Title
Computer-generated laryngoscopy profiles to assess competence in airway management.

Author
Moitoza, Jennifer

Publication Date
2012

Jennifer Moitoza

COMPUTER-GENERATED LARYNGOSCOPY PROFILES TO ASSESS COMPETENCE IN AIRWAY MANAGEMENT

ABSTRACT

***Background**:* A method to discern a trainee's expertise with direct laryngoscopy would be useful for following progress and determining readiness for call or advanced assignments. We have developed methodology to record laryngoscopy motion and performance. The goal of this study was to determine whether laryngoscopy skill could be evaluated by electronic assessment of technique.

***Methods**:* Three anesthesia faculty with three to 23 years experience performed laryngoscopy five times each on a Medical Plastics Intubation Mannequin. Blade path and force were measured with a Mini-Bird magnetic position sensor (Ascension) and a 6-axis force transducer (ATI), respectively, attached to the laryngoscope handle and recorded on a Dell laptop computer. A Matlab (Mathworks) program digitally aligned the separate laryngoscopy trajectories and a virtual curvilinear tube was calculated within the mannequin airway that encompassed all the expert trajectories.

Twelve residents were studied on the first day of their anesthesia residency. They performed laryngoscopy three times on the same mannequin with the instrumented handle. The percentage of the trajectory that fell within the expert tube was measured for each laryngoscopy. Residents were divided into high and low experience groups, high being > 20 previous laryngoscopy attempts in an airway model, and into upper and lower 50^{th} percentiles based on laryngoscopy success in their first 10 patient attempts. Trajectory percentages, force, torque, and jerks were compared between groups by T-test.

***Results**:* A trajectory from one expert generally fell within the tube derived from the other two experts. Resident laryngoscopies fell within the expert tube an average 74% ± 6% (mean ± SE) of the path length ($P < 0.001$ vs. 100%). Conformity to the expert path was greater in the upper third of the airway (93% ± 3%) than near the larynx (61% ± 10%, $P < 0.01$). Conformity did not appear to differ among residents as a function of previous experience or patient laryngoscopy success. The average number of jerks for

expert laryngoscopy trajectories was 16.5 ± 1.8 jerks per laryngoscopy, while residents averaged 45.8 ± 11.8 (P = 0.01). Laryngoscopy duration averaged 3.6 ± 0.3 seconds for experts and 8.0 ± 0.6 seconds for residents (P = 0.01). There was a trend towards significance for expert vs. resident trajectory path lengths, and laryngoscopy force and torque did not differ among the groups. None of the motion parameters differed among the residents as a function of experience or success.

***Discussion**:* Digital analysis of laryngoscopy trajectories with engineering technology is a novel and objective way to measure skill. Beginners and experts can be differentiated with the methods described here. Cross sectional and longitudinal studies will be necessary to determine whether assessment of technique can distinguish finer differences in skill as trainees progress in developing expertise. Supported by APSF.

INTRODUCTION

Laryngoscopy is an important procedure carried out in every hospital. It allows a patient's airway to be visualized and controlled by endotracheal intubation during surgery or life-threatening situations. It is a highly critical skill, but there have been few advances in the teaching and evaluation of this essential procedure. Training programs for endotracheal intubation for healthcare personnel, such as residents and paramedics, are not standardized and in many cases may be inadequate. There is also little information regarding how much training is required to become competent in airway management[1].

The current methods of laryngoscopy training can include introduction to the procedure on plastic mannequins, either partial task trainers or patient simulators, followed by practice on real patients. Konrad et. al. determined that a 90% success rate for endotracheal intubation could be achieved after an average of 57 attempts on patients[2]. This volume of patient contact is impractical in most specialties and fields other than anesthesiology, and improvements in teaching methods must be directed at training on mannequins. Simulation-based methods could provide trainees with safe, unlimited training opportunities, but they currently allow only limited feedback beyond "success" or "failure." Incorporating the key aspects of a successful intubation, such as proper head positioning, laryngoscope blade positioning and lift, and a timely performance[1], into training feedback may improve the quality of laryngoscopy training.

We have developed an electronic method of evaluating trainee competence in hopes of addressing these issues and advancing the teaching and evaluation of laryngoscopy. This is a relatively new concept, especially in the field of anesthesia. Skill simulators have been used frequently to assess the technical skills of surgical residents before they operate unsupervised[3]. Haptic feedback used during robot-assisted endoscopic surgery has been observed to improve task completion time, accuracy, and number of errors made[5]. Similar methods could be applied in laryngoscopy training.

Our strategy was to outfit a laryngoscope handle with a magnetic position sensor and a force transducer. This device could electronically track laryngoscopy trajectories within a mannequin's airway. Engineering collaborators at UCSD have developed a method with the instrumented handle to map the zone for laryngoscopy trajectories by experts. The map is a curved tube coursing through the mannequin airway from mouth to larynx. It is a benchmark for expert performance. Trajectories of expert anesthetists are expected to be within this tube close to 100% of the time whereas trajectories of novices might stray outside the expert space. Having this gold standard for intubation creates a novel and objective way of measuring skill.

