Predicting Difficult Intubation in Emergency Department by Intubation Assessment Score

Winchana Srivilaithon\textsuperscript{a, c}, Sombat Muengtaweepongs\textsuperscript{b}, Yuwares Sittichanbuncha\textsuperscript{c}, Jayanton Patumanon\textsuperscript{d}

Abstract

Background: The difficult intubation is associated with failure of emergency tracheal intubation. This study aimed to develop and validate a model for predicting difficult intubation in emergency department (ED).

Methods: A cross-sectional study was conducted in the ED. We collected data from all consecutive adult patients who underwent emergency tracheal intubation. Patients were excluded if they were intubated by low experience intubator. The difficult intubation was defined by grade III or IV of Cormack and Lehane classification. We used multivariable regression model to identify significant predictors of difficult intubation and weighted points proportional to the beta coefficient values. The ability to discriminate was quantified by using the area under receiver operating characteristics curve (AuROC). The bootstrapping method was used to validate the performance.

Results: A total of 1,212 intubations were analyzed. One hundred and fifty-seven intubations were enrolled in difficult intubation group. Five independence predictors were identified, and each was assigned a number of points proportional to its beta coefficient: male gender (one), large tongue (two), limit mouth opening (two), poor neck mobility (two), and presence of obstructed airway (three). Intubation assessment score model was created and applied to all subjects. The AuROC was 0.81 (95% confidence interval (CI): 0.77 - 0.85) for the development dataset, and 0.80 (95% CI: 0.76 - 0.85) for the validation dataset. We defined three risk groups: low risk (zero to one points), intermediate risk (two to three points), and high risk (above three points), and the difficult intubation rate was 4.7%, 22.5%, and 53.6%, respectively.

Conclusions: Intubation assessment score model was constructed from patients’ simple characteristics and performed well in predicting difficult intubation and can discriminate between with and without difficult intubation.

Keywords: Airway assessment; Difficult intubation; Difficult intubation indicator; Emergency department intubation

Introduction

In the emergency departments (EDs), emergency physicians are responsible for managing the airway problems. The need for intubation in the acutely ill patients in the ED is unpredictable and is often promptly required [1]. When compared with the same procedure in the operating room, a higher incidence of airway management failure had been suggested [1, 2]. The difficult intubation is associated with an increasing number of intubation attempts [3]. Repeated attempts in intubation are also associated with an increasing risk of complications such as cardiac arrest, hypoxemia, arrhythmia, regurgitation, and airway trauma [4].

Many large multicenter studies had reported information on emergency airway management in the ED [4-7]. The previous study showed that the predicted airway difficulty was the major factor associated with first attempt success of intubation [5]. Many airway assessment methods had been proposed and evaluated, but each of these had limitation in sensitivity and specificity [2, 5, 8-10]. The combined assessment of multiple predictors successfully predicted difficulty for intubation in preoperative patients [11, 12]. However, the utility of these combinations in the ED setting has not been demonstrated.

The main objective of this study was to develop and validate a model for predicting difficult intubation in patients who underwent emergency tracheal intubation in the ED.

Materials and Methods

Study design and study patients

The retrospective cross-sectional study was conducted in the
ED of tertiary care, Thammasat Hospital, between September 2012 and February 2016. We implemented the airway registry system for all consecutive patients who presented to the ED and underwent emergency tracheal intubation. The data collection was based on a protocol from National Emergency Airway Registry (NEAR).

The study included all patients who were intubated with conventional tracheal intubation in the ED. Patients were excluded if they were younger than 15 years old. For the most accuracy of the difficult intubation assessment, patients who were intubated and collected data by low experience intubators were also excluded.

Ethical approval

This study and the airway registry were approved by Human Research Ethics Committee of Thammasat University (Faculty of Medicine).

Definition of difficult intubation (the reference standard)

The definition of difficult intubation was used in the same definition of difficult laryngoscopy, according to the structures that can be visualized and identified by direct laryngoscopy. By four-grade laryngoscopic view defined by Cormack and Lehane [13], intubation was easy for grade I or grade II and intubation was difficult for grade III or grade IV. The measure of laryngoscopic view was assessed and recorded by the final intubator after each intubation was finished.

