A taxonomy of key performance errors for emergency intubation

Scott D. Weingart, MD,⁎, Ryan N. Barnicle, MD, Alexander Janke, Sabrina D. Bhagwan, Matthew Tanzi, MD, Peter J. McKenna, MD, Alexander Bracey, MD for the Resuscitationists Research Group

Study objective: Currently the videographic review of emergency intubations is an unstructured, qualitative process. We created a taxonomy of errors that impede the optimal procedural performance of emergency intubation.

Methods: This was a prospective, observational, study reviewing a convenience sample of deidentified laryngoscope recordings of emergency department intubations that were qualitatively flagged before the study as demonstrating suboptimal technique. These videos were coded for the presence of 13 predetermined performance errors. Our primary outcome was the incidence of each of these specified errors during emergency intubation. Errors fell into 3 categories: errors of structure recognition during laryngoscope insertion, errors of vallecula manipulation, and errors of device delivery.

Results: A total of 100 intubation attempts were reviewed. The most common error was inadequate lifting force with the blade tip in the vallecula which lowered the percent of glottic opening, occurring in 45% of the attempts. The least common performance error was the premature removal of the laryngoscope during bougie placement, occurring in only 9% of the videos.

Conclusion: We developed a taxonomy of 13 performance errors of laryngoscopy. Further study is warranted to determine how to best incorporate these into emergency airway training and the airway review process.

© 2023 Elsevier Inc. All rights reserved.

Keywords: Laryngoscopy Airway Resuscitation Critical care Video laryngoscopy Bougie

1. Introduction

Emergent endotracheal intubation is among the most common high-risk procedures in emergency medicine. Many departments have transitioned to video laryngoscopy to optimize first pass success (FPS), improve glottic views, decrease esophageal intubations, and reduce airway trauma [1-5]. In New York, a video device was present in 97.8% of emergency departments with established residencies and 84.3% of non-residency EDs [6]. Most of these devices can record the entirety of an intubation attempt for later review. In programs where this option is available, it presents an opportunity for airway quality improvement and teaching [7]. Unfortunately, many programs do not take advantage of these opportunities by either not recording or having no established system to review these recordings. The need for review and feedback on emergency ultrasound is a well-established component of EM quality improvement to allow optimization of technique and accuracy of diagnosis [8]. Arguably, providing a similar process for airway management is just as crucial to increase first pass success and avoid recurrent intubation errors.

If a program does create a path for continuous quality improvement (CQI) for ED airway management, currently there are no established parameters to describe performance errors. As an expert, the airway CQI lead can often state qualitatively that an intubation contained performance issues. However, this lack of a universal language for evaluating laryngoscopy attempts makes the feedback on less-than-optimal intubation attempts less generalizable and more difficult to communicate to practitioners.

We aimed to create a taxonomy of performance errors observed during the practice of ED intubation. The goals for this development are to provide a tool for airway quality assurance, a method of providing specific feedback to airway learners, and to encourage specific interventions and solutions to avoid these errors in the future. Hopefully, more standardized communication will allow improvement in objective clinical endpoints such as first pass success.

⁎ Corresponding author.
E-mail address: scottweingart@gmail.com (S.D. Weingart).

https://doi.org/10.1016/j.ajem.2023.08.035
0735-6757/© 2023 Elsevier Inc. All rights reserved.
2. Methods

2.1. Study design and setting

This study took place at a suburban Level 1 academic ED with over 110,000 annual patient encounters with approximately 1 ED intubation per day. ED policy is for all intubations to be performed with a video laryngoscope whether the intubator was utilizing the device for direct visualization or using the video screen. Use of a bougie device was standard operating procedure for all first-pass attempts. All intubations were recorded per departmental policy. Every recorded intubation was reviewed for quality by an expert in airway management with 20 years of experience, as part of a continuous quality improvement initiative [7]. The study was compliant with the STROBE guidelines for the performance of observational research [9].

2.2. Selection of participants

Any airway recordings meeting the following criteria were flagged: rapid-sequence-intubation in patients ≥18-years-old, with standard geometry blade [CMAC, Size 3 or 4, disposable], and at least one qualitative, perceived error in technique during routine quality assurance review. These flagged recordings were entered into the study protocol. Videos with multiple attempts were edited into separate recordings for each attempt, assigned a unique study number, and treated as separate laryngoscopy attempts. All videos were de-identified prior to review and the original videos were destroyed.

2.3. Intervention

A target of 100 videos was chosen as a representative sample before commencing the study. All flagged videos from February to November 2020 were included. 102 attempts containing performance errors were re-reviewed and these performance errors were separated into 13 distinct categories. One subject was later excluded after being identified as an awake intubation and another was found to be a duplication of a prior recording allowing the 100 videos included in final analysis.

