

國立臺灣大學公共衛生學院流行病學與預防醫學研究所
碩士論文



Institute of Epidemiology and Preventive Medicine
College of Public Health
National Taiwan University
Master's Thesis

評估台灣到院前五級檢傷系統在緊急醫療救護中的有效
性：評分者間信度與患者結果預測
Evaluating the Effectiveness of a Five-Level Taiwan Prehospital
Triage System (TPTS) in Emergency Medical Services:
Inter-Rater Reliability and Patient Outcome Prediction

張詠翔
Yung-Hsiang Chang

指導教授：杜裕康 博士
Advisor: Yu-Kang Tu, DDS, Ph.D.

中華民國 114 年 1 月
January, 2025

國立臺灣大學碩(博)士學位論文 口試委員會審定書



國立臺灣大學碩士學位論文 口試委員會審定書

論文中文題目:評估台灣到院前五級檢傷系統在緊急醫療救護中的有效性:評分者間信度與患者結果預測

論文英文題目:Evaluating the Effectiveness of a Five-Level Taiwan Prehospital Triage System (TPTS) in Emergency Medical Services: Inter-Rater Reliability and Patient Outcome Prediction

本論文係 張詠翔 君(學號 p11849001)在國立臺灣大學流行病學與預防醫學研究所完成之碩士學位論文,於民國 114 年 1 月 10 日承下列考試委員審查通過及口試及格,特此證明。

口試委員:

杜祿康

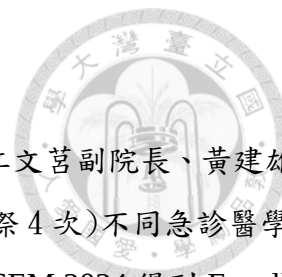
(簽名)

(指導教授)

張淑惠

江文萱

誌謝



首先，要謝謝帶領我進入 EMS 研究領域的老師們，黃集仁院長、江文莒副院長、黃建雄主任等許多貴人，在碩士班的兩年半間，參與了共 6 次(國內 2 次、國際 4 次)不同急診醫學會舉辦的研討會，其中一篇關於 OHCA 腎上腺素給藥最佳時間研究在 ICEM 2024 得到 Excellent Award，另有兩篇關於骨針給藥的研究，刊登在 AJEM，在此表達我最深的謝意。

特別感謝指導教授杜裕康老師，在攻讀碩士班的日子，啟發我對統計的興趣，並在 lab meeting 中引領學生們去思考研究背後的脈絡及方法學上的議題。批判性思考一直是臺灣學生缺乏的東西，過去我們被教導要服從師長，卻忘卻了多數情境下並沒有標準答案，且應該勇於發表自己的想法才是，好像從眾才不至於被討厭。老師獨樹一幟的教學風格及永遠不直接告訴學生答案，才能使學生去細細咀嚼科學背後嚴謹的推論思維。每次老師提及過去歷史，並在課堂報告中要求學生穿越過去使用現代方法去回答前人的難題，會驚訝地發現，人類仍不斷地重蹈覆轍，不停地在試錯中進步，且我們也沒人敢保證這就是真相。

在流預所，最印象深刻是張淑惠老師的進階統計課程及李文宗老師的流病課程，雖然過去我完全沒有學過微積分及線性代數，且博班學長姐告訴我老師的課非常扎實，求學路途上最大罩門就是數學，因為不知為何而學，所以只會背誦公式。張老師不厭其煩的向我解釋背後的邏輯，引領學生的學習動機，都是為了尋求假設分布之下的真實值。而李老師幽默風趣的教學，雖然常常會跳 tone 離開課程主線任務，但又在不知不覺下拉回，快樂的時間總是過得特別快，上完一堂課後不感覺到疲勞，卻還是對上課內容印象深刻，尤其是旁聽高等流病時，終於搞懂因果推論中 G-methods 的概念，也很好奇為什麼老師可以有那麼高的道行，融會貫通不同的方法學，其實都是換湯不換藥。

最後，特別感謝之前的同事，楊尚樵學長，從新北市調回家鄉桃園，一心一意地想改變當地的 EMS 文化及系統，雖然聽起來遙不可及，但學長就是一位 Role Model，不論是救護品管、救護競賽及學術研究都是身體力行。若沒有他的鼓勵，我想我也不會來讀臺大流預所。還有感謝我敬愛的家人們，總是鼓勵我去做自己想做的事。"A Dream is Not Just a Dream."