The handle records other data also related to laryngoscopy expertise, including force, torque, time, path length, and jerks. The combination of this information with the trajectory data could be used to track the progress of medical personnel as they proceed in airway management training. Furthermore, the comparison with experts could be used to assess the expertise of medical personnel and determine if they are competent in airway management. As an initial step in developing these assessment instruments, the goal of this study was to compare the trajectories of laryngoscopy novices with the expert laryngoscopy tube and to determine whether other motion parameters differed significantly between novices and experts. The novices were first year anesthesiology residents at the start of their training, while the experts were anesthesiology faculty with many years of experience.

METHODS

Subjects. The subjects involved in this study consisted of three anesthesiology professors with 3-23 years of experience and twelve first-year anesthesiology residents with minimal experience. After obtaining IRB approval, each subject provided written informed consent to participate in the study. Additionally, the residents filled out a survey to assess their prior experience with direct laryngoscopy on mannequins and/or patients. Residents were divided into three groups with little prior experience (6-21 model or patient intubations), intermediate experience (21-60 intubations) and greatest experience (>60 intubations).

Assessment of resident expertise. We measured residents' success on patient intubations as an independent method of classifying their expertise. This part of the study was also approved by the IRB and the patients gave written informed consent. Residents were observed performing intubations in the operating room during their initial month of anesthesia residency, generally on 6-10 patients. Patient data related to anticipated difficulty for laryngoscopy (Mallampati score, hyomental distance, and mouth opening) were recorded prior to intubation. A record was made on success or failure of each intubation attempt (determined by the attending faculty anesthesiologist) and the number of intubation attempts on each patient. The residents were divided into two groups representing the upper and bottom 50th percentile in terms of overall patient intubation success or success on the first intubation attempt.

Experimental protocol. All subjects used a Macintosh 3 blade to perform direct laryngoscopy on a Medical Plastics Airway model (Mass Group, Inc, Miami, FL). The laryngoscope handle was equipped with a miniBIRD Model 800 magnetic position sensor (Ascension Technology Corp, Burlington, VT; http://www.ascension-tech.com/) and a 6-axis (force plus torque in 3D) force transducer (ATI Industrial

Automation, Apex, NC). In this study, a direct laryngoscopy procedure was defined as starting at the mouth of the mannequin and ending once the subject declared a view of the glottis and vocal cords. The position trajectories of each direct laryngoscopy procedure were processed with a MATLAB program (MathWorks, Natick, MA; http://www.mathowrks.com/) written by the UCSD engineering group and running on a 2.5-GHz Dell Latitude D630 computer. Data were also recorded for maximum force, torque, and total time of the procedure. The position of the mannequin was calibrated before each intubation attempt. Anesthesiology faculty performed direct laryngoscopy 5 times on the mannequin, while anesthesiology residents performed direct laryngoscopy 3 times.

Estimates of expertise from the motion analysis. Using the MATLAB program mentioned above, we digitally aligned faculty laryngoscopy trajectories to create a three-dimensional (3D) curvilinear tube that followed the mannequin's airway (see **Figure 1**). The tube represented the possible space traversed by the laryngoscope blade when experts performed laryngoscopy. We aligned resident laryngoscopy trajectories with this tube and then calculated the percentage that each resident trajectory fell within it. An additional motion parameter, number of jerks, was derived from the laryngoscopy trajectories. Any change in the direction of the laryngoscope blade tip was noted and called a "jerk" (see **Figure 2**). Expert benchmarks for this and the other motion parameters were established from the faculty data.

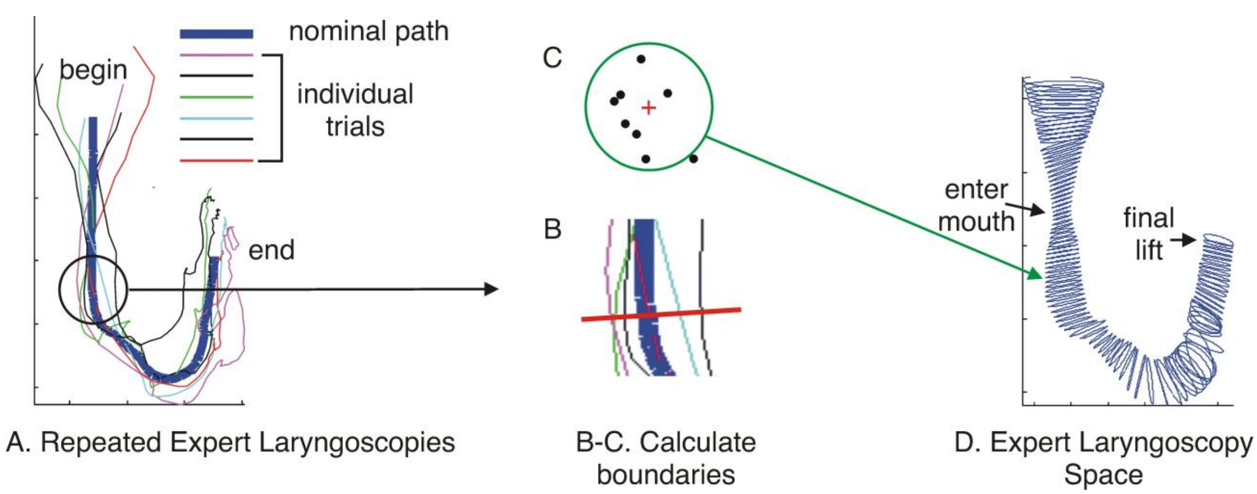

FIGURE 1. **A.** Numerous expert trajectories were aligned together and averaged to create a nominal path. **B-C.** Calculation of the boundaries of the expert space was based on an average radius around the nominal path. **D.** The 2D representation of the space occupied by expert laryngoscopy trajectories.

