'This is the peer reviewed version of the following article: Szczesniak MM, Maclean J, Zhang T, Liu R, Cock C, Rommel N, Omari TI, Cook IJ. Inter-rater reliability and validity of automated impedance manometry analysis and fluoroscopy in dysphagic patients after head and neck cancer radiotherapy. Neurogastroenterol Motil. 2015 Aug;27(8):1183-9.

which has been published in final form at
http://dx.doi.org/10.1111/nmo.12610

Inter-rater reliability and validity of automated impedance-manometry (AIM) analysis and fluoroscopy in dysphagic patients after head and neck cancer radiotherapy.

M.M. Szczesniak[1,2], J. Maclean[3], T. Zhang[1,2], R. Liu[1, 4], N. Rommel[5], T.I.Omari[6], I.J, Cook[1,2]

1. Department of Gastroenterology and Hepatology, St George Hospital, Sydney, NSW, Australia
2. St George Clinical School, University of New South Wales, Sydney, NSW, Australia
3. Speech Pathology Department, St George Hospital, Sydney, NSW
4. Department of Geriatric Medicine, The First Hospital of Lanzhou University, Lanzhou, Gansu, China
5. Translational Research Center for Gastrointestinal Diseases, University of Leuven, Leuven, Belgium
6.

Human Physiology, Medical Science and Technology, School of Medicine, Flinders University, South Australia

Corresponding Author:
Dr Michal Szczesniak
Department of Gastroenterology and Hepatology
St George Hospital, Kogarah , NSW, 2217, Australia
email: m.szczesniak@unsw.edu
t: +61 2 91132940
f:+61 2 91133993

ABSTRACT

Introduction: Automated Impedance manometry (AIM) pressure-flow analysis is novel non-radiological method to analyse swallowing function based on impedance-pressure recordings of pharyngeal swallows. In a population of dysphagic head and neck cancer patients, we evaluated the reliability and validity of the AIM-derived swallow risk index (SRI) and a novel measure of post-swallow residue (iZn/Z) by comparing it against videofluoroscopy as the gold standard for assessing aspiration and post-swallow residue risk.

Materials and Methods: Three blinded experts classified 88 videofluoroscopic swallows from 16 patients for aspiration and the degree of post-swallow residue using validated videofluroscopy scales. Pressure-impedance recordings of the swallows were also analysed using automated analysis software by one expert and two novice observers who derived the SRI and iZn/Z. Inter-observer concordance for videofluoroscopic and AIM measures was assessed using intraclass correlation coefficients (ICC). Patient SRI and iZn/Z measurements were compared with videofluoroscopy scores and control subjects to determine validity for detecting clinically relevant swallowing dysfunction. .

Results: Among individual swallows, agreement among observers assessing presence of penetration and aspiration on videofluoroscopy was modest (ICC 0.57). Agreement among observers for AIM-derived swallow risk index (SRI) and the iZn/Z was good (ICC of 0.71 and ICC of 0.82 respectively). When compared with age-matched controls the SRI was higher in patients with aspiration (mean diff. 28.6, 95% CI [55.85 1.355], $p<0.05$). The iZn/Z was increased, suggesting greater post-swallow residues, in both patients with aspiration (Δ244 [419.7, 69.52, $p<0.05$]) and penetration (Δ240 [394.3, 85.77, $p<0.05$]) compared to controls.

Discussion: AIM based measures of swallowing function have better inter-rater reliability than comparable fluoroscopically-derived measures. These measures can be easily determined and are objective markers of clinically relevant features of disordered swallowing following head and neck cancer therapy.

INTRODUCTION

When evaluating the swallowing of any patient with dysphagia, aspiration and the presence of post-swallow pharyngeal residues are of primary interest. Aspiration carries a significant risk of pulmonary sequelae while the presence of residue can signify pharyngeal propulsive deficiency or a restrictive defect at the pharyngo-esophageal junction. Videofluoroscopy is the gold standard for evaluating these clinically relevant features. Although videofluoroscopy remains the gold standard for swallowing assessment,it is subjectively applied in routine clinical practise. An objective, non-radiological and easily administrable test of swallowing function could be a useful adjunct to radiology, especially when patient monitoring over longer periods of time is needed in order to assess responses to intervention or to monitor the trajectory of swallow function in a progressive disease, as is the case with post-radiotherapy dysphagia in head and neck cancer patients.

