



Bedside voice assessments cannot be used as a screening test for laryngeal injury following prolonged intubation in an intensive care population

Lucy Huang^{1,2^}, Theodore Athanasiadis^{1,2}, Mistyka Schar^{2,3^}, Charmaine Woods^{1,2^}, Ahmed Bassiouni^{1^}, Stephanie Martin^{3,4}, Jane Bickford^{3,4^}, Shailesh Bihari^{2,5^}, Eng Hooi Ooi^{1,2^}

¹Department of Otorhinolaryngology Head and Neck Surgery, Flinders Medical Centre, Bedford Park, SA, Australia; ²Flinders Health and Medical Research Institute, College of Medicine and Public Health, Flinders University, Bedford Park, SA, Australia; ³Department of Speech Pathology and Audiology, Flinders Medical Centre, Bedford Park, SA, Australia; ⁴College of Nursing and Health Sciences, Flinders University, Bedford Park, SA, Australia; ⁵Department of Intensive & Critical Care Unit, Flinders Medical Centre, Bedford Park, SA, Australia

Contributions: (I) Conception and design: L Huang, M Schar, C Woods, S Bihari, EH Ooi; (II) Administrative support: S Bihari, C Woods, EH Ooi; (III) Provision of study materials or patients: S Bihari, EH Ooi; (IV) Collection and assembly of data: L Huang, M Schar, T Athanasiadis, EH Ooi, J Bickford, S Martin; (V) Data analysis and interpretation: A Bassiouni, C Woods, S Bihari, T Athanasiadis, EH Ooi; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Dr. Lucy Huang. Department of Otorhinolaryngology Head and Neck Surgery, Flinders Medical Centre, Bedford Park, SA, Australia. Email: Lucy.huang.tw@gmail.com.

Background: Prolonged intubation in intensive care patients is associated with laryngeal injury. The recognition of laryngeal injury by nasoendoscopy examination is not routine clinical practice following extubation. This study aimed to determine if bedside voice assessments provide clinical utility as a simple screening tool to identify laryngeal injury in patients following extubation in the intensive care setting.

Methods: Patients intubated for longer than 24 hours were recruited from a tertiary hospital intensive care unit. Bedside voice assessments [S/Z ratio, GRBAS, Consensus Auditory Perceptual Evaluation-Voice (CAPE-V)] and laryngeal endoscopic examination were conducted initially at 24–48 hours and repeated at 5–7 days after extubation. Voice outcomes were correlated with endoscopic findings.

Results: Initial assessments were conducted in 60 participants, with 37 participants completing the second assessment. The median [interquartile range (IQR)] duration of intubation was 60.1 (38.2–136.5) hours. The prevalence of endoscopic laryngeal injury was 92% at 24 hours and remained high (78%) at 5 days post extubation. The most common injury type was vocal process ulceration/granuloma (in 84% of patients). Bedside voice assessments detected 100% of patients had dysphonia at 24 hours, but did not correlate with the presence of laryngeal injury.

Conclusions: Laryngeal injury and dysphonia are prevalent following a prolonged period of intubation. Bedside voice assessments do not correlate with the type or severity of laryngeal injury. It is important to be aware of the high prevalence of laryngeal injury in this population following prolonged intubation.

Keywords: Laryngeal injury; intensive care; intubation; voice; dysphonia

Received: 16 October 2021; Accepted: 22 March 2022; Published: 28 April 2022.

doi: 10.21037/ajo-21-42

View this article at: <https://dx.doi.org/10.21037/ajo-21-42>

[^] ORCID: Lucy Huang, 0000-0001-7910-1477; Mistyka Schar, 0000-0002-7440-592X; Charmaine Woods, 0000-0001-7898-9689; Ahmed Bassiouni, 0000-0002-5545-0194; Jane Bickford, 0000-0001-6080-4279; Shailesh Bihari, 0000-0001-7553-8620; Eng Hooi Ooi, 0000-0002-7201-815X.

Introduction

Intubation in the intensive care setting provides essential oxygenation and ventilation for the critically ill patient. The endotracheal tube (ETT) used for intubation has been described to cause short- and long-term effects on the larynx and its functions (1). Subsequently, impaired laryngeal function poses consequences to the recovery process and quality of life, with associated increase in healthcare expenditure (2).

During mechanical ventilation, the ETT exerts pressure on the posterior larynx due to the supine position of the patient, the natural cervical lordosis, and the V-shape of the glottis (3,4). Persistent tissue ischemia leads to ulceration of the tissue with superficial injury occurring in as little as three hours of intubation (5). Once the insult is removed, re-epithelisation occurs with the formation of granulation tissue (6). Injury involving the perichondrium or cartilage can result in fibrosis which may lead to laryngeal stenosis (4). In patients intubated for three days or more, the reported incidence of laryngeal injury is as high as 97% (7). Injury to the larynx can impair laryngeal functions such as respiration, phonation and swallowing (8).

A systematic review identified a lack of screening tools and guidelines to assist in identifying patients who may be at high risk of laryngeal injury post-extubation (2). There are limited studies that assess both the voice and the laryngeal examination findings in the critical care population following extubation (2). It is not routine, nor practical, to perform endoscopic laryngeal examination in all patients post-extubation, with observation remaining the most commonly employed method of monitoring upper airway symptoms after extubation (2). Bedside screening assessments may have clinical utility in the identification of those patients at high risk of injury to prompt early diagnostic investigation and intervention by an Otolaryngologist or Speech Pathologist. The S/Z ratio is a simple bedside voice assessment proposed as a screening tool to diagnose laryngeal pathology post-extubation (3). The idea that a simple voice assessment can identify patients of concern is attractive, especially if it reduces the need for aerosol generating procedures associated with endoscopic laryngeal examination.

