Peer Review File

Article Information: https://dx.doi.org/10.21037/ajo-21-42

Reviewer A:

Comment 1: The supplementary appendices were not provided for review which limited my ability to comment on the validity of some of the results.
Reply 1: The two supplementary appendices were submitted with the original manuscript under AJO-21-42-OTHER1-3459 and AJO-21-42-OTHER1-8318. It is uncertain why they were not provided with the rest of the manuscript. However, I have attached them again with this reply and will hopefully be available for you to view.

Comment 2: This paper also describes a unique endoscopic evaluation tool designed within the authors centre (FLIS score) I feel the authors need to spend more time discussing the relevance of their scoring system and why they felt it necessary to design this scoring system.
Reply 2: The necessity of the scoring system is discussed and expanded in the Discussion, fifth paragraph (lines 297-303).

Comment 3: As the FLIS score is new and not validated is a secondary aim of this paper to assess its validity against the PELS evaluation tool?
Reply 3: The development of a scoring system requires assessment of its validity and reliability before being widely accepted for use. Given there are no currently validated scoring systems, a simple assessment using concurrent validity was performed against another unvalidated scoring system (PELLS). The validity and reliability of FLIS is beyond the scope of this study and will constitute another study in itself. This has been addressed in the limitations and future directions section (lines 314-315, 345-346).

Comment 4: Would like to see images illustrating the scoring system within the paper.
Reply 4: This is now addressed in line 196 and Supplementary appendix 1.

Comment 5: The lack of long term followup of this cohort I feel limits the usefulness of the results as well. How relevant is review at 5 days and does this correlate with long term outcomes?
Reply 5: We agree the data from long term follow up would be useful and interesting. It is an acknowledged limitation in this cohort study and now added in lines 337-338). After the conclusion of this study, ethical approval was provided to recall these patients 3-6 months following discharge. However, attempts at contacting these patients following discharge proved to be difficult, with very few consenting to represent to the hospital for repeat examination.

Comment 6: Line 40 - prevalence of ENDOSCOPIC laryngeal injury
Reply 6: ‘Endoscopic’ is added to line 40.

Comment 7: Line 42-44 - bedside voice assessment didn’t correlate but S/Z did?? S/Z is a bedside test
Reply 7: Thank you for pointing this out. The sentence regarding S/Z ratio correlation has been removed to simplify the results in the abstract as the correlation is only present when analysed as a continuous variable and not in the conventional category (using 1.4 as the cutoff).

Comment 8: bedside tests don't correlate? But stated S/Z does correlate to severity??
Reply 8: See reply 7. To not confuse the readers, the interpretation is that S/Z ratio does not correlate when used in the conventional way.

Comment 9: Line 49 - early assessment and treatment may be beneficial - does the article support this ? Little in discussion correlating to this statement.
Reply 9: The literature has documented significant healthcare costs and reduced quality of life in patients with intubation related laryngeal injuries (line 269-271). We agree that our study was not aimed specifically at early intervention. Hence, the speculation that it may be beneficial. We have removed this sentence from the abstract's conclusion.

Comment 10: I am unsure about the relevance of the COVID statements within the introduction. This study was designed and performed preCOVID and thus the inference that this paper was designed to look at alternative non aerosol generating airway assessments is somewhat misleading. This may have some validity in the discussion but would suggest the authors don't use COVID as a crutch.
Reply 10: We agree with the reviewers and thank them for bringing this to our attention. It was not meant to be as a crutch and was included because the pandemic was an issue towards the end of this study. This sentence has been removed from the introduction.

Comment 11: Line 74 - referencing a COVID paper (2), may not be the best reference in the introduction. It may be better within the discussion if the authors want to discuss the relevance within the COVID environment (and the laryngeal injuries associated with intubation within this cohort) There are better classic papers on laryngeal injury eg the Benjamin papers. Benjamin B. Prolonged intubation injuries of the larynx: endoscopic diagnosis, classification, and treatment. Ann Otol Rhinol Laryngol Suppl 1993;102(suppl 160):1-15.
And the updated version
Benjamin B, F.R.A.C.S., Holinger LD, M.D. Laryngeal Complications of Endotracheal Intubation. The Annals of Otology, Rhinology & Laryngology 2008 09;117(9):2-20.
Reply 11: The reference has been replaced by the one suggested above.

Comment 12: Line 78 - reference 4 seems an unusual choice to reference this statement - on a quick read it is not about laryngeal injuries associated with intubation.
Reply 12: Reference 4 is an error. It has been removed and replaced with the correct reference.

Comment 13: How was the patient assessed as able to give consent within the 48hour window?
Reply 13: Patients were assessed as they would if they were to undergo any procedure or treatment. They were determined to have the capacity (line 98-99) to understand that this is a research project and how they would participate in it. These patients have been recently sedated and needed time to regain their capacity, and this is another reason to wait for 24-48 hours before inviting them to participate (line 327-329).