Definition of predictors

Intubation methods

We describe each encounter by “method” and number of “attempts”. We define a “method” as a single set of medication or devices, such as rapid sequence intubation with a Macintosh laryngoscope.

Intubation attempts

We define an “attempt” as a single effort to place an airway. Each encounter could have one or more methods and each method could have one or more attempts. After each intubation was finished, intubator completed the data including patients’ baseline characteristics, main indication for intubation, initial method of intubation, difficult intubation indicators, operator level of training, number of attempt, success or failure, dosage and name of medication use, and structures identified by direct laryngoscopy.

Operator levels

We divided operator level of training into three groups: 1) last year of medical student and first-year internist in general practice were in low experience group; 2) second-year to third-year internist in general practice and first-year residency in emergency medicine were in moderate experience group; and 3) second-year to third-year residency in emergency medicine and emergency attending staff were in high experience group.

Difficult intubation indicators

The characteristics originally described in LEMON method were used to assess the patients who undergo tracheal intubation [6, 8]. For the “look” criteria, we assessed any significant facial injury, large incisors, significant beard or mustache, and large tongue. For the unfavorable “evaluate” criteria, we noted the patients with mouth opening less than three finger breadths, hypo-mental distance less than three finger breadths, and thyro-hyoid distance less than two finger breadths. For the presence of “obstructed” airway, we noted any condition that can make laryngoscopy and ventilation difficult; such conditions were hematoma, infection, and abnormal mass. For the “neck” criteria, we noted the patients with limit neck movement, such as collar neck immobilization. We did not record the “Mallampati” criteria, because it needed patient cooperation and had limitation in critical ill patients in the ED.

Study size estimation and statistical analysis

The sample size was calculated by two-sample comparison of proportions with alpha 0.05, 80% power, and one-side test, to detect the difference of difficult intubation indicators between patients.

Study patients were divided into two groups based on the difficult intubation (difficult intubation group and non-difficult intubation group). Baseline patients’ characteristics were compared using exact probability test for categorical variables and Student’s t-test for continuous variables. To develop the predicting model, a multivariable stepwise backward logistic regression model was constructed. Predictors with a cut level P-value of 0.1 after univariable analysis were included in the model and eliminated with significant level 0.05. We assigned the predictors identified by multivariable analysis weighted points proportional to the beta coefficient values. A risk score was then calculated for each patient.

Goodness-of-fit (GOF) was quantified by using the Hosmer-Lemeshow test. For the internal validation of the predicting model, we used bootstrap to random the dataset. Two-thirds of study enrolls were used to construct the validation dataset and involved in the analysis. The ability of the development dataset and validation dataset to discriminate the difficult intubation was quantified by using the area under receiver operating characteristics curve (AuROC) with 95% confidence interval (CI).

The predictive model was stratified into three risk groups (low, intermediate, and high) based on the probability of difficult intubation. All data analysis was performed using STATA.
Results

Model construction

During the 42-month study period, 1,641 patients required emergency tracheal intubation in the ED. Among these, the data from airway registry recoded 1,632 intubations (captured rate 99.45%). We excluded patients younger than 15 years old (n = 34), intubated by low experience intubator (n = 379), and not intubated by direct laryngoscope (n = 7). The remaining 1,212 intubations were included in the analysis. One hundred and fifty-seven intubations (12.95%) were enrolled in difficult intubation group (Fig. 1).

Table 1 summarizes patients’ baseline characteristics, result of univariable analysis between the difficult intubation group and the non-difficult intubation group, and AuROC for predicting difficult intubation. In comparison between groups, the male gender was found in the difficult intubation group much more than the non-difficult intubation group (74.5% and 62.7%, P = 0.004). The patients in the difficult intubation group were younger than the non-difficult intubation group (mean age 56.5 years and 60.4 years, P = 0.02). The patients with traumatic as cause of intubation had higher number in the difficult intubation group (26.1% and 14.2%, P < 0.001). For the method of intubation, the non-difficult intubation group had higher number of patients who undergo rapid sequence intubation with neuromuscular blockage agent before intubation (22.3% and 17.2%, P = 0.053). The difficult intubation group needed more number of attempts than the non-difficult intubation group (mean number of attempts; 2 and 1.3, P < 0.001) and had higher failed intubation (needed more than three intubation attempts by experience intubator) (11.5% and 1.42%, P < 0.001).

