



## Prediction of Difficult Tracheal Intubation by Artificial Intelligence: A Prospective Observational Study

Yapay Zekâ ile Zor Trakeal Entübasyon Tahmini: Prospektif Gözlemsel Bir Çalışma

Fatma ÇELİK<sup>1</sup>

 0000-0003-0192-0151

Emrah AYDEMİR<sup>2</sup>

 0000-0002-8380-7891

<sup>1</sup>Kırşehir Ahi Evran University Faculty of Medicine Department of Anesthesiology and Reanimation, Kırşehir, Turkey

<sup>2</sup>Kırşehir Ahi Evran University Faculty of Engineering and Architecture Department of Computer Engineering, Kırşehir, Turkey

Corresponding Author

Sorumlu Yazar

Emrah AYDEMİR

aydemir.emrah23@gmail.com

Received / Geliş Tarihi : 16.01.2021

Accepted / Kabul Tarihi : 22.03.2021

Available Online /

Çevrimiçi Yayın Tarihi : 07.04.2021

### ABSTRACT

**Aim:** Many predictive clinical tests are used together for preoperative detection of patients with difficult airway risk. In this study, we aimed to predict difficult intubation with different artificial intelligence algorithms using various clinical tests and anthropometric measurements, besides, to evaluate the accuracy performance of Cormack and Lehane (C-L) classification with artificial intelligence.

**Material and Methods:** This study was conducted as a single-blind prospective observational study between 2016 and 2019. A total of 1486 patients with American Society of Anesthesiologists physical status I-III, scheduled to undergo elective surgery and requiring endotracheal intubation, were included. Demographic variables, clinical tests and anthropometric measurements of the patients were recorded. Difficult intubation was evaluated using the 4-grade C-L system according to the easy and difficult intubation criteria. Difficult intubation was tried to predict using 16 different artificial intelligence algorithms.

**Results:** The highest success rate among artificial intelligence algorithms was obtained by the RandomForest method. With this method, difficult intubation was predicted with 92.85% sensitivity, 96.94% specificity, 93.69% positive predictive value and 96.52% negative predictive value. C-L classification accuracy performance also determined as 95.60%.

**Conclusion:** Artificial intelligence has been considerably successful in predicting difficult intubation. Besides, C-L classifications of easy and difficult intubated patients were successfully predicted with artificial intelligence algorithms. Using a 6-grade modified C-L classification for laryngeal view may provide stronger difficult intubation prediction. A safer and more potent prediction in training artificial intelligence can be achieved by adding individual differences and clinical features that support the definition of difficult intubation.

**Keywords:** Tracheal intubation prediction; difficult intubation; artificial intelligence; Cormack-Lehane; intubation, anesthesia.

### ÖZ

**Amaç:** Zor hava yolu riski olan hastaların preoperatif tespiti için birçok prediktif klinik test birlikte kullanılmaktadır. Bu çalışmada, çeşitli klinik testler ve antropometrik ölçümler kullanarak farklı yapay zekâ algoritmaları ile zor entübasyonun tahmin edilmesi, ayrıca Cormack ve Lehane (C-L) sınıflandırmasının doğruluk performansının yapay zekâ ile değerlendirilmesi amaçlanmıştır.

**Gereç ve Yöntemler:** Bu çalışma, 2016 ve 2019 yılları arasında tek kör prospektif gözlemsel bir çalışma olarak gerçekleştirildi. Elektif cerrahi planlanan ve endotrakeal entübasyon gerektiren, Amerikan Anesteziyologlar Derneği fiziksel durumu I-III olan toplam 1486 hasta dahil edildi. Hastaların demografik değişkenleri, klinik testleri ve antropometrik ölçümleri kaydedildi. Zor entübasyon 4 dereceli C-L sistemi ile kolay ve zor entübasyon kriterlerine göre değerlendirildi. Zor entübasyon, 16 farklı yapay zekâ algoritması kullanılarak tahmin edilmeye çalışıldı.

**Bulgular:** Yapay zekâ algoritmaları arasında en yüksek başarı oranı RandomForest yöntemi ile elde edilmiştir. Bu yöntemle zor entübasyon %92,85 duyarlılık, %96,94 özgüllük, %93,69 pozitif öngörü değeri ve %96,52 negatif öngörü değeri ile tahmin edildi. C-L sınıflandırması doğruluk performansı ise %95,60 olarak belirlendi.

**Sonuç:** Yapay zekâ, zor entübasyonu tahmin etmede oldukça başarılı olmuştur. Ayrıca yapay zekâ algoritmaları ile kolay ve zor entübe hastaların C-L sınıflandırmaları başarıyla tahmin edilmiştir. Laringeal görünüm için 6 dereceli modifiye C-L sınıflandırması kullanmak, daha güçlü zor entübasyon tahmini sağlayabilir. Yapay zekâ eğitiminde daha güvenli ve daha güçlü bir tahmin, zor entübasyon tanımını destekleyen bireysel farklılıklar ve klinik özellikler eklenerek elde edilebilir.