2.4. Measurement and outcome

Each video was assessed for the presence of a set of predetermined performance errors (see https://emcrit.org/airway-error-taxonomy). Our primary outcome was the incidence of each of these specified errors during emergency intubation. The list of performance errors was created iteratively from the prior >1000 airway videos gathered between January 2017 to December 2019. These videos were reviewed by the airway quality assurance (QA) committee as part of the department’s airway quality project. The airway QA committee comprised 3 emergency physicians with additional training in airway management. Each time a new error was noted by a reviewer, it was taken to the 3-person airway QA and agreed on as an additional error in our QA process. These errors were then grouped into 3 categories by the airway QA committee.

Secondary outcomes included interrater agreement on the errors of the taxonomy, as well as correlations of the performance errors with each other, the number of errors with duration of the intubation attempt, and the number of errors with intubation success.

List of Performance Errors

Fig. 4 displays the performance errors based on the phase of laryngoscopy and tube delivery.

2.5. Errors of structure recognition during laryngoscope insertion

1. Insertion Off Midline Leading to Esophageal Visualization

This error was coded when the tip of the blade is inserted lateral of midline (most commonly to the right) with advancement initially resulting in visualization lateral to the glottic structures (e.g., piriform recess, aryepiglottic fold) with further advancement resulting in esophageal visualization.

2. Overly Deep Insertion Leading to Esophageal Visualization

This error was coded when a midline insertion was used, but lack of structure recognition or too rapid an insertion led to overrunning the glottic structures and visualizing esophagus.

3. Missed Anatomical Structure Recognition

This error was coded when the video reviewers were able to identify glottic structures during the laryngoscopy without visualization of esophagus, but the intubator did not recognize these structures as they did not correct tip position to the vallecula or glottic inlet, nor did they attempt tube insertion. This error is distinct from the prior two because the intubator did not approach the esophagus.

4. Inadequate Suction

This error was coded when tube insertion was attempted without adequate structure visualization due to secretions, hemorrhage, or debris and without adequate suctioning.

2.6. Errors of vallecula manipulation

5. Inadequate Lifting Force

This error was coded if the tip of the laryngoscope was properly seated in the vallecula indicated by translation of the epiglottis, but inadequate lifting force was applied to expose the glottis sufficiently for bougie or tube passage.

6. Failure to Engage Midline of Vallecula

This error was coded if the tip of the laryngoscope did not engage the median glossopiglottic fold resulting in inadequate epiglottic translation.

7. Lost Seating in Vallecula

This error was coded if during the intubation attempt, loss of vallecular engagement caused loss of glottic visualization.

8. Not Fully Seated in Vallecula

This error was coded if the blade tip was midline in the vallecula but was not advanced forward sufficiently to cause epiglottic translation.

9. Too Much Force in Vallecula

This error was coded if there was over-insertion into the vallecula causing downward movement of the epiglottis and preventing glottic visualization.

10. Bougie Delivery Issue

This error was coded if there was bougie hung up on anterior tracheal rings without correction or intubator was unable to manipulate bougie through vocal cords.

11. Over-Rotated Insertion (Kovacs Sign)

This error was coded if the operator was unable to pass device while the glottis filled >50% of the screen and cricoid cartilage visualized indicating over-rotation/too deep an insertion of laryngoscope.

12. Tube Delivery Issues

This error was coded if the endotracheal tube passage prevented by hangup on arytenoid cartilage without correction.
13. Premature Withdrawal of Camera

This error was coded if the laryngoscope was withdrawn from mouth prior to passage of the ETT over the bougie.

Each video was reviewed independently by two emergency physicians (EPs) using a standardized scoring sheet and data entry was performed using REDCap software. A third EP served as a tiebreaker when scoring disagreement occurred. Each laryngoscopy attempt was dichotomously scored as success or failure. The length of each laryngoscopy attempt was also captured—the duration of an attempt was defined as the time the tip of the laryngoscope blade passed the teeth to the time of passage of the endotracheal tube through the vocal cords or in the event of an unsuccessful attempt, removal of the laryngoscope tip past the teeth. This study was reviewed by the institutional review board and given exempt status.

