup> (< .50) to moderate<sup>4</sup> (.50 – .75).
- + ICC values for online and offline measurement of the same images were moderate<sup>4</sup> (.50 – .75) for dynamic measures and excellent<sup>4</sup> (>.90) for static measures.
- + Inter- and intra-rater reliability for offline measures was good<sup>4</sup> (>.75) to excellent<sup>4</sup> (>.90) for hyoid excursion and static measures and moderate<sup>4</sup> (.50 – .75) for thyrohyoid approximation.

Table 1:

Measure	Inter-rater ICC online acquisition (95% CI)	Intra-rater ICC online/ offline measurement (95% CI)	Intra-rater (R1) ICC offline measurement (95% CI)	Intra-rater (R2) ICC offline measurement (95% CI)	Inter-rater ICC offline measurement (95% CI)
Hyoid excursion	.33 (.00, .66)	.67 (.27, .87)	.83 (.52, .94)	.84 (.51, .94)	.83 (.49, .94)
Thyrohyoid approximation	.56 (.12, .82)	[.63] (.19, .83)	[.75] (.30, .91)	.78 (.40, .92)	.78 (.35, .91)
Tongue thickness	* (.00, .65)	.64 (.00, .93)	.98 (.90, 1.00)	.99 (.98, 1.00)	.85 (.44, .97)
Geniohyoid muscles	[.60] (.00, .90)	[.97] (.89, 1.00)	.98 (.93, 1.00)	.99 (.96, 1.00)	[.99] (.94, 1.00)
LAB	.78 (.25, .94)	.95 (.81, .99)	[.99] (.97, 1.00)	.97 (.89, .99)	[.84] (.44, .96)
RAB	[.42] (.00, .85)	[.99] (.96, 1.00)	[.91] (.93, 1.00)	.99 (.96, 1.00)	[.95] (.77, .99)

Note: CI= confidence interval, LAB/RAB = left/right anterior belly of digastric muscles

[ ]= assumptions for analysis not met \* For this measure, data is based on estimates of variance only as the model is over fitted

## Discussion

High reliability of offline measurement of US images obtained with Clarius™ is comparable to previous studies using sophisticated instrumentation<sup>1,2,3</sup>. Reliability is reduced when acquiring and measuring the images online in a clinical environment. This may be due to different techniques by examiners as well as variance in patient performance. Online data analyses may be affected by the pressure and lighting of a clinical environment paired with lower resolution of the device. This finding suggests that it is important to explore methods of improving reliability of data acquisition as well as online immediate analysis to achieve clinical translation for US assessment of swallowing.

<sup>1</sup> Macrae, et al. (2012). *Journal of clinical ultrasound*, 40(2), 74-78. | <sup>2</sup> Hsiao, et al. (2012). *Ultrasound Med Biol*, 38(9), 1522-1528. | <sup>3</sup> Huang, et al. (2009). *Ultrasound in medicine & biology*, 35(7), 1103-1108. |

<sup>4</sup> Portney, L. G., & Watkins, M. P. (2000). Validity of measurements. *Foundations of clinical research: Applications to practice*, 2.