

Development, Reliability, and Validation of an Infant Mammalian Penetration–Aspiration Scale

Shaina Devi Holman · Regina Campbell-Malone ·
Peng Ding · Estela M. Gierbolini-Norat · Anne M. Griffioen ·
Haruhi Inokuchi · Stacey L. Lukasik · Rebecca Z. German

Received: 31 May 2012 / Accepted: 27 September 2012 / Published online: 7 November 2012
© Springer Science+Business Media New York 2012

Abstract A penetration–aspiration scale exists for assessing airway protection in adult videofluoroscopy and fiberoptic endoscopic swallowing studies; however, no such scale exists for animal models. The aim of this study was threefold: (1) develop a penetration–aspiration scale (PAS) for infant mammals, (2) test the scale’s intra- and interrater reliabilities, and (3) validate the use of the scale for distinguishing between abnormal and normal animals. After discussion and reviewing many videos, the result was a 7-point infant mammal PAS. Reliability was tested by having five judges score 90 swallows recorded with videofluoroscopy across two time points. In these videos, the frame rate was either 30 or 60 frames per second and the animals were either normal, had a unilateral superior laryngeal nerve (SLN) lesion, or had hard palate local anesthesia. The scale was validated by having one judge score videos of both normal and SLN lesioned pigs and testing the difference using a *t* test. Raters had a high intrarater reliability [average $\kappa = 0.82$, intraclass correlation coefficient (ICC) = 0.92] and high interrater reliability (average $\kappa = 0.68$, ICC = 0.66). There was a significant difference in reliability for videos captured at 30 and 60 frames per second for scores of 3 and 7 ($P < 0.001$). The scale was

also validated for distinguishing between normal and abnormal pigs ($P < 0.001$). Given the increasing number of animal studies using videofluoroscopy to study dysphagia, this scale provides a valid and reliable measure of airway protection during swallowing in infant pigs that will give these animal models increased translational significance.

Keywords Aspiration · Penetration · Airway protection · Videofluoroscopy · Deglutition · Deglutition disorders · Animal models

Penetration and aspiration can occur in the absence of adequate airway protection during swallowing. Penetration occurs when material passes into the airway but remains at or above the vocal folds in the supraglottic space. Aspiration occurs when material passes below the vocal folds and into the trachea. In clinical studies of swallowing dysfunction, the incidence of both penetration and aspiration is associated with increased risk of developing pneumonia, which can be life-threatening [1]. The 8-point penetration–aspiration scale (PAS) [2], used in clinical research studies, is an established and validated instrument for measuring the airway protection during swallowing [3–7]. No such scale, however, currently exists for animal models of swallowing dysfunction.

Animal models are an important tool for understanding both normal and abnormal swallowing. Such models are valuable for collecting data on swallowing dysfunction that is not possible in humans for logistic and ethical reasons [8–11]. For example, we can study swallow function in animals using videofluoroscopy without the time limitation that is necessary in clinical studies due to limits of radiation exposure. This model also permits multiple and

S. D. Holman (✉) · R. Campbell-Malone · P. Ding ·
E. M. Gierbolini-Norat · A. M. Griffioen · H. Inokuchi ·
S. L. Lukasik · R. Z. German
Department of Physical Medicine and Rehabilitation, Johns
Hopkins University School of Medicine, 98 N Broadway,
Suite 409, Baltimore, MD 21231, USA
e-mail: swahl001@umaryland.edu

S. D. Holman · R. Z. German
Department of Neural and Pain Sciences, University
of Maryland School of Dentistry, 650 W Baltimore Street,
Baltimore, MD 21201, USA

unlimited videofluoroscopy recordings over several days or even months. In animal models, researchers can generate pathological conditions through nerve lesions, anesthesia, or other means and then test the effect of that specific condition on function in the same individual [12, 13]. The other advantage of animal models is the ability to control extraneous and confounding factors that frequently co-occur in human clinical conditions [14, 15].

A valuable animal model of studying infant swallowing is the infant pig (*Sus scrofa*). The main reason for studying infant pigs in swallowing dysfunction studies is because the anatomy of the infant pig parallels other infant mammals. This enables us to draw clinically relevant conclusions to the treatment of human infants with swallowing dysfunction [16]. For example, a model of superior laryngeal nerve (SLN) lesion in the infant pig is currently being used to understand how loss of sensory information carried by this nerve affects the motor control of swallowing [17, 18]. In this model we have documented both penetration and aspiration in the infant pig [17].

The aim of this study was to (1) develop a novel PAS for the infant pig model of swallowing dysfunction based on the standard clinical research scale, (2) assess the intra- and interrater reliabilities, and (3) test the validity of this scale for differentiating abnormal versus normal animals.

Methods

Summary of Methods Used for Unilateral SLN Lesion and Palatal Local Anesthesia

The videofluoroscopic recordings utilized to develop this scale and assess reliability and validity were from ongoing studies in the lab. These methods are summarized below.

The pigs that had a unilateral SLN lesion had undergone two surgeries. The pigs were 10–16 days old and 5–6 kg in weight. During the first surgery they were given general anesthesia (5 % isoflurane) and intubated and then underwent surgery that lasted 3–5 h. During this surgical procedure, electromyographic (EMG) electrodes were implanted into hyolaryngeal muscles and a radiopaque marker was sutured to the hyoid bone and thyroid cartilage. These hyolaryngeal muscles include mylohyoid, genioglossus, geniohyoid, digastric, thyrohyoid, sternothyroid, sternohyoid, and cricothyroid. After this surgical procedure was completed, radiopaque markers were placed into the tongue, gingiva, hard palate, and soft palate, and another radiopaque marker was placed with an intraoral approach on the epiglottis. The animal recovered from surgery after 1–5 h and was then fed swine milk formula (Land O'Lakes Solustart Pig Milk Replacer) containing barium using a bottle with a "pig nipple" (Nasco, Fort Atkinson, WI, USA). While the

pigs drank the formula, investigators recorded both lateral videofluoroscopy at 30 or 60 frames per second and EMG signals from the electrodes placed into the hyoid musculature. This first recording was for control data. After 1–2 days the pigs underwent another surgery where the SLN was cut on the right side before it branches into the internal and external branches that supply the larynx and cricothyroid muscles. After the animal recovered, the animal was again fed using procedures identical to those for control data collection. After all the necessary recordings were completed, the animal was euthanized and the locations of the EMG electrodes and markers were confirmed by dissection.