**Anahtar kelimeler:** Trakeal entübasyon tahmini; zor entübasyon; yapay zekâ; Cormack-Lehane, entübasyon, anestezi.

## INTRODUCTION

Airway management is the most important clinical skill that anesthesiologists must append. Although major complications are rare during airway management, they are among the most life-threatening causes in medicine (1). This status may lead to major financial medical cases accompanied by catastrophic sequelae such as irreversible brain injury and death (1-4). There is a lack of reliable information about the frequency and nature of major adverse events related to airway management (3). In the UK, airway and respiratory complications have been reported to account for 12% of anesthesia-related claims but, these were described to account for 53% of deaths, 27% of the cost, and 10 of the 50 most costly claims (2). Successful intubation is not always possible due to the patient's anatomical features and systemic diseases. Although the incidence of difficult intubation varies between 1-13%, severe intubation difficulty is generally encountered in 2-3% of patients. Preoperative evaluation can help to identify the difficulty of intubation and take the necessary precautions to deal with the problem. The real danger and risk are that the intubation difficulty is unpredictable (5).

The unpredictable difficult airway is still a serious cause of concern (6). Published algorithms for unexpected difficult or failed tracheal intubation management, devices such as gum elastic bougie, videolaryngoscope and fiberoptic bronchoscope are widely used (7,8). However, the unpredictable difficult airway can lead to significant complications and up to 30% of anesthesia-related deaths (6). Several clinical tests and anthropometric features have been identified during preoperative evaluation to find safe airway management strategies (e.g. Modified Mallampati test [MMT], sternomental distance [SMD], thyromental distance [TMD] and neck circumference). However, the accuracy of difficult intubation prediction is not possible with the evaluation of a single parameter but can be improved by evaluating many other parameters (5,6,9-15). Artificial intelligence is currently used in many fields of medicine to create programs that can perform clinical diagnostic procedures and offer treatment recommendations (16,17). For the literature review, we asked the question "Can artificial intelligence help us predict difficult intubation?". In fact, there was insufficient published literature to provide this answer.

The main aim of our study is to predict difficult intubation with different artificial intelligence algorithms using various clinical tests and anthropometric measurements. Our second aim is to evaluate the accuracy performance of Cormack and Lehane (C-L) classification with artificial intelligence.

## MATERIALS AND METHODS

### Study Design and Ethical Considerations

This single-blind prospective observational study was conducted after the approval of the Ethics Committee of Kırşehir Ahi Evran University (21.09.2016, 10/02; 11.12.2019, 01/01). We planned our research according to current Helsinki guidelines. Before starting the study, we informed the volunteers about the research and received their signed informed consent.

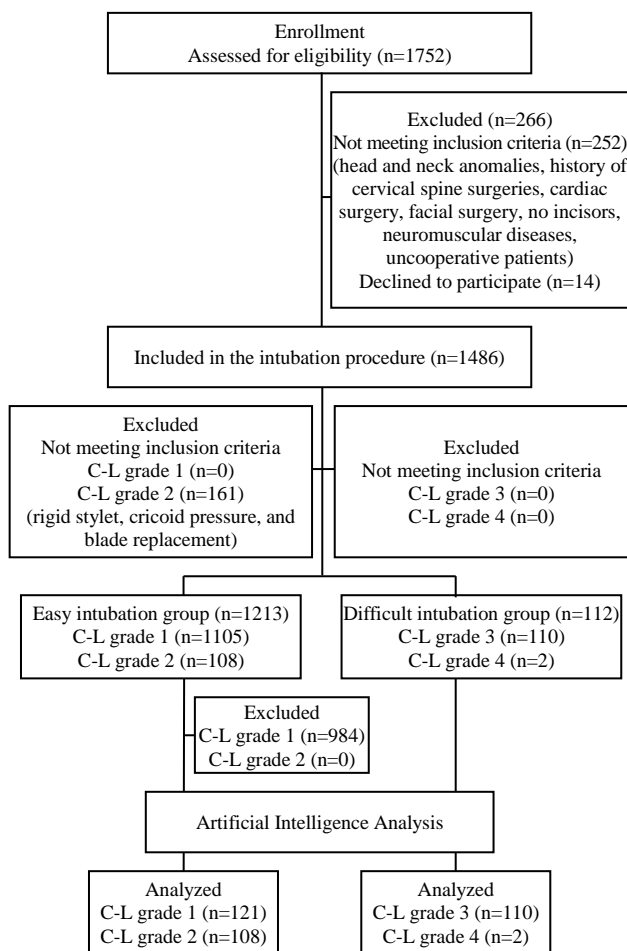
### Patients, Inclusion and Exclusion Criteria

This study included 1486 patients with American Anesthesiologists Association (ASA) physical status I-III,

aged 18-70 years who were scheduled to undergo elective surgery and requiring endotracheal intubation between 2016 and 2019 at Kırşehir Ahi Evran University Training and Research Hospital.

