

Comparison of the Success of Emergency Intubation with the Heaven Score and The Lemon Score in Critically Ill Patients Who Require Mechanical Ventilation

T. Abdurrahman Johan, Andriamuri Primaputra Lubis[✉], Muhammad Ihsan

Department of Anesthesiology and Intensive Therapy, Faculty of Medicine, Universitas Sumatera Utara/Adam Malik General Hospital, Medan, Indonesia

[✉]Correspondence: andriamuri@usu.ac.id

ABSTRACT

Background: The LEMON and HEAVEN scores are currently used only to assess airway difficulties in emergency patients. This research was conducted to evaluate the success of emergency intubation using the LEMON and HEAVEN scores in critically ill patients who require mechanical ventilation.

Objective: The objective of this study is to compare the success of emergency intubation with the HEAVEN score with the LEMON score in critically ill patients requiring mechanical ventilation.

Methods: This research employed a cross-sectional observational analytical design and took place at Adam Malik General Hospital, Medan, Dr. Pirngadi Medan, and RSU Haji Medan from January to March 2024. All critically ill patients aged 18-64 years requiring emergency intubation were included, excluding those with prior tracheotomy or intubation experience who refused to participate. A minimum sample size of 70 participants was obtained, with statistical analysis planned using SPSS version 26 software.

Results: A total of 76 patients were included in this study. According to the LEMON score, 5 patients (6.6%) were identified as having difficult intubation, whereas the HEAVEN score categorized 30 patients (39.5%) as difficult cases. Following emergency intubation, only 3 patients (3.9%) experienced initial intubation failure. The LEMON score demonstrated an area under the curve (AUC) of 0.984, indicating excellent accuracy, while the HEAVEN score had an AUC of 0.911, also demonstrating excellent accuracy. Both scores proved equally effective in assessing intubation success, with statistical analysis favoring the superiority of the LEMON score. However, the HEAVEN score can serve as a viable alternative in critically ill patients.

Conclusion: The LEMON score and HEAVEN score can be used to evaluate the success of emergency department intubation in critically ill patients who require mechanical ventilation.

Keywords: critical ill patients; emergency intubation; HEAVEN score; LEMON score; mechanical ventilation

INTRODUCTION

Airway management outside the operating room is used to resuscitate critically ill and seriously ill patients who are physiologically unstable or to secure an emergency airway when necessary.¹ In a study conducted by Amaha et al in 2021, the prevalence of patients experiencing difficult intubation was 11% in patients who were carried out outside the operating room.²

Critically ill patients have less physiological reserve oxygen capacity, and the risk of complications related to endotracheal intubation is higher in the intensive care unit (ICU) than in the operating room. Adverse events may occur during the induction of anesthesia and after the patient is intubated.³ Critically ill patients, especially those treated in the ICU, have various conditions that make it difficult to carry out emergency intubation. Emergency conditions and decreased physiologic oxygen reserves contribute significantly to the increased risk of severe peri-intubation hypoxemia, hypotension, arrhythmia, cardiac arrest, and death.⁴ Delay in tracheal intubation and multiple attempts at laryngoscopy are associated with increased complications, including death.⁵

For many years, anesthesiologists have tried to predict difficult airways using a variety of clinical and predictive signs. Of the 3,391 difficult intubations, 3,154 (93%) were unanticipated. When difficult intubation was anticipated, only 229/929 (25%) actually experienced difficult intubation. Existing difficult airway assessment tools, such as LEMON (Look externally, evaluate 3-3-2 rule, mallampati, obstruction, neck mobility), are designed for use in elective surgical settings that require patient cooperation. Therefore, such

prediction tools may have limited use in emergency airway management. In addition, the prediction tool does not include physiological factors. Most difficult airway prediction tools only consider anatomical factors without physiological factors, such as the LEMON score. From previous studies, it is known that the LEMON score can be effectively used in emergencies to predict difficult intubation because it can be determined only by the patient's clinical appearance and the observer's finger count, and does not require a specific cut-off value or additional examination tools. However, the Mallampati component of the LEMON score is difficult to observe in trauma cases requiring emergency intubation.^{5,6}