Comment 14: Figure 1 states notify ICU if any airway concerns - this isn't mentioned again in results or discussion
Reply 14: We have added this to the methods and also results. Patients with stridor and airway concerns on examination (e.g severe laryngeal oedema) were immediately notified to

the ICU team (line 118-119). Even though patients had varying severity of laryngeal injuries, none were severe enough to cause airway obstruction or to warrant immediate intervention. (line 181-183).

Comment 15: acute physiology - Line 121-122 what does this mean and where is this data presented?
Reply 15:
Acute physiology and chronic health evaluation II (APACHE II) form an acronym and a score used in ICU that is calculated based on physiological measurements to represent the severity of illness in the ICU population. It is mentioned again under 'Patient characteristics', stating the captured population are a group of patients who are considered moderately unwell (APACHE II of 19) (line 176-177) and Table 3. It is also assessed as a risk factor of laryngeal injury (Supplementary appendix 2).

Comment 16: Line 125 - self extubated - this data is not presented within the results
Reply 16: No patient self-extubated. This is now added to line 179.

Comment 17: Line 127 - who performed FNE ?
Reply 17: FNE was conducted by the first author. This is now addressed in line 141-142.

Comment 18: What size scope and who performed the scope?
Reply 18: FNE was performed by the first author (L.H), using the Karl Storz Tele Pack x LED TP100 attached to a 3.7mm Strobo Video Rhino-Laryngoscope. This is now addressed in line 141-143.

Comment 19: Line 154 - the PELLS scores 0-5, it is not clear in this section that this is a 5 point score. The PELLS score may be better presented in a table?
Reply 19: The presentation of PELLS has now been updated to communicate the 5-categories within the text (lines 149-154). To present it in a table requires extra fee for copyright permission from the Publisher.

Comment 20: Line 185 is grammatically poor - It was noted that patients undergoing the second laryngeal assessment had a longer hospital admission (first assessment median duration 16.1 [IQR 8.4-29.9] days vs second assessment median duration 25.6 [IQR 11.5-45.4] days) but not statistically different (p=.09), and neither was the duration of endotracheal intubation (p=.19).
Reply 20: Thank you for bringing this to our attention. We agree with the reviewer that the sentence is confusion. We have now removed it.

Comment 21: Line 191-192 - how specifically was this determined?? - is this data available within the article for the reader to self assess?
Reply 21: Concurrent validity is determined using the Spearman's Rank Order Correlation. This is described and now further clarified in the 'Statistical Analysis section' under 'Methods' (line 166-167). The raw data for each patient has now been provided as Supplementary data 4. The median values for FLIS and PELLS are presented in Table 4.

Comment 22: How do they correlate? Pells 1 correlates to FLIS 1-5??. This needs more explanation.
Reply 22: FLIS and PELLS correlated well, producing a spearman rho coefficient of 0.745 (line 190-192). 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 clarified in line 190-192, 308-313).

Comment 23:
In the discussion this requires interpretation to explain why a more complicated scoring system (FLIS) is better than the less (PELLS) If they correlate - what is the benefit of the FLIS?.
Reply 23: This has been expanded and discussed in lines 297-303.

Comment 24: Also given FLIS and PELLS scores were reasonably low (PELLS 3 and FLIS < 14) do we know that worse endoscopic scores still correlate?
Reply 24: This is an acknowledged limitation (line 331-332) as no patients had FLIS scored 15 and above. The scores used for correlation are now included in lines 190-192

Comment 25: Line 194 - Given the lack of raw data provided to the reader it is impossible to check the authors results. The FLIS and PELLS scores are only given for the patients who completed both examinations. I would like to see more results provided. The patient with the highest FLIS score appears not to have undergone the second examination thus no long term results regarding the consequences of their supraglottic trauma. What was this patients PELLS score and how did this correlate? We also can't compare to their vocal results.
Reply 25: Further details about the patient with the highest FLIS has been provided in lines 199-201. Additionally, raw data is now provided in Supplementary results 4 to allow readers to evaluate the results in detail themselves.

Comment 26: Line 207 - the 25% with abnormal S/Z scores. How does this correlate to the endoscopic findings?
Reply 26: The abnormal S/Z scores did not demonstrate association when analysed categorically (using threshold of 1.4) to predict whether there is presence or absence of injury (lines 242-245). However, it did demonstrate significant correlation when analysed in a continuous manner ($p=0.013$) with FLIS at the first assessment (lines 247-249). Words of caution were provided following this (lines 249-251) as the S/Z >1.4 is the widely accepted cut-off value and should be interpreted in this context.