The pigs that had palatal local anesthesia underwent surgery to place EMG electrodes into their hyoid musculature and radiopaque markers were placed into the same structures as in the SLN lesion pigs. Starting a day after surgery, these pigs were fed while recording lateral videofluoroscopy and recording from the EMG electrodes at four time points, 2 h apart over a day for control data. Early the next day, the pigs were given a 0.5-ml injection of 0.5 % bupivacaine, a long-lasting local anesthetic, at three locations: the nasopalatine foramen and left and right greater palatine foramina. The technique used was a standard dental local anesthesia nerve block. Bupivacaine starts being effective 15 min post injection and lasts for 3–5 h in small dogs, which are considered comparable to infant pigs because of their similar size [19]. Starting 1 h after the nerve blocks, the animals were fed using the same procedures used for the SLN lesion animals. They were fed four times, every 2 h. Animals were counterbalanced so that half of the pigs were recorded on day 1 post surgery with no local anesthesia and on day 2 with local anesthesia, while the other half were recorded on day 1 post surgery with local anesthesia and on day 2 without local anesthesia. After all the recordings were completed, the animals were euthanized and the location of the electrodes and markers were confirmed by dissection. All of these procedures were approved by the Johns Hopkins Medical Institute IACUC.

Development of the Scale

In order to adapt the clinical PAS to the infant mammal, we examined a number of infant pig videofluoroscopic images of normal swallows as well as swallows with clear penetration and aspiration. The infant pigs studied were all from the previously described studies.

Our infant mammal PAS was based on the current clinical 8-Point PAS [1] (Table 1). The result was a 7-point infant mammal PAS (IMPAS) (Table 2). As with the 8-Point PAS, this scale was multidimensional. It took into account the depth of the bolus into the airway, whether it was above or below the vocal folds, and the animal's response to the bolus, whether it was expelled passively,

Table 1 The 8-point penetration–aspiration scale (from Rosenbek et al. [2])

Score	Description
1	Material does not enter the airway
2	Material enters the airway, remains above the vocal folds, and is ejected from the airway
3	Material enters the airway, remains above the vocal folds, and is not ejected from the airway
4	Material enters the airway, contacts the vocal folds, and is ejected from the airway
5	Material enters the airway, contacts the vocal folds, and is not ejected from the airway
6	Material enters the airway, passes below the vocal folds, and is ejected into the larynx or out of the airway
7	Material enters the airway, passes below the vocal folds, and is not ejected from the trachea despite effort
8	Material enters the airway, passes below the vocal folds, and no effort is made to eject

Table 2 The 7-point infant mammal penetration–aspiration scale

Score	Description
1	Material does not enter the airway
2	Material is in the supraglottic space, remains above the vocal folds, and passively leaves the airway before the epiglottis returns to rest position
3	Material is in the supraglottic space, a small amount remains above the vocal folds after epiglottis in rest position
4	Material is in the supraglottic space, a larger amount remains above the vocal folds after epiglottis in rest position
5	New material is in the supraglottic space, then passes below the vocal folds, and is actively ejected above the vocal folds
6	New material is in the supraglottic space, then passes below the vocal folds and is not ejected from the trachea despite effort
7	New material is in the supraglottic space, then passes below the vocal folds, and no effort is made to eject (silent aspiration)

actively, or not at all. Similar to the PAS used in clinical research, our scale was ordinal, meaning that lower scores represent less severe conditions (more airway protection during swallowing), and higher scores reflect more severe conditions (less airway protection during swallowing). A description of each score is given below.

No Penetration

A score of 1 on the IMPAS is equivalent to the score of 1 on the clinical PAS (Table 1). During these swallows no material enters the airway. The material, or in this case milk, flows over the epiglottis as it moves caudally to protect the airway (Fig. 1b).

Penetration

A score of 2 on the IMPAS is similar to the score of 2 on the clinical 8-Point PAS. The score of 2 on the clinical scale is when material enters the airway, remains above the vocal folds, and then is ejected either passively or by a cough. In the infant pig we observed a similar case where material would enter the airway, remain above the vocal

folds, and passively leave the airway, often before the epiglottis returned to its upright position (Fig. 1c).

Scores of 3 and 4 on the IMPAS are similar to the score of 3 on the clinical PAS. In the clinical scale, material enters the airway and remains above the vocal folds. This was also seen in infant pigs; however, we observed two distinct conditions where milk remained above the vocal folds. Occasionally, a very small amount of milk was seen on the caudal side of the epiglottis by its base, forming a small triangle (Fig. 1d). In other cases, a substantial amount of material remained in the airway, above the vocal folds; however, it was traveling toward the vocal folds (Fig. 1e). Thus, we divided this category into two scores because having more material closer to the vocal folds was perceived as a more severe condition that would increase the risk of aspiration as it does in human infants [20]. The score of 3 indicates a small amount of material is above the vocal folds on the inverse side of the epiglottis, forming a small triangle by the base of the epiglottis, and a score of 4 indicates a large amount of material.