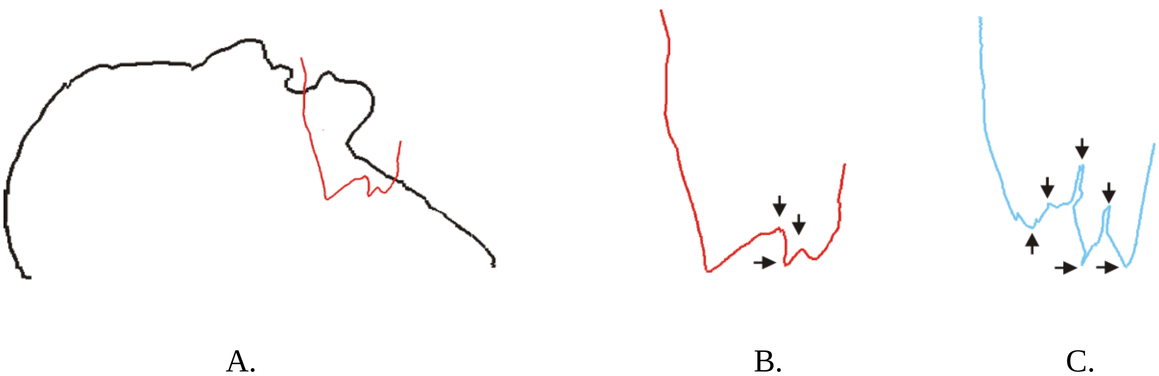

FIGURE 2. **A.** Head profile with an overlaid laryngoscopy trajectory. **B.** Laryngoscopy trajectory with a low number of jerks, as denoted by black arrows. **C.** Laryngoscopy trajectory with a higher number of jerks, as denoted by black arrows.

Data analysis. Motion parameters did not vary systematically from the first laryngoscopy attempt to the last. Therefore, data for the replicate attempts by each subject were averaged and used in the subsequent analyses. Data are presented as mean ± standard error (SE). Resident trajectory percentages and the other motion parameters were compared to expert benchmarks using Student's T-tests. Data were compared among resident groups divided by experience or patient success by analysis of variance. Significance was accepted for $P < 0.05$.

RESULTS

Laryngoscopy Characteristics for Residents vs. Experts

Conformity to Expert Tube. **Figure 3** demonstrates that a trajectory from one expert generally fell within the tube derived from the other two experts. **Figure 4** shows the shape of the expert laryngoscopy tube with superimposed resident trajectories. The expert trajectory looks like an upside-down letter 'r': the initial downward trajectory corresponds to the blade moving through the oropharynx toward the back of the throat. The downard movement reaches a turning point where the trajectory begins to curve upwards and laterally to the right. This portion corresponds to rounding the base of the tongue, reaching the vallecula and finally lifting the jaw to expose the larynx to view. Resident trajectories are superimposed in blue on the expert tube in black. The examples in this figure show resident trajectories that followed the expert tube with varying degrees of success with greatest success on the right, least on the left. Jerks are visible in the resident trajectories, particularly in the left-most example. The residents stray outside the expert path in all dimensions and in these examples push beyond the expert stopping point. The residents do move through the expert tube for much of the procedure, however, but the fraction of their paths within the tube is significantly less than 100%. On average, resident laryngoscopies fell within the expert tube over 74% ± 6% of the path length ($p < 0.001$ vs. 100%). To investigate whether the residents strayed from the expert path throughout the trajectory or only at specific points, we separated the expert tube into thirds representing different portions of the airway. The first third represents the mouth, the second third represents the oropharynx, and the final third represents the laryngopharynx. Residents conformed 93% ± 3% to the first third of expert tube, 74% ± 7% to the second third, and 61% ± 10% to the final third, with the difference between the first third and the final third being statistically significant ($p = 0.004$, see **Figure 5**).

QuickTime™ and a
decompressor
are needed to see this picture.

FIGURE 3. This figure shows an expert tube made from the averages of 2 experts (pink and black circles). The green trajectory belongs to the 3rd expert. We observed that a trajectory from one expert generally fell within the tube derived from the other two experts, suggesting that the experts were overall consistent with each other.