We defined two groups as easy intubation group and difficult intubation group. Intubation was considered as easy when it was performed at the first intubation attempt without the use of any additional intubation aid. Intubation was considered as difficult when there was a need for a gum elastic bougie and/or three or more attempts and/or any additional intubation aid in patients with C-L grades 3 or 4. If the trachea could not be intubated after three attempts at least, it was considered as a failed intubation. Failed intubation patients were included in the difficult intubation group. It was aimed to strengthen artificial intelligence learning. The selection of patients was described in the flow diagram (Figure 1).

Patients with significant head and neck anomalies, history of cervical spine surgeries, cardiac surgery, facial surgery, no incisors, neuromuscular diseases, and uncooperative patients were excluded from the study. Patients with C-L grade 1 and 2 laryngeal views were excluded from the easy intubation group if there was a rigid stylet, cricoid pressure, and blade replacement during intubation. Because, according to our definition of easy intubation, it is accepted as easy intubation when performed at the first intubation attempt without using any additional intubation assistance.



**Figure 1.** Flow diagram of study

Age, sex, height, weight, body mass index, ASA physical status, pregnancy and diabetes mellitus data of patients were recorded during their preoperative visits. Clinical tests and anthropometric measurements (MMT, upper-lip bite test, mouth opening, TMD, SMD, neck circumference, head circumference, neck length, atlanto-occipital joint movement, horizontal mandible length and bigonial width) were recorded by the investigator in charge before the operation.

The patients were evaluated as easy and difficult intubation. Then, the data obtained were taken to the stage of analysis with artificial intelligence. For artificial intelligence learning, analysis of a similar number of patients from C-L grade 1,2,3,4 groups is required. Due to the nature of the study, the number of patients with C-L grade 1 was high. Therefore, patients in the easy intubation group were excluded using a computer-generated random number list. It was paid attention that the number of patients in the C-L grade 1 group was not less than the number of the patients in the C-L grade 1 and 2 groups. Because the number of individuals with C-L grade 1 in the natural distribution is higher in the population.

#### **Predictive Clinical Tests and Anthropometric Measurements**

For anthropometric measurements, depth gauge (0-6 inch, 0-150 mm, ASIMETO® Electronic Depth Gauge) with a standard error of 0.01, standard 12-inch plastic goniometer and flexible tape measure were used.

Airway assessment tests and anthropometric measurements for each patient were defined as follows (Figure 2):

**Modified Mallampati Test:** It is used to determine difficult intubation (18). When the patient was in the sitting position, he was asked to open his mouth as wide as possible and to protrude his tongue as much as possible without phonation. The observer sitting at eye level examined the pharyngeal structures with a light source. According to the examination they were evaluated in four degrees as follows (5,19).

- **Class I:** Uvula, soft palate, tonsils, anterior and posterior pleats are easily visible
- **Class II:** Uvula and soft palate are visible
- **Class III:** Soft palate and uvula base are visible
- **Class IV:** Uvula is completely covered by tongue root, soft palate cannot be seen, the hard palate is visible

**Sternomental Distance:** It was measured as the distance between the upper limit of the manubrium sterni and the mental protuberance in the supine position when the head is in full extension and the mouth is closed.

**Thyromental Distance:** The distance between the superior thyroid notch and mental protuberance was measured in the supine position, while the head is in full extension and the mouth is closed (5,20).

**Mouth Opening:** The patient was asked to open his mouth as much as possible with the head in the neutral position while sitting and the distance between the upper and lower incisors was measured (15).

**Upper-Lip Bite Test:** The patient's ability to bite his upper-lip with his lower incisors was evaluated in three degrees while he was in the sitting position (21).

- **Class I:** The vermilion line of the upper-lip is not completely visible when the lower incisors bite the upper-lip.
- **Class II:** With the same biting maneuver, the upper-lip mucosa appears partially.
- **Class III:** The lower incisors cannot bite the upper-lip.

**Atlanto-Occipital Joint Movement:** The line passing through the tragus and the mouth corner was determined. In the supine patient, the angle between this line and the horizontal line was measured when the head was in full extension (5).

**Neck Circumference:** The neck circumference was measured at the level of the cricoid cartilage while the head was in the neutral position (22).

**Horizontal Mandible Length:** Measured by taking the distance between gonion and mental protuberance.

**Bigonial Width:** Measured by taking the distance between two gonions.

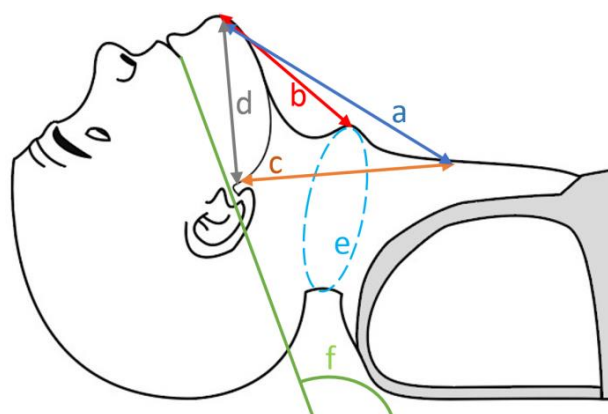
**Neck Length:** Distance between the processus mastoideus and upper medial point of the manubrium sterni was measured while the patient was in the supine position, the head at full extension and the mouth closed (13).

