

# An alternative perspective on triage systems: the Progressive Real-world Optimization of Triage System (PROGRESS) study

Arian Zaboli,<sup>1</sup> Francesco Brigo,<sup>1</sup> Serena Sibilio,<sup>2</sup> Gloria Brigiari,<sup>3</sup> Magdalena Massar,<sup>1</sup> Norbert Pfeifer,<sup>4,5</sup> Marta Parodi,<sup>6</sup> Eleonora Rella,<sup>4,5</sup> Gianni Turcato<sup>6</sup>

<sup>1</sup>Innovation, Research and Teaching Service (SABES-ASDAA), Teaching Hospital of the Paracelsus Medical Private University (PMU), Bolzano, Italy; <sup>2</sup>Universität Basel Department Public Health, Institute of Nursing Science, Basel, BS, Switzerland; <sup>3</sup>Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Padova, Italy; <sup>4</sup>Department of Emergency Medicine, Hospital of Merano-Meran (SABES-ASDAA), Merano-Meran, Italy; <sup>5</sup>Lehrkrankenhaus der Paracelsus Medizinischen Privatuniversität, Salzburg, Austria; <sup>6</sup>Department of Internal Medicine, Intermediate Care Unit, Hospital Alto Vicentino (AULSS-7), Santorso, Italy

## Abstract

Triage systems have remained largely unchanged since the 1990s and rely on expert consensus, with no single system consistently outperforming others in accurately identifying critically ill or urgent patients. This study aimed to determine whether

incorporating additional tools improves the predictive accuracy of the Manchester Triage System (MTS). A prospective, monocentric study was conducted at Merano Hospital (Italy) from June 1st to December 31st, 2023. A triage nurse and two ED physicians assigned patient priorities. The cohort was split for model derivation and validation. An ordinal logistic regression model was developed using MTS, the National Early Warning Score, and the Charlson Comorbidity Index, then tested on a validation cohort, bootstrapped to 5000 cases. Of the 1270 patients enrolled, 821 were in the derivation cohort and 449 in the validation cohort. The model outperformed MTS alone in most outcomes, except for predicting death at 72 hours and 7 days. Decision Curve Analysis confirmed its superiority in identifying urgent cases. Integrating multiple tools into triage models can enhance their performance, improving patient prioritization accuracy.

Correspondence: Arian Zaboli, Innovation, Research and Teaching Service (SABES-ASDAA), Teaching Hospital of the Paracelsus Medical Private University (PMU), Bolzano, Italy.  
E-mail: zaboliarian@gmail.com

Key words: triage; accuracy; patient prioritization; early warning score; emergency nurse; triage model; logistic models.

Contributions: AZ, conceptualization, methodology, investigation, formal analysis, data curation, writing-original draft preparation; SS, conceptualization, investigation; MM, investigation, data curation; GB, formal analysis; ER, investigation; MP, investigation; NP, supervision, resources; FB, supervision, writing-original draft preparation, writing - review & editing; GT, conceptualization, methodology, formal analysis, writing-original draft preparation.

Conflict of interest: the authors declare no conflict of interest.

Funding statement: Not applicable.

Ethics approval: the study was conducted with the approval of local ethic committees (Ethics Committee Clinical Experiments of Bolzano, Italy, approval number 86-2023), according to the ethical principles for medical research involving human subjects of the Declaration of Helsinki.

Availability of data and materials: all data generated or analyzed during this study are included in this published article.

Received: 19 October 2024.

Accepted: 4 February 2025.

Early view: 27 February 2025.

This work is licensed under a Creative Commons Attribution 4.0 License (by-nc 4.0).

©Copyright: the Author(s), 2025

Licensee PAGEPress, Italy

Emergency Care Journal 2025; 21:13269

doi:10.4081/ecj.2025.13269

*Publisher's note: all claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.*

## Introduction

The current triage systems, conceived in the 1990s,<sup>1,2</sup> were developed through consensus among experts in the field.<sup>3,4</sup> At the time of their inception, this approach proved to be effective and functional.<sup>1,5</sup> Since then, triage systems have remained largely unchanged, except for the increase in risk stratification levels from 3 or 4 to 5.<sup>6</sup> However, currently, none of the triage systems stands out in terms of performance compared to the others, indicating their suboptimal ability to reliably identify patients with critical and time-dependent conditions.<sup>7,8</sup>