2.8. Analysis

We calculated Cohen’s kappa and associated 95% confidence intervals to assess interrater reliability for the two video adjudicators across all videos and all categories of laryngoscopy/intubation performance error. Then, after utilizing the third review to adjudicate disagreements, we calculated Pearson correlation coefficients to assess for correlation among different performance errors. We plotted the number of mistakes against time to intubation, and modeled time to intubation in a simple, unadjusted linear model to test the hypothesis that the number of mistakes is related to prolonged intubation. We then modeled intubation success or failure in a simple, binary logistic regression model to test the hypothesis that an additional laryngoscopy/intubation performance error is associated with increased odds of intubation failure. While we report statistical significance at $p < 0.05$ in the manuscript otherwise, in the case of pairwise comparisons among error types (where we made 78 pairwise comparisons) we adjusted our cutoff for statistical significance based on the conservative Bonferroni correction factor ($0.05/78$ or $p < 0.0006$). Data were analyzed using the R software package.

3. Results

From the 100 videos included in final analysis gathered from January 2018 to June 2021, we scored 13 distinct performance errors. All videos contained more than one error. The most common error was inadequate lifting force with the blade tip in the vallecula which lowered the percent of glottic opening (POGO), occurring in 45% of the attempts. The least common performance error was the premature removal of the laryngoscope during bougie placement, occurring in only 9% of the videos. 98 of the 100 videos were performed by resident physicians, 1 by a physician assistant, and 1 by an ED attending physician.

82% of the reviewed attempts were successful. The average length of each attempt was 81 s, but this varied significantly from 35 s to 231 s. Although not included as a performance error, the epiglottis was overrun 40 times with the curved standard-geometry video laryngoscope (SGVL) blade akin to the traditional Miller straight blade usage. This was not considered a performance error as part of the airway teaching at this program is the option to use this technique deliberately (Mac as Miller), i.e., to overrun the epiglottis with a curved blade to allow better visualization.

3.1. Observed performance errors

The thirteen errors were divided into 3 categories. The incidence of each performance error is shown in Table 1.

<table>
<thead>
<tr>
<th>Performance Error</th>
<th># of Occurrences</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inadequate Suctioning</td>
<td>35</td>
<td>0.913</td>
</tr>
<tr>
<td>Insertion Off Midline Leading to Esophageal Visualization</td>
<td>30</td>
<td>0.884</td>
</tr>
<tr>
<td>Overly Deep Insertion Leading to Esophageal Visualization</td>
<td>25</td>
<td>0.921</td>
</tr>
<tr>
<td>Missed Anatomical Structure Recognition</td>
<td>17</td>
<td>0.809</td>
</tr>
<tr>
<td>Inadequate Lifting Force</td>
<td>45</td>
<td>0.839</td>
</tr>
<tr>
<td>Failure to Engage Midline of Valleculla</td>
<td>40</td>
<td>0.937</td>
</tr>
<tr>
<td>Lost Seating in Valleculla</td>
<td>31</td>
<td>0.907</td>
</tr>
<tr>
<td>Not Fully Seated in Valleculla</td>
<td>26</td>
<td>0.869</td>
</tr>
<tr>
<td>Too Much Force in Valleculla</td>
<td>11</td>
<td>0.951</td>
</tr>
<tr>
<td>Bougie Delivery Issues</td>
<td>39</td>
<td>0.958</td>
</tr>
<tr>
<td>Over-Rotated Insertion (Kovacs Sign)</td>
<td>25</td>
<td>0.896</td>
</tr>
<tr>
<td>Tube Delivery Issues</td>
<td>24</td>
<td>0.972</td>
</tr>
<tr>
<td>Premature Withdrawal of Camera</td>
<td>9</td>
<td>1.000</td>
</tr>
</tbody>
</table>

3.2. Errors of structure recognition during laryngoscope insertion

In this series of videos, the mutually exclusive errors: insertion off midline and overly deep insertion occurred frequently, 30% and 25% of the time respectively. Both errors can result in the performance error of no anatomical structure recognition leading to esophageal visualization and may result in a failed intubation attempt. Of all the performance error pairs, the two strongest direct correlations were overly deep insertion:no anatomical structure recognition ($r = 0.35, p = 0.003$, not statistically significant) and insertion off midline:no anatomical structure recognition ($r = 0.34, p = 0.005$, not statistically significant).

3.3. Errors of vallecula manipulation

We observed 5 distinct performance errors related to vallecula manipulation. The most common overall error was failure to provide adequate lifting force for the epiglottis even when properly seated in the vallecula preventing the passage of the airway device, occurring in 45 of 100 videos.

Failure to engage the median glossoepiglottic fold at the midline was the second most common performance error occurring in 40% of the recorded attempts.

Maintaining vallecula engagement until an airway device is passed is also a necessary component to intubation. Engagement was lost in 31% of our videos. Related distinct performance errors that were identified include inadequate insertion into the vallecula preventing translation of the epiglottis and obscuration of the glottic opening and too much force in vallecula causing an iatrogenic Cormack-Lehane Grade IIIb view [10].