It is important to note that for scores 2–4, “material” may be old or new material (Table 2). In clinical video-fluoroscopic swallowing studies (VFSS), swallows are often isolated; however, in the infant pig videofluoroscopy recordings are made across a feeding session where there

are multiple swallows per second with no break. As a result, there is often milk residue in the airway from previous swallows. Since any milk in the airway above the vocal folds was due to a failed swallow, it was determined that it should be scored regardless of whether it was from a previous swallow or the current swallow.

Aspiration

A score of a 5 or 6 on the IMPAS is similar to a score of 6 or 7 on the clinical 8-point PAS. A score of 6 on the clinical PAS describes material passing below the vocal folds and then being ejected into the larynx or out of the airway. A score of 7 means material passed below the vocal folds and was not ejected despite effort. Although we did not observe a score of 6 or 7 in our videofluoroscopic recordings, past observations of coughing in pigs with SLN lesions meant that these two scores are possible. Further, aspiration following sensory or motor nerve lesions could trigger a coughing reflex that may or may not be successful. For this reason a score of 5 on the IMPAS is when material passes below the vocal folds and is ejected into the larynx or out of the airway. A score of 6 is when material passes below the vocal folds and is not ejected despite effort. Silent aspiration is described as a score of 7 on the IMPAS (Fig. 1f). This score is the equivalent of a score of 8 on the clinical 8-Point PAS. This was often seen after nerve lesions in the infant pig model.

With scores of 5, 6, and 7, *new* milk must be visualized moving from the supraglottic space to below the vocal folds. There were some instances where milk from a previously failed swallow was visualized in the trachea, moving above and then back below the vocal folds. This was not scored a 5, 6, or 7 because for those scores new material must be in the supraglottic space and then move below the vocal folds. If material is in the supraglottic space and joins milk residue from below the vocal folds, forming a solid stream of milk, then it is aspiration and is scored a 5, 6, or 7 depending on whether there is effort to eject that material.

No equivalent score of 4 and 5 on the clinical PAS was included in the infant pig scale because milk was always clearly either above or below the vocal folds and never just contacted them.

Guidelines for Users of the Scale

To ensure consistency and higher reliability, specific directions were developed for the judges who were to score videos. The swallow being scored begins at the start of the epiglottal flipping (posterior/inferior movement) and ends

immediately before the next epiglottal posterior/inferior movement. The rationale for this definition was that in some cases the events leading to a score of 3, 4, or 7 did not develop until after the epiglottis had returned to its upright position but before the next swallow occurred. Judging was also based only on a score where material was clearly visible in the airway. Sometimes due to the poor quality of the video, judges would describe seeing very small amounts of residue of material in the airway. This resulted in variable scoring of the videos. Judges were not permitted to alter the contrast or brightness of the image since that could alter the amount of material visualized or create artifacts. When scoring videos, judges were given as much time as needed to review the swallow and they could view it multiple times and in slow motion.

Determining the Reliability of the Scale

To test the reliability of the IMPAS, five judges scored 90 videofluoroscopic recordings. Images of swallows were randomly selected from ten infant pigs and 30 total feeding sessions from the previously described ongoing studies. From each feeding session, three swallows were randomly clipped out of the sequence and separated into their own video file.

The judges had various levels of experience and education with respect to analysis of VFSS data, but all had experience with either animal or clinical swallowing research. The judges were given the 90 videofluoroscopic recordings of individual swallows in a randomized order to score using IMPAS. They were instructed to score all of the videos within 48 h and then, as was done with the development of the clinical PAS, to score them all again after a 2-week hiatus, also within a 48-h period [2]. The judges were given the same set of videos for their second scoring but in a different, randomized order. All videos were viewed using MaxTRAQ ver. 2.2.4.1 (Innovision Systems, Inc., Columbiaville, MI, USA). All judges also were given the same five videos that were examples of scores 1, 2, 3, 4, and 7 to reference during their scoring of the videos.

To assess inter- and intrarater reliabilities, Cohen's κ and two-way mixed intraclass correlation coefficient (ICC) with absolute agreement were calculated using SYSTAT 13 (Systat Software, Inc., Chicago, IL, USA) and IBM SPSS Statistics 20 (SPSS, Inc., Chicago, IL, USA), respectively. Cohen's κ was also calculated by scale score in order to determine if there were differences in the intra- and interrater reliabilities based on each score. Rosenbek et al. [2] used κ to calculate reliability; however, we also calculated ICC since the scale is ordinal and that calculation takes into account the degree of difference. Percentage of agreement and the stratification of the scores were also calculated. We also tested for statistically significant

differences in the Cohen's κ score given for videos captured at 30 versus 60 frames per second using an analysis of variance test (ANOVA) and post-hoc Tukey's test. The statistical analysis was carried out by an independent investigator who had not judged the videos.

Assessing Validity of the Scale

A separate study was conducted to assess if the scale could distinguish normal and abnormal pigs. A blinded judge, who was not one of the five judges used to assess reliability, scored 39 swallows from six intact infant pigs and 39 swallows from three abnormal infant pigs. The abnormal pigs had a unilateral SLN lesion. Each swallow was from a different randomly selected feeding session. The scores were graphed to see if there was a distinct difference in the scores between normal and abnormal infant pigs. A two-sample t test was calculated to determine if the normal and abnormal pigs were statistically significantly different with α set at 0.05.

Results

Reliability of the Scale

In an initial attempt to score the 90 videos using the IMPAS, the intra- and interrater reliabilities were low and deemed unacceptable. The intrarater reliability measured using Cohen's κ averaged 0.65 and the interrater reliability ranged from 0.36 to 0.67 with an average of 0.58. Following that preliminary analysis, difficulties and problems with the scale were discussed. Ten videos that were a source of significant variation were reviewed by our group in order to better define the seven categories. The subsequent clarification and revision to the scale resulted in the version presented here.