Comment 27: Line 219 - 221 - 3 patients with laryngeal granulomas required further otolaryngology review - I would have hoped these patients were not sent home or discharged from hospital without these being addressed - both could cause significant airway compromise. It may be useful to know these patients endoscopic and voice scores and if they correlated and whether these injuries were predicted from the first examination. Long term outcomes would also be helpful.
Reply 27: These three patients were reviewed by the ENT team while they were inpatients. None of the injuries resulted in symptomatic airway obstruction and were determined safe to be discharged. More information about these patients are provided in lines 223-228. The subglottic granulomas were only seen on flexible nasoendoscopy at the second assessment as the swelling of the supraglottis and glottis subsided and therefore not previously seen during the first assessment. The endoscopic findings and voice scores did not correlate (lines 251-253)

Comment 28: I was not provided with either supplementary appendices.

Reply 28: refer to Reply 1.

Comment 29: Particularly interested in the 3 patients who needed Ent followup. Were their S/Z scores > 1.4 either first or 2nd examination??
Reply 29: Refer to reply 27. Additionally, the details of the 3 patients information are as follows:

Injury	First S/Z	Second S/Z
Subglottic granuloma 1 (ID 195)	1.4	1.77
Subglottic granuloma 2 (ID 159)	1.2	0.94
Glottic granuloma 1 (ID 197)	1.19	0.93

Comment 30: Line 292 - subglottic scores aren't included in the the FLIS however the authors identified 2 patients with subglottic granulomas on the second followup - thus should it be? Or could be be included as a supplementary score when identified. Subglottic pathology holds significant risk for airway compromise post extubation
Reply 30: The subglottic findings are not included in the FLIS scoring system with the reasoning discussed in lines 303-305. The inclusion of it was thought to contribute to inaccurate total scores when comparing between the two assessments, especially as swelling subsides and the subglottis becomes more visible. However, when they were seen during FNE, they were referred appropriately. It is a good idea for it 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 (line 305-307).

Comment 31: Line 311 - should this therefore be mentioned in the methods? Ie patients successfully extubated
Reply 31: This has now been added to line 98.

Comment 32: Line 315 - no mention in the methods or results of previous voice issues being asked about.
Reply 32: This has now been added to line 103.

Comment 33: Another limitation is the significant proportion lost to followup for second evaluation. This may have biased the results. The lack of long term followup also limits conclusions regarding long term consequences of laryngeal trauma from intubation. Are the endoscopic findings observed relevant to long term outcomes?
Reply 33: We agree this is a limitation – this had now been added to line 336-337. This study addresses short term outcome of laryngeal injuries from intubation and is unable to comment on the long term outcomes. Also, refer to Response 5 regarding this topic.

Comment 34: 3/60 (5%) patients required further laryngeal intervention. How does this compare to the literature?
Reply 34: Upon re-reviewing the literature, there is no clear documentation of the prevalence of intubation related injuries that require intervention.

Comment 35: Given that endoscopic and voice reviews are not routine post extubation how does the author suggest these patients are identified and treated appropriately. I would like to

see an algorithm suggested by the authors with regards to how they suggest laryngeal injuries are identified in their patient cohorts moving forward,
Reply 35: An algorithm has been developed and incorporated into the discussion (lines 271-273) (Figure 3).

Comment 36: Do the authors feel the S/Z scores have utility?
Reply 36: Given the lack of correlation when S/Z ratio is used in its defined parameters (abnormal when greater than 1.4), we don't think S/Z ratio has a utility in this situation. This may be due to factors such as reduced respiratory capacity in the critically unwell population (now added to line 339-340).

Comment 37: Do the authors feel no review required until day 5?
Reply 37: The majority of patients improved despite the high prevalence of laryngeal injuries and therefore we believe sufficient time should be given to patients before subjecting them to additional examination and investigation. However, as the new algorithm suggests, if there are concerning signs of airway obstruction or worsening of laryngeal functions (such as dysphonia and swallow), review by Otolaryngology is warranted.

Comment 38: Why do the authors feel speech pathology review is helpful at day 5 when their results suggested that vocal scores couldn't predict the endoscopic findings? Perhaps the conclusion is FNE should be performed for persistent dysphonia rather than ent or ST review?
Reply 38: We believe the role of speech pathologists goes beyond voice assessment. They have valuable roles in the assessment of swallowing function, performing FEES and raising awareness for further Otolaryngology referral. However, given swallowing is not discussed in this study, we agree it is best to remove speech pathology in this setting and we have replaced it with the suggestion of performing flexible laryngoscopy for persistent dysphonia (line 354-356).

Reviewer B:
Comment 1: Future research could compare these high rates of laryngeal injury with international standards and the actual outcome of laryngeal injuries later than 5-7 days.
Reply 1: We agree, and this has been added to line 337-338.