For difficult intubation indicators including facial trauma, large incisors, beard or mustache, large tongue, limit mouth opening below three fingerbreadths, short hypo-mental distance below three fingerbreadths, short thyro-hyoid distance below two fingerbreadths, poor neck mobility, and presence of obstructed airway, the difficult intubation group had higher number of patients than the non-difficult intubation group in...
Table 1. Characteristics of Patients With Difficult and Non-Difficult Intubation, and Discrimination Power (AuROC) Under Univariable Analysis

| Characteristics                        | Difficult intubation (n = 157), n (%) | Non-difficult intubation (n = 1,055), n (%) | P-value  | AuROC (95% confidence interval) |
|----------------------------------------|---------------------------------------|--------------------------------------------|---------|---------------------------------|
| Male gender                            | 117 (74.5)                            | 661 (62.7)                                 | 0.004   | 0.56 (0.52 - 0.60)              |
| Age (year), mean (±SD)                 | 56.5 (± 18.7)                         | 60.4 (± 19.8)                              | 0.020   | 0.43 (0.39 - 0.48)              |
| Glasgow coma scales                     |                                       |                                            |         |                                 |
| 3 - 8                                  | 86 (62.8)                             | 575 (61.9)                                 | 0.421   | 0.49 (0.45 - 0.53)              |
| 9 - 13                                 | 23 (16.8)                             | 126 (13.6)                                 |         |                                 |
| 14 - 15                                | 28 (20.4)                             | 228 (24.5)                                 |         |                                 |
| Traumatic causes                       | 41 (26.1)                             | 150 (14.2)                                 | < 0.001 | 0.56 (0.52 - 0.60)              |
| Method of intubation                   |                                       |                                            |         |                                 |
| Rapid sequence                         | 27 (17.2)                             | 235 (22.3)                                 | 0.053   | 0.50 (0.45 - 0.54)              |
| Sedation only, without paralysis       | 63 (40.1)                             | 323 (30.6)                                 |         |                                 |
| No-medication assisted                 | 67 (42.7)                             | 497 (47.1)                                 |         |                                 |
| Used of induction agents               | 87 (55.4)                             | 555 (52.6)                                 | 0.549   | 0.51 (0.47 - 0.56)              |
| Mean number of attempts (± SD)         | 2.0 (± 1.2)                           | 1.3 (± 0.7)                                | < 0.001 | n/a                             |
| Failed intubation                      | 18 (11.5)                             | 15 (1.42)                                  | < 0.001 | n/a                             |
| Difficult intubation indicator         |                                       |                                            |         |                                 |
| Facial trauma                          | 15 (9.6)                              | 27 (2.6)                                   | < 0.001 | 0.54 (0.51 - 0.56)              |
| Large incisors                         | 12 (7.6)                              | 28 (2.7)                                   | 0.003   | 0.52 (0.50 - 0.55)              |
| Beard or mustache                      | 8 (5.1)                               | 33 (3.1)                                   | 0.232   | 0.51 (0.49 - 0.53)              |
| Large tongue                           | 50 (31.9)                             | 64 (6.1)                                   | < 0.001 | 0.63 (0.59 - 0.67)              |
| Limited mouth opening                  | 57 (36.3)                             | 71 (6.7)                                   | < 0.001 | 0.65 (0.61 - 0.69)              |
| Short hypo-mental distance             | 24 (15.3)                             | 70 (6.6)                                   | 0.001   | 0.54 (0.51 - 0.57)              |
| Short thyro-hyoid distance             | 22 (14.0)                             | 61 (5.8)                                   | 0.001   | 0.54 (0.51 - 0.57)              |
| Poor neck mobility                     | 63 (40.1)                             | 156 (14.8)                                 | < 0.001 | 0.63 (0.59 - 0.67)              |
| Presence of obstructed airway          | 27 (17.2)                             | 15 (1.4)                                   | < 0.001 | 0.58 (0.55 - 0.61)              |

AuROC: area under receiver operating characteristic curve.