When the scoring was performed again, the results showed much higher inter- and intrarater reliability scores. Intrarater reliability calculations showed an average Cohen's κ of 0.82 and an average ICC of 0.92 with 86 % agreement (Table 3). The intrarater reliability by score was calculated using Cohen's κ (Table 3) and showed higher reliability for scores 1, 2, and 7 (0.90, 0.82, and 0.83, respectively) compared to reliability for scores 3 and 4 (0.74 and 0.75, respectively). Interrater reliability, as expected, was lower than the intrarater reliability. For the first scoring, the average Cohen's κ value was 0.70 and the ICC was 0.89. For the second scoring, the average Cohen's κ value was 0.66 and the ICC was 0.87 (Tables 4, 5). Cohen's κ values ranged from 0.65 to 0.76 for the first ranking and from 0.58 to 0.84 for the second ranking. No scores of 5 or 6 were observed by the judges.

In the first scoring, Cohen's κ for each score showed higher reliability for scores 1, 2, 4, and 7 (0.86, 0.68, 0.67,

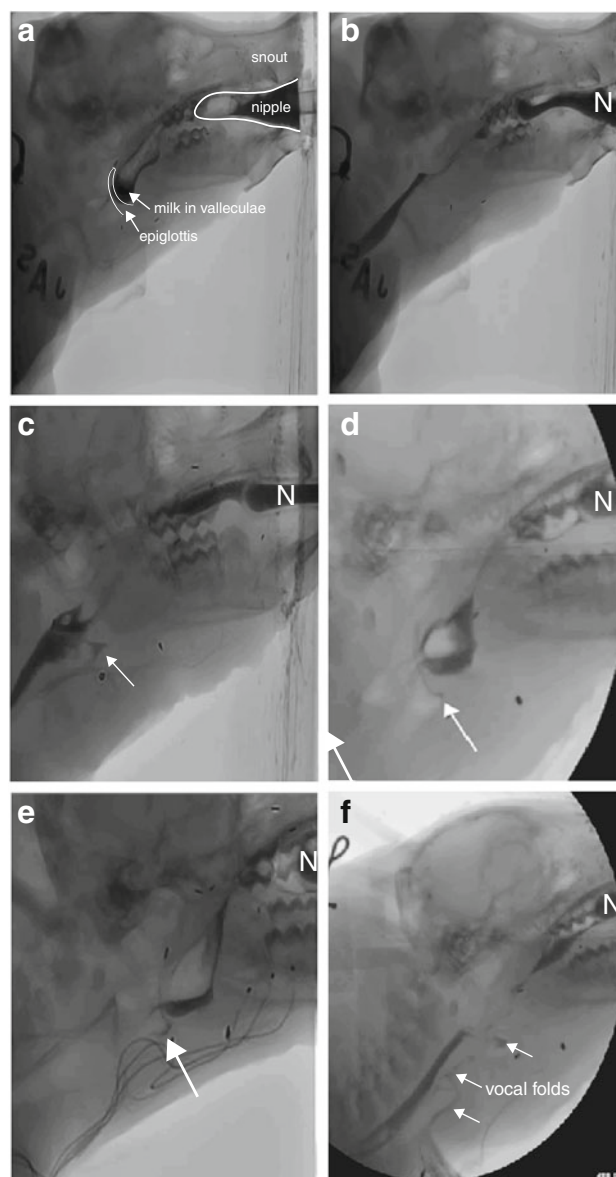


Fig. 1 Videofluoroscopic images from each score on the 7-point infant mammal PAS. **a** This videofluoroscopic image shows that the epiglottis is in the *upright position* with milk in the valleculae *right* before a swallow is initiated. The anatomy is clearly outlined and labeled for orientation. **b** Milk is being emptied out of the valleculae, the epiglottis is fully flipped to protect the airway, and milk continues into the esophagus. This is a score of 1. **c** The epiglottis has flipped fully to protect the airway; however, milk penetrates the airway as identified by the *arrow*. When the epiglottis returns to its rest position, there is no milk left in the supraglottic space. This is a score of 2. **d** The epiglottis is in its *upright position* after a swallow. The *arrow* points to a small amount of milk visible on the caudal side of the epiglottis forming a *triangle* at the *base*. When this occurs it is scored as a 3. **e** The epiglottis has returned to its rest position after a swallow. The *arrow* points to a large amount of milk residue left on the caudal side of the epiglottis but above the vocal folds. This is scored as a 4. **f** The epiglottis has flipped to protect the airway during the swallow. Milk is clearly visible crossing the vocal folds and entering the larynx and is identified by the *arrows*. The nipple is in the *far right* of all the images and is labeled “N”

Table 3 Intrarater reliability by score

Judge	Agreement		ICC	Cohen's κ	Cohen's κ by scale score						
	<i>n</i>	%			1	2	3	4	5	6	7
1	75	83.33	0.87	0.78	0.90	0.82	0.68	0.65	–	–	0.71
2	79	87.78	0.97	0.84	0.80	0.80	0.84	0.85	–	–	0.96
3	82	91.11	0.96	0.89	0.92	0.87	0.86	0.87	–	–	0.88
4	79	87.78	0.95	0.84	0.91	0.85	0.73	0.82	–	–	0.86
5	72	80.00	0.87	0.74	0.94	0.77	0.57	0.58	–	–	0.73
Average	77.4	86.00	0.92	0.82	0.90	0.82	0.74	0.75	–	–	0.83