Recently, Davis and Olvera proposed the HEAVEN criteria as a new difficult airway prediction tool for emergency airway management.⁷ HEAVEN is an airway prediction tool that combines physiological and anatomical factors. In addition, the HEAVEN criteria are more relevant and can be applied in emergency airway assessment because they take into account physiological factors, namely hypoxemia (oxygen saturation $\leq 93\%$ at the start of laryngoscopy) and exsanguination-suspected anemia, both chronic and acute.⁸ HEAVEN criteria were also associated with a reduced incidence of success on the first attempt with OR 0.10 (05% CI 0.02–0.43, $p<0.01$).⁹ The study also found that intubation complications occurred more frequently in the group with hypoxemia and vomit/blood/fluid.¹⁰ Another study of 2,419 patients undergoing rapid sequence intubation found that success on the first attempt was lower for each HEAVEN criterion, where intubation success without the HEAVEN criteria was 94%, and with 5 criteria, the success was 43%.¹²

This study aimed to assess the difference in the success of emergency department intubation with the HEAVEN score and the LEMON score in critically ill patients who require mechanical ventilation.

METHOD

This cross-sectional observational analytical study aimed to compare the efficacy of emergency intubation using the HEAVEN and LEMON scores among critically ill patients requiring mechanical ventilation at Adam Malik General Hospital, Medan, Dr. Pirngadi Medan, and RSU Haji Medan from January to March 2024. To minimize bias, incidental data collection was employed, ensuring all potential subjects underwent consistent diagnostic procedures. The study included all critically ill patients aged 18-64 years undergoing emergency intubation, excluding those with prior tracheostomy or intubation experience, as well as those declining to participate.

The sample size was calculated using the formula for analytical correlational numeric research within a single population, and the minimum number of samples required was 70. The research proceeded only after obtaining approval from the Ethics Committee and consent from the participants.

In the data collection process, each standardized data collector obtained patient observation sheets, which included patient data, pre-intubation hemodynamic data, post-intubation hemodynamic data, and intubation success data. The LEMON score and HEAVEN score were observed for each participant. The HEAVEN score, a method for assessing intubation difficulty based on evaluating the airway with scores ranging from 0 to 6, was

evaluated. The LEMON score, one of the tools used to evaluate the airway with scores ranging from a minimum of 0 to a maximum of 10, was also assessed.

Statistical analyses were performed using SPSS version 26. Categorical data were summarized using frequencies and percentages, while numerical data were presented as mean \pm standard deviation. To evaluate the success of intubation on the first attempt, inferential analyses utilized the Chi-Square test. Fisher's Exact test was employed if any expected cell count was less than 5. The choice between parametric (t-test) and non-parametric (Mann-Whitney test) methods for comparing means between two groups depends on the results of the Kolmogorov-Smirnov normality test. Statistical significance was set at a p-value less than 0.05.

Research methods are research procedures and techniques. Written clearly and concisely so that it is easy for readers to understand and can repeat similar research (reproduced). Writing formulas using mathematical equations. The format of tables and figures can be seen in the example below.

RESULTS

A total of 76 subjects met the study criteria. Table 1 summarizes their baseline characteristics, with most patients being female (65.8%), with an average age of 43.97 ± 15.66 years, and an average heart rate of 80.36 ± 10.75 times/minute, with a mean MAP of 89.30 ± 8.64 mmHg. Based on the success rate, emergency intubation was successful in 73 subjects (96.1%) and only failed in 3 subjects (3.9%). Based on the interpretation of the LEMON score, 5 subjects (6.6%) were categorized as having difficult intubation, while 30 subjects (39.5%) were categorized as

having difficult intubation based on the interpretation of the HEAVEN score. Based on complications, 5 subjects were found to have dysrhythmia (6.6%), 1 with aspiration (1.3%), 2 with hypoxemia (2.6%), and 5 with other complications (6.6%).

The HEAVEN criteria parameters in Table 2, it was found that anatomical challenge, neck mobility abnormalities, and vomit/blood/fluid were significantly related to the success of emergency intubation in critically ill patients who required mechanical ventilation ($p<0.05$).