**Head Circumference:** The head circumference was measured by encircling the protuberantia occipitalis externa and the superciliary arches with tape measure (20).

#### **Anesthesia Management**

After the acceptance of the patient to the operation room, patients were monitored with end-tidal carbon dioxide, peripheral oxygen saturation, electrocardiogram, and non-invasive arterial blood pressure. After 5 minutes of preoxygenation, induction was performed with intravenous propofol 2 mg kg<sup>-1</sup>, rocuronium bromide 0.6 mg kg<sup>-1</sup>, fentanyl citrate 2 µg kg<sup>-1</sup>. Later, the patient was ventilated using a standard face mask for 90 seconds. The physician who performed the intubation procedure was blinded to all preoperative measurement information. After evaluating the C-L classification and intubation difficulty, the physician performing the intubation was asked to inform the responsible investigator verbally. Intubation was performed by the same experienced anesthesiologist using the standard Macintosh 4 or 5 blades (HEINE Classic® Macintosh Fiber Optic Blades, Germany) to achieve standardization. The laryngeal view was evaluated according to the C-L scale without applying cricoid pressure while the patient was in sniffing position (23):

- **Grade 1:** Glottis is fully visible
- **Grade 2:** Glottis is partially visible, anterior commissure of the glottis is not visible
- **Grade 3:** No part of the glottis is visible. Only epiglottis is visible
- **Grade 4:** Epiglottis is not visible



**Figure 2.** Anthropometric measurements. **a.** sternomental distance, **b.** thyromental distance, **c.** neck length, **d.** horizontal mandible length, **e.** neck circumference, **f.** atlanto-occipital joint movement

**Artificial Intelligence**

Many different programs are used for artificial intelligence analysis. In this study, the Waikato Environment for Knowledge Analysis (WEKA) Version 3.8 program was used to evaluate the data. WEKA includes several classifier algorithms that can be accessed on the internet developed at the University of Waikato in New Zealand (24). In this study, RandomForest, REPTree, RandomTree, LMT, J48, HoeffdingTree, DecisionStump, JRip, DecisionTable, LWL, KStar, IBk, SMO, SimpleLogistic, MultiLayerPerception, Logistic algorithms were used in WEKA program. Data were separated as training and test data using the 10-fold cross validation method. Correctly Classified Instances, F-Measure and ROC area values were used for the success of the methods. The formulas used to calculate these values are as follows:

$$Sensitivity = \frac{tp}{tp+fn}$$

$$Specificity = \frac{tn}{tn+fp}$$

$$F - Measure = 2 \frac{pr}{p+r}$$

$$Correctly\ Classified\ Instances = \frac{tp+tn}{fp+fn+fp+fn}$$

where, tp represents the true positive, whose actual value is predicted to be positive and positive; fn represents the false negative, whose actual value is positive but predicted as negative, fp represents the false positive, predicted to be positive, but the actual value is negative, and tn represents the true negative, whose actual value is predicted to be negative and negative.

The reported incidence of difficult intubation ranges from 1 to 13% (5). In artificial intelligence learning, a similar number of patients is expected to be taken for each class. Therefore, it was planned to include a similar number of patients from each C-L grade to be able to evaluate the accuracy performance of the C-L classification. In determining this number, the number of difficult intubated patients with low incidence was mainly considered. Thus, a similar number of patient data from each C-L class was analyzed with artificial intelligence. A computer-generated random numbers list was used to select patients with C-L grade 1 from the easy intubation group for artificial intelligence analysis.

**RESULTS**

The number of difficult intubated patients was low by the nature of the study. Therefore, only 112 of 1752 patients were considered as difficult intubation in three years. Among 1752 patients whose eligibility for the study was evaluated, 252 patients were excluded due to head and neck anomalies, history of cervical spine surgery, cardiac surgery, facial surgery, no incisors, neuromuscular diseases, and uncooperative. Among the 1752 patients, 14 refused to participate in the study.

One hundred and sixty-one patients with C-L grade 2 laryngeal views were excluded from the easy intubation group due to the application of a rigid stylet, cricoid pressure, and blade replacement during intubation.

For artificial intelligence learning, analysis of a similar number of patients from C-L grade 1,2,3,4 groups is

necessary. The number of C-L grade 1 patients was high. Therefore, in the easy intubation group, 121 patients with a total of 1105 C-L grade 1 were selected using a computer-generated random number list, and 984 patients were excluded. A total of 229 patients were recorded as easy intubation (Figure 1). Two patients with C-L grade 4 in the difficult intubation class were intubated with flexible fiberoptic bronchoscopy. All 341 patients were successfully intubated. Eventually, data obtained from 341 patients (212, 62.17% male and 129, 37.83% female, aged 18-70 years) were included in the study. The categorical demographic and clinical features of the participants are shown in Table 1. The numerical characteristics of the demographic and anthropometric measurements of the patients are shown in Table 2 as mean, standard deviation and minimum-maximum values.