Percentage of agreement is shown for each judge between the first time they did the scoring and when they did scoring 2 weeks later. Cohen's κ is also calculated overall and by score for each judge

Table 4 Interrater reliability for each pair of judges at first scoring

Judge pair	1–2	1–3	1–4	1–5	2–3	2–4	2–5	3–4	3–5	4–5	Average
Two scores agree											
<i>n</i>	66	69	70	71	67	67	67	70	73	70	69.0
%	73.3	76.7	77.8	78.9	74.4	74.4	74.4	77.8	81.1	77.8	76.7
Number of scores that differ by											
1	20	17	18	19	18	22	21	20	16	19	19.0
2	3	4	1	0	5	1	1	0	1	0	1.6
3	1	0	1	0	0	0	1	0	0	1	0.4
>3	0	0	0	0	0	0	0	0	0	0	0.0
Cohen's κ	0.65	0.70	0.71	0.73	0.67	0.67	0.67	0.71	0.76	0.71	0.70

This round of scoring was completed within 1 week of a training session for the judges where examples were reviewed. The average Cohen's κ was 0.70 and the average percentage of agreement was 76.7 %. Most scores that did not agree differed by only one score

Table 5 Interrater reliability for each pair of judges at second scoring

Judge pair	1–2	1–3	1–4	1–5	2–3	2–4	2–5	3–4	3–5	4–5	average
Two scores agree											
<i>n</i>	66	69	68	71	62	63	61	78	65	65	66.8
%	73.3	76.7	75.6	78.9	68.9	70.0	67.8	86.7	72.2	72.2	74.2
Number of scores that differ by											
1	20	19	22	14	22	24	24	11	20	21	19.7
2	4	2	0	3	5	3	3	1	4	3	2.8
3	0	0	0	2	1	0	2	0	1	1	0.7
>3	0	0	0	0	0	0	0	0	0	0	0.0
Cohen's κ	0.65	0.70	0.68	0.72	0.60	0.61	0.58	0.83	0.64	0.64	0.66

This scoring was completed 2 weeks after the first scoring. The average Cohen's κ is lower but still acceptable at 0.66, with a percentage of agreement of 74.2 %. As before, most scores that were not in agreement differed by one score

and 0.80, respectively) than for 3 (0.50) (Table 6). In the second scoring, reliability was lower for scores 3 and 4 (0.46 and 0.59, respectively) than for scores 1, 2, and 7 (Table 7). An examination of the distribution and differences of scores between the first and second scoring showed that if there was not agreement, they usually differed by only one score (Table 8).

There was no significant difference in Cohen's κ for videos captured at 30 versus 60 frames per second ($P = 0.16$). When comparisons were made by score, there were no statistically significant differences in frame rate for scores 1, 2, and 4; however, there was a statistically significant difference for scores 3 and 7 ($P < 0.001$; Fig. 2). For a score of 3, the videos captured at 60 frames per second had a significantly lower

Table 6 Interjudge Cohen's κ by scale score for first scoring

Judge pair	Scale score						
	1	2	3	4	5	6	7
1–2	0.94	0.77	0.57	0.58	–	–	0.73
1–3	0.90	0.64	0.46	0.63	–	–	0.81
1–4	0.83	0.67	0.46	0.75	–	–	0.81
1–5	0.86	0.70	0.58	0.70	–	–	0.75
2–3	0.84	0.64	0.49	0.60	–	–	0.73
2–4	0.88	0.62	0.31	0.71	–	–	0.73
2–5	0.85	0.59	0.44	0.66	–	–	0.79
3–4	0.84	0.72	0.49	0.65	–	–	0.86
3–5	0.81	0.69	0.68	0.75	–	–	0.94
4–5	0.85	0.72	0.52	0.65	–	–	0.81
Average	0.86	0.68	0.50	0.67	–	–	0.80

The interjudge κ by score shows high reliability (0.60+) for scores, except for a score of 3, which is more variable and has a lower average Cohen's κ value. There were no scores of 5 or 6 observed

Table 7 Interjudge Cohen's κ by scale score for second scoring

Judge pair	Scale score						
	1	2	3	4	5	6	7
1–2	0.77	0.61	0.45	0.52	–	–	0.81
1–3	0.92	0.65	0.33	0.58	–	–	0.89
1–4	0.82	0.62	0.40	0.68	–	–	0.88
1–5	0.77	0.64	0.75	0.71	–	–	0.81
2–3	0.79	0.49	0.30	0.53	–	–	0.81
2–4	0.78	0.56	0.38	0.56	–	–	0.69
2–5	0.72	0.48	0.49	0.55	–	–	0.63
3–4	0.89	0.77	0.75	0.84	–	–	0.88
3–5	0.84	0.63	0.37	0.45	–	–	0.68
4–5	0.89	0.65	0.35	0.47	–	–	0.78
Average	0.82	0.61	0.46	0.59	–	–	0.78

The interjudge Cohen's κ by score shows high reliability (0.60+) for scores, except for scores of 3 and 4, which are more variable and have lower average Cohen's κ values. This indicates that the training may not be retained over 2 weeks and is important for scoring 3's and 4's. There were no scores of 5 or 6 observed

reliability than those captured at 30 frames per second (0.30 vs. 0.56). For a score of 7, the reliability of videos captured at 60 frames per second was significantly higher than that at 30 frames per second (1.00 vs. 0.693).