The LEMON criteria parameters in Table 3, it was found that Malampati score >3 , obstruction, and neck mobility abnormalities were significantly related to the success of emergency intubation in critically ill patients who required mechanical ventilation ($p<0.05$).

Based on Table 4, the mean LEMON score in the group that was successful with emergency intubation was 1.89 ± 1.17 , while the mean in the group that was unsuccessful was 4.33 ± 2.08 . The difference in scores between the two groups was considered statistically significant ($p<0.001$). The comparison of the HEAVEN score in the group that was successful with emergency intubation was 0.55 ± 0.82 , while for the group that was unsuccessful, it was

1.33 ± 1.53 . Statistically, the difference in HEAVEN scores between the successful group and the unsuccessful group was considered significant ($p=0.009$).

The receiver operating characteristic (ROC) curve for the LEMON and HEAVEN scores as predictors of successful emergency intubation in critically ill patients who require mechanical ventilation (Figure 1). Based on Table 5, the cut-off value, sensitivity, and specificity of the LEMON score are 3, 100.0%, and 95.9%. The cut-off values, sensitivity, and specificity of the HEAVEN score are 2, 66.7%, and 97.3%. The LEMON ($AUC = 0.984$; $p = 0.005$) and HEAVEN ($AUC = 0.911$; $p = 0.016$) scores were found to have excellent accuracy as predictors of successful emergency intubation.

Based on Table 6, it was found that 35.2% of intubations were declared difficult with the HEAVEN score but were declared easy with the LEMON score, while 100% of intubations declared difficult with the HEAVEN score were also declared difficult with the LEMON score. There was a significant difference between determining intubation difficulty using the HEAVEN score and the LEMON score ($p<0.05$).

Table 1. Characteristics of a critically ill sample requiring mechanical ventilation

Characteristics	Amount
Gender	
Man	26 (34.2%)
Woman	50 (65.8%)
Age (years)	
Mean ± SD	43.97 ± 15.66
Heart rate (times/minute)	
Mean ± SD	80.36 ± 10.75
Mean arterial pressure (mmHg)	
Mean ± SD	89.30 ± 8.64
Body mass index (kg/m ²)	
≤25	17 (22.4%)
>25	59 (77.6%)
Scores and Outcomes	
LEMONS	
Easy	71 (93.4%)
Difficult	5 (6.6%)
HEAVEN	
Easy	46 (60.5%)
Difficult	30 (39.5%)
Success	
Succeed	73 (96.1%)
Not successful	3 (3.9%)
Complications	
There isn't any	63 (82.9%)
Dysrhythmia	5 (6.6%)
Aspiration	1 (1.3%)
Hypoxemia	2 (2.6%)
Other complications	5 (6.6%)

SD: standard deviation

Table 2. Correlation of HEAVEN parameters with the success of emergency intubation

HEAVEN Criteria	Emergency Intubation		p-value
	Succeed	Fail	
Hypoxemia			0.848
• Yes	4 (100.0%)	0 (0.0%)	
• No	69 (95.8%)	3 (4.2%)	
Anatomical Challenge			<0.001*
• Yes	3 (50.0%)	3 (50.0%)	
• No	70 (100.0%)	0 (0.0%)	
Neck Mobility			<0.001*
• Yes	4 (57.1%)	3 (42.9%)	
• No	69 (100.0%)	0 (0.0%)	
Vomit/blood/fluid			0.021*
• Yes	5 (71.4%)	2 (28.6%)	
• No	68 (98.6%)	1 (1.4%)	
Extreme of Size			0.124
• Yes	15 (88.2%)	2 (11.8%)	
• No	58 (98.3%)	1 (1.7%)	
Exanguination			0.813
• Yes	5 (100.0%)	0 (0.0%)	
• No	68 (95.8%)	3 (4.2%)	

*Fisher's Exact Test

Table 3. Relationship between LEMON parameters and the success of emergency intubation