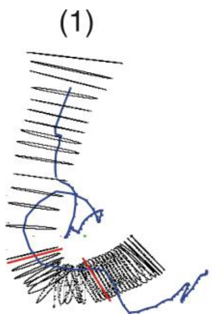

FIGURE 4. This figure shows the expert tube (black circles) containing the trajectories of three different residents in blue with varying degrees of percent conformity to the tube. (1) shows a resident who strays considerably from the tube whereas (3) shows a resident who is almost entirely within the tube. This figure also demonstrates how we separated the tube into thirds, as denoted by red lines. It is evident that resident trajectories stray from the expert tube in a number of different areas and contain various numbers of jerks.

FIGURE 5. This graph shows the percent conformity of resident trajectories to the whole expert tube and its subsequent thirds. Percent conformity decreases as residents move down the length of the expert tube: percent conformity is highest at the upper third of the tube and lowest at the last third of the tube. ** denotes statistical significance with $p < 0.05$.

Motion Parameters. **Table 1** compares the data for laryngoscopy force and motion parameters between experts and first-year residents. The study groups differed significantly in number of jerks and time. Experts had an average of 16.5 ± 1.8 jerks per laryngoscopy, while residents averaged 45.8 ± 11.8 ($p = 0.01$, **Figure 6**). Laryngoscopy duration averaged 3.6 ± 0.3 seconds for experts and 8.0 ± 0.6 seconds for residents ($p = 0.01$, **Figure 7**). The distance traveled by the laryngoscope during each procedure (i.e. path length) was greater for residents than for residents, but the effect did not quite reach significance with $p = 0.058$. Maximum force and torque did not vary between the groups.

Table 1 Motion and force parameters for residents vs. experts.

	Residents (Avg ± SE) (N = 12)	Experts (Avg ± SE) (N = 3)	p value
Maximum Force (N)	67.3 ± 3.6	66.8 ± 2.6	0.97
Torque (N-M)	3.0 ± 0.2	3.5 ± 0.2	0.47
Path Length (cm)	204 ± 8	155 ± 4	0.058
*Number of Jerks	45.8 ± 11.8	16.5 ± 1.8	0.01*
*Time (sec)	8.0 ± 0.6	3.6 ± 0.3	0.01*

* denotes a statistically significant difference when $p < 0.05$.

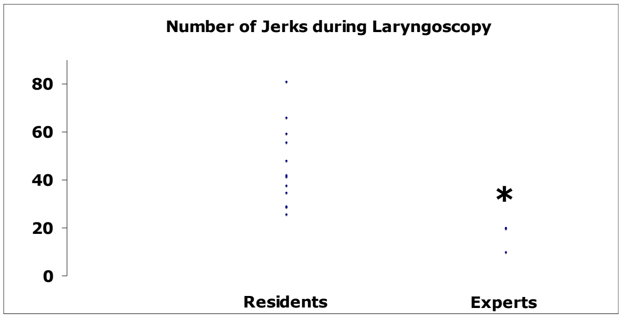

FIGURE 6. This graph shows the average number of jerks for all twelve residents and the three experts. There appears to be a clear delineation between resident averages and expert averages, with resident averages being greater than expert averages. *P < 0.01 residents vs. experts.

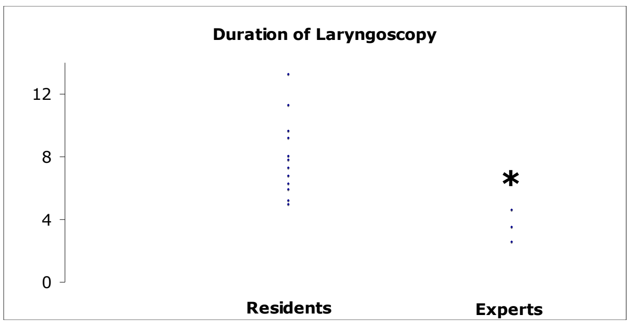

FIGURE 7. This graph shows the average duration of a laryngoscopy for all twelve residents and the three experts. As with number of jerks, there appears to be a clear delineation between resident averages and expert averages, with resident averages being greater than expert averages. *P < 0.01 residents vs. experts.

Motion Analysis vs. Minor Variation in Resident Experience

Based on the initial survey data, all of the twelve first-year anesthesia residents had some degree of experience with laryngoscopy on mannequins and intubation in patients. Six residents had made 6-20 intubation attempts in the past, 4 had made 21-60 attempts, and 2 had >120 past attempts. One could surmise from their varying degrees of experience that they would possess varying degrees of expertise. We divided the residents into three different "experience groups," corresponding to the categories just listed in order to determine if our methods could distinguish between intermediate degrees of expertise.

Conformity to the whole expert tube did not differentiate between the most experienced and the least experienced residents: 71% ± 7%, 75% ± 11%, 74% ± 10%, respectively (p = 0.85 for most experienced vs. least experienced). However, there was a significant difference when we looked at just the last third of the expert tube: 71% ± 17%, 61% ± 17%, 58% ± 16%, respectively (p = 0.004 for most experienced vs. least experienced). The other motion parameters did not differ between the groups (**Table 2**).

Table 2 Motion and force parameters for residents according to experience level.