This study aimed to assess the clinical utility of the S/Z ratio and other voice assessments [GRBAS (9) and Consensus Auditory Perceptual Evaluation-Voice (CAPE-V) (10)] as screening tools for the presence and type of laryngeal injury following prolonged intubation

of at least 24 hours or more in the intensive care setting. A secondary aim was to assess short-term changes of the laryngeal injuries in this cohort. We present the following article in accordance with the STROBE reporting checklist (available at <https://www.theajo.com/article/view/10.21037/ajo-21-42/rc>).

Methods

Study design

This is a prospective observational study conducted at the Flinders Medical Centre Intensive and Critical Care unit (ICCU) between October 2018–October 2019. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Ethical approval was sought from the Human Research Ethics Committee at Southern Adelaide Local Health Network. Ethics approval code is OFR 208.17. Informed consent was obtained from patients prior to enrolment in the study.

Inclusion and exclusion criteria

The inclusion criteria were patients in the ICCU who were 18 years of age or above, translaryngeal intubation for greater than 24 hours, successfully extubated for at least 24 hours, with the capacity to consent and follow instructions.

The exclusion criteria were patients who underwent a tracheostomy, those with head and neck injury, inhalation injury, thyroid, neck, cervical spine surgery, cognitive impairment, inability to follow instructions, dysphasic, dysarthric, unable to phonate, those with previous voice concerns requiring intervention and those who the ICCU team deem not medically appropriate.

Study protocol

Patients underwent laryngeal assessment and voice assessments at two timepoints following extubation (24–48 hours and at 5–7 days; *Figure 1*).

The following clinical data were collected: patient characteristics (age, gender, height, weight), significant medical history [smoking history, use of proton pump inhibitor or steroids, acute physiology and chronic health evaluation II (APACHE II) score], indications for ICCU admission and endotracheal intubation, intubation information (date and time of intubation, site of intubation, number of attempts, Cormack-Lehane score, ETT

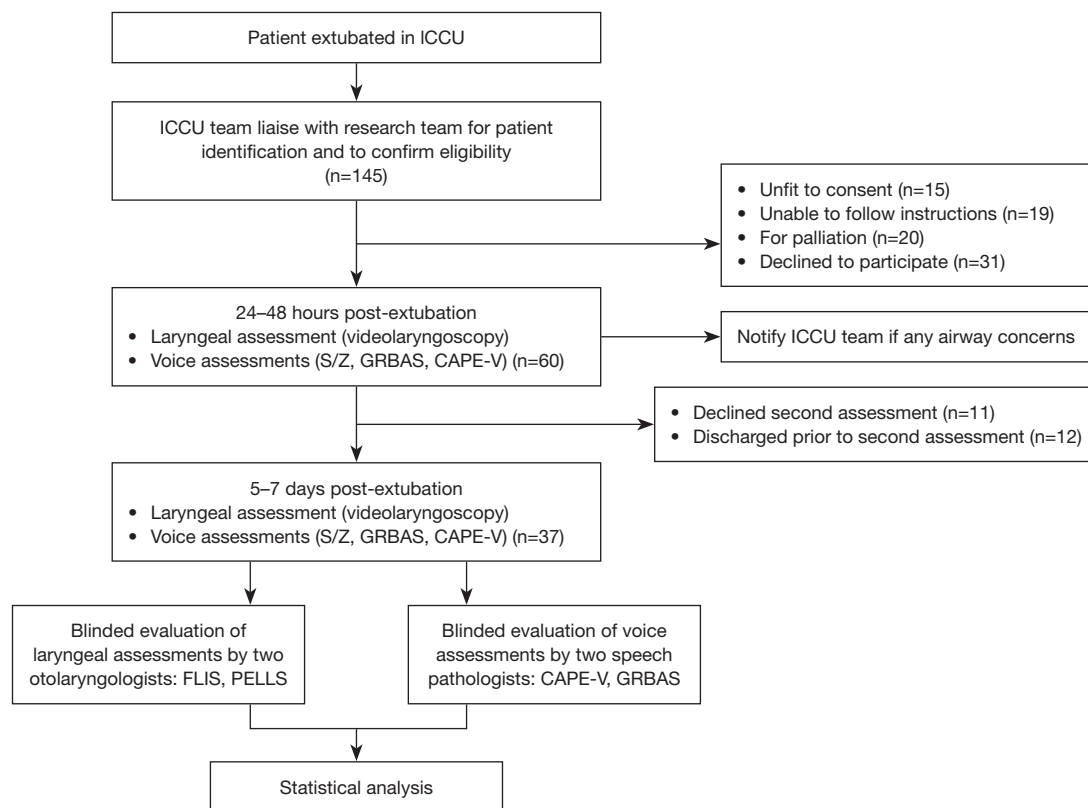


Figure 1 Study protocol of bedside voice assessments versus endoscopic laryngeal evaluation in recently extubated intensive care patients. ICU, Intensive and Critical Care unit; CAPE-V, Consensus Auditory Perceptual Evaluation of Voice; FLIS, Flinders Laryngeal Injury Score; PELLs, Post-Extubation Laryngeal Lesions Scale.

size, neuromuscular blockers, mean cuff pressure, self-extubation), presence of nasogastric tube, length of ICU and hospital stay.

Examination of the larynx and the voice were conducted at the bedside within the ICU or on the ward and were recorded for subsequent analysis. Patients with stridor or airway concerns on examination were immediately communicated with the ICU team.

Voice assessment and analysis

Voice recordings were performed using a Shure SM48 microphone attached to a Zoom H5 Handy Recorder and placed at approximately 15 cm from the patient’s mouth. The patient was asked to speak directly into the microphone. The recordings were anonymised.