10-fold cross validation method was used in the separation of data as training and test. Accordingly, the data was divided into 10 parts, the first of which was used for testing and the remaining nine for training. In the second application, the second part was used for testing and the remaining nine parts were used for training. This continued with the application of all parts. Thus, all data were used for both training and testing. This results in more realistic outcomes than randomly allocating data for training and testing. After analyzing the data with artificial intelligence algorithms, the highest classification success was obtained with the RandomForest algorithm. Table 3 shows the C-L classification success of different algorithms.

**Table 1.** Categorical data of demographic and clinical features of the patients, n (%)

	Difficult Intubation (n=112)	Easy Intubation (n=229)	Total (n=341)
<b>Sex</b>			
Male	88 (25.81%)	124 (36.36%)	212 (62.17%)
Female	24 (7.04%)	105 (30.79%)	129 (37.83%)
<b>ASA</b>			
I	20 (5.87%)	102 (29.91%)	122 (35.78%)
II	72 (21.11%)	95 (27.86%)	167 (48.97%)
III	20 (5.87%)	32 (9.38%)	52 (15.25%)
<b>Pregnancy</b>			
Yes	9 (2.64%)	22 (6.45%)	31 (9.09%)
No	103 (30.21%)	207 (60.7%)	310 (90.91%)
<b>DM</b>			
Yes	15 (4.4%)	8 (2.35%)	23 (6.74%)
No	97 (28.45%)	221 (64.81%)	318 (93.26%)
<b>MMT</b>			
I	0 (0.00%)	58 (17.01%)	58 (17.01%)
II	3 (0.88%)	78 (22.87%)	81 (23.75%)
III	52 (15.25%)	68 (19.94%)	120 (35.19%)
IV	57 (16.72%)	25 (7.33%)	82 (24.05%)
<b>ULBT</b>			
I	42 (12.32%)	88 (25.81%)	130 (38.12%)
II	63 (18.48%)	124 (36.36%)	187 (54.84%)
III	7 (2.05%)	17 (4.99%)	24 (7.04%)
<b>C-L</b>			
1	110 (32.26%)	-	112 (32.84%)
2	2 (0.59%)	-	-
3	-	121 (35.48%)	121 (35.48%)
4	-	108 (31.67%)	108 (31.67%)

ASA: American Society of Anesthesiologist physical status, DM: Diabetes Mellitus, MMT: Modified Mallampati Test, ULBT: Upper-Lip Bite Test, C-L: Cormack and Lehane classification

**Table 2.** Descriptive statistics for numerical data of the demographic and anthropometric measurements of the patients

	Difficult Intubation (n=112)		Easy Intubation (n=229)		Total (n=341)	
	Mean±SD	Min-Max	Mean±SD	Min-Max	Mean±SD	Min-Max
Age (y)	46.04±12.00	22-69	40.58±14.42	18-70	42.37±13.91	18-70
Weight (kg)	88.04±11.08	63-123	81.11±14.54	50-140	83.39±13.89	50-140
Height (cm)	169.73±6.87	155-185	168.79±8.93	150-193	169.10±8.32	150-193
BMI (kg/m <sup>2</sup> )	30.61±3.93	21.80-48.05	28.57±5.28	18.71-46.88	29.24±4.97	18.71-48.05
TMD (mm)	65.12±5.23	50.70-75.95	74.94±8.52	50.50-89.56	71.72±8.89	50.50-89.56
SMD (cm)	15.35±1.16	12-18	17.36±1.77	12-21.5	16.70±1.85	12-21.5
Mouth opening (mm)	43.87±5.67	30.25-59.20	46.28±5.14	25.20-59.22	45.49±5.44	25.20-59.22
Neck circumference (cm)	44.21±3.33	38-49.9	43.40±4.01	31-49	43.66±3.82	31-49.9
Neck length (cm)	15.62±1.29	12-18.5	16.66±1.25	12-19	16.32±1.35	12-19
Head circumference (cm)	58.98±1.76	55.5-62	58.79±1.79	56-65	58.85±1.78	55.5-65
AOJM (degree)	88.42±4.23	78-99	101.90±6.85	86-119	97.47±8.80	78-119
HML (mm)	106.04±5.27	92.18-113.89	104.13±6.81	80.25-116.39	104.76±6.41	80.25-116.39
Bigonial Width (mm)	122.32±4.02	110.03-129.90	116.90±5.52	100.92-129.40	118.68±5.68	100.92-129.90

SD: Standard deviation, Min-Max: Minimum-maximum, BMI: Body mass index (weight/(height)<sup>2</sup>), TMD: Thyromental distance, SMD: Sternomental distance, AOJM: Atlanto-occipital joint movement, HML: Horizontal mandible length

When Table 3 is examined, it is seen that Logistic, LMT and SimpleLogistic algorithms have similar classification success to RandomForest algorithm. The results obtained with the RandomForest algorithm, which has the highest success rate, are given in Table 4. The F-Measure value was 0.967 for the classification success rate of the patients grouped as easy intubate classification. According to this,