A challenging issue is the transformation undergone by Emergency Departments (EDs) in recent decades, marked by an increase in attendances for non-urgent conditions and a higher prevalence of elderly patients and individuals with comorbidities.<sup>9,10</sup> This trend results from population aging and the improved management of various chronic conditions.<sup>11</sup> While these changes are widely recognized by ED healthcare personnel and researchers, they have not been actively considered for modifying or enhancing current triage systems. Instead, there has been a tendency to explore scales or scores aimed at improving risk prediction during triage assessment.<sup>12,13</sup>

There persists a necessity to evaluate whether tools measuring frailty, functionality, acuity, and comorbidity could be effectively integrated into a new predictive model. The present study aimed to assess whether the incorporation of these additional tools improves the predictive capability of the Manchester Triage System (MTS), one of the most widely used triage systems.

## Materials and Methods

### Study design and setting

The prospective observational study was conducted at the ED of Merano Hospital, Italy, from June 1, 2023, to December 31, 2023. This ED utilizes the Manchester Triage System (MTS) as its triage protocol since 2014.<sup>14</sup> The MTS is a standardized system that categorizes patients into five levels of severity, each associated with a maximum waiting time before receiving medical attention.

All 26 nurses assigned to triage duties in the ED have undergone a comprehensive two-day MTS course and completed a period of apprenticeship under the guidance of experienced nurses. Among them, 12 nurses who had previously participated in studies on risk stratification agreed to take part in this study.

### Study protocol and data collection

The study was conducted over 80 non-consecutive daytime working shifts, each lasting 8 hours, during which an experienced triage nurse and two emergency physicians were present. All patients who did not require immediate assessment through the fast-track system were considered for inclusion in the study. Additional exclusion criteria included: i) age under 18 years; ii) possible or confirmed pregnancy; iii) direct admission to the shock room due to a severe condition that prevented the collection of informed consent; iv) absence of written consent for participation in the study.

At the time of patient enrollment, after risk stratification assessment via MTS,<sup>4</sup> the 12 nurses recorded the following variables on a predetermined form: gender; age; reason for ED attendance; method of ED arrival; MTS score; vital signs incorporated in the National Early Warning Score (NEWS); and past medical conditions and comorbidities as outlined in the Charlson Comorbidity Index (CCI). The presence of frailty was evaluated utilizing the Clinical Frailty Scale (CFS), Identification of Seniors At Risk (ISAR), and Program of Research to Integrate Services for the Maintenance of Autonomy (PRISMA-7), while the level of functionality was assessed using the Barthel Index. The triage nurse followed the standard triage model outlined by the MTS and subsequently gathered all additional data required to develop the nomogram.

Following triage assessment, the patient underwent evaluation by two emergency physicians, each possessing a minimum of 5 years of experience in Emergency Medicine, with expertise in risk stratification and MTS. These physicians had been previously involved in clinical and educational activities related to the triage system. After conducting an initial assessment of the patient's medical history, symptom evaluation, and physical examination, which occasionally involved bedside ultrasound, the emergency physicians were asked to determine clinical urgency (urgency status) using the following scale: 0 (non-urgent; indicating patients can wait for over 60 minutes before accessing medical assessment), 1 (urgent; patients should receive medical assessment within 60 minutes), and 2 (very urgent; requiring medical assessment within 10 minutes). Although this evaluation may appear unfair compared to the triage process; given that the two physicians involved could collaborate, take more time, and utilize additional tools; its purpose was not to directly compare it to triage systems. Instead, it aimed to provide a clear and accurate of the patient's clinical urgency, thereby ensuring a reliable outcome measure for the study. This assessment of urgency status following the initial evaluation was documented in a prespecified chart. Any discrepancies or lack of agreement between the two physicians were resolved through discussion with a third emergency physician, who possessed over 15 years of clinical experience in the ED.

Upon completion of the entire diagnostic-therapeutic process, including the availability of instrumental and laboratory test results, the two emergency physicians were again asked to confirm or modify their previous judgment of clinical urgency. In cases of discrepancies, the involvement of the third emergency physician with over 15 years of experience was sought. Additionally, the two physicians provided a dichotomous judgment of the medical severity of the condition (0 = no serious medical condition; 1 = serious medical condition). As this is the first study to prospectively evaluate medical urgency using a three-tiered classification, a formal sample calculation was not feasible. A convenience sample was adopted, with consecutive enrollment. To validate the findings, patients included in the first 50 days represented the derivation cohort, while those enrolled in the subsequent 30 days comprised the validation cohort.