Three bedside voice evaluation tools were used: S/Z ratio (11), CAPE-V (10) and GRBAS (9). S/Z ratio is a simple way of measuring glottic function where vocal fold pathology

reduces the phonation of /z/ but not /s/, increasing the s and z ratio (11). Participants were asked to repeat and sustain /s/ and /z/ three times each, with the duration measured using a stopwatch. The longest /s/ and /z/ were then used for the calculation. An S/Z ratio greater than 1.4 indicates the presence of vocal fold abnormality (11). CAPE-V and GRBAS are both perceptual voice assessments. Assessment of CAPE-V required the patient to repeat a standardised set of sentences as detailed by the American Speech-Language-Hearing Association (12). It is rated on a continuous scale from 0–100; giving a score to the overall severity, roughness, breathiness, strain, pitch and loudness. GRBAS scores were determined from the same recording as CAPE-V, rated using an ordinal scale from 0–3; giving a score to the overall grade, roughness, breathiness, asthenia, and strain (10). Two senior speech pathologists with special interests in voice disorders assessed the voice recordings independently using the 2 tools. Disagreements, more than 10 points with CAPE-V and 1 point with GRBAS, were resolved by consensus.

Table 1 Flinders laryngeal injury score (FLIS)

Subsites	Injuries	Right	Left
Supraglottis	Ulceration/granuloma		
	Erythema		
Glottis	Ulceration/granuloma		
	Erythema		
	Mobility		
	Stenosis		
Subtotal			
Total			

Ulceration/granuloma: 0: none; 1: mild; 2: moderate; 3: severe. Erythema: 0: none; 1: hyperaemia; 2: haematoma. Stenosis: 0: none; 1: <25%; 2: 25–50%; 3: >50%. Mobility: 0: normal; 1: reduced mobility; 2: immobility. Range of 0–27 with higher scores indicating increasing severity.

Laryngeal assessments and analysis

Topical anaesthesia with Co-Phenylcaine Forte™ (5% lignocaine, 0.5% phenylephrine) was applied to the nasal cavity prior to the laryngeal examination. Digital video nasoendoscopy was performed (by LH, 3-year experience performing this) with continuous white light using Karl Storz Tele Pack x LED TP100 (3.7 mm Strobe Video Rhino-Laryngoscope, Tuttlingen, Germany) and recorded. Patients were asked to whistle, inhale gently, repeat /i/, gliding /i/, prolonged /a/, and count to ten to assess the larynx.

Two evaluation tools were used to assess the larynx: the Post-Extubation Laryngeal Lesions Scale (PELLS) (13), and the newly developed Flinders Laryngeal Injury Score (FLIS). PELLS is a scale for assessment of laryngeal injury following intubation (13), however it is simplified into 5 categories:

- 0: no oedema, injury, or hyperaemia;
- 1: supraglottic oedema;
- 2: both supraglottic and glottic oedema and/or presence of vocal process granuloma, vocal cord ulcers and/or arytenoid luxation;
- 3: more intense supraglottic and glottic oedema with or without haematomas;
- 4: involvement of subglottic lesion.

Therefore, the FLIS was developed to detail the type (ulceration/granuloma, erythema, stenosis, mobility), and severity (detailed in *Table 1*) of an injury at each of the two laryngeal subsites visibly accessible via nasoendoscopic

examination (supraglottis and glottis; subglottis was not evaluated as the view is often obstructed by the glottis) in the awake patient. The total FLIS scores range from 0 (no injury) to 27 (most severe injury in every subsite). Laryngeal videos were evaluated by two senior otolaryngologists blinded to the participants' identities and clinical presentations, with disagreements resolved by consensus.

Statistical analysis

Statistical analysis was performed using R version 3.6.1 (R Project for Statistical Computing, Vienna, Austria). Data are presented as median and interquartile range (IQR). The Mann-Whitney U and the Wilcoxon signed-rank test were used when comparing un-paired and paired data, respectively. Chi-squared tests were used for unpaired categorical data. Spearman's Rank Order Correlation was used for the assessment of correlation in determining the concurrent validity. Regression modelling (logistic regression for dichotomous outcomes, and linear regression for continuous outcomes) was used to explore whether variables were significantly associated with outcomes. One way analysis of variance (ANOVA) was used for assessing the risk factors of predictive laryngeal injury. An alpha level of 0.05 was used for statistical significance.

Results

One hundred and forty-five ICCU patients were screened for eligibility. 60 patients consented to participate in this study, with 37 of these patients consenting to the second examination at 5–7 days post extubation (*Figure 1*).

Patient characteristics

The population had a median age of 65 (IQR: 57–75) years old, were predominantly male, with an APACHE II score of 19 (IQR: 17–23) representing a moderately unwell group of ICCU patients. Patient characteristics, admission and intubation details are presented in *Table 2*. Patients were all intubated via the orotracheal route and mechanically ventilated for a median duration of 60.1 (IQR: 38.2–136.5) hours. No patient self-extubated.