**Table 3.** C-L classification success with different artificial intelligence algorithms

Algorithm	Classification Success (%)
<b>Trees</b>	
RandomForest	95.60
REPTree	91.49
RandomTree	91.78
LMT	94.13
J48	91.49
Hoeffding Tree	92.66
Decision Stump	86.80
<b>Rules</b>	
JRip	91.49
DecisonTable	88.86
<b>Lazy</b>	
LWL	87.68
KStar	90.32
IBk	88.56
<b>Functions</b>	
SMO	93.84
Simple Logistic	94.42
MultiLayer Perceptron	94.42
Logistic	94.13

the classification success rate is 96.7% in the patients evaluated as easy intubation. Difficult intubation classification success is 93.3%. The average success rate of C-L classification was 95.6%.

The Confusion Matrix table of the RandomForest algorithm with the highest success rate is shown in Table 5. Two hundred and twenty-two of the 229 patients in the easy intubation group were classified as easy intubation and correctly recognized. However, 7 easy intubations were misclassified as difficult intubation. One hundred four of the 112 patients in the difficult intubation group were classified as difficult intubation and correctly recognized. However, 8 difficult intubations were misclassified as easy intubation. It is demonstrated in the confusion matrix that difficult intubation was predicted with 92.86% sensitivity, 96.94% specificity, 93.69% positive predictive value, and 96.52% negative predictive value. Overall accuracy rate was 95.60%.

**Table 5.** The confusion matrix of intubation prediction

		True Value		
		Difficult Intubation	Easy Intubation	
Predicted Value	Difficult Intubation	104	7	93.69%
	Easy Intubation	8	222	96.52%
		92.86%	96.94%	95.60%

**Table 4.** Easy and difficult intubation classification results for RandomForest algorithm

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
Difficult Intubation	0.929	0.031	0.937	0.929	0.933	0.900	0.986	0.968
Easy Intubation	0.969	0.071	0.965	0.969	0.967	0.900	0.986	0.994
Weighted Average	0.956	0.058	0.956	0.956	0.956	0.900	0.986	0.985

## DISCUSSION

Our study demonstrates that difficult intubation can be considerably predicted (92.86% sensitivity) with different artificial intelligence algorithms using various clinical tests and anthropometric measurements. Besides, in this study, C-L classification with artificial intelligence was calculated with a high accuracy rate (95.60%).

A limited number of studies have reported predictions of difficult intubation through artificial intelligence (25-27). The results of our study are consistent with these studies. Yan et al. (25) evaluated the C-L classification prediction only with a multi-layer perceptron network-based medical decision support system. The database was created to train and test the system using 13 features of 824 patients. While they reported C-L classification accuracy as 91.9%, we found 94.42% for the same algorithm. In another study with 13 physical features of 264 patients by Yan et al. (26), C-L classification accuracy was determined as 90.53% with the support vector machine (SVM) based decision support system. In another study with 10 features of 1200 patients by Lazouni et al. (27), the C-L classification success rate was reported as 97.26% with the SVM algorithm. In the previous three studies, a single artificial intelligence algorithm was evaluated. In our study, data of 19 features and 341 patients were assessed by the WEKA program in 16 different artificial intelligence algorithms. The success rates of these studies seem to be high. On the other hand, the selection of patients, measurement of data and standardization of methods were not clearly explained in these three studies. However, the most considerable aspect of our study is it being a single-blind clinical trial. Standardization of our study was achieved by specifying the criteria for inclusion and exclusion of patients, performing intubation by the same experienced physician who was unaware of the patient's clinical test and anthropometric measurement information. Anesthesia induction and tracheal intubation were also standardized. Thus, the average success rate of C-L classification accuracy of easy and difficult intubations, which is our second goal in the study, seems to be stronger than previous studies, despite the possibility of 95.60%.

There is no accepted universal definition for difficult intubation. It is known that definitions vary widely. (e.g. C-L grade 3 and 4 classifications, the need for changing the equipment and the physician who is performing the intubation, number of intubation attempts [more than two or three attempts], duration of intubation lasting more than 10 minutes and the presence of failed intubation; 14,28). Difficult laryngoscopy does not mean difficult intubation. In previous studies, difficult laryngoscopy was predicted according to the laryngeal view. Generally, C-L grade 2 is considered as easy laryngoscopy, but the endotracheal tube may not be inserted into the trachea without the use of any intubation aid. A relative difficulty can be mentioned in patients with some C-L grades 2. For this reason, not all patients classified as C-L grade 2 can be considered easy intubation. In our study, rigid stylet, cricoid pressure, blade replacement was used for 161 patients evaluated as C-L grade 2, so they could not be included in the easy intubation group (Figure 1). Since these patients did not meet our diagnostic criteria for difficult intubation, they could not be identified in the difficult intubation group. In the light of these definitions, not each difficult intubation