### Outcomes

The primary outcome of the study was the assessment of clinical priority (urgency status) utilizing the following scale: 0 (non-urgent; indicating patients can wait for more than 60 minutes before accessing medical assessment), 1 (urgent; patients should undergo medical assessment within 60 minutes), and 2 (very urgent; patients require medical assessment within 10 minutes). Secondary outcomes were: mortality at 72 hours, 7 days, and 30 days post-assessment, direct hospital admissions from the Emergency Department (ED), admissions to the Intensive Care Unit (ICU), necessity for Life-Saving Interventions (LSI), adherence to reference standard criteria, and the presence of a severe medical condition (0 = no serious medical condition; 1 = serious medical condition) as judged by the ED physicians.<sup>14-16</sup> The study employed both objective secondary endpoints, such as mortality and LSIs, and more subjective measures, such as the assessment of severe medical conditions. This comprehensive approach aligns with previous studies and offers a more holistic understanding of how triage systems function overall. Moreover, by treating these as secondary outcomes, the study was able to enrich its results without interfering with the development and functionality of either the new model or the traditional triage system (the MTS).

### Statistical analysis

The enrolled patients were assigned to a derivation and a validation cohort based on their order of attendance to the ED. Patients who were enrolled during the initial 50 days were included in the derivation dataset, while those enrolled in the subsequent 30 days were allocated to the validation dataset.

Continuous variables were reported as mean and Standard Deviation (SD) or median and Interquartile Range (IQR), depending on the underlying distribution. Univariate comparisons were conducted using Student's t-test, Mann-Whitney test, or Kruskal-Wallis test, as appropriate. Categorical variables were expressed as proportions and percentages, and comparisons were made using Fisher's exact test or chi-square test. Within the derivation cohort, a univariate comparison was performed, and significant variables were subsequently entered into a multivariable analysis carried out using ordinal logistic regression. The results were reported as Odds Ratios (OR) with their respective 95% confidence intervals (95% CI). The obtained coefficients were used to construct a nomogram for risk prediction. The stratification provided by the developed model enabled a second level of priority and severity classification, informed directly by the study's statistical coefficients. This approach depended solely on objectively analyzed data, rather than the subjective judgment of the triage operator who initially collected

the information. For comparisons, we performed an ordinal logistic regression analysis of MTS. Subsequently, in the validation cohort, we assessed the predictive ability of the model compared to MTS, analysing the predictive ability for each level of medical priority; we then compared them using the Areas Under the Receiver Operating Characteristic Curve (AUROC) for all study outcomes. Finally, a Decision Curve Analysis (DCA) was performed to compare the predictive ability of the obtained model with that of the MTS. All analyses in the validation cohort were conducted using bootstrap sampling with 5000 samples. Statistical analyses were carried out using STATA 16.1 and R 4.3.3.

### Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

### Ethical aspects

The study was conducted with the approval of local ethic committees (Ethics Committee Clinical Experiments of Bolzano, Italy, approval number 86-2023), according to the ethical principles for medical research involving human subjects of the Declaration of Helsinki.

## Results

We included a total of 1270 patients. Among these, 64.6% (821/1270) were enrolled during the initial 50 days and were included in the derivation dataset, while 35.4% (449/1270) were enrolled in the subsequent 30 days and were assigned to the validation dataset. The demographic and clinical characteristics of the included patients are summarized in Table 1.

The patients included in both the validation and derivation datasets exhibited comparable clinical characteristics and outcomes (Table 1).

In the derivation cohort, 821 patients were enrolled; among them, 657 (80.0%) were classified as non-urgent by ED physicians, 120 (14.6%) as urgent, and 44 (5.4%) as very urgent (Table 2).