LEMON Criteria	Emergency Intubation		p-value
	Successful (n = 73)	Failed (n = 3)	
External Abnormalities			0.111
• Yes	34 (91.9%)	3 (8.1%)	
• No	39 (100.0%)	0 (0.0%)	
Abnormality Rule 3-3-2			0.071
• Yes	29 (90.6%)	3 (9.4%)	
• No	44 (100.0%)	0 (0.0%)	
Mallampati >3			0.029*
• Yes	21 (87.5%)	3 (12.5%)	
• No	52 (100.0%)	0 (0.0%)	
Obstruction			0.002*
• Yes	7 (70.0%)	3 (30.0%)	
• No	66 (100.0%)	0 (0.0%)	
Neck Mobility			<0.001*
• Yes	4 (57.1%)	3 (42.9%)	
• No	69 (100.0%)	0 (0.0%)	

*Fisher's Exact Test

Table 4. Comparison of LEMON and HEAVEN scores based on emergency intubation success rates

Parameter	Succeed (n = 73)	Not successful (n = 3)	p-value
LEMONS			
Mean ± SD	1.89 ± 1.17	4.33 ± 2.08	<0.001*
Median (min-max)	2 (1 - 7)	5 (2 - 6)	
HEAVEN			
Mean ± SD	0.55 ± 0.82	1.33 ± 1.53	0.009*
Median (min-max)	0 (0 - 3)	1 (0 - 3)	

*Mann-Whitney test

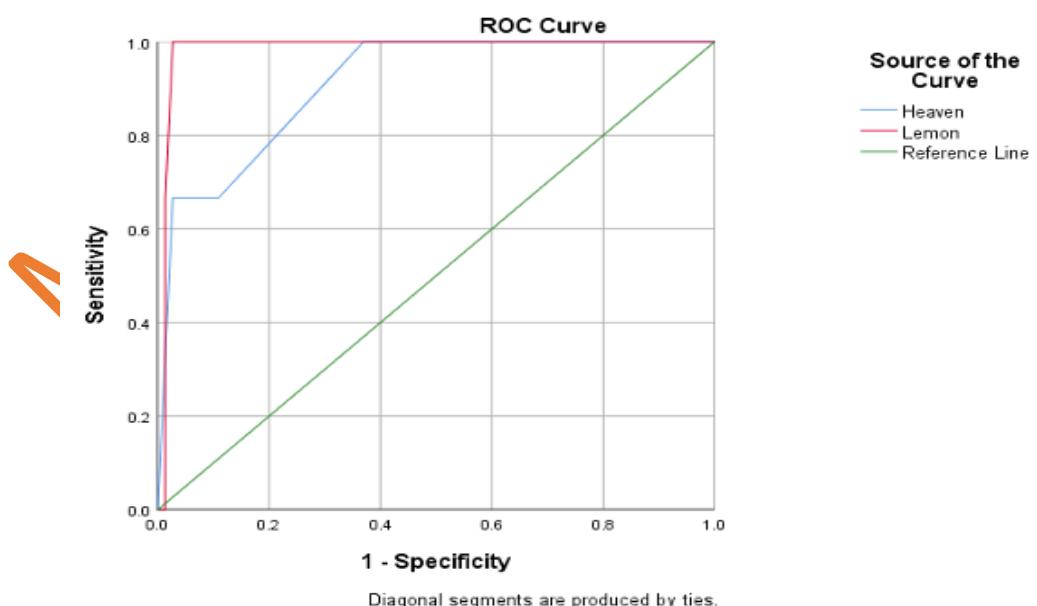
**Figure 1.** ROC curves for LEMON and HEAVEN scores as predictors of successful emergency intubation

Table 5. Cut-off, sensitivity, and specificity of LEMON and HEAVEN scores as predictors of successful emergency intubation

Score	AUC	Cut-off	Sensitivity	Specificity	p-value
LEMONS	0.984	3	100.0%	95.9%	0.005*
HEAVEN	0.911	2	66.7%	97.3%	0.016*