Validity of the Scale

When the IMPAS was used to score 39 swallows from intact animals and 39 swallows from abnormal animals, there was a clear difference in the distribution of scores (Fig. 3). In intact pigs, 61.6 % scored a 1, 33.3 % scored a 2, 5.1 % scored a 3, and none scored a 4 or a 7. Again, there were no scores of 5 or 6. This was in stark contrast to what was found for the abnormal pigs where 46.2 % scored a 1, 2.6 % scored a 2, 7.7 % scored a 3, 2.6 % scored a 4, and 41.0 % scored a 7. In addition, a two-sample *t* test determined that the normal and abnormal pigs were significantly different ($t = -4.89$, $P < 0.001$). The mean for normal pigs was 1.42 with a standard deviation of 0.60 and the mean for abnormal pigs was 3.72 with a standard deviation of 2.86.

Discussion

The IMPAS will allow researchers to use the infant pig model and other infant mammalian models for assessing the pathophysiology of swallowing dysfunction and outcomes of rehabilitation. The Cohen's κ result was interpreted as follows: 0–0.20 = slight, 0.20–0.40 = fair, 0.40–0.60 = moderate, 0.60–0.80 = substantial, and 0.80–1.00 = almost perfect strength of agreement [21]. The interrater reliability assessment demonstrated substantial strength of agreement and the intrarater reliability demonstrated “almost perfect strength” of agreement [21]. The ICCs for intra- and interrater reliabilities were also very high (0.92 and 0.88, respectively). The interrater reliability by score demonstrated that scores 3 and 4 were harder to score reliably and that reliability decreased after the 2-week break. This underscores the importance of the training of judges, especially when multiple judges are used. When the scale is used, inter- and intrarater reliabilities should be calculated and assessed to determine functional significance of results.

Table 8 Distribution of first and second grading scores for all raters collectively

Second grading score	First grading score							Total	% of total	% Agree
	1	2	3	4	5	6	7			
1	75	4						79	29.3	94.9
2	9	68	1				1	79	29.3	86.1
3	1	3	29	4				37	13.7	78.4
4		1	3	36			5	45	16.7	80.0
5					0			0	0	0
6						0		0	0	0
7			1	1			28	30	11.1	93.3
Total	85	76	34	41	0	0	34	270		
% of total	31.5	28.1	12.6	15.2	0	0	12.6			
% Agree	88.2	89.5	85.3	87.8	0	0	82.35			

Table shows the frequency of scores at the first scoring and 2 weeks later at the second scoring. No scores of 5 or 6 were observed

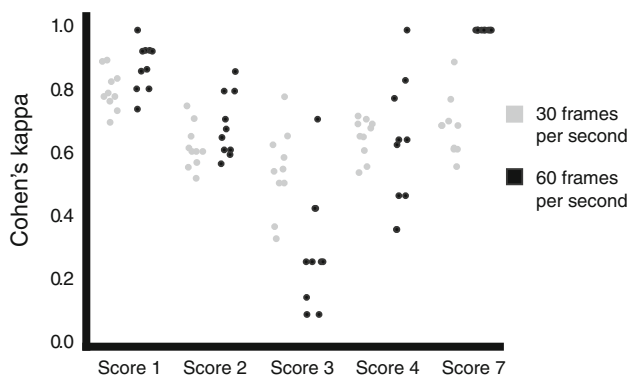


Fig. 2 Distribution of intraclass correlation coefficients (ICCs) for videos captured at 30 and 60 frames per second by 7-point infant mammalian penetration–aspiration scale score and by video capture rate. There is no significant difference in ICC between videos captured at 30 and 60 frames per second for scores 1, 2, and 4. There is a significant difference for scores 3 and 7 ($P < 0.001$)

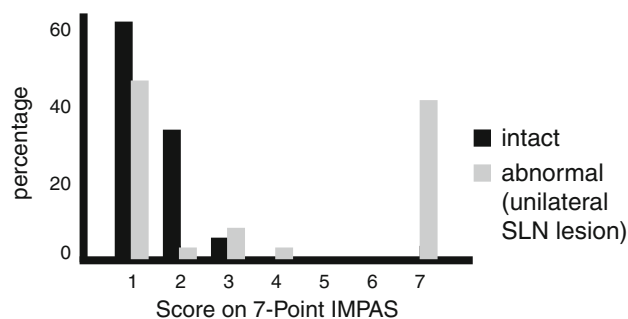


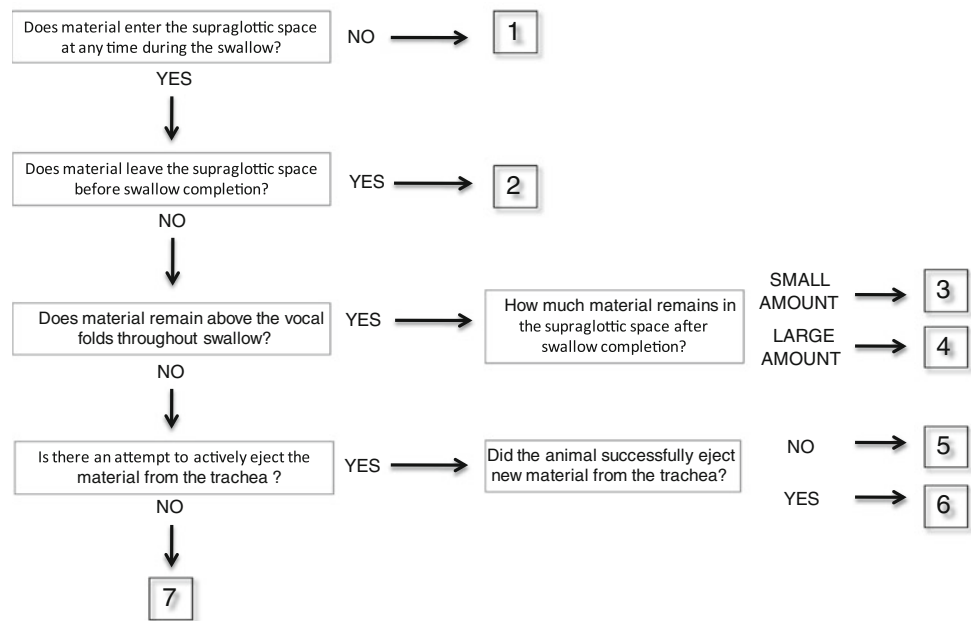
Fig. 3 Distribution of scores given to control animals and animals with a unilateral superior laryngeal nerve lesion. The graph shows a clear distinction between scores given to normal and abnormal animals