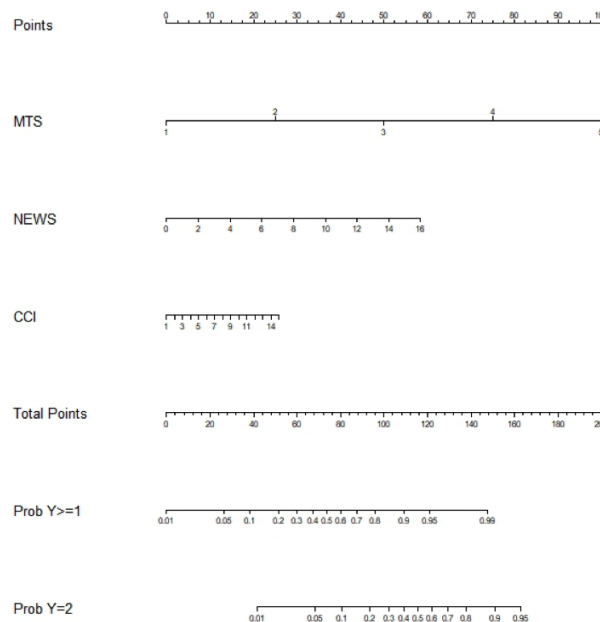
Urgent patients were older and more likely to be male compared to non-urgent and very urgent patients (Table 2). Moreover, patients with higher severity had elevated CCI scores (4.4 vs 4.3 vs 2.7;  $p < 0.001$ ) and increased frailty, as indicated by the CFS (3.2 vs 3.0 vs 2.1;  $p < 0.001$ ), PRISMA-7 (3.1 vs 2.9 vs 2.2;  $p < 0.001$ ), and ISAR (1.8 vs 1.8 vs 0.9;  $p < 0.001$ ). The Barthel Index was 80.2 for very urgent patients, 82.9 for urgent patients, and 92.3 for non-urgent patients ( $p < 0.001$ ). Very urgent patients had higher NEWS scores compared to those in lower categories (5.7 vs 3.2 vs 1.4;  $p < 0.001$ ; Table 2).

The results of ordinal logistic regression models, including MTS alone and MTS integrated with variables identified as significant in the previous univariate analysis (MTS, NEWS, and CCI), are presented in Table 3.

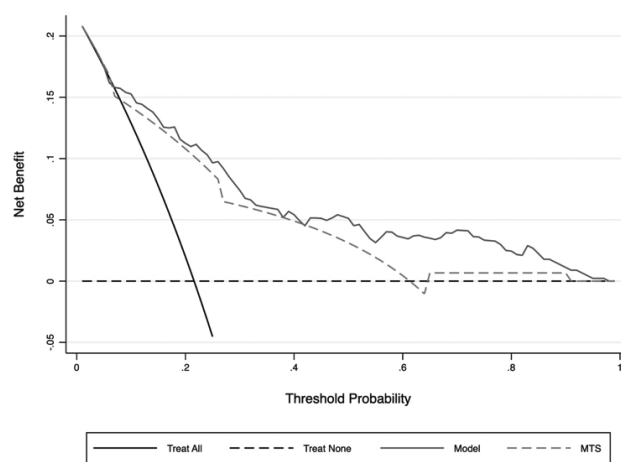
This model (with MTS, NEWS and CCI) was used to create a nomogram predicting the presence of an urgent status classified as urgent or very urgent (Figure 1).

In the validation dataset ( $n=449$ ), physicians classified 78.4% (352/449) of patients as non-urgent, 17.6% (79/449) as urgent, and 4% (18/449) as very urgent. The predictions generated by the multivariable model, which combined urgency, comorbidity, and acuity, estimated urgency levels significantly differently compared to those predicted by MTS alone (Likelihood Ratio Test  $< 0.001$  and Wald Test  $< 0.001$ ) (Supplementary materials, Table 1).

Among the 352 patients judged as non-urgent by ED physicians, 6.8% (24/352) were classified differently by MTS, with only 2.6% (9/352) being classified differently by the model. For the 18 patients in the validation cohort judged as very urgent, MTS correctly predicted urgency in 11.1% (2/18) of cases, while the model correctly predicted 22.2% of cases (4/18). Utilizing a bootstrap with 5000 cases, the discriminatory ability to identify non-urgent patients was 0.645 (95% CI 0.573-0.717) for MTS and 0.811 (95% CI 0.757-



**Figure 1.** Nomogram constructed using the result of the ordinal logistic regression, capable of calculating the probability of being urgent or higher and of being very urgent.