Table 6. Comparison of intubation difficulty using HEAVEN score and LEMON score

Criteria	HEAVEN				p-value
	Easy		Difficult		
	n	%	n	%	
LEMONS	Easy	46	64.7%	25	35.2%
	Difficult	0	0%	5	100%

DISCUSSION

From the results of this assessment, we found that there were 5 subjects (6.6%) who were categorized as difficult intubation cases based on the LEMON score. As a result, we found that only 3 patients (3.9%) experienced failed intubation on the first attempt. From several previous studies, the sensitivity and specificity of the LEMON score for assessing intubation difficulty are quite high. Sengel et al's study reported that for the cut-off LEMONS \geq 4, sensitivity and specificity for predicting difficult intubation are 100% and 91.8%. On the other hand, for cut-off LEMON \geq 5, sensitivity and specificity for prediction are 90.9% and 98.4%.¹¹ In this study, we also assessed the relationship of parameters in the LEMON score to the success of emergency department intubation in critically ill patients. As a result, we found that only mallampati >3 , obstruction, and impaired neck mobility had a significant relationship with the success of emergency intubation ($p<0.05$).

In a study conducted by Derakhshan et al., only impaired neck mobility was reported as a clinical finding that influenced the success of intubation ($OR = 6.152$, 95% CI [1.909-19.821], $p = 0.002$).¹⁰ On the other hand, in a study conducted by Jung and Kim¹³ on patients who received emergency intubation in

the ER, it was found that restriction of mouth opening ($OR=0.47$ [95%CI] [0.31-0.72]), restricted neck mobility ($OR= 0.57$ [95%CI] [0.39-0.85]), and swollen tongue ($OR=0.46$ [95%CI] [0.28-0.77]) as factors that influence the success of emergency intubation. To date, the use of LEMON to predict the success of intubation in critical patients requiring mechanical ventilation has not been reported. Based on the findings in this study, it can be concluded that the success of emergency intubation in critically ill patients who require mechanical ventilation will be higher if the overall LEMON score is ≤ 3 .

The HEAVEN score is one of the scores to assess the success of emergency intubation in this study. HEAVEN Score ≥ 1 was defined as difficult intubation cases. We found that there were at least 30 subjects (39.5%) who were categorized as difficult intubation based on the HEAVEN score assessment. As a result, only 3 patients experienced failed intubation on the first attempt. Based on the results of the analysis of the parameters in the HEAVEN criteria, we found that anatomical challenges ($p<0.001$), neck mobility abnormalities ($p<0.001$), and vomit/blood/fluid ($p=0.021$) were parameters that had a significant relationship with the success of emergency intubation in critically ill patients requiring mechanical ventilation.

Unfortunately, studies documenting the sensitivity and specificity of the HEAVEN criteria are still very limited. One includes the study by Kuzmack et al.¹⁴ In that study, a cut-off value of 2 was reported to provide sensitivity and specificity of 47% and 75%. In our study, an analysis of the cut-off of HEAVEN was also carried out. As a result, cut-off value 2 has a sensitivity and specificity of 66.7% and 97.3%. On the other hand, the AUC value obtained from the HEAVEN score ROC curve is 0.911 (excellent accuracy). We found that the mean of the HEAVEN score in the group who successfully received emergency intubation was 0.55 ± 0.82 , while the mean in the group who failed was 1.33 ± 1.53 . The difference between the two groups is also statistically significant ($p=0.009$). Based on the findings in our study, it can be concluded that the HEAVEN value ≤ 2 can be used as a reference in determining the success of emergency intubation in critically ill patients who require mechanical ventilation.

This study represents the first documented use of the LEMON score and HEAVEN score to evaluate the success of emergency department intubation in critically ill patients requiring mechanical ventilation. Additionally, the study presents thresholds, sensitivity, and specificity of these scores as predictors of successful emergency intubation. Based on our findings, it can be concluded that utilizing the LEMON and HEAVEN scores to assess the success of emergency intubation in critically ill patients requiring mechanical ventilation demonstrates a high level of accuracy. However, our study encountered several limitations. Initially, the sample size was relatively small. Additionally, there was an imbalance in group comparisons, with

only three patients included in the failure group. Such substantial differences between groups can introduce bias in the statistical analysis and interpretation of results. Furthermore, we observed that the cutoff value obtained was lower compared to previous studies. This discrepancy was likely influenced by the disproportionate distribution of samples between the successful and unsuccessful groups, leading to a decrease in the cutoff values for both the LEMON and HEAVEN scores. Given these limitations, it is essential to conduct additional studies to confirm the findings of our research.