The initial assessments were conducted at 1.1 (IQR: 0.9–1.2) days after extubation (n=60), with a second assessment performed at 6.1 (IQR: 5.1–7.1) days after extubation (n=37). Even though patients had varying severity of laryngeal injuries, none were severe enough to cause airway

Table 2 Patient demographic and characteristics

Characteristics	Patients who completed first assessment (n=60)	Patients who completed both assessments (n=37)
Age (years)	65 [57–75]	65 [57–78]
Gender		
Male	39 (65%)	22 (59%)
Female	21 (35%)	15 (41%)
APACHE II score	19 [17–23]	19 [17–23]
Length of ICCU stay (days)	3.9 [2.8–8.8]	4.3 [2.6–9.0]
Length of hospital admission (days)	16.1 [8.4–29.9]	25.6 [11.5–45.4]
Body mass index (kg/m ²)	26 [24–33]	26 [23–30]
Height/tube ratio (cm)	22.7 [21.9–23.3]	22.8 [21.8–23.6]
Smoking history		
Ex-smoker	30 (50%)	17 (35%)
Non-smoker	19 (32%)	13 (46%)
Active smoker	11 (18%)	7 (19%)
Admission specialty		
Medicine	33 (55%)	20 (54%)
Surgery	27 (45%)	17 (28%)
Nasogastric tube inserted	48 (80%)	32 (86%)
Proton pump inhibitor use	42 (70%)	29 (78%)
Steroid (oral or parenteral) use	20 (33%)	14 (38%)
Diabetes	19 (32%)	13 (35%)
Intubation duration (hours)	60 [38–136]	68 [37–160]
Intubation location		
OT	30 (50%)	21 (57%)
ICCU	15(25%)	9 (24%)
Retrieval team	5 (8%)	2 (5%)
Ward	2 (3%)	1 (3%)
ED	8 (14%)	4 (11%)
Cormack Lehane		
1	36 (60%)	23 (62%)
2	7 (11%)	3 (8%)
3	3 (5%)	0 (0%)
4	1 (2%)	1 (3%)

Table 2 (continued)

Table 2 (continued)

Characteristics	Patients who completed first assessment (n=60)	Patients who completed both assessments (n=37)
VLS with good view	7 (10%)	6 (16%)
Not documented	6 (12%)	4 (11%)
Number of intubation attempts		
1	53 (88%)	35 (94%)
2	4 (7%)	1 (3%)
Not documented	3 (5%)	1 (3%)
Use of muscle relaxant		
Rocuronium	41 (68%)	26 (70%)
Suxamethonium	8 (13%)	4 (11%)
Vecuronium	6 (10%)	4 (11%)
Pancuronium	4 (7%)	2 (5%)
Not documented	1 (2%)	1 (3%)
Size of endotracheal tube		
7 mm	20 (33%)	14 (38%)
7.5 mm	2 (3%)	2 (5%)
8 mm	38 (63%)	21 (57%)
Mean cuff pressure (cmH ₂ O)	27 [26–28]	27 [26–28]

Data presented as median [IQR] or number of patients (%). First assessment is within 24–48 hours post extubation. Second assessment is at 5–7 days post extubation. APACHE II, Acute Physiology, Age, Chronic Health Evaluation II; ICCU, intensive and critical care unit; OT, operating theatre; ED, emergency department; VLS, videolaryngoscopy.

obstruction or to warrant immediate intervention.

Concurrent validity of FLIS

As FLIS is a newly developed scoring system we tested its concurrent validity against PELLs (13). FLIS (scores ranging from 0–14) demonstrated good correlation with PELLs (scores ranging 0–4) (spearman rho correlation 0.745, P value <0.001).

Prevalence and severity of laryngeal and voice injury

The FLIS identified presence of laryngeal injury in 92% (55/60) of patients when assessed at 24–48 h post-

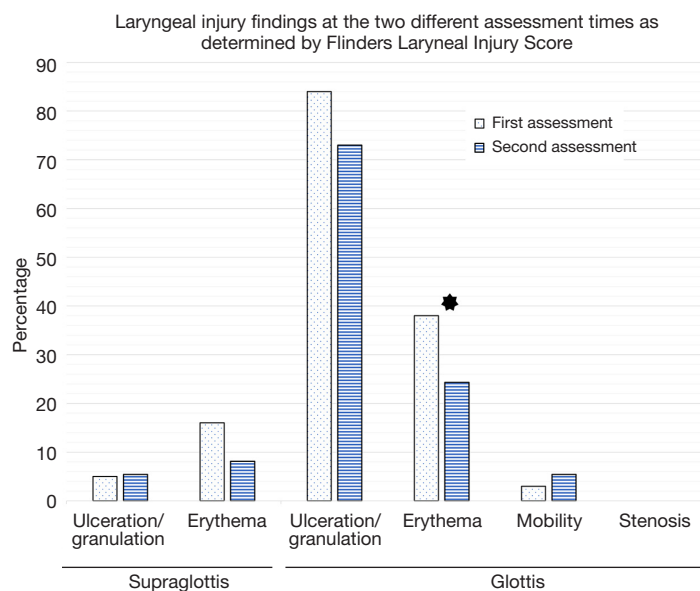


Figure 2 Laryngeal injury findings at the two different assessment times as determined by Flinders Laryngeal Injury Score. Star indicates a statistically significant difference between the two time points.

Table 3 Outcome of voice assessments (CAPE-V & GRBAS)

Voice parameters	CAPE-V		GRBAS	
	First assessment	Second assessment	First assessment	Second assessment
Overall severity (CAPE-V)/grade (GRBAS)*	34 [25–42]	25 [19–35]	2 [1–2]	1 [1–2]
Roughness	25 [20–40]	21 [13–26]	1 [1–2]	1 [1–1]
Breathiness	25 [20–32]	20 [14–30]	1 [1–2]	1 [1–2]
Asthenia	N/A	N/A	1 [1–2]	1 [0–1]
Strain	31 [21–40]	23 [15–29]	1 [1–2]	1 [1–1]
Pitch	17 [0–27]	0 [0–15]	N/A	N/A
Loudness	20 [3–30]	13 [0–19]	N/A	N/A