**Figure 2.** Decision Curve Analysis comparing the model developed in the study (solid gray line) with the MTS (dashed gray line). The x-axis represents the threshold probability, and the y-axis represents the net clinical benefit. The decision curve analysis compares the ability of the two systems with two opposing strategies

0.863) for the model,  $p < 0.001$  (*Supplementary materials, Table 2*). The discriminatory ability to identify urgent patients was 0.523 (95% CI 0.502-0.663) for MTS and 0.755 (95% CI 0.693-0.818) for the model,  $p < 0.001$  (*Supplementary materials, Table 3*). Lastly, the discriminatory ability to identify very urgent patients was 0.773 (95% CI 0.662-0.882) for MTS and 0.889 (95% CI 0.831-0.947) for the model,  $p < 0.001$  (*Supplementary materials, Table 4*). We compared the ROCs for study outcomes using a bootstrap with 5000 cases (Table 3). No significant difference was observed in mortality within 72 hours and 7 days between the model and MTS. However, for all other study outcomes, the model showed superior performance.

In the DCAs comparing MTS with the model for the presence of severe conditions as assessed by the ED physicians (Figure 2), the model consistently exhibited superior performance compared to MTS.

## Discussion

This innovative study assessed the impact and utility of integrating scales measuring functionality, frailty, acuity, and comorbidity within the MTS. Our results showed that the model incorporating CCI and NEWS is more effective than MTS alone.

**Table 1.** Clinical and anamnestic characteristics of patients enrolled in the study, divided according to the derivation or validation cohort.

Variables	Total	Derivation	Validation	p
Patients, n (%)	1270 (100)	821 (64.6)	449 (35.4)	
Age, years, mean (SD)	62 (21)	63 (20)	62 (22)	0.083
Sex, n (%)				0.391
Male	643 (50.6)	421 (51.3)	222 (49.4)	
Female	627 (49.4)	400 (48.7)	227 (50.5)	
Mode of arrival in ED, n (%)				0.891
Autonomous	797 (62.7)	523 (63.7)	274 (61.0)	
Rescue vehicles	473 (37.2)	298 (36.3)	175 (39.0)	
MTS triage code, n (%)				0.769
Red	9 (0.7)	6 (0.7)	3 (0.7)	
Orange	179 (14.1)	120 (14.6)	59 (13.1)	
Yellow	328 (25.8)	204 (24.8)	124 (27.6)	
Green	719 (56.6)	470 (57.2)	249 (55.4)	
Blue	35 (2.7)	21 (2.5)	14 (3.1)	
Vital parameters, median (IQR)				
Heart rate (bpm)	80 (71-92)	80 (71-91)	80 (73-94)	0.301
Peripheral Oxygen Saturation (%)	98 (96-98)	98 (96-98)	98 (96-98)	0.578
Respiratory rate (breaths per minute)	16 (16-17)	16 (16-17)	16 (16-17)	0.836
Systolic blood pressure (mmHg)	135 (120-154)	136 (120-155)	135 (120-150)	0.369
Temperature (°C)	36.1 (36.0-36.5)	36.1 (36.0-36.5)	36.0 (36.0-36.4)	0.395
Comorbidity				
CCI, value, mean (SD)	2.9 (2.7)	3.0 (2.7)	2.9 (2.7)	0.170
Frailty				
CFS, value, media (SD)	2.2 (1.8)	2.3 (1.9)	2.2 (1.7)	0.784
Prisma-7, value, media (SD)	2.3 (1.6)	2.3 (1.6)	2.2 (1.6)	0.819
ISAR, value, media (SD)	1.0 (1.5)	1.1 (1.6)	1.0 (1.5)	0.065
Functionality				
Barthel Index, value, median (IQR)	100 (100-100)	100 (100-100)	100 (100-100)	0.939
Acuity				
NEWS, value, median (IQR)	1 (1-2)	1 (1-2)	1 (1-2)	0.339
Hospitalisation	274 (21.6)	172 (20.9)	102 (22.7)	0.535
Hospitalisation in the ICU	33 (2.6)	13 (2.9)	20 (2.4)	0.623
Death				
within 72 hours	5 (0.4)	4 (0.5)	1 (0.2)	0.472
within 7 days	11 (0.9)	8 (1.0)	3 (0.7)	0.573
within 30 days	26 (2.0)	18 (2.2)	8 (1.8)	0.621
Life-saving intervention	35 (2.7)	23 (2.8)	12 (2.7)	0.893
Medical priority				0.244
Non urgent	1,009 (79.4)	657 (80.0)	352 (78.4)	
Urgent	199 (15.7)	120 (14.6)	79 (17.6)	
Very urgent	62 (4.9)	44 (5.4)	18 (4.0)	
Reference standard				0.966
R1 – Immediate and Very urgent	67 (5.3)	44 (5.4)	23 (5.1)	
R2 – Urgent	711 (56.1)	461 (56.2)	250 (55.8)	
R3 – Standard and Non-urgent	490 (38.6)	315 (38.4)	175 (39.1)	