The ability to assess the difficulty of intubation is influenced by factors such as the number and ability of operators and the measurement technique. The differences in results in this study were also due to different intubation operators for each patient. This causes discrepancies in assessment and implementation. In research, Nakazawa et al stated that the localization of the cricothyroid membrane identified by more skilled operators did not change when the neck was extended.¹⁵

CONCLUSION

The LEMON score and HEAVEN score can be used to assess the success of emergency department intubation in critically ill patients who require mechanical ventilation.

REFERENCES

1. Admass BA, Endalew NS, Tawye HY, et al. Evidence-based airway management protocol for a critically ill patient in a medical intensive care unit: Systematic review. Ann Med Surg 2022; 80: 1042-84.

2. Cabrini L, Landoni G, Baiardo Redaelli M, et al. Correction to: Tracheal intubation in critically ill patients: A comprehensive systematic review of randomized trials. *Crit Care* 2019; 23: 1–9.
3. Natt B, Mosier J. Airway Management in the Critically Ill Patient. *Curr Anesthesiol Rep* 2021; 11: 116–127.
4. Cumberworth A, Lewith H, Sud A, et al. Major complications of airway management: a prospective multicentre observational study. *Anaesthesia* 2022; 77: 640–648.
5. Srivilaithon W, Muengtaweepongsa S, Sittichanbuncha Y, et al. Predicting difficult intubation in the emergency department by the intubation assessment score. *J Clin Med Res* 2018; 10: 247–253.
6. Crewdson K, Fragoso-Iniguez M, Lockey DJ. Requirement for urgent tracheal intubation after traumatic injury: a retrospective analysis of 11,010 patients in the Trauma Audit Research Network database. *Anaesthesia* 2019; 74: 1158–1164.
7. Davis DP, Olvera DJ. HEAVEN criteria: derivation of a new difficult airway prediction tool. *Air Med J* 2017; 36: 195–197.
8. Jose S, Pulicken M. Predictability of Heaven Criteria for Difficult Airway in a Tertiary Care ICU. *Indian J Crit Care Med.* 2025;29(Suppl 1):S60.
9. Hayes-Bradley C, et al. Predictive and protective factors for failing first pass success in emergency airway management: cohort analysis. *Br J Anaesth.* 2024;133(4):442–450.
10. Tan NE, Yoong KPY, Yahya HMF. Use of HEAVEN criteria for predicting difficult intubation in the emergency department. *Clin Exp Emerg Med* 2022; 9: 29–35.
11. Şengel N, Karabulut G, Toprak ME, et al. Predictive values of LEMON and LEON scores in elective oral and maxillofacial surgery- correlation with intubation difficulty scale. *Yüksek İhtisas Üniversitesi Sağlık Bilim Derg* 2022; 3: 33–38.
12. Derakhshan P, Nikoubakht N, Alimian M, et al. Relationship between airway examination with LEMON criteria and difficulty of tracheal intubation with IDS criteria. *Anesthesiol Pain Med*; 13.
13. Jung W, Kim J. Factors associated with first-pass success of emergency endotracheal intubation. *Am J Emerg Med* 2020; 38: 109–113.
14. Kuzmack E, Inglis T, Olvera D, et al. A novel difficult-airway prediction tool for emergency airway management: validation of the HEAVEN criteria in a large air medical cohort. *J Emerg Med* 2018; 54: 395–401.
15. Nakazawa H, Uzawa K, Tokumine J, Lefor AK, Motoyasu A, Yorozu T. Airway ultrasound for patients anticipated to have a difficult airway: Perspective for personalized medicine. *World Journal of Clinical Cases.* 2023 Mar 26; 11(9): 1951–1962.