Data presented as median [IQR] in 60 patients. *, CAPE-V overall severity and GRBAS grade have the same meaning. CAPE-V is on a range from 0–100 continuous scale, with 100 being severely deviant from normal. GRBAS range from 0–3 ordinal scale with increasing severity. CAPE-V, Consensus Auditory Perceptual Evaluation of Voice; N/A, not applicable.

extubation. Example photographs of laryngeal injuries are presented in the [Figure S1](#). Ulceration/granuloma at the glottis was the most common injury, occurring in 84% of patients (*Figure 2*). Total FLIS scores ranged from 0–14, with median score 2.5 (IQR: 2–4). Most patients (96%) had a low score of 0–6. The patient with the highest score (FLIS =14) had severe supraglottic ulceration, hyperaemia and moderately severe glottic granuloma. This patient had a PELLs score of 3. The patient recovered well after discharge from ICU and did not consent to a second

assessment as they felt their voice was better. There was no difference in injury type or severity between the left or the right sides of the larynx ($P>0.05$). The median PELLs score was 2.0 (IQR: 2.0–3.0). The majority of patients (70%) scored between 0–2, largely with mild laryngeal injuries. 15% of patients scored 3 and 15% of patients scored the highest score of 4.

All 60 patients (100%) had abnormal voice when assessed with CAPE-V and GRBAS (*Table 3*) at the first voice assessment (24–48 h). The voice parameter most

Table 4 Laryngeal and voice assessment for the subgroup of patients who completed both assessments (n=37)

Variables	First assessment	Second assessment	P value	Post hoc effect size
FLIS total	5.5 (4.0–7.3)	4.0 (2.0–6.0)	<0.001*	0.5
PELLS	2.0 (2.0–2.5)	2.0 (2.0–2.0)	0.63	0.3
CAPE-V overall severity	33.8 (25.0–44.8)	24.5 (19.4–35.0)	0.002*	0.6
GRBAS grade	2 [1–2]	1 [1–2]	0.02*	0.5
S/Z	1.2 (1.0–1.4)	1.0 (0.9–1.2)	0.15	0.3

Analysed using Wilcoxon signed rank test. *, statistical significance with an alpha level of 0.05. FLIS, Flinders Laryngeal Injury Score; PELLs, Post-Extubation Laryngeal Lesions Scale; CAPE-V, Consensus Auditory Perceptual Evaluation of Voice.

impacted by intubation was ‘strain’ (assessed with CAPE-V) signifying mild-moderate abnormality [median 30.5 (IQR 21.4–40.0)] (Table 3). All parameters assessed with GRBAS had a median score of 1 (IQR 1–2); 25% (15/60) of patients had an abnormal S/Z ratio of above 1.4 (11).

Short term progress of laryngeal and voice injury

A subset of 37 patients consented to a repeat assessment prior to discharge [median time between assessments was 5.0 (IQR: 4.0–6.1) days; median duration of endotracheal intubation was 68.5 (IQR 36.8–160.3)] hours. In this subset, the prevalence of laryngeal injury was 92% (34/37) at the first assessment and 78% (29/37) at the second assessment. Laryngeal injury type and location as identified using FLIS are presented in Figure 2. Glottic ulceration/granuloma remained the most common laryngeal injury, present in 73% patients. There was a statistically significant decrease in the number of patients with glottic erythema from 38% to 24% ($P < 0.01$). The severity decreased significantly over time when assessed using the total FLIS score ($P < 0.001$; Table 4). The PELLs detected laryngeal injury in 86% patients, however the score was not statistically different from the first assessment $P = 0.63$; Table 4). At the second assessment, two patients had significant subglottic granulomas and one had severe glottic granulomas seen on examination and were subsequently referred to the laryngology clinic for further review after determining they were safe to be discharged from the hospital. The subglottic granulomas were only seen at the second assessment following resolution of laryngeal oedema. The patient with the glottic granuloma was stable since the first assessment with no evidence of worsening. None had clinical signs or symptoms of airway obstruction. One patient with the subglottic granuloma had complete resolution after 3 months of follow-up with

regular monitoring. The two other patients did not attend further follow up and expressed they felt well upon phone communication.

Voice assessments demonstrated a significant improvement in voice quality (CAPE-V overall severity, $P < 0.05$; GRBAS grade, $P < 0.05$). There was no statistical difference in S/Z ratio scores between the time points ($P > 0.05$).

Predictors of laryngeal injury

Patient and treatment factors were tested to determine if they could predict laryngeal injury following mechanical ventilation. Patient characteristics (gender, body mass index), medical history (smoking status) history of voice overuse [requiring previous ear, nose, and throat (ENT) or speech intervention], intervention (nasogastric tube), medications (proton pump inhibitor, steroids), and intubation details (number of attempts, size of tube, height tube ratio, Cormack Lehane score, mean cuff pressure) demonstrated no significant statistical association with the severity of a laryngeal injury (graded using FLIS or PELLs scoring systems) using one way ANOVA, at either assessment time point (Table S1).