The aim of our study was to evaluate the possibility of developing a triage model different from those previously adopted by integrating other scales commonly used alongside the triage system, without directly incorporating them into it.<sup>12,13</sup> This approach proved feasible and warrants further investigation, although it was not aimed to create a tool for routine clinical practice but to emphasize the need to revise and update triage systems that have remained unchanged for decades.<sup>1</sup> We acknowledge that the developed model may not always be optimal for triage due to the extensive data it requires. However, our primary objective was to demonstrate that alternative, viable pathways exist for improving the predictive ability of triage systems through specific statistical methods. This

underscores the need to move away from the current trend of continuously adding multiple parameters to triage systems, such as frailty assessments for patients over 65. These additions are often introduced arbitrarily, without evaluating their true impact on triage performance. Instead, it is essential to pursue scientifically robust and actionable methods grounded in statistical modeling, rather than relying solely on expert consensus. Evidence accumulated over the years has shown that the predictive ability of triage has not significantly improved, despite numerous updated versions of triage systems. It is time to adopt innovative approaches that truly enhance both the predictive accuracy and the practical utility of triage.

The only two scales found to be significant in the multivariable

**Table 2.** Characteristics of patients enrolled in the model derivation cohort, divided based on the priority assigned by the physician.

Variables	Non urgent	Urgent	Very urgent	p
Patients, n (%)	657 (80.0)	120 (14.6)	44 (5.4)	
Age, years, mean (SD)	60 (21)	71 (17)	70 (14)	<0.001
Sex, n (%)				0.047
Male	326 (49.6)	74 (61.7)	21 (47.7)	
Female	331 (50.4)	46 (38.3)	23 (52.3)	
Mode of arrival in the ED, n (%)				<0.001
Autonomous	458 (69.7)	49 (40.8)	16 (36.3)	
Rescue vehicles	199 (30.3)	71 (59.2)	28 (63.6)	
Comorbidity				
CCI, value, mean (SD)	2.7 (2.5)	4.3 (2.9)	4.4 (3.1)	<0.001
Frailty				
CFS, value, mean (SD)	2.1 (1.7)	3.0 (2.2)	3.2 (2.1)	<0.001
PRISMA-7, value, mean (SD)	2.2 (1.5)	2.9 (1.9)	3.1 (1.8)	<0.001
ISAR, value, mean (SD)	0.9 (1.4)	1.8 (1.9)	1.8 (2.1)	<0.001
Functionality				
Barthel Index, value, mean (SD)	92.3 (20.3)	82.9 (30.2)	80.2 (31.1)	<0.001
Acuity				
NEWS, value, mean (SD)	1.4 (1.6)	3.2 (2.9)	5.7 (4.2)	<0.001
MTS, value, mean (SD)	3.6 (0.6)	2.9 (0.8)	2.1 (0.5)	<0.001

**Table 3.** Comparison between the ROC curves of MTS and the model in the validation cohort after bootstrap resampling of 5000 cases.

Outcome	AUC	95% CI	p
Death within 72 hours			0.081
MTS	0.995	0.989-1.000	
Model	0.988	0.979-0.998	
Death within 7 days			0.088
MTS	0.910	0.828-0.992	
Model	0.967	0.934-1.000	
Death within 30 days			0.016
MTS	0.678	0.400-0.955	
Model	0.903	0.802-1.000	
Life-saving intervention			0.002
MTS	0.709	0.516-0.902	
Model	0.881	0.790-0.972	
Hospitalization			<0.001
MTS	0.542	0.467-0.616	
Model	0.744	0.685-0.803	
Hospitalisation in the ICU			<0.001
MTS	0.546	0.336-0.757	
Model	0.783	0.661-0.906	
Severe condition in the ED			<0.001