	Most Experience: >120 prior intubations (Avg ± SE) (N = 2)	Middle Experience: 21-60 prior intubations (Avg ± SE) (N = 4)	Least Experience: 6-20 prior intubations (Avg ± SE) (N = 6)	p value (most vs. least experience)
Maximum Force (N)	58 ± 1.4	64 ± 2.5	72 ± 3	0.19
Torque (N-M)	2.7 ± 0.2	3.0 ± 0.1	3.1 ± 0.4	0.47
Path Length (cm)	204 ± 15	238 ± 36	203 ± 12	0.99
Number of Jerks	47 ± 11.3	45.2 ± 9.2	45.7 ± 4.9	0.90
Time (sec)	8.7 ± 1.1	8.0 ± 1.3	7.7 ± 0.7	0.50

Success Groups

Residents' expertise could also be classified by the measurement of their success over a short series of patient laryngoscopies in the period immediately following the motion testing. Intubation success rates for the first month (also corresponding to the residents first month of training in anesthesiology) ranged from 36% to 100% with an average of 82% ± 5%. Intubation success on the first attempt, a more stringent measure of expertise, ranged from 27% to 90% and averaged 71% ± 5%. Although lower success could indicate an individual with less skill than others, a resident with average or superior expertise might have worse results if asked to intubate more difficult patients. Specific anatomic factors, such as oropharyngeal view, thyromental distance and mouth opening[5], may be markers for patients who are difficult to intubate. We investigated the frequency of those markers in patients for the most successful and the least successful residents to test the possibility that success in our group of residents varied because of differences in laryngoscopy difficulty. In fact, the residents who performed below the median for intubation success on the average had patients with more favorable characteristics: better oropharyngeal views, thyromental distance, and/or greater mouth opening (**Table 3**). This finding argues against patient factors being a cause of differences in success.

Table 3. Frequency of factors associated with ease of laryngoscopy in patients intubated by the most successful residents and by the least successful residents.

Success Group	OPV ≥ 2 % of pts ± SE	P	TMD ≥ 5 cm % of pts ± SE	P	MO ≥ 5 cm % of pts ± SE	P

Top 50th Percentile	77.7 ± 6.5	0.21	91.7 ± 5.4	0.16	71.8 ± 9.8	0.17
Bottom 50th Percentile	89.7 ± 6.3		100		88.8 ± 6.1	

OPV = oropharyngeal view. Values of 1 or 2 are associated with easy laryngoscopy.
TMD = thyromental distance. Values ≥ 5 cm signify easy laryngoscopy.
MO = mouth opening. Laryngoscopy ease depends on values ≥ 5 cm.
P values refer to the comparison between the top and bottom 50th percentile groups.

We next compared conformity to the expert tube and motion parameters between the most successful and least successful resident groups. Conformity to the whole expert tube did not differentiate between the two success groups. The most successful residents had trajectories that fell within the tube an average of 74% ± 6% and the least successful residents fell within the tube an average of 75% ± 6% (p = 0.95). Looking at the averages for conformity to the last third of the expert tube did not help to distinguish the two study groups either: 62.2% ± 9% and 59% ± 11% (p = 0.78). **Tables 4 and 5** summarize the comparison for the motion and force parameters. In general, there were no detectable differences between the overall success groups for any of the parameters, except for path length. However, this difference was a paradoxical result in the opposite direction than expected, and the difference was small, only about 10% (see **Table 4**). When looking at only the success on the 1st intubation attempt, there were significant differences between the groups for maximum force and time, but these, too, were paradoxical results and occurred in the opposite direction than expected. There were no detectable differences for torque, number of jerks, or path length when looking at success on the 1st intubation attempt (see **Table 5**).

Table 4 Motion and force parameters for residents according to overall intubation success.

	Most Successful Residents (Avg ± SE) (N = 6)	Least Successful Residents (Avg ± SE) (N = 6)	p value
Maximum Force (N)	65.5 ± 4	69.1 ± 1.2	0.44
Torque (N-M)	2.8 ± 0.2	3.2 ± 0.2	0.32
*Path Length (cm)	226 ± 6	203 ± 6	0.05*
Number of Jerks	42.8 ± 7.3	48.7 ± 0.4	0.46
Time (sec)	7.7 ± 1	8.3 ± 0.2	0.64

* denotes a statistically significant difference when p < 0.05.

Table 5 Motion and force parameters for residents according to success on 1st intubation attempt.

	Most Successful Residents on	Least Successful Residents on 1st	p value

	1st intubation attempt (N = 6)	intubation attempt (N = 6)	
*Maximum Force (N)	59.6 ± 1.2	75 ± 4	0.02*
Torque (N-M)	2.9 ± 0.1	3.13 ± 0.3	0.5
Path Length (centimeters)	245 ± 12	184 ± 8	0.14
Number of Jerks	54.1 ± 5.3	37.4 ± 5.4	0.09
*Time (seconds)	9.6 ± 0.6	6.4 ± 0.8	0.03*

* denotes a statistically significant difference when $p < 0.05$.