We explored if voice assessments could be used as a screen for laryngeal injury. Initially, a regression analysis was conducted using each of the voice assessment scores (S/Z ratio > 1.4 , CAPE-V overall severity and GRBAS grade) and presence of laryngeal injury (FLIS 0 = no injury, 1 or above = presence of injury). We could not find a significant association (Table S2). This is likely a result of the high prevalence of laryngeal injury at 92%. Next, linear regression was used to assess voice assessment scores (S/Z, CAPE-V and GRBAS) as predictors of laryngeal injury severity (PELLs and FLIS total scores). The S/Z ratio, when assessed in a continuous manner, appeared

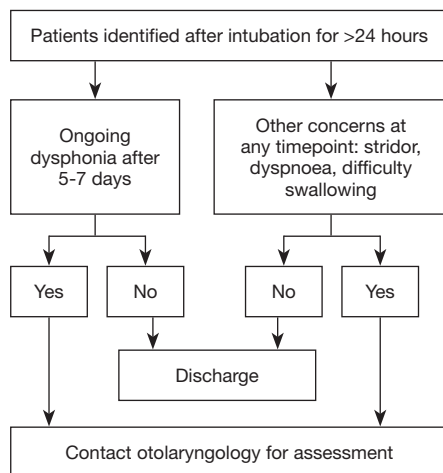


Figure 3 Guideline for otolaryngology referral post prolonged intubation.

to demonstrate a significant correlation with FLIS at the first assessment ($P=0.013$, regression coefficient 0.076). However, this was not demonstrated when assessing S/Z ratio as dichotomous data of greater or lesser than 1.4, and multiple testing correction was not performed. Therefore, this result should be treated with caution as S/Z ratio greater than 1.4 is accepted as the threshold for abnormality. No other significant association between voice assessments and endoscopic examination was identified, at either time point (Table S2).

Discussion

The findings from this study have demonstrated that prolonged intubation for 24 hours or more is associated with a high prevalence of laryngeal injury (92%) and dysphonia which persists for at least 5 days in the majority (78%) of patients. Bedside voice assessments were able to detect post-extubation dysphonia but did not correlate with the presence of laryngeal injury.

The high prevalence of laryngeal injury in patients who have been intubated for more than 24 hours is consistent with that reported in the literature (2,7,14,15). However, the standard clinical practice following extubation does not involve examination by a Speech Pathologist or an Otolaryngologist as assessments require specialised equipment and skilled personnel. Furthermore, considering the number of patients extubated every day in intensive care units, it is not practical to do so. Reassuringly, short term follow-up at 5–7 days demonstrated clinical improvement

in severity of these laryngeal injuries. However, 78% of patients still had some form of laryngeal injury and mild dysphonia at 5–7 days following extubation. The significance of this is supported by a long-term follow-up study that identified some patients have persistent laryngeal injury with ongoing dysphonia or dyspnoea after 10 weeks (15). A small subset of patients may require treatment as a result of the injury sustained from intubation. The cost associated with treating intubation-related moderate to severe laryngeal injuries have been reported up to US \$6,000 (2). In addition to the health care expenditure, dysphonia also has significant impact on the patient's quality of life (16). Despite the lack of correlation between voice assessments and examination findings, we recommend seeking Otolaryngology input if the patient has ongoing dysphonia after 5–7 days (Figure 3).

This study investigates post-intubation laryngeal injuries in the modern intensive care setting using comprehensive bedside voice assessment tools. Patient characteristics, medical history, history of voice overuse, intervention, medications, intubation details did not contribute to laryngeal injury severity. The lack of significance is contradictory to other studies (2,7,17-19), many reported more than 10 years ago, but is supported by a study by Colton House (14). The lack of significance with previously identified risk factors is likely multifactorial: the evaluation tool used, lack of comparison group, small sample size, shorter intubation duration, optimisation of modifiable risk factors and the improvement in the care provided in the modern intensive care setting such as prevalent use of neuromuscular blockers and regular cuff checks.

There is a need for non-invasive methods to identify laryngeal injuries that may persist in this population. A range of bedside voice assessments varying in complexity were employed in this study. Auditory perceptual assessments using CAPE-V and GRBAS are considered the gold standard assessments for voice disorders (20). The current study found that even though all the patients were dysphonic to varying degrees, the two bedside perceptual voice assessments did not correlate with the presence or severity of laryngeal injury. These findings are supported by other studies that found voice assessments poorly reflected the endoscopic findings for other laryngeal pathology (21-23). The patient's ability to compensate for their laryngeal pathology using the contralateral vocal fold, the ventricular folds or the arytenoids may have contributed to this result (22,24,25). The S/Z ratio is a quick and simple method of screening for vocal fold pathology (11) and has

previously been suggested to be helpful in the evaluation of laryngeal function in patients after intubation (3). The S/Z ratio has been shown to correlate with the FLIS as a continuous variable in this study but not when analysed as a dichotomous variable, therefore, future study with a larger study population should be done.

There are currently no widely accepted validated scoring systems to describe laryngeal injury secondary to intubation trauma. The simplicity of PELLs reduces its utility for reporting on complex laryngeal injuries and makes it a less sensitive tool to detect small changes. We initially used it for scoring but realised that the majority of patients had a PELLs score of 2 at both assessments. PELLs is only suitable for assessing larynges with severe oedema at risk of re-intubation in a categorical manner. Therefore, FLIS was developed after identifying the common intubation-related laryngeal injuries, to score them in a way that allows a wider range of data points than PELLs. Subglottic pathology was not included in FLIS due to poor visibility especially with the frequently observed oedema superiorly or supraglottic constriction limiting accurate assessment of the subglottis. It would provide more information for subglottic findings to be included as a supplementary score when identified and can be included in the future study of FLIS as we seek to improve this scoring system and validate it for use. Oedema was excluded from FLIS due to subjectivity in its assessment with most patients having at least mild oedema following extubation. A simple method of assessing a tool's validity is the use of concurrent validity to determine the agreement between two different tools (26). While there is good correlation, PELLs categories cannot be directly translated to a range of numbers in FLIS as FLIS is a sum of different injuries. The correlation seen here is a way to communicate that those assessed with a higher score in PELLs also had a higher total score in FLIS, therefore, providing FLIS with concurrent validity. Further study is required to assess the reliability and prognostic validity of FLIS with the potential to group the scores into mild, moderate, severe categories to increase its clinical utility.