3.4. Errors of device delivery

The last group of performance errors relate to the passage of a device through the vocal cords: either bougie or endotracheal tube.

Bougie delivery issues occurred in 39% of the videos. The most common issue was inadequate bougie control despite adequate glottic visualization. The other significant bougie delivery issue was the coudé tip of the bougie becoming stuck on the anterior tracheal rings without maneuvers to correct the issue.

A common error seen in hyperangulated video laryngoscopy is over-rotation causing the glottis to fill a majority of the screen and creating an angle for tube insertion that is difficult or impossible—the so called
Kovacs Sign [11]. We observed a similar situation when CMAC Size 4 disposable standard geometry blades [Storz, Germany] were used.

Improper passage of the endotracheal tube was also the cause of performance errors and occurred 24 times in this study. In each case, despite the tip of the tube being placed at the glottic opening with or without a bougie, passage through the cords was delayed because the tube was stuck on the arytenoid cartilage without appropriate maneuvers to relieve this obstruction.

The least common performance error was premature removal of the camera after successful placement of a bougie, but before the endotracheal tube had been passed through the cords preventing successful placement of the endotracheal tube. This occurred nine times in our study.

There was a high inter-rater reliability between the two initial reviewers. There were only 53 disagreements that required tie breaking. Overall, the kappa was 0.901 (95% CI, 0.882–0.911). There was 100% agreement regarding pulling out the camera too early and not visualizing the endotracheal tube passing through the vocal cords, which is a very objective and distinct error. The next strongest agreement occurred regarding failure to engage the midline in the vallecula at the median glossoepiglottic fold which had a kappa of 0.951 (95% CI, 0.855–1.00). The performance error with the weakest agreement was for lack of anatomical structure recognition with a kappa of 0.809 (95% CI, 0.648–0.971).

We performed an exploratory analysis on the correlation of performance error with each other. Fig. 1 shows the correlation between each of the performance errors. The strongest positive correlation noted (0.35) was between overly deep insertion and lack of anatomical visualization. The strongest negative correlation (−0.30) was between lack of anatomical visualization and lost seating in the vallecula, suggesting these are unrelated errors. After correcting the cutoff for statistical significance for the multiple comparisons involved, none of these pairwise correlations in error were statistically significant.

We noted a positive correlation trend between the number of performance errors committed and the overall length of the intubation attempt (Fig. 2); however, the number of errors is not significantly associated with intubation success (Fig. 3). In an unadjusted logistic regression model for the outcome of intubation success, an additional performance error identified by video was associated with a nonsignificant 0.97 lower odds of success (95% confidence interval 0.76 to 1.27).

4. Limitations

There are numerous limitations to our study. First and foremost, these intubations were performed at a single-center, tertiary-care, university hospital. This ED has a comprehensive airway quality initiative, including pre-intubation checklists, mandatory video laryngoscopy, nearly exclusive bougie-first technique, required recording that allows for quality assurance and improvement, and monthly airway video review during resident didactics. This created a culture of high reliability that the authors recognize is not standard for EDs and the study findings may not be generalizable to all departments with trainees that have different practices. The FPS success rate at this institution was 98.1% and
the 100 videos included were screened from over 1300 intubations in which no qualitative performance errors were observed. In many departments, the actual incidence of each error may significantly differ depending on the characteristics of airway teaching and quality improvement.

Further, extensive preoxygenation optimization was used in all intubations including apneic oxygenation and PEEP—which allowed for a longer time for each intubation attempt [12]. In programs with different pre-intubation optimization, the error frequency may change due to the need for earlier termination of each attempt.

We recognize that there is a subjective nature to the initial video selection. The clips included in this study were specifically screened and selected because each included what was perceived to be one or more findings that deviated from acceptable (not optimal) performance. Optimal or near-optimal intubations, though far more common at our program, were specifically excluded from this study. This obviously raises the incidence of the performance errors discussed in this study and readers may inappropriately infer that they are this common with routine intubations. This is not a study of performance error prevalence. This study was intended to be descriptive to develop a common language surrounding airway performance and, therefore, the inclusion of the attempts deemed good practice would not benefit this endeavor. However, the results of this study may generate the framework for future quantitative research on performance errors in video laryngoscopy.

Additionally, since these videos were intentionally deidentified and disassociated from formal patient records, identifying the level of experience for each operator (program year and number of prior intubations) was not possible for this study. It is impossible to tell if these opportunities for optimization were more common for junior residents compared to senior residents and fellows or attendings.