will be of equal difficulty. It will also vary if the artificial intelligence recognizes and distinguishes the difficulty spectrum of the intubation. Artificial intelligence predicted difficult intubation with a sensitivity of 92.86%, specificity of 96.94%, positive predictive value of 93.69% and negative predictive value of 96.52% according to our easy and difficult intubation criteria. When we classify the intubation difficulty as easy, difficult and very difficult, we can provide significant improvements for artificial intelligence learning. This classification we recommend can make a distinction beyond easy or difficult classification. On the other hand, tracheal intubation is generally considered to be "easy" or "difficult" in practice. Currently, there are no intubation definitions that classify difficulty as easy, difficult and very difficult. Also, recognizing and distinguishing difficulty in intubation can be improved by adding some individual differences and clinical features. These features may include questioning the snoring and sleep apnea history, rigid stylet or gum elastic bougie use, blade replacement, presence of cricoid pressure, requirement for changing the physician performing intubation, intubation times, etc. Laryngeal view of easy and difficult intubated patients was evaluated with 4-grade C-L classification, which is widely accepted. Recently modified C-L grades are offered to achieve more sensitive classifications. Yentis et al. (29) modified the C-L classification to 5 grades by separating C-L grade 2 into two subclasses: 2a (glottis partial visible) and 2b (posterior part of vocal cords or only arytenoids are visible). Cook TM (28) defined the laryngeal view with a 6-grade modified C-L classification. C-L grade 2 and grade 3 were divided into two subclasses: C-L grade 2a (posterior part of the vocal cords is visible) and 2b (only arytenoids are visible), C-L grade 3a (epiglottis can be seen and lifted) and 3b (epiglottis adherent to pharynx). According to this classification, C-L grade 1 was defined as easy laryngeal view, grade 2a and 3a as restricted, grade 3b and 4 as difficult. The laryngeal view is defined as C-L grade 1 of easy, grade 2a and 3a of restricted, grade 3b and 4 of difficult in this classification. Thus, the use of modified C-L classifications will provide more sensitive data for artificial intelligence algorithms (e.g. artificial intelligence can better distinguish the degree of difficulty between C-L grade 3a and grade 3b in a patient difficult intubated). Although 4-grade C-L classification was used in our study, the accuracy performance of C-L classification was achieved with a substantial rate of 95.60%.

One of the limitations of the present study is that it was a single-center study. If the study can be a multi-center study and carried out in a specific population likely to include a high proportion of difficult intubation, there would be higher estimation success. We are planning a multi-center study for this purpose in the future. Another limitation of the study was the low sample size. There were only two patients in the C-L grade 4 class. Therefore, C-L grade 4 classification accuracy performance could not be assessed, it was evaluated within the C-L grade 3 class.

## CONCLUSIONS

In conclusion, artificial intelligence provided a remarkable distinctive prediction about predicting difficult intubation. Besides, C-L classifications of easy and difficult intubated

patients were successfully predicted with artificial intelligence algorithms. A safer and more potent prediction in training artificial intelligence can be achieved by adding individual differences and clinical features that support the definition of difficult intubation. In further studies, modified 6-grade C-L classifications (2a, 2b, 3a, and 3b) can be used to predict difficult intubation. This will provide us with more detailed and powerful defined patient data. Future, we believe that artificial intelligence can safely achieve difficult intubation prediction and management with an anesthetist's perspective.

**Ethics Committee Approval:** The study was approved by the Ethics Committee of Kırşehir Ahi Evran University (21.09.2016, 10/02; 11.12.2019, 01/01).

**Conflict of Interest:** None declared by the authors.

**Financial Disclosure:** None declared by the authors.

**Acknowledgements:** For his support in the creation of this study, we would like to thank Recai DAĞLI, MD from Kırşehir Ahi Evran University Faculty of Medicine Department of Anesthesiology and Reanimation.

**Author Contributions:** Idea/Concept: FÇ, EA; Design: FÇ, EA; Data Collection/Processing: FÇ; Analysis/Interpretation: FÇ, EA; Literature Review: FÇ; Drafting/Writing: FÇ, EA; Critical Review: FÇ, EA.