analysis were CCI and NEWS, while none of the frailty and functionality scores were significant. This suggests that triage system predictions aimed at improving performance regarding multiple outcomes require a careful consideration of all vital patient parameters and comorbidity assessment. In a recent study, which aimed to analyze variables recorded in the ED and their association with hospital admission and the presence of critical illness through machine learning, vital signs (except for body temperature) were identified as main outcome predictors. This emphasizes the importance and necessity of integrating vital sign assessment in all patients accessing the ED.<sup>18</sup> A previous multicenter study conducted on the pediatric population showed that adding indicators based on vital signs in MTS could increase the predictive ability of this triage system, significantly enhancing its sensitivity for high levels of urgency from 0.66 to 0.71.<sup>19</sup> Another multicenter observational study revealed that including vital signs changed the classification in 7.9% of patients, thereby improving risk prediction. However, despite these findings, vital signs have not been systematically integrated into triage systems for all patient categories.<sup>20</sup>

Another significant factor for enhancing the predictive capability of triage systems is comorbidity, which is currently not incorporated into the most validated and widely used triage systems, despite the increasing number of patients with chronic illnesses and comorbidities accessing EDs.<sup>21</sup> As demonstrated by the current study, integrating various scales into triage systems may offer additional value in risk stratification and prediction.<sup>22,23</sup>

Our study used an ordinal outcome to reflect the level of clinical priority (urgency status) determined by ED physicians. This type of outcome has previously been identified as one of the most effective for triage systems, as it enables qualified and trained personnel to assess each case in a structured manner.<sup>6,24</sup> While objective outcomes such as death or admission are commonly used, they do not fully convey the priority and severity of the patient's condition, as critically ill or urgent patients may be discharged after ED evaluation.<sup>6,24</sup> Furthermore, our study showed that the obtained model performed significantly better than MTS alone for all outcomes, except for mortality at 72 hours and 7 days. This outcome could primarily be attributed to the low incidence of death cases. Despite utilizing bootstrap sampling with 5000 cases, the incidence of this outcome remains low, as previously reported in the literature.<sup>6,14,24</sup>

In the DCA we found that for lower threshold probabilities, the model's performance is slightly superior to that of the MTS. Thus, while the MTS remains a valid tool for outcome prediction and risk stratification, its performance deteriorates at other levels of threshold probability.

There is a continued need to develop or update specific tools for assessing comorbidity within the context of the ED. The CCI, the most widely used tool for assessing chronic conditions, was developed in 1987 for a different context than that of the ED and triage.<sup>25</sup> The same applies to frailty scores, which require adaptation for the use in EDs.<sup>9</sup>

This single-center study has some limitations. The decision to divide the patient cohort based on a temporal criterion was made a priori, although it followed consecutive and chronological enrollment over time. Moreover, it was not possible to perform a formal sample calculation as the study is the first of its kind; nevertheless, we attempted to address this issue by conducting consecutive enrollment. Another limitation of the study concerns the use of the NEWS scale, despite the availability of NEWS2 scale. We chose NEWS because we believe that rapidly identifying patients with hypercapnic respiratory failure at triage is not always straightforward. Nonetheless, the study's results might differ if

NEWS2 were used. Additionally, the physicians who assessed clinical urgency were aware of the MTS-assigned triage codes, although they were explicitly instructed not to consider this information in their evaluations. While this may have introduced some bias in the assignment of clinical priorities, it did not impact the development of the study's model. A further limitation is that the model we developed produced only three priority levels, instead of the five commonly used in triage systems. This may reduce its direct clinical applicability. However, our aim was not to create a tool for immediate clinical use, but rather to highlight the need for new strategies to enhance triage system performance. Notably, previous studies evaluating triage system performance have frequently grouped the five triage level codes into three or even two categories. In this regard, our study remains consistent with standard approaches to triage system evaluation.

## Conclusions

Our study integrated widely used clinical practice scores such as NEWS and CCI into MTS, increasing its predictive performance. Our results emphasize the importance of updating and modifying triage systems by incorporating available tools to enhance their predictive ability. Similar approaches could be used in future studies with other triage systems to achieve better and more accurate risk stratification of ED patients.