張詠翔 謹誌於

臺灣大學公共衛生學院流行病學與預防醫學研究所

中華民國 113 年 12 月 14 日

中文摘要



背景：

臺灣院前五級檢傷系統是一項新開發的檢傷工具，目的為提高院前檢傷的準確性和一致性。相較於過去的二級檢傷系統，臺灣院前五級檢傷系統能更詳細地分類患者的病情嚴重程度。本研究包含兩部分，第一部分評估緊急救護技術員與急診檢傷護理師之間的檢傷一致性。第二部分則是探討臺灣院前五級檢傷系統對病人到急診的院後結果的區分能力。

研究目的：

本研究旨在評估臺灣院前五級檢傷系統在以下兩個方面的臨床效能：（1）緊急救護技術員與急診檢傷護理師之間的檢傷一致性。（2）區分病人到急診的院後結果（如離開急診、住院、死亡及急診停留時間）的能力

方法：

研究一為橫斷性研究，納入 10 名緊急救護技術員與 10 名急診檢傷護理師對 50 個到院前電子病歷進行檢傷，並使用組內相關係數及多層次模型進行分析。研究二為前瞻性世代研究，分析 2023 年 1 月至 12 月期間，在臺灣桃園市由緊急醫療服務運送的 10,713 名患者。

結果：

研究一顯示，臺灣院前五級檢傷系統在考量多位評分者間的平均分數差異具有卓越的一致性（ $ICC = 0.96$ ），在考量單一評分者間的絕對分數差異則具有中等一致性（ $ICC = 0.57$ ）。雖然救護技術員的檢傷分數普遍高於檢傷護理師，但整體一致性仍非常高（ $r = 0.909$ ）。研究二發現，臺灣院前五級檢傷系統能有效預測急診結果，分數越低的患者死亡率及資源使用需求越高。離開急診病人的中位停留時間為 2.73 小時，而住院病人則為 13.90 小時。

結論：

臺灣院前五級檢傷系統是一項可靠且有效的院前檢傷工具，可準確區分急診結果，並在醫療人員之間展現出高度的一致性，整合院前檢傷系統與院內檢傷系統有望優化患者流向及急診資源管理。

關鍵字：臺灣院前五級檢傷系統, 台灣急診五級檢傷系統, 到院前緊急救護, 院前檢傷

英文摘要



Introduction:

The five-level Taiwan Prehospital Triage System (TPTS) is a newly developed tool aimed at improving prehospital triage accuracy and consistency. Unlike the previous binary system, 2-level TPTS provides a more detailed categorization of patient severity. Study 1 examined inter-rater reliability among Emergency Medical Technicians (EMTs) and Triage Registered Nurses (TRNs), while Study 2 explored the system's ability to discriminate ability patient outcomes.

Aims:

This research aims to evaluate the clinical effectiveness of the 5-level TPTS in two key areas: (1) its inter-rater reliability between EMTs and TRNs ; and (2) its ability to discriminate emergency department (ED) outcomes such as discharge, admission, mortality, and length of stay (LOS).

Methods:

In Study 1 assessed inter-rater reliability through a prospective design involving 10 EMTs and 10 TRNs rating 50 prehospital electronic medical records. Intraclass correlation coefficients (ICCs) and multilevel models were used for analysis. Study 2, prospective cohort study, conducted in Taoyuan, Taiwan, from January 1 to December 31, 2023, analyzed 10,713 EMS-transported patients using the five-level TPTS.

Results:

Study 1 demonstrated excellent inter-rater reliability for multiple raters ($ICC = 0.96$) and moderate reliability for single raters ($ICC = 0.57$). EMTs generally scored higher than TRNs, though agreement was strong overall ($r = 0.909$). Study 2 revealed that 5-level TPTS effectively discriminated ED outcomes, with lower scores correlating with higher mortality and resource

utilization. Discharged patients had a median LOS of 2.73 hours, compared to 13.90 hours for admitted patients.

Conclusions:

The five-level TPTS is a reliable and effective tool for prehospital triage. It accurately discriminates ED outcomes and demonstrates strong alignment among healthcare providers, making it a valuable addition to emergency medical services. Further integration with hospital triage systems could enhance patient flow and resource management.