Pressure flow analysis using Automated impedance manometry (AIM) is a new non-radiological method to assess swallowing function and biomechanics based on impedance and pressure measurements recorded during pharyngeal bolus swallows. AIM analysis derives objective pressure-flow variables with high intra- and inter-rater reliability in other dysphagia populations [Omari, 2011 #13771]. Pressure-flow variables are combined to derive a Swallow Risk Index (SRI), which is a composite score predictive of levels of swallowing dysfunction that predispose to aspiration risk [Omari, 2011 #13635]. Furthermore, a novel

AIM-derived measure of post-swallow residue, the integral of the ratio of the nadir impedance to post-swallow impedance (iZn/Z), has been shown to be a potential non-radiological predictor of ineffective pharyngeal bolus clearance [Szczesniak, 2008 #11775;Szczesniak, 2009 #12050], [Omari, 2012 #13829].

To date AIM has been broadly applied to dysphagia patients [Omari, 2011 #13635] [Omari, 2011 #14092], however its utility and reliability still remains to be examined in specific dysphagia sub-groups. Dysphagia is currently the most common, serious and disabling non cancer-related complications of head and neck radiotherapy, resulting in malnutrition, aspiration pneumonia and impaired quality of life [Mittal, 2003 #12470;Langendijk, 2008 #12259;Maclean, 2009 #13961]. In the current study we aimed to determine the reliability and validity of the AIM analysis method for the assessment of pharyngeal function in patients who develop dysphagia after radiotherapy treatment for head and neck cancer. The hypotheses tested were that AIM-derived SRI and iZn/Z: (1) can be determined reliably by different observers with varying levels of experience; and (2) that these global measures correlate with levels of swallowing dysfunction, specifically radiological evidence of aspiration and post-swallow residue respectively.

METHODS

Patients and controls

The study cohort comprised 16 patients (mean age 63, SD 9.8, range 47 – 86 years) complaining of dysphagia after radiotherapy for head and neck cancer who were enrolled consecutively. Radiotherapy was delivered between 0.8 – 14 years previously (mean 3.95, SD 3.87). Data from age-matched controls (n=16, mean age 63, SD 9.8, range 47-86 years) were randomly selected from an existing database of normal asymptomatic subjects aged 20-91 year studies with an identical system and catheter (Charles Cock, Repatriation General Hospital).

Experimental protocol

Patients' swallow function was evaluated using videofluoroscopy combined with concurrent pressure and impedance recording using a 3.6 mm diameter combined solid state manometry and impedance catheter incorporating 25 1 cm-spaced pressure sensors and 12 impedance segments spaced 2cm apart (Unisensor USA inc, Portmouth, NH, USA). Subjects were intubated after topical nasal anaesthesia with lignocaine (10%) and the catheter was positioned trans-nasally with the sensors spanning from velopharynx to proximal oesophagus. Subjects were seated upright and swallowed triplicate boluses of 2, 5 and 10ml of EZ-HD barium (Bracco UK Limited, Woodburn Green, High Wycombe, UK). Barium suspension was made up with NaCl to give a 1% final concentration. Fluoroscopy cineloops (MultiDiagnost Eleva, Philips, Best, Netherlands) as well as pressure and impedance signals were acquired concurrently on MMS Solar GI system (Software Version 8.21o, MMS, Enschede, Netherlands). Fluoroscopy pulse rate was 12Hz and images were acquired at 25Hz. Manometry and impedance data were acquired at 40Hz.

Following data acquisition, pressure-impedance and video data (MPEG-4) were exported as single swallows and de-identified for blind analysis.

Fluoroscopy Analysis

Three expert observers (geriatrician, radiologist, speech pathologist) independently scored the de-identified videos. Observers scored the presence of aspiration using an 8-point penetration-aspiration scale (PAS) [Rosenbek, 1996 #7697] and the presence of post-swallow residue using a 6-point bolus residue scale (BRS) [Omari, 2012 #13829].

PAS scores reflect the extent of contrast entry into the airway, and whether or not material in the airway is expelled during the swallow sequence. Expert consensus PAS score was determined for each swallow based upon the mode of the three observers' scores. Furthermore, the consolidated aspiration status of each patient was categorized overall based on the highest consensus PAS score recorded for their swallows. Patients were then grouped as demonstrating: 1) no aspiration (score 1); 2) penetration (Score 2-5) or 3) aspiration (score 6-8).