## REFERENCES

- Cook TM, MacDougall-Davis SR. Complications and failure of airway management. *Br J Anaesth.* 2012;109(Suppl 1):i68-i85.
- Cook TM, Scott S, Mihai R. Litigation related to airway and respiratory complications of anaesthesia: an analysis of claims against the NHS in England 1995-2007. *Anaesthesia.* 2010;65(6):556-63.
- Cook TM, Woodall N, Frerk C, Fourth National Audit Project. Major complications of airway management in the UK: results of the Fourth National Audit Project of the Royal College of Anaesthetists and the Difficult Airway Society. Part 1: anaesthesia. *Br J Anaesth.* 2011;106(5):617-31.
- Peterson GN, Domino KB, Caplan RA, Posner KL, Lee LA, Cheney FW. Management of the difficult airway: a closed claims analysis. *Anesthesiology.* 2005;103(1):33-9.
- Kayhan Z. Endotrakeal Entübasyon. In: *Klinik Anestezi Genişletilmiş 3. Baskı.* İstanbul: Logos Yayıncılık; 2004. p.243-73.
- de Carvalho CC, da Silva DM, de Carvalho Junior AD, Santos Neto JM, Rio BR, Neto CN, et al. Pre-operative voice evaluation as a hypothetical predictor of difficult laryngoscopy. *Anaesthesia.* 2019;74(9):1147-52.
- Frerk C, Mitchell VS, McNarry AF, Mendonca C, Bhagrath R, Patel A, et al. Difficult Airway Society 2015 guidelines for management of unanticipated difficult intubation in adults. *Br J Anaesth.* 2015;115(6):827-48.
- Marshall SD, Pandit JJ. Radical evolution: the 2015 Difficult Airway Society guidelines for managing unanticipated difficult or failed tracheal intubation. *Anaesthesia* 2016 71(2):131-7.
- Roth D, Pace NL, Lee A, Hovhannisyan K, Warenits AM, Arrich J, et al. Bedside tests for predicting difficult airways: an abridged Cochrane diagnostic test accuracy systematic review. *Anaesthesia.* 2019;74(7):915-28.
- Karkouti K, Rose DK, Wrigglesworth D, Cohen MM. Predicting difficult intubation; a multivariable analysis. *Can J Anesth.* 2000;47(8):730-9.
- Srivilaithon W, Muengtawepongsa S, Sittichanbuncha Y, Patumanond J. Predicting difficult intubation in emergency department by intubation assessment score. *J Clin Med Res.* 2018;10(3):247-53.
- Cattano D, Panicucci E, Paolicchi A, Forfori F, Giunta F, Hagberg C. Risk factors assessment of the difficult airway: an Italian survey of 1956 patients. *Anesth Analg.* 2004;99(6):1774-9.
- Wilson ME, Spiegelhalter D, Robertson JA, Lesser P. Predicting difficult intubation. *Br J Anaesth.* 1988;61(2):211-6.
- Yildiz TS, Korkmaz F, Solak M, Toker K, Erciyes N, Bayrak F, et al. Prediction of difficult tracheal intubation in Turkish patients: a multi-center methodological study. *Eur J Anaesthesiol* 2007;24(12):1034-40.
- Savva D. Prediction of difficult tracheal intubation. *Br J Anaesth.* 1994;73(2):149-53.
- Demirhan A, Kılıç YA, Güler İ. Artificial intelligence applications in medicine. *Yoğun Bakım Dergisi.* 2010;9(1):31-41.
- Connor CW. Artificial intelligence and machine learning in anesthesiology. *Anesthesiology.* 2019;131(6):1346-59.
- Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraksa B, Freiburger D, et al. A clinical sign to predict difficult tracheal intubation: a prospective study. *Can Anaesth Soc J.* 1985;32(4):429-34.
- Samsoon GL, Young JR. Difficult tracheal intubation: a retrospective study. *Anaesthesia.* 1987;42(5):487-90.
- Kandemir T, Şavlı S, Ünver S, Kandemir E. Sensitivity of the combination of mallampati scores with anthropometric measurements and the presence of malignancy to predict difficult intubation. *Turk J Anaesthesiol Reanim.* 2015;43(1):7-12.
- Khan ZH, Kashfi A, Ebrahimkhani E. A comparison of the upper lip bite test (a simple new technique) with modified Mallampati classification in predicting difficulty in endotracheal intubation: a prospective blinded study. *Anesth Analg.* 2003;96(2):595-9.
- Acer N, Akkaya A, Tugay BU, Öztürk A. A comparison of corneck-lehane and mallampati tests with mandibular and neck measurements for predicting difficult intubation. *Balkan Med J.* 2011;28(2):157-63.
- Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. *Anaesthesia* 1984;39(11):1105-11.
- Aydemir E. *Weka ile yapay zeka.* Ankara: Seçkin Yayınevi; 2018.
- Yan HM, Wei XC, Zhang H, Chen XF, Luo EQ. Predicting Cormack classification based on neural network with multiple anthropometric features. *The 2010 International Conference on Apperceiving Computing and Intelligence Analysis Proceeding.* 2010:52-5. doi: 10.1109/ICACIA.2010.5709849.

26. Yan Q, Yan H, Han F, Wei X, Zhu T. SVM-based decision support system for clinic aided tracheal intubation predication with multiple features. *Expert Syst Appl.* 2009;36(3):6588-92.
27. Lazouni MEA, Settouti N, Daho MEH, Mahmoudi S, Chikh, A. Automatic detection of difficult tracheal intubation. 2014 International Conference on Multimedia Computing and Systems (ICMCS). 2014:453-8. doi: 10.1109/ICMCS.2014.6911235.
28. Cook TM. A new practical classification of laryngeal view. *Anaesthesia.* 2000;55(3):274-9.
29. Yentis SM, Lee DJ. Evaluation of an improved scoring system for the grading of direct laryngoscopy. *Anaesthesia.* 1998;53(11):1041-4.