Table 2. Odds Ratio of Predictors and Difficult Intubation From Multivariable Stepwise Backward Elimination Logistic Regression Analysis, and Item Score Assignment

| Predictor                        | Odds ratio | 95% confidence interval | P-value | Beta coefficient | Score |
|----------------------------------|------------|-------------------------|---------|------------------|-------|
| Male gender                      | 1.82       | 1.17 - 2.83             | 0.008   | 0.59             | 1     |
| Large tongue                     | 4.47       | 2.72 - 7.35             | < 0.001 | 1.49             | 2     |
| Limited mouth opening            | 3.93       | 2.44 - 6.31             | < 0.001 | 1.37             | 2     |
| Poor neck mobility               | 2.76       | 1.81 - 4.19             | < 0.001 | 1.01             | 2     |
| Presence of obstructed airway    | 7.54       | 3.57 - 15.93            | < 0.001 | 2.02             | 3     |

Estimation of the accuracy of the model

The model’s ability to discriminate is shown in Figure 2. The
AuROC of the development dataset was 0.81 (95% CI: 0.77 - 0.85). The internal validation of the model coefficient was analyzed with the validation dataset after bootstrapping of two-thirds of the all dataset. Eight hundred patients were included in the analysis, showing similar discrimination ability (AuROC: 0.80; 95% CI: 0.76 - 0.85).

The Hosmer-Lemeshow GOF test showed no statistically significant difference between predicted and observed probability (P = 0.19).

Probability categories of difficult intubation

Table 3 shows the distribution of intubation assessment score according to the difficult intubation in development dataset. The higher score is associated with a higher prevalence of difficult intubation. The patients were divided into three groups according to intubation assessment score: 0 - 1, 2 - 3, and above 3, and were categorized into low risk group, intermediate risk group, and high risk group, respectively.

In the development dataset, the difficult intubation rate in the low, intermediate, and high risk groups was 4.7%, 22.5%, and 53.6%, respectively. For the low risk group, the likelihood ratio for difficult intubation was 0.33 (95% CI: 0.25 - 0.44, P < 0.001). For the intermediate risk group, the likelihood ratio for difficult intubation was 1.95 (95% CI: 1.53 - 2.49, P < 0.001). For the high risk group, the likelihood ratio for difficult intubation was 7.75 (95% CI: 5.56 - 10.79, P < 0.001).

Table 3. Distribution of Score Categorized by Difficult and Non-Difficult Intubation

| Score | Total (n = 1,212) | Prevalence (%) | Difficult intubation (n = 157) | Non-difficult intubation (n = 1,055) |
|-------|------------------|----------------|--------------------------------|------------------------------------|
|       | n                | %              | n                             | %                                  |
| 0     | 331              | 27.3           | 10                            | 3.0                                | 321 | 97.0 |
| 1     | 516              | 42.6           | 30                            | 5.8                                | 486 | 94.2 |
| 2     | 65               | 5.3            | 12                            | 18.5                               | 53  | 81.5 |
| 3     | 188              | 15.5           | 45                            | 23.9                               | 143 | 76.1 |
| 4     | 28               | 2.3            | 13                            | 46.4                               | 15  | 53.6 |
| 5     | 47               | 3.9            | 22                            | 46.8                               | 25  | 53.2 |
| 6     | 11               | 0.9            | 6                             | 54.6                               | 5   | 45.4 |
| 7     | 14               | 1.1            | 9                             | 64.3                               | 5   | 35.7 |
| 8     | 6                | 0.5            | 5                             | 83.3                               | 1   | 16.7 |
| 9     | 3                | 0.3            | 2                             | 66.7                               | 1   | 33.3 |
| 10    | 3                | 0.3            | 3                             | 100                                | 0   | 0    |
Discussion

The simple model, intubation assessment score model, was performed well in predicting difficult intubation in our study population (AuROC: 0.81; 95% CI: 0.77 - 0.85) as in the internal validation dataset (AuROC: 0.80; 95% CI: 0.76 - 0.85). Our finding showed that the weight combination of five independent predictors (male gender, large tongue, limited mouth opening, poor neck mobility, and presence of obstructed airway) can help discriminate between patients with and without difficult intubation in adults who were intubated by conventional direct laryngoscope tracheal intubation in the ED. These simple predictors can be assessed easily by ED personals before initiation of intubation.