The bolus residue scale (BRS) score (between 1 and 6) was assigned for each swallow according to the number of structures showing evidence of residue: No residue in any of these structures was assigned a BRS score of 1. If residue was present, then additional scores were weighted toward the anatomical regions where residue posed an aspiration risk (+1 for valleculae, +2 for piriform sinus, and +2 for posterior pharyngeal wall). Hence residue in valleculae only derives BRS 2, posterior pharyngeal wall or piriform sinus only = BRS 3, valleculae and posterior pharyngeal wall or piriform sinus = BRS 4, posterior pharyngeal wall and piriform sinus = BRS 5, and all structures = BRS 6 [Omari, 2012 #13829]. Expert consensus BRS score was determined for each swallow based upon the mode of the scores for that swallow, or median if mode could not be determined. Based on previous studies [Omari, 2012 #13829] a BRS score of 4–6 was considered abnormal for the purposes of this analysis.

AIM analysis

One expert and two novice observes performed analysis of pressure-impedance data with AIMplot software, a purpose-designed MATLAB-based analysis program developed to increase the applicability of the methodology for routine use [Omari, 2011 #14092]. To operate AIMplot the observer was required to define three space-time landmarks from a standard pharyngeal pressure topography plot [Citation??].

These were:

- The time of onset of pharyngeal swallow; defined by the onset of upper esophageal sphincter (UES) relaxation often, but not always, associated with a proximal excursion of the UES high pressure zone.
- The position of the UES proximal margin immediately post pharyngeal swallow.
- The position of the velopharynx; defined as the pressure zone immediately superior to the propagated pharyngeal stripping wave.

Guided by the landmarks above, AIMplot software then automatically derived values for pharyngeal pressure-flow variables, the Swallow Risk Index (SRI) and integral of the ratio of the nadir impedance to postswallow impedance (iZn/Z).

Calculation of the iZn/Z has been described previously [Omari, 2012 #13829] but briefly, for each array of impedance values, nadir impedance (Zn) to impedance (Z) ratio (Zn/Z ratio) is calculated and then values of Zn/Z ratio residing within posts-wallow region of interest (ROI) were numerically integrated to generate a single value reflecting the overall intensity of iZn/Z ratio. The ROI is defined as 0.25s from peak of pharyngeal pressure wave for duration of 1 second (Figure 1).

Statistical Analysis

Concordance between videofluoroscopic observations and AIM measurements made by multiple observers was assessed using intraclass correlation coefficients (ICC). Comparisons of AIM measurements between groups of subjects (historical controls, patients with/without aspiration penetration) were performed with one-way ANOVA. Association of AIM residue scores (iZn/Z) with BRS (ordinal scale) was assessed using Spearman rank correlation coefficient.

RESULTS

Interobserver agreement of videoflouroscopic analysis

Three observers independently analysed fluoroscopy videos from 88 of 91 swallows; three swallows were excluded due to poor image quality. The concordance among observer scores for individual swallows was modest with an intra-class correlation (ICC) of 0.57 95% CI [0.46, 0.68]. For 14/16 patients there was complete agreement amongst observers with respect to the overall aspiration status. The concordance between observer's scores for bolus residue in individual swallows was also modest with an intra-class correlation of 0.53, 95% CI [0.41 0.64] (Table 1) and for 71% of swallows, all three observers agreed on clinically significant BRS (BRS $\geq$ 4).

Interobserver agreement in AIM analysis

The AIM-derived measures, Swallow Risk Index (SRI) and impedance ratio (iZn/Z) had higher inter-rater agreement among the three observers than did videofluoroscopic analysis of aspiration and residue. The intra-class correlations for the SRI score and iZn/Z were 0.78 and 0.75 respectively. The intra-class correlations for component variables used to derive the SRI ranged from 0.91 (PeakP) to 0.68 (FI) (Table 1).