Further, we do not know if the video laryngoscopy was used as a direct visualization device or if the operator was using the video screen for any given attempt. It is possible that performance errors may differ predicate on which method was chosen.

Additionally, all the intubations were performed using CMAC video disposable laryngoscopes [Storz, Germany], which may have mechanical differences compared with to other video laryngoscope manufacturers. While this may result in different frequencies of performance errors, the nature of these errors is unlikely to be different when using any manufacturer’s standard geometry blades.

Finally, many departments use hyperangulated video laryngoscopy (HAVL) for all attempts, while this taxonomy dealt exclusively with standard geometry blades. We believe that a study of hyperangulated blades will reveal the same performance errors, but that study must be performed to verify this taxonomy.

5. Discussion

Our review of 100 SGLV attempts serves as the foundation of describing and teaching the most common pitfalls when using this technique for endotracheal intubation. Prior to this investigation, there was no unified description of these common performance errors and,
therefore, no common approach to teaching those performing SGVL for endotracheal intubation on how to avoid these missteps. The authors have analyzed suboptimal intubation attempts in order to provide a comprehensive taxonomy of imperfect technique laryngoscopic errors.

Other groups have utilized and researched the use of VL recordings to improve airway management. Miller et al. use video recordings to find the association of technical factors of laryngoscopy to FPS in a pediatric ED [13]. They identified ideal tip placement in the vallecula, and the ability to obtain a Grade 1 or 2a Cormack-Lehane view to be associated with FPS. Miller et al. also identified the benefits of multiple participants being able to see the intubation on the video screen during emergency airway management [14].

Some of the errors described have been studied in the emergency airway literature. Driver et al. found that engaging the median glossoepiglottic fold with a SGVL blade tip during orotracheal intubation was associated with improved laryngeal visualization, with 87% glottic opening with the midline vallecular fold engaged vs 78% when not engaged [15]. Furthermore, a modified Cormack-Lehane grade 1 or 2a was obtained 96% of the time with engagement vs 87% when there was no engagement. Driver concluded that this anatomical structure should be the target of a video laryngoscopy blade tip positioning. Similar findings were reported by Miller et al. for pediatric emergency intubations in a study noting technical factors associated with increased FPS included placement of VL blade tip into the vallecula regardless of blade type and sufficient glottic visualization (grade 1 or 2a) on the VL screen [13].

We feel that the performance errors identified in this study may be beneficial to the airway management curriculum of EM trainees. Educators should be familiar with these specific performance errors so that each can be mitigated both prior to, during and after supervised intubations. The EM community at large may benefit from a universal classification of these pitfalls as well and provide a guide to both educators and quality assurance studies in the future. Anticipating, planning for, and responding to these technical pitfalls can theoretically improve FPS, increase speed of intubation, avoid peri-intubation complications, and improve patient-centered outcomes, but that remains to be proven. Further study is required to accurately quantify the incidence of these performance errors in more generalized settings, and to evaluate whether incorporation of airway review highlighting these performance errors ultimately improves patient-oriented intubation metrics. Another future direction will be to develop and describe the microskills necessary to overcome or prevent the commitment of these performance errors. An EM training program can also develop a video library for resident education within the limits of HIPAA compliance. We foresee this taxonomy system being applicable both to didactic review at residence conference and individual operator-specific feedback as part of quality assurance initiatives. Validation of the range and objective assessment of these errors at multiple centers with varied intubation skill levels would bolster the usefulness of this taxonomy.

In summary, after reviewing imperfect VL attempts we developed a taxonomy of 13 performance errors of laryngoscopy. These performance errors fall into three main categories: errors of structure recognition...
Fig. 4. Performance errors categorized by phase of intubation procedure.
during laryngoscope insertion, errors of vallecula manipulation, and errors of device delivery. We feel that such a system may be useful when educating EM personnel on the optimal means of performing ETI, rather than the more common dichotomy of success versus failure.

Note

A vodcast accompanies this paper with video examples of the taxonomy as well as a more in-depth discussion of each error and its mitigation. It can be found at [emcrit.org/airway-error-taxonomy].

Financial support

None.

Contributions

SDW conceived the study and designed the trial. SDW, RNB, & AB supervised the conduct of the trial and data collection. MT & PJM helped with data collection. AJ provided statistical advice on study design and analyzed the data. RB drafted the manuscript, and all authors contributed substantially to its revision. SDB reviewed the paper and provided cross-specialty review. SDW takes responsibility for the paper as a whole.

CRediT authorship contribution statement


Declaration of Competing Interest

The authors have no conflicts of interest or personal funding to report.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajem.2023.08.035.

References


144