Keywords:TPTS, TTAS , EMS , Prehospital Triage

List of Abbreviations



ACEP: American College of Emergency Physicians

AIC: Akaike Information Criterion

ALS: Advanced Life Support

AUROC: Area Under the Receiver Operating Characteristic curve

AVPU: Alert, Voice, Pain, Unresponsive

ATS: Australasian Triage Scale

AUC: Area Under the Curve

BIC: Bayesian Information Criterion

BLS: Basic Life Support

BP: Blood Pressure

BT: Body Temperature

CIF: Cumulative Incidence Function

CIs: Confidence Intervals

CME: Continuing Medical Education

CRT: Capillary Refill Time

CT: Computerized Tomography

CTAS: Canadian Triage and Acuity Scale

DBP: Diastolic Blood Pressure

eCTAS: electronic Canadian Triage and Acuity Scale

ED: Emergency Department

EHRs: Electronic Health Records

EMS: Emergency Medical Services



EMRs: Electronic Medical Records

EMTs: Emergency Medical Technicians

ENA: Emergency Nurses Association

ESI: Emergency Severity Index

ETS: Electronic Triage System

EWS: Early Warning Scores

GCS: Glasgow Coma Scale

HR: Heart Rate

ICC: Intraclass Correlation Coefficients

ICU: Intensive Care Unit

IQRs: Interquartile Ranges

JTAS: Japanese Triage and Acuity Scale

KTAS: Korean Triage and Acuity Scale

LCGMH: Linkou Chang Gung Memorial Hospital

LMICs: Low- and Middle-Income Countries

LOS: Length of Stay

MCIs: Mass Casualty Incidents

MTS: Manchester Triage System

NHS: National Health Service

NEWS: National Early Warning Score

NSTEMI: Non-ST Elevation Myocardial Infarction

NTS: National Triage Scale

O₂: Oxygen Saturation

OHCA: Out-of-Hospital Cardiac Arrest

EMT-Ps: Emergency Medical Technician-Paramedics

PedKTAS: Pediatric version of KTAS

RR: Respiratory Rate

SATS: South African Triage Scale

SBP: Systolic Blood Pressure

SETS: Swiss Emergency Triage Scale

START: Simple Triage and Rapid Treatment

TEWS: Triage Early Warning Score

TRNs: Triage Registered Nurses

TTAS: Taiwan Triage and Acuity Scale

TPTS: Taiwan Prehospital Triage System

TYFD: Taoyuan Fire Department



Contents



口試委員會審定書	i
誌謝	ii
中文摘要	iii
英文摘要	iv
List of Abbreviations	vi
Contents	ix
List of Figures	xii
List of Tables	xiv
List of Appendices	xv
Chapter 1. Introduction	1
1.1 Background	1
1.2 Definitions	3
Chapter 2. Literature Review	5
2.1 A Global Overview of EMS System	5
2.2 Evolution and Development of Taiwan's EMS	9
2.3 The Evolution of Triage	10
2.4 Global Emergency Department Triage Systems	12
2.5 Global Prehospital Triage Systems	27
2.6 Triage System in EDs in Asia	32
2.7 Prehospital Triage System in Asia	40
2.8 Impact of ED LOS and Resource Use on Patient Outcomes	45
2.9 Research Gap	48
2.10 Research Aims	49
Chapter 3. Materials and Methods	50
3.1.1 Study Design and Methodology	50

3.1.2 Participants and Setting.....	50
3.1.3 The Two-Level and Five-Level TPTS	51
3.1.4 Data Collection and Participant Procedure.....	52
3.1.5 Outcomes.....	54
3.1.6 Statistical Analysis.....	54
3.2 Methods of Study 2.....	57
3.2.1 Study Design	57
3.2.2 Participants and Setting.....	57
3.2.3 Five-Level Taiwan Prehospital Triage System (TPTS).....	58
3.2.4 Data Collection	61
3.2.5 Outcomes.....	62
3.2.6 Statistical Analysis.....	63
4.1 Results of Study 1	65
4.1.1 Patient Flow and Enrollment	65
4.1.2 Characteristics of the study population	67
4.1.3 Main results.....	69
4.1.4 Variance Components Model.....	77
4.1.5 Two-Level Multilevel analysis.....	81
4.1.6 Three-Level Multilevel analysis	87
4.2 Results of Study 2	94
4.2.1 Patient Flow and Enrollment	94
4.2.2 Characteristics of the study population	96
4.2.3 Main results.....	100
4.2.4 Cumulative Incidence for Different Events	102
4.2.5 Proportional Cause-Specific Hazards Models for different events	126
4.2.6 Exploring the Distribution of ESI Score through Stratification	

.....	132
Chapter 5. Discussion.....	137
5.1 Discussion of Study 1.....	137
5.1.1 Major findings	137
5.1.2 Comparison with previous studies.....	138
5.1.3 Strengths.....	141
5.1.4 Limitations and future work	142
5.2 Discussion of Study 2.....	144
5.2.1 Major findings	144
5.2.2 Comparison with previous studies.....	145
5.2.3 Strengths.....	149
5.2.4 Limitations and future work	150
Chapter 6. Conclusion.....	152
Appendix	154
I. Supplemental figures	154
II. Supplemental tables	157
III. Supplementary R and SAS Code for Data Analysis	159
References	172

List of Figures



Figure 1 Global Distribution of Triage System.....	25
Figure 2 Flow chart of the studied patients.....	66
Figure 3 Distribution of 5-level TPTS Scores between EMT vs. RN: Scatterplot	71
Figure 4 Distribution of 5-level TPTS Scores Across Different Groups: Scatterplot.	72
Figure 5 Distribution of 5-level TPTS Scores between EMT vs. RN: Bland-Altman Plot.....