Correlation of Swallow Risk Index (SRI) and iZn/Z with videofluoroscopic measures of aspiration and residue

The mean SRI of age-matched controls was 4.8, 95% CI [2.1 7.5]. In patients without radiological evidence of penetration or aspiration the mean SRI was 7.7, 95% CI [5.6 9.8]. In those with penetration mean SRI was 14.9, 95% CI [5.9 23.9] and 33.41, 95% CI [0 74.71] in those with aspiration. The difference in SRI between patients with aspiration and controls was statistically significant (mean diff. 28.6, 95% CI [55.85 1.355], $p<0.05$) (Fig. 2)

The mean iZn/Z of controls was 113.7, 95% CI [77.66 149.7]. In patients without penetration or aspiration the mean iZn/Z was 222.7, 95% CI [82.76 362.6] in those with penetration or aspiration mean iZn/Z was 358.3, 95% CI [186.5 530.1] and 353.7 95%, CI [160.5 547] respectively. The difference between controls and patients with either aspiration or penetration was statistically significant (penetration: mean diff. 244.6, 95% CI [419.7 to 69.52]; aspiration: 240.1, 95% CI [394.3 to 85.77]) (Fig. 2). The bolus residue quantified by iZn/Z had moderate positive correlation with bolus residue on fluoroscopy (BRS score) ($r_s(86) = 0.4120$, $p< 0.0001$) (Fig. 3)

Discussion

In the current study we aimed to evaluate the reliability of SRI, a global index of swallowing dysfunction, and iZn/Z, an AIM-derived measure of residue, in dysphagic post head and neck radiotherapy patients. The interobserver consistencies of AIM variables (SRI and iZn/Z) were considerably higher than that of videofluoroscopic PAS and BRS scores. This result was anticipated as AIM is a mostly automated analysis method where the operator is only required to determine landmarks from a pressure plot such as the onset of the swallow and position of the UOS high pressure zone. Furthermore, personnel performing AIM analysis in this study included two novices demonstrating that AIM analysis is consistent even when the operators have little or no prior experience with pressure-impedance based methods of analysis.

In the current study the concordance between scorer's PAS scores was modest with ICC of 0.57. Reports in the literature of inter-observer reliability of videofluoroscopic assessment of penetration/aspiration vary greatly. Near perfect agreement (ICC 0.8-0.87) is reported by some [Stoeckli, 2003 #13786; Omari, 2011 #14092] while others report much poorer agreement of (ICC 0.1-0.6)[McCullough, 2001 #11368;]. In the current study, the disagreement of radiological scoring of aspiration occurred predominantly when there was contrast remaining from previous swallows in the larynx or trachea. In these circumstances it was difficult for observers to reliably ascertain whether any further contrast has entered the airway during the swallow.

Our BRS scores with ICC of 0.53 also had lower concordance between observers than previously reported ICC of 0.78 [Omari, 2012 #13829]. Fluoroscopy imaging system used in the current study allowed capture of cineloops with a higher resolution (1280x1024) than previously [Omari, 2012 #13829]. This increased resolution enabled observation of even trace amounts of barium contrast coating the pharyngeal mucosa in all swallows. We believe that this increased imaging quality resulted in the expert observers reporting what they considered to be a noteworthy amount of residue and this consequently increased the variability between their scores.

The mean SRI was significantly increased in the dysphagic patients with aspiration, however three of the seven aspirating patients had SRI values that were within the normal range (SRI <15) [Omari, 2011 #13635]. Additionally, all of these patients had an iZn/Z that was below the level optimally predictive for post-swallow residue (iZn/Z<300) [Omari, 2012 #13829]. Our study was predominantly designed to evaluate the reliability of AIM measures and under powered for assessing prognostic value.

The contrasting findings between objective pressure-flow measurements and the video-fluoroscopic assessments of some individuals requires consideration. It could be concluded that the AIM analysis method lacks the sensitivity needed to detect clinically significant swallowing dysfunction within the specific cohort studied. The SRI was empirically derived based on data from a predominantly neurological cohort [Omari, 2011 #13635] and therefore may have less predicative value in our patients who had a different pathology underlying their dysphagia. A possible interpretation of the contrasting findings between objective functional measures and aspiration is that some aspirating patients may still be swallowing well. Aspiration is defined when swallow material passes the vocal cords, however one of the limitations of the PAS is that 'trace' vs. volume aspiration are both scored equally. Correlating AIM measures to findings on videofluoroscopy is an important first step, however it is also important to examine the predictive value of our objective measures in relation to other important clinical outcomes. Whilst it seems intuitive that aspiration on videofluoroscopy should predict clinical sequelae such a pneumonia, the evidence that it does so is in fact poor [Purkey, 2009 #14099; Feinberg, 1996 #7669]. Further studies are required to examine the utility of AIM measures to predict clinically relevant outcomes, and to objectively document longitudinal change and over time and following therapy.