A significant difference was found in reliability of the IMPAS for videos captured at 30 versus 60 frames per second only for scores 3 and 7. For scores of 7, although the difference

was statistically significant, it was most likely not functionally significant since the reliability was still high (0.60+). For a score of 3, there was lower reliability for videos at 60 frames per second, which was not expected since these videos can capture behaviors with a higher time resolution. A score of 3 may appear differently depending on the capture rate or resolution of the video. While clinical videofluoroscopic swallowing studies utilize 30 frames per second for recording, animal studies are able to take advantage of the 60-frames-per-second setting that is an option on most videofluoroscopic units. This suggests that the score of 3 needs to be defined clearly to raters by having a detailed description of the size of the bolus and having examples.

The data also showed that this scale can be used to distinguish normal swallowing from abnormal swallowing. The distribution of scores for normal swallows indicates that scores of both 1 and 2 are seen and do not indicate a pathological condition. A score of 2 occurred during approximately one third of all swallows. Very rarely was there a score of 3. Although a score of 2 is penetration, it is often seen in normal pigs and may be a result of their developing coordination between sucking, swallowing, and breathing, although it is not observed in normal infant feeding [22]. Both infant humans and pigs feed in an upright position, so posture, or gravity, should not affect the rate of penetration. The rate of scores of 2 in the normal, healthy infant pigs is a notable difference between infant human and pig swallows. The rate of scores of 2 in infant pigs is actually more comparable to that of adult humans, who have scores of 2 normally in about 20 % of swallows, and it does not indicate abnormal swallowing [5]. In the experience of the authors, scores of 4 and 7 also may occur in a normal, healthy animal, but these are extremely rare conditions. The distribution of scores for infant pigs with a unilateral SLN lesion shows more frequent scores of 4 and many scores of 7, indicating silent

Fig. 4 Schematic representation of the 7-point infant mammal penetration–aspiration scale. This schematic can be used by the judges as a systematic way to approach scoring the videos



aspiration. This difference was statistically significant ($P < 0.001$) when tested using a two-way *t*-test. As the infant pig model is used to model different causes of swallowing dysfunction, this scale can be used to describe the extent of airway protection.

We found that it is important for all judges to first review the data and have a group discussion before doing any scoring in order to maximize the reliability of the results. After the judges scored the videos and the inter- and intrarater reliability scores were low, we gave further training based on that data in order to achieve high inter- and intrarater reliability scores. After concluding the assessment of inter- and intrarater reliabilities, a schematic was developed to help train judges who will use this scale in the future (Fig. 4). Because the IMPAS is a multidimensional scale, it may help judges to score the videos.

It is important to note that this scale does not quantify all possible events during swallowing; rather it groups them into categories based on extent of airway protection. Instead of grouping swallows as (a) normal, (b) having penetration, or (c) having aspiration, the seven categories described allow for a more precise description of the swallow observed. While this study focused on using videofluoroscopy to assess swallow function, it should be emphasized that videofluoroscopy should be used along with other methods. Videofluoroscopy is essential for visualizing penetration and aspiration; however, ultrasound and electromyography are other important tools that can allow researchers to understand the mechanisms of swallowing dysfunction in animal models [23–25]. Endoscopy,

a valuable research tool in humans, is not possible in many animal models because of the extensive nasal conchae.

There are many opportunities for swallowing research to advance by using the 8-Point PAS for the infant pig model. This scale could be adapted for other infant mammal models such as cats, rabbits, and monkeys, which are already being used to study feeding and swallowing [26–28]. Mammals, both infants and adults, share a common pharyngeal anatomy, with an intranasal larynx and a soft palate and epiglottis that contact [16]. Despite differences in chewing and oral transport, it is expected that the same scores described in the IMPAS would also be seen in other mammals since they share a common pharyngeal anatomy. Using this scale, new studies can be designed to further understand swallowing neurophysiology by using a pathological animal model. Along with other advanced technology, we can further understand what causes penetration and aspiration.

Acknowledgments We acknowledge the Johns Hopkins University School of Medicine animal support staff which includes Dr. Dawn Ruben, Dr. Eric Hutchinson, Kristy Koenig, and Melanie Albano. We also thank Dr. Ianessa Humbert for her helpful review of the manuscript. This project was funded by a 2012 American Association for Dental Research Student Research Fellowship (T32 DE07309) to the University of Maryland School of Dentistry and National Institutes of Health DC03604 to RZG.

Funding This project was funded by a 2012 AADR Student Research Fellowship, T32 DE07309 to the University of Maryland School of Dentistry and NIH DC03604 to RZG.

Conflict of interest The authors have no conflicts of interest or financial ties to disclose.