If experience and laryngoscopy success were both measures of expertise in our group of residents, then one might expect greater levels of laryngoscopy success in the residents with the most experience. In fact, the two residents with the most prior experience were successful at laryngoscopy and intubation 100% of the time. Those with mid-range experience were successful 85% of the time, and those with the least amount of experience were successful only 74% of the time ($p = 0.16$ for most experienced vs. least experienced). The average percent of success on the 1st attempt was also calculated: 83% ± 3%, 77% ± 5%, and 62% ± 8% for the three groups, respectively ($p = 0.23$ for most experienced vs. least experienced). Thus, there was a trend for more experienced residents to have greater overall success with patient intubations as well as greater success on the 1st intubation attempt, but the differences were not significant.

DISCUSSION

The goal of this study was to determine if expertise in direct laryngoscopy can be distinguished by motion analysis. We particularly looked at the differences between expert laryngoscopists and novice anesthesia residents in potential measures of laryngoscopy accuracy and fluency. By establishing the expert 3D curvilinear tube, we were able to successfully differentiate between experts and novices based on the percent conformity to the tube. Resident laryngoscopy trajectories fell outside of the expert tube a significant amount of the time, enough to clearly distinguish between a novice trajectory and an expert one. Interestingly, when the expert tube was divided into thirds to represent different anatomical sections of the airway, the difference between an expert and a novice became even more striking in the last third of the trajectory, the portion moving through the laryngopharynx. The radius of the tube appreciably narrows from the top third to the bottom third, which indicates that placement of the laryngoscope blade by expert laryngoscopists varies less as they approach the vocal cords. Beginning residents, on the other hand, exhibited quite a wide range of variability at the bottom third of the tube. The path for the first portion of the laryngoscopy procedure is defined by the mouth opening and tongue, but landmarks are not discerned as easily deeper in the airway. This suggests that accurate blade placement at the laryngopharynx requires a certain amount of expertise. We speculate that the position of the laryngoscope tip likely needs to be very accurate to get the proper angle when lifting; the correct angle would offer the laryngoscopist the best view of the vocal cords and glottis.

In addition to the expert tube, we looked at several other motion parameters that are associated with the direct laryngoscopy procedure. Experts and novice residents significantly differed with two of the parameters, number of jerks and duration of a laryngoscopy, and there was a trend towards significance for path length. These parameters are all measures of laryngoscopy fluency: the “smoothness” of a DL procedure and the speed of a DL procedure. Expert laryngoscopists would be expected to be “smoother” and faster in their procedures than novices, just as a concert pianist would play a difficult piece more smoothly and effortlessly than a student. Interestingly, jerks, duration and path length are also associated with expertise in laparoscopic and micro-surgery[6]. Thus, we have established that motion analysis can clearly distinguish an expert laryngoscopist from a novice. The classification parameters have face validity because they measure accuracy or fluency, factors that should mature as a practitioner attains expertise.

In addition to face validity, our motion analysis method also has context validity. Context validity is demonstrated if the method discriminates between individuals with differing levels of experience and thus presumably different levels of expertise. This project showed that motion analysis measurements differed between first-year anesthesiology residents and anesthesiology faculty. Even though our method has face validity and context validity, further work is necessary to demonstrate its practical utility. Identifying such a large difference in skill would not require a sophisticated assessment technique and motion analysis should distinguish smaller differences in skill to be useful.

We tested the method's ability to resolve small skill differences in our study population by analyzing differences among first-year anesthesia residents with small differences in experience with direct laryngoscopy. Conformity to the last third of the expert tube did vary in a consistent manner among the groups, with the most experienced residents performing better than the least. On the other hand, number of jerks, duration, and path length did not vary with experience. These findings suggest that expert conformity may resolve smaller differences in expertise than do the fluency measures, jerks, duration, and path length. The latter parameters might resolve larger differences in expertise than found within a group of first-year residents. In order to better determine the relationship between those motion parameters and expertise, it may be worthwhile to conduct a future study that compares fluency measures between individuals with bigger differences in laryngoscopy training, perhaps groups of residents who are separated by 6-12 months of experience.

Ultimately, we would like to know whether motion analysis has predictive validity for laryngoscopy expertise. Predictive validity means that individuals who score higher by motion analysis would demonstrate greater expertise when performing laryngoscopy in a real situation than do practitioners with lower scores. In other words, motion analysis would generally predict which individuals would have greater success rates when intubating patients. We tested this proposition on a small scale by comparing first-year residents based on their success in patient laryngoscopy. When divided into 2 groups according to intubation success in the first month of their training, we did find some significant differences among the motion parameters, but these differences were all paradoxical and in the opposite direction than expected. The inability to demonstrate predictive validity may result from the small difference in expertise across the first year resident group and/or the small sample size. Again, a larger study with appropriate subject groups will be necessary for an adequate assessment of predictive validity.