In summary, in a head and neck cancer cohort, AIM-analysis pressure-flow metrics and global measures can be derived with high inter-observer agreement and are therefore reliable amongst analysts. These measures are also altered in relation to increasing levels of swallowing dysfunction as defined by radiological scores and therefore appear valid. Further studies are needed to establish the prognostic value of this novel methodology for assessing swallowing function in the HNC patient population.

Table 1: Inter-rater reliability of videofluoroscopic variables and impedance/pressure metrics derived from AIMplot analysis(ICC [95% CI]).

	2ml	**5ml**	**10ml**	**All Volumes [95% CI]**
Videofluoroscopic Measures				
Aspiration	**0.76** [0.51 0.91]	**0.61** [0.45 0.76]	**0.45** [0.24 0.66]	**0.57** [0.46 0 .68]
Residue	**0.53** [0.21 0.80]	**0.49** [0.31 0.66]	**0.64** [0.45 0.79]	**0.53** [0.41 0.64]
AIM metrics				
SRI	**0.71** [0.44 0.89]	**0.74** [0.61 0.84]	**0.59** [0.40 0.75]	**0.78** [0.73 0.81]
PeakP	**0.93** [0.83 0.98]	**0.93** [0.88 0.96]	**0.93** [0.88 0.96]	**0.91** [0.88 0.93]
PNadImp	**0.94** [0.87 0.98]	**0.88** [0.76 0.94]	**0.77** [0.51 0.89]	**0.84** [0.73 0.90]
TNI_PP	**0.75** [0.49 0.91]	**0.81** [0.70 0.89]	**0.6** [0.41 0.76]	**0.86** [0.83 0.89]
FI	**0.40** [0.08 0.72]	**0.58** [0.41 0 .73]	**0.67** [0.49 0.81]	**0.68** [0.62 0.73]
iZn/Z	**0.70** [0.24 0.90]	**0.77** [0.62 0.87]	**0.73** [0.53 0.85]	**0.75** [0.65 0.82]

FIGURES

Figure 1. Examples of swallows with and without post-swallow bolus residue. High-resolution spatio-temporal pressure plots (top) and iZn/Z plots with pressure as isocontours (bottom). Plots show the region used to calculate the iZn/Z and the resultant value for each swallow.