Management of the difficult intubation in the ED has not been as well studied as that in the operating room. Several previous small studies in Thailand had reported the patients’ characteristics, intubation methods, and success rates within single institution; however, they had limitation on predictor of multiple attempts at intubation in the ED and how to predict difficult intubation before starting intubation [14-16]. The proportion of difficult intubation in this study was 12.95%, comparable with other studies in the ED setting (5.4-27%) [2, 3, 6, 7, 17], whereas the rate seemed to be higher than the incidence of elective intubation in the operating room (0.3-13.3%) [18-20]. Some predictors such as LEMON criteria were used to predict difficult intubation before starting intubation [21]; however, most of the assessments lack accuracy. Poor correlation between the LEMON criteria and Cormack and Lehane laryngoscopic view grade in the ED had been proposed, demonstrating a sensitivity of 23.3% and a specificity of 91.3% [5]. Moreover, some assessment predictors had been derived from studies in which cooperative, composed patients were examined for preoperative preparation by anesthesiologists rather than by emergency physicians for evaluating patients in the ED. The validation of the modified LEMON criteria to predict difficult intubation in ED had demonstrated high sensitivity and negative predictive values (85.7% and 98.2%, respectively) but poor specificity and positive predictive values (47.6% and 8.9%, respectively) [2]. The combination of multiple predictors or multivariable risk scores, such as intubation difficulty scale (IDS) and Naguib score, had been proposed and evaluated in preoperative setting to improve the predictive capability of bedside test; however, many predictors cannot evaluate the critical ill patients in the ED and these composite scores tend to be time consuming and difficult to use in clinical practice [11, 12, 22].

There are a few limitations in this study. First, it is retrospective in nature, patient’s characteristics and clinical indicators were only derived from patient airway record form, and the incomplete records had to be excluded and would limit the accuracy of the study. However, due to high capture rate (99.45%), this limitation might have had little effect on the result of the study. Second, each airway record form was completed by intubator after finishing each intubation, and the observer bias might have occurred if intubators had incorrect assessment and record of patient’s difficult intubation indicators. We minimized this bias by controlling patient selection in the exclusion process, and patients were excluded if they were intubated by low experience intubator. Third, the patients who were not intubated with conventional direct laryngoscope tracheal intubation (video-assisted intubation, fiberoptic-assisted intubation and surgical cricothyroidotomy) should also be excluded, and these patients may have some characteristics of difficult intubation and susceptible to selection bias. Finally, this study was limited to adults. Because of differences in anatomical appearances, the results of the current study cannot be applied to children.

Conclusions

Small subset of emergency intubation in ED will cause difficulty in laryngoscopy. The difficult intubation patients may receive repeated attempts at intubation and lead to greater likelihood of adverse events. The intubation assessment score model was constructed from patients’ simple characteristics and their visible anatomy, performed well in predicting difficult intubation in our study population and can help discriminate between patients with and without difficult intubation in adults who were intubated by conventional tracheal intubation in the ED. It can classify patients into subgroups at low, intermediate, and high risk according to their difficulty risk, which in turn will help clinicians optimize decision making on preparation process, select method for intubation, and rescue airway decision.

Conflict of Interest

None.

Funding

This research did not receive any specific grant from funding
agencies in the public, commercial, or not-for-profit sectors.