Clinical implications

Laryngoscopy training usually begins on patients with normal anatomy, and it is estimated that beginners reach 90% success after a mean of 57 attempts[2]. Since the learning curves have a relatively flat slope at this point, improving above 90% success rate still requires a large number of additional patient attempts, orders of magnitude more experience than what the trainee has amassed to that point. In addition, the aspiring laryngoscopy expert must develop expertise for intubating patients with more difficult anatomy. However, only about 1-3% of surgical patients are considered to have difficult airways[1], so access to these types of patients is limited and infrequent. Thus, the development of expert laryngoscopy skills may proceed slowly and vary depending on how frequently rare, difficult patients are encountered. To date, there remains no comprehensive method to assess the procedural skills associated with anesthesia[6]. With airway mishaps being the most common adverse event in the field of anesthesia[5], an objective method to assess laryngoscopy skill and determine level of expertise would be ideal to decrease the number of mishaps. Thus, one of our long-term goals is to create benchmarks with levels corresponding to “novice,” “intermediate,” and “expert.” In this way, our model could be an accurate assessment method that would assist in teaching and perfecting laryngoscopy skills on a mannequin. Benchmarks could also be used to determine when an individual was ready to graduate to independent, unsupervised practice.

Our expertise assessment model and sensor equipment could also be used to improve laryngoscopy training methods. Our system of electronic sensors records and displays laryngoscopy trajectories in real time, which could give trainees immediate feedback and allow them to track their progress through varying degrees of difficulty. Alternatively, instructors could use laryngoscopy trajectories to guide in the oral feedback they give trainees.

According to Bould, et. al., there are 4 essential aspects of training program: 1) validity, 2) reliability, 3) feasibility, and 4) comprehensiveness. With this study we have proven the validity of our model, although more studies with a larger number of subjects and greater variation in experience and expertise would be helpful in further establishing its predictive validity. Repeated studies are also needed to ascertain the reliability of our study. We believe that our model is feasible and would be easy to put into practice[6]. Because it allows trainees to practice laryngoscopy multiple times in a short period of time and allows for instant feedback, it is perhaps an ideal training scenario[7]. Additionally, it could easily be incorporated into any type of intubation training for medical personnel, from Emergency Medicine physicians to Paramedics and other medical staff.

Questions

Several questions arose in the course of analyzing the results. First, which benchmarks would prove useful in evaluating expertise? Also, could different ranges in the motion parameters define the levels of novice, intermediate, and expert? The utility of motion parameters would depend on the variability and the effect size when looking across varying levels of expertise. A combination of parameters might work better than a single measure. For example, some subjects might be able to score well on 1 or 2 of the parameters, even though they had not achieved expertise for the procedure as a whole. Furthermore, results might be improved by broadening the scope of what is described here. There may be other potential parameters that we didn't measure that might be better suited for assessing expertise, such as laryngoscope blade angle or head angle. Finally, the results suggest that motion analysis may be useful in discerning the factors that are important in performing the procedure. For example, the last third of the airway and the laryngoscopy trajectory should receive greater scrutiny because this region appears to be so crucial to the direct laryngoscopy procedure.

Review of Literature

Forrest et. al. used a subjective rating scale to assess the overall technical performance of novice anesthetists, including intubation. Their scoring system appeared to be valid as novices improved their scores over time, and it successfully detected performance differences between experts and novices. However, there may have been a problem with inter-rater reliability as scores were subjective in nature. This study showed that simulation and a scoring system can be used to assess technical performance in anesthesia. It did not focus on the details of the procedure, such as duration of the laryngoscopy or the smoothness of the procedure, but it can potentially be tailored to focus on specific aspects of anesthesia skill sets[8]. A study recently published by T. Rahman, et. al. used motion analysis to track direct laryngoscopy in pediatric mannequins. Although they measured similar parameters to our study, namely success or failure of intubation, blade-tip motion path length, laryngoscope handle angle, and time of the full intubation, their motion analysis system was placed in a different location than ours. Their system (called the "Flock of Birds" EM system from Ascension Technology Corporation, Burlington, Vt) was attached to the base of the handle of a laryngoscope, and our system was placed on the tip of the laryngoscope blade. Thus, their system recorded non-specific swings from the end of the instrument, whereas our method required translating changes in motion to movement of the blade tip. Nonetheless, we did share one similar result with this study: they demonstrated that experts had greater overall success than novices at direct laryngoscopy just as we did. However, most of the other results contradicted our own. For example, their experts had longer path lengths and durations of intubation than novices, whereas in our study, experts had shorter path lengths and laryngoscopy times than novices. The researchers speculated that the experts were perhaps being more careful with their intubations to minimize known laryngoscopy risks, such as edema or bleeding. Novices do not have enough experience to know how to avoid the dangers of intubation, and so they were perhaps haphazardly, and thus speedily, performing intubations[9]. In a 1996 study, Hastings et. al. studied the force and torque associated with direct laryngoscopy. They observed that force varied widely amongst experienced laryngoscopists and depended on technique and equipment. Similar to our results, force and torque were found to not be predictive of expertise[10].