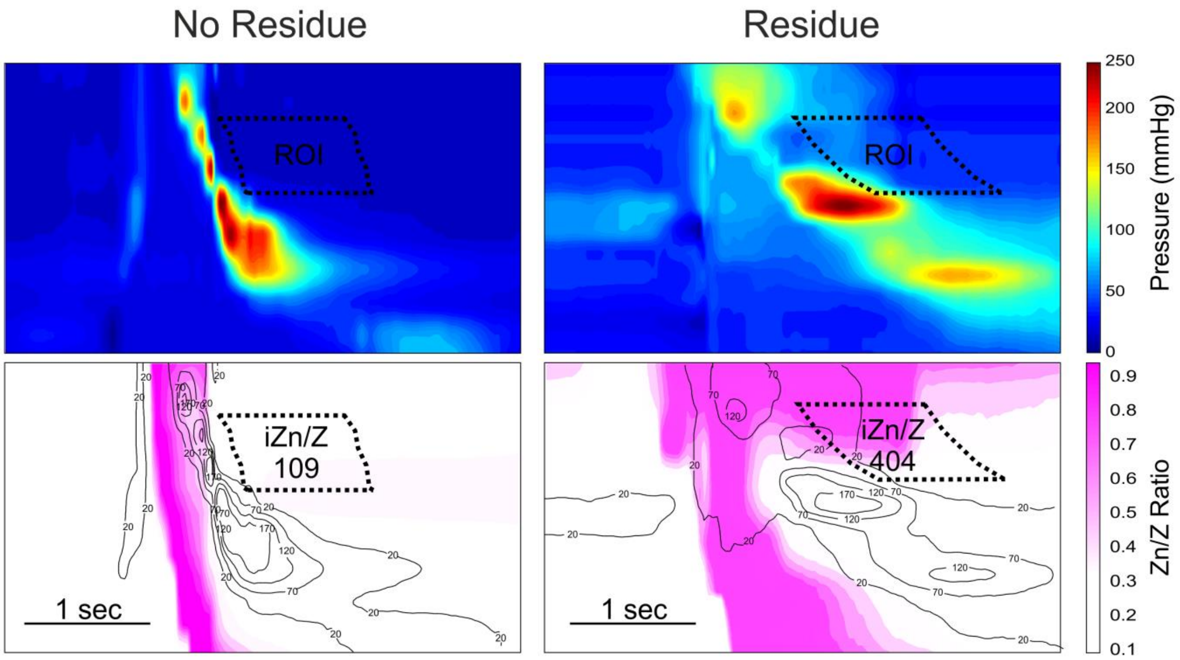

Figure 2: Box plots showing median and interquartile ranges for Swallow Risk Index with component variables (PeakP, PNadImp, TnadImp-PeakP, FI) and iZn/Z in controls and patients. Patient data are further stratified based on aspiration score (No aspiration: score 1, penetration: score 2–5, and aspiration: score 6–8), * p < 0.05.

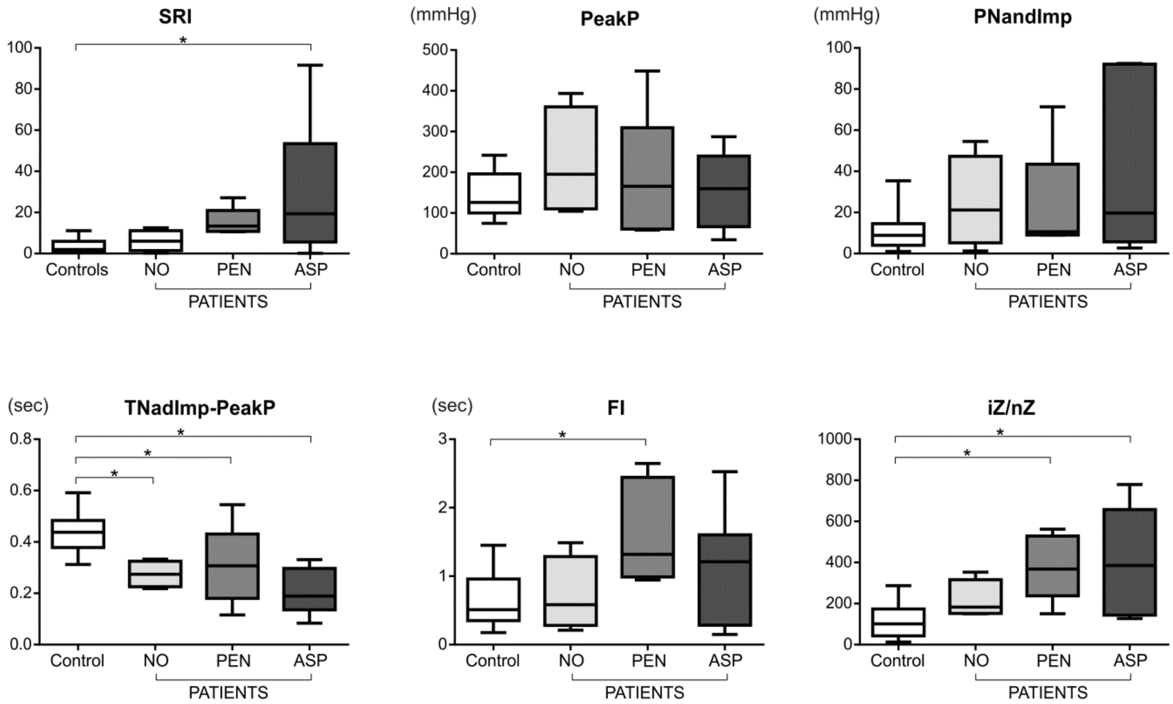

Figure 3: Correlation of consensus Bolus Residue Score (BRS) with AIM derived marker of residue, impedance ratio (iZ/Z). Shouldn't this graph have R^2 and p values on it?)