References

1. Pikus L, Levine M, Yang YX, Rubesin S, Katzka D, Laufer I, Gefter W. Videofluoroscopic studies of swallowing dysfunction and the relative risk of pneumonia. *AJR Am J Roentgenol*. 2003; 180:1613–6.
2. Rosenbek JC, Robbins JA, Roecker EB, Coyle JL, Wood JL. A penetration–aspiration scale. *Dysphagia*. 1996;11:93–8.
3. Rofes L, Arreola V, Almíral J, Cabre M, Campins L, Garcia-Peris P, Speyer R, Clave P. Diagnosis and management of oropharyngeal dysphagia and its nutritional and respiratory complications in the elderly. *Gastroenterol Res Pract*. 2011;2011.
4. Bulter S, Stuart A, Leng X, Wilhelm E, Rees C, Williamson J, Kritchevsky S. The relationship of aspiration status with tongue and handgrip strength in healthy older adults. *J Gerontol A Biol Sci Med Sci*. 2011;66A:452–8.
5. Robbins J, Coyle J, Rosenbek J, Roecker E, Wood J. Differentiation of normal and abnormal airway protection during swallowing using the penetration–aspiration scale. *Dysphagia*. 1999; 14:228–32.
6. Baijens LW, Speyer R, Passos VL, Pilz W, Roodenburg N, Clave P. Swallowing in Parkinson patients versus healthy controls: reliability of measurements in videofluoroscopy. *Gastroenterol Res Pract*. 2011;2011:380682.
7. Hey C, Pluschinski P, Stanschus S, Euler HA, Sader RA, Langmore S, Neumann K. A documentation system to save time and ensure proper application of the fiberoptic endoscopic evaluation of swallowing (FEES®). *Folia Phoniatr Logop*. 2011;63: 201–8.
8. Ciucci M, Russell J, Schaser A, Doll E, Vinney L, Conner N. Tongue force and timing deficits in a rat model of Parkinson disease. *Behav Brain Res*. 2011;222(2):315–20.
9. Lever T, Simon E, Cox K, Capra N, O'Brien K, Hough M, Murashov A. A mouse model of pharyngeal dysphagia in amyotrophic lateral sclerosis. *Dysphagia*. 2012;25:112–26.
10. Watrous BJ, Suter PF. Oropharyngeal dysphagias in the dog: a cinefluorographic analysis of experimentally induced and spontaneously occurring swallowing disorders. *Vet Radiol*. 1983;24: 11–24.
11. German RZ, Crompton AW, Thexton AJ. Integration of the reflex pharyngeal swallow into rhythmic oral activity in a neurologically intact pig model. *J Neurophysiol*. 2009;102:1017–25.
12. Klein BG, Duffin JR, Kraje B. Neonatal infraorbital nerve damage and the development of eating behavior in the rat. *Behav Brain Res*. 1994;60:25–33.
13. Bennett SW, Lanovaz JL, Muir GD. The biomechanics of locomotor compensation after peripheral nerve lesion in the rat. *Behav Brain Res*. 2012;229:391–400.
14. Aluffi P, Policarpo M, Cherovac C, Olina M, Dosdegani R, Pia F. Post-thyroidectomy superior laryngeal nerve injury. *Eur Arch Otorhinolaryngol*. 2001;258:451–4.
15. Kark AE, Kissin MW. Superior laryngeal nerve injury. *Head Neck*. 1995;17:542–3.
16. Hiimae KM. Feeding in mammals. In: Schwenk K, editor. *Feeding: form, function and evolution in tetrapod vertebrates*. San Diego: Academic Press; 2000. p. 411–48.
17. Ding P, Campbell-Malone R, Wahl S, Fukuhara T, Feng W, German R. Aspiration mechanism for unilateral superior laryngeal nerve lesion in animal model. In: *Dysphagia Research Society Annual Meeting*, San Antonio, TX; 2011 (abstract).
18. Ding P, Chen Y, Lukasik S, German R. Mechanism of epiglottal movement during the infant pharyngeal swallow. In: *Dysphagia Research Society Annual Meeting*, Toronto, ON, Canada; 2012 (abstract).
19. Reuss-Lamky H. Administering dental nerve blocks. *J Am Anim Hosp Assoc*. 2007;43:298–305.
20. Friedman B, Frazier JB. Deep laryngeal penetration as a predictor of aspiration. *Dysphagia*. 2000;15:153–8.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
22. Newman L, Cleveland R, Blickman J, Hillman R, Jaramillo D. Videofluoroscopic analysis of the infant swallow. *Invest Radiol*. 1991;26:870–3.
23. Macrae PR, Doeltgen SH, Jones RD, Huckabee ML. Intra- and inter-rater reliability for analysis of hyoid displacement measured with sonography. *J Clin Ultrasound*. 2012;40:74–8.
24. Kuo P, Holloway RH, Nguyen NQ. Current and future techniques in the evaluation of dysphagia. *J Gastroenterol Hepatol*. 2012;27:873–81.
25. Konow N, Thexton AJ, Crompton AW, German RZ. Regional differences in length-change and electromyographic heterogeneity in the sternohyoid muscle during infant mammalian swallowing. *J Appl Physiol*. 2010;109:439–48.
26. Iriarte-Diaz J, Reed DA, Ross CF. Sources of variance in temporal and spatial aspects of jaw kinematics in two species of primates feeding on foods of different properties. *Integr Comp Biol*. 2011;51:307–19.
27. Tsujimura T, Yamada A, Nakamura Y, Fukuhara T, Yamamura K, Inoue M. The digastric muscle is less involved in pharyngeal swallowing in rabbits. *Dysphagia*. 2012;27:271–6.
28. Medda B, Lang I, Dodds W, Christl M, Kern M, Hogan W, Shaker R. Correlation of electrical and contractile activities of the cricopharyngeus muscle in the cat. *Am J Physiol*. 1997;273:G470–9.

Shaina Devi Holman BS

Regina Campbell-Malone PhD

Peng Ding MD, PhD

Estela M. Gierbolini-Norat BS

Anne M. Griffioen BA

Haruhi Inokuchi MD

Stacey L. Lukasik BA

Rebecca Z. German PhD