## References

1. Tam HL, Chung SF, Lou CK. A review of triage accuracy and future direction. *BMC Emerg Med.* 2018;18:58.
2. Estrada EG. Triage systems. *Nurs Clin North Am* 1981;16:13-24.
3. González J, Soltero R. Emergency Severity Index (ESI) triage algorithm: trends after implementation in the emergency department. *Bol Asoc Med P R* 2009;101:7-10.
4. Mackaway-Jones K, Marsden J, Windle J. *Emergency Triage: Manchester triage group*, Third edition. John Wiley & Sons, Ltd, West Sussex. 2014.
5. Azeredo TR, Guedes HM, Rebelo de Almeida RA, Chianca TC, Martins JC. Efficacy of the Manchester Triage System: a systematic review. *Int Emerg Nurs* 2015;23:47-52.
6. Kuriyama A, Urushidani S, Nakayama T. Five-level emergency triage systems: variation in assessment of validity. *Emerg Med J* 2017;34:703-710.
7. Hinson JS, Martinez DA, Cabral S, et al. Triage Performance in Emergency Medicine: A Systematic Review. *Ann Emerg Med* 2019;74:140-152.
8. Zachariasse JM, van der Hagen V, Seiger N, et al. Performance of triage systems in emergency care: a systematic review and meta-analysis. *BMJ Open* 2019;9:e026471.
9. Fehlmann CA, Nickel CH, Cino E, et al. Frailty assessment in emergency medicine using the Clinical Frailty Scale: a scoping review. *Intern Emerg Med* 2022;17:2407-2418.
10. Zaboli A, Brigo F, Sibilio S, et al. The impact of COVID-19 pandemic on the urgency of patients admitted to the emergency department. *Int Emerg Nurs* 2022;65:101229.
11. van Dam CS, Labuschagne HA, van Keulen K, et al. Polypharmacy, comorbidity and frailty: a complex interplay in older patients at the emergency department. *Eur Geriatr Med* 2022;13:849-57.

12. Elliott A, Taub N, Banerjee J, et al. Does the clinical frailty scale at triage predict outcomes from emergency care for older people? *Ann Emerg Med* 2021;77:620-7.
13. Rashiti P, Rashiti-Bytyçi A, Haradinaj K, et al. Using the HEART score in patients with chest pain in the emergency department of Kosovo as an important triage criterion for treatment and hospitalization. *Ann Med Surg (Lond)* 2023;85:3318-22.
14. Zaboli A, Sibilio S, Cipriano A, et al. Italian validation of the Manchester Triage System towards short-term mortality: a prospective observational study. *Emerg Care J* 2023;19:11443
15. Platts-Mills TF, Travers D, Biese K, et al. Accuracy of the Emergency Severity Index triage instrument for identifying elder emergency department patients receiving an immediate life-saving intervention. *Acad Emerg Med* 2010;17:238-43.
16. Zachariasse JM, Seiger N, Rood PP, et al. Validity of the Manchester Triage System in emergency care: A prospective observational study. *PLoS One* 2017;12:e0170811.
17. Zaboli A, Ausserhofer D, Sibilio S, et al. Effect of the emergency department assessment of chest pain score on the triage performance in patients with chest pain. *Am J Cardiol* 2021;161:12-18.
18. Raita Y, Goto T, Faridi MK, et al. Emergency department triage prediction of clinical outcomes using machine learning models. *Crit Care* 2019;23:64.
19. Zachariasse JM, Maconochie IK, Nijman RG, et al. Improving the prioritization of children at the emergency department: Updating the Manchester Triage System using vital signs. *PLoS One* 2021;16:e0246324.
20. Cooper RJ, Schriger DL, Flaherty HL, et al. Effect of vital signs on triage decisions. *Ann Emerg Med* 2002;39:223-32.
21. Pellerin G, Gao K, Kaminsky L. Predicting 72-hour emergency department revisits. *Am J Emerg Med* 2018;36:420-4.
22. Usman OA, Usman AA, Ward MA. Comparison of SIRS, qSOFA, and NEWS for the early identification of sepsis in the Emergency Department. *Am J Emerg Med* 2019;37:1490-7.
23. Almutary A, Althunayyan S, Alenazi K, et al. National Early Warning Score (NEWS) as prognostic triage tool for septic patients. *Infect Drug Resist* 2020;13:3843-51.
24. Challen K. How good is triage, and what is it good for? *Emerg Med J* 2017;34:702.
25. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373-83.