Strengths and limitations

This study population is representative of the Australian ICU patient population. The study undertook a comprehensive examination of the larynx and utilised gold standard voice assessments. Several methods were employed to reduce the bias of subjectivity when using perceptual

ratings, and included following a standardized protocol for both voice and laryngeal assessments, recording all the examinations and anonymising them prior to scoring, and blinding the raters to the identity and history of the patients to reduce cognitive bias (27). Scoring was done independently by two experts from each field where differences were resolved with consensus.

There are several limitations in this study. The critically ill population presents challenges for research which limited the number of participants resulting in a small sample size. As delirium and the use of sedatives are common in the immediate post-extubation period, the decision to evaluate these patients at least 24 hours post-extubation was made. Those who were re-intubated due to laryngospasm, presence of post-extubation stridor, or respiratory failure were excluded from recruitment as these presentations often occur within minutes to hours after extubation. This, along with the other exclusion criteria introduced a selection bias where patients with the most severe laryngeal injury may have been excluded. This study also assumed there was no laryngeal injury prior to intubation. To address this issue case notes were reviewed, and patients were excluded if they had previous voice issues requiring ENT or speech intervention. Given the high prevalence of laryngeal injury of 92% and all patients having dysphonia, there were not enough patients to form a comparative "normal larynx" group, thereby limiting the statistical analyses that could be performed. Some patients declined a second assessment, potentially presenting another selection bias. A further limitation is the lack of longer-term follow-up in these patients to assess the resolution of their voice and laryngeal injury, warranting further research. The critically ill population frequently have reduced respiratory capacity which has potential to impact on the maximum phonation duration required for the calculation of the S/Z ratio. The rationale behind choosing perceptual assessment tools in this study was for their practicality and ease of use by the bedside. Although auditory perceptual assessments are subjective, they are considered gold standard as the voice is perceptual in nature (28). However, voice assessment is multi-dimensional and perhaps other tools requiring specialist analysis such as acoustic analysis, aerodynamics and patient reported outcomes may provide more insight into the relationship between voice and endoscopic findings of laryngeal injury after prolonged intubation. The newly developed scoring system FLIS we used in this study is yet to be validated and we are planning to do this in a future study.

Conclusions

This study demonstrated a very high prevalence of laryngeal injury in the critically ill population following intubation and mechanical ventilation. Despite clinical improvement, a high prevalence of laryngeal injury and the presence of mild voice abnormality remained at 5 days post extubation. Simple bedside voice assessment was not able to identify the presence, type, or severity of laryngeal injury. Therefore, this study concludes there is limited utility for their use as a screening tool for laryngeal injury following prolonged endotracheal intubation. Routine referral for an assessment with flexible video laryngoscopy is recommended for further assessment in this patient group should the patient have persistent symptoms of dysphonia after 5 days post-extubation.

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at <https://www.theajo.com/article/view/10.21037/ajo-21-42/rc>

Data Sharing Statement: Available at <https://www.theajo.com/article/view/10.21037/ajo-21-42/dss>

Peer Review File: Available at <https://www.theajo.com/article/view/10.21037/ajo-21-42/prf>

Funding: Department Research Budget for equipment.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://www.theajo.com/article/view/10.21037/ajo-21-42/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Ethical approval was sought from the Human Research Ethics Committee at Southern Adelaide Local Health Network. Ethics approval code is OFR

208.17. Informed consent was obtained from patients prior to enrolment in the study.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Benjamin BF, Holinger LD. Laryngeal Complications of Endotracheal Intubation. *Ann Otol Rhinol Laryngol* 2008;117:2-20.
2. Brodsky MB, Levy MJ, Jedlanek E, et al. Laryngeal Injury and Upper Airway Symptoms After Oral Endotracheal Intubation With Mechanical Ventilation During Critical Care: A Systematic Review. *Crit Care Med* 2018;46:2010-7.
3. Van der Meer G, Ferreira Y, Loock JW. The S/Z ratio: a simple and reliable clinical method of evaluating laryngeal function in patients after intubation. *J Crit Care* 2010;25:489-92.
4. Gordin A, Chadha NK, Campisi P, et al. Effect of a novel anatomically shaped endotracheal tube on intubation-related injury. *Arch Otolaryngol Head Neck Surg* 2010;136:54-9.
5. Donnelly WH. Histopathology of endotracheal intubation. An autopsy study of 99 cases. *Arch Pathol* 1969;88:511-20.
6. Harrison GA, Tonkin JP. Prolonged (therapeutic) endotracheal intubation. *Br J Anaesth* 1968;40:241-9.
7. Santos PM, Afrassiabi A, Weymuller EA Jr. Risk factors associated with prolonged intubation and laryngeal injury. *Otolaryngol Head Neck Surg* 1994;111:453-9.
8. Bishop MJ, Weymuller EA Jr, Fink BR. Laryngeal effects of prolonged intubation. *Anesth Analg* 1984;63:335-42.
9. Hirano M. *Clinical examination of voice*. Wien New York: Wien New York: Springer-Verlag; 1981.
10. Kempster GB, Gerratt BR, Verdolini Abbott K, et al. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol* 2009;18:124-32.
11. Eckel FC, Boone DR. The S/Z ratio as an indicator of laryngeal pathology. *J Speech Hear Disord* 1981;46:147-9.
12. Kempster GB, Gerratt BR, Verdolini Abbott K, et al.

- Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol* 2009;18:124-32.
13. Antonaglia V, Vergolini A, Pascotto S, et al. Cuff-leak test predicts the severity of postextubation acute laryngeal lesions: a preliminary study. *Eur J Anaesthesiol* 2010;27:534-41.
 14. Colton House J, Noordzij JP, Murgia B, et al. Laryngeal injury from prolonged intubation: a prospective analysis of contributing factors. *Laryngoscope* 2011;121:596-600.
 15. Shinn JR, Kimura KS, Campbell BR, et al. Incidence and Outcomes of Acute Laryngeal Injury After Prolonged Mechanical Ventilation. *Crit Care Med* 2019;47:1699-706.
 16. Stachler RJ, Francis DO, Schwartz SR, et al. Clinical Practice Guideline: Hoarseness (Dysphonia) (Update). *Otolaryngol Head Neck Surg* 2018;158:S1-S42.
 17. Bishop MJ. Mechanisms of laryngotracheal injury following prolonged tracheal intubation. *Chest* 1989;96:185-6.
 18. Whited RE. A prospective study of laryngotracheal sequelae in long-term intubation. *Laryngoscope* 1984;94:367-77.
 19. Mencke T, Echternach M, Kleinschmidt S, et al. Laryngeal morbidity and quality of tracheal intubation: a randomized controlled trial. *Anesthesiology* 2003;98:1049-56.
 20. Nemr K, Simões-Zenari M, Cordeiro GF, et al. GRBAS and Cape-V scales: high reliability and consensus when applied at different times. *J Voice* 2012;26:812.e17-22.
 21. Chang J, Fang TJ, Yung K, et al. Clinical and histologic predictors of voice and disease outcome in patients with early glottic cancer. *Laryngoscope* 2012;122:2240-7.
 22. Rzepakowska A, Sielska-Badurek E, Osuch-Wójcikiewicz E, et al. Multiparametric Assessment of Voice Quality and Quality of Life in Patients Undergoing Microlaryngeal Surgery—Correlation Between Subjective and Objective Methods. *J Voice* 2018;32:257.e21-30.
 23. Colice GL, Stukel TA, Dain B. Laryngeal complications of prolonged intubation. *Chest* 1989;96:877-84.
 24. Dewan K, Vahabzadeh-Hagh A, Soofer D, et al. Neuromuscular compensation mechanisms in vocal fold paralysis and paresis. *Laryngoscope* 2017;127:1633-8.
 25. Soliman Z, Hosny SM, El-Anwar MW, et al. Laryngeal Compensation for Voice Production After CO2 Laser Cordectomy. *Clin Exp Otorhinolaryngol* 2015;8:402-8.
 26. Boateng GO, Neilands TB, Frongillo EA, et al. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health* 2018;6:149.
 27. Sauder C, Eadie T. Does the Accuracy of Medical Diagnoses Affect Novice Listeners' Auditory-Perceptual Judgments of Dysphonia Severity? *J Voice* 2020;34:197-207.
 28. Oates J. Auditory-perceptual evaluation of disordered voice quality: pros, cons and future directions. *Folia Phoniatr Logop* 2009;61:49-56.

doi: 10.21037/ajo-21-42

Cite this article as: Huang L, Athanasiadis T, Schar M, Woods C, Bassiouni A, Martin S, Bickford J, Bihari S, Ooi EH. Bedside voice assessments cannot be used as a screening test for laryngeal injury following prolonged intubation in an intensive care population. *Aust J Otolaryngol* 2022;5:9.

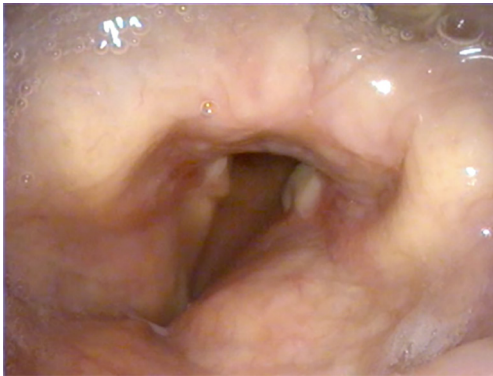


Figure S1 Glottis granuloma.

Table S1 Risk factors of laryngeal injury

Factors	P value
Gender	0.396
Age	0.244
Weight	0.364
Height	0.692
BMI	0.462
Smoking history	0.365
APACHE II	0.478
Admission specialty	0.657
Site of intubation	0.331
Height tube ratio	0.501
Use of muscle relaxant	0.831
Cormack Lehane score	0.234
Number of intubation attempts	0.327
Duration of intubation	0.99
Steroid use	0.065
NGT use	0.954
PPI use	0.290
Mean cuff pressure	0.571
Tube size	0.387

Statistical test used: one way ANOVA. BMI, body mass index; NGT, nasogastric tube; PPI, proton pump inhibitor.

Table S2 Correlation analysis of the voice assessments and laryngeal findings

		First assessment		Second assessment	
		P value	Regression co-efficient	P value	Regression co-efficient
Logistic regression					
S/Z >1.4 S/Z <1.4	FLIS =0, FLIS >1	0.167	0.195	0.354	0.156
Linear regression					
CAPE-V	FLIS =0, FLIS >1	0.945	0.428	0.416	4.039
GRBAS	FLIS =0, FLIS >1	0.232	-0.294	0.27	0.25
S/Z	PELLS	0.769	-0.025	0.619	0.019
S/Z	FLIS total score	0.013	0.076	0.364	0.014
CAPE-V overall severity	PELLS	0.335	1.894	0.09	3.031
CAPE-V overall severity	FLIS total score	0.136	1.162	0.211	0.942
GRBAS grade	PELLS	0.629	-0.038	0.067	0.150
GRBAS grade	FLIS total score	0.868	-0.005	0.118	0.053

FLIS, Flinders Laryngeal Injury Score; CAPE-V, Consensus Auditory Perceptual Evaluation-Voice; PELLs, Post-Extubation Laryngeal Lesions Scale.