Alternatives

The ability to use motion analysis to determine expertise in direct laryngoscopy is a relatively novel concept for laryngoscopy training. Other common methods could be used to assess laryngoscopy skill, such as simple checklists or a Global Rating Scale (GSR). GSR's use a multiple point, semi-quantitative

Likert scale to score an individual on specific issues thought related to procedural expertise. Checklists provide a considerate amount of internal validity but do not necessarily assess the performance of a procedure and GSRs are somewhat more quantitative. For example, the University of Toronto developed a GRS that can differentiate between junior and senior trainees performing an interscalene block. It can also discriminate between various levels of experience at performing epidural anesthesia[6]. GRSs can be used for different procedural skills, so in theory, one could be developed for the procedural skill of direct laryngoscopy. The Imperial College Surgical Assessment Device (ISCAD) is a motion analysis device that has been validated in various surgical fields as an objective measure of technical ability[3]. It can only assess the process rather than the outcome of the procedure, and it requires further validation in Anesthesia. Part-task trainers (PTT) are modeled after a particular anatomical area or procedure. One example is the Human Patient Simulator (METI, Sarasota, FL, USA) that showed good predictive validity for tracheal intubation, but it cannot simulate a wide range of difficult airway scenarios[6].

Critiques and Limitations

There were some limitations with this study. The first was that the number of novice residents in our study was somewhat small. Also, the number of faculty to which the residents were compared was also small and not enough to prove that experts do not show as much variability as residents. Furthermore, the first-year anesthesia residents were not all raw beginners, as several of them had previous experience with intubation in either patients and/or mannequins. Thus, the starting baseline for the novices was not equal for everyone. All subjects did use the same mannequin and while we attempted to calibrate the position sensor before each intubation attempt, the calibration remains a potential source of variability. The position of the mannequin could also be a source of variability, although it was brought to its neutral position after every intubation attempt. Other confounding factors included the differences among patients' Mallampati scores, thyromental distances, and mouth opening distances. Patients with higher Mallampati scores and shorter thyromental and mouth opening distances would likely be more difficult to intubate. Because of the random assignment of patients to residents, there was no way to control for the number of difficult patients a resident received, and this may have potentially affected the measured success for some residents. One last limitation of this study is that it may be hard to extrapolate the results for those with intermediate levels of experience, as there were no differences detected in motion parameters amongst the residents.

Conclusion

This study sought to demonstrate expertise in the technique of direct laryngoscopy using motion analysis. We were able to successfully distinguish between expert laryngoscopists and novice anesthesia residents using this form of analysis. Looking at the 3D curvilinear tube that represents the space occupied by expert laryngoscopy trajectories, residents conformed to only a percentage of the expert tube. Thus, there was a clear difference between expert trajectories and novice trajectories. Furthermore, significant differences between experts and novices were found in certain motion parameters, namely number of jerks and duration of a laryngoscopy. Our model was also able to detect smaller differences in experience between the residents themselves when looking at the expert tube but not the motion parameters. Other parameters associated with laryngoscopy, such as the angle of the laryngoscope or the head angle, may be associated with expertise and would perhaps be helpful in better differentiating between intermediate levels of experience. Further studies using motion analysis for direct laryngoscopy may also possibly benefit from the addition of haptic feedback. Procedural simulation with haptic feedback has been shown to lead to improved performances in minimally invasive surgery and robot-assisted endoscopic surgery[4]. Our instrumented laryngoscope and electronic motion analysis system have the potential to expand into a system of haptic feedback. In theory, we could provide trainees with sensations related to applied force and tissue damage with the hope of further avoiding airway mishaps and improving laryngoscopy accuracy and fluency.

RESOURCES

1. Mulcaster, J., et. al. "Laryngoscopic Intubation: Learning and Performance." *Anesthesiology* 2003; Volume 98: 23-27.
2. Konrad, C., et. al. "Learning Manual Skills in Anesthesiology: In There a Recommended Number of Cases for Anesthetic Procedures?" *Anesth Analg* 1998; Volume 86: 635-9.
3. Tavakol, M., et. al. "Assessing the Skills of Surgical Residents Using Simulation." *Journal of Surgical Education* 2008; Volume 65: 77-83.
4. van der Meijden, O., et. al. "The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review." *Surg Endosc* 2009; Volume 23: 1180-1190.
5. Benumof, J. "Management of the difficult adult airway with special emphasis on awake tracheal intubation." *Anesthesiology* 1991; Volume 75: 1087-1110.
6. Bould, M., et. al. "Assessment of procedural skills in anaesthesia." *British Journal of Anaesthesia* 2009; Volume 103: 472-83.
7. Ericsson, K. "Deliberate Practice and Acquisition of Expert Performance: A General Overview." *Academic Emergency Medicine* 2008; Volume 15: 988-994.
8. Forrest, F., et. al. "Use of a high-fidelity simulator to develop testing of the technical performance of novice anaesthetists." *British Journal of Anaesthesia* 2002: Volume 88: 338-44.
9. Rahman, T., et. al. "Tracking Manikin Tracheal Intubation Using Motion Analysis." *Pediatric Emergency Care* 2011; Volume 27: 701-705.
10. Hastings, R., et. al. "Force and Torque Vary Between Laryngoscopists and Laryngoscope Blades." *Anesth Analg* 1996; Volume 82: 462-8.