References

1. Orebaugh SL. Difficult airway management in the emergency department. J Emerg Med. 2002;22(1):31-48.
2. Hagiwara Y, Watase H, Okamoto H, Goto T, Hasegawa K, Japanese Emergency Medicine Network Investigators. Prospective validation of the modified LEMON criteria to predict difficult intubation in the ED. Am J Emerg Med. 2015;33(10):1492-1496.
3. Duchynski R, Brauer K, Hutton K, Jones S, Rosen P. The quick look airway classification. A useful tool in predicting the difficult out-of-hospital intubation: experience in an air medical transport program. Air Med J. 1998;17(2):46-50.
4. Hasegawa K, Shigemitsu K, Hagiwara Y, Chiba T, Watase H, Brown CA, 3rd, Brown DF, et al. Association between repeated intubation attempts and adverse events in emergency departments: an analysis of a multicenter prospective observational study. Ann Emerg Med. 2012;60(6):749-754 e742.
5. Kim C, Kang H, Lim T, et al. What factors affect the success rate of the first attempt at endotracheal intubation in emergency departments? Emerg Med J. 2013;30:88-92.
6. Reed MJ, Dunn MJ, McKeown DW. Can an airway assessment score predict difficulty at intubation in the emergency department? Emerg Med J. 2005;22(2):99-102.
7. Sagarin MJ, Barton ED, Chng YM, Walls RM, National Emergency Airway Registry I. Airway management by US and Canadian emergency medicine residents: a multicenter analysis of more than 6,000 endotracheal intubation attempts. Ann Emerg Med. 2005;46(4):328-336.
8. Calder I. Identification of the Difficult Airway. Anesthesiology. 2011;12(8):340-342.
9. Calder I. Acromegaly, the Mallampati, and difficult intubation. Anesthesiology. 2001;94(6):1149-1150.
10. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently normal patients: a meta-analysis of bedside screening test performance. Anesthesiology. 2005;103(2):429-437.
11. Adnet F, Borron SW, Racine SX, Clemency JL, Fournier JL, Plaisance P, Lapandry C. The intubation difficulty scale (IDS): proposal and evaluation of a new score characterizing the complexity of endotracheal intubation. Anesthesiology. 1997;87(6):1290-1297.
12. Naguib M, Scamman FL, O’Sullivan C, Aker J, Ross AF, Kosmach S, Ensor JE. Predictive performance of three multivariate difficult tracheal intubation models: a double-blind, case-controlled study. Anesth Analg. 2006;102(3):818-824.
13. Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. Anaesthesia. 1984;39(11):1105-1111.
14. Srivilaithon W. A prospective observational study of emergency airway management in emergency department. J Med Assoc Thai. 2016;99(Suppl 4):S131-S137.
15. Wongyingsin M, Songarj P, Assawinvinijkul T. A prospective observational study of tracheal intubation in an emergency department in a 2300-bed hospital of a developing country in a one-year period. Emerg Med J. 2009;26(8):604-608.
16. Trakulsrichai S, Sundarathithi P, Chalermdichai S, Palasatien I, Rattanasiri S, Chatichaipun P, Hathirat S. An observation study of rapid sequence, awake and sedation-only intubations in an emergency department in Thai patients. J Med Assoc Thai. 2009;92(8):1022-1027.
17. Martin LD, Mhyre JM, Shanks AM, Tremper KK, Kheterpal S. 3,423 emergency tracheal intubations at a university hospital: airway outcomes and complications. Anesthesiology. 2011;114(1):42-48.
18. Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraksa B, Freiberger D, Liu PL. A clinical sign to predict difficult tracheal intubation: a prospective study. Can Anaesth Soc J. 1985;32(4):429-434.
19. Williams KN, Carli F, Cormack RS. Unexpected, difficult laryngoscopy: a prospective survey in routine general surgery. Br J Anaesth. 1991;66(1):38-44.
20. Bellhouse CP, Dore C. Criteria for estimating likelihood of difficulty of endotracheal intubation with the Macintosh laryngoscope. Anaesth Intensive Care. 1988;16(3):329-337.
21. Murphy F, Walls M. Identification of the difficult and failed airway. In: Walls RM, Murphy MF, Luten RC, eds. Manual of Emergency Airway Management, 3rd ed. Philadelphia: Lippincott, Williams & Wilkins/ Wolters Kluwer Health; 2008. p. 81-93.
22. Baker P. Assessment before airway management. Anesthesiol Clin. 2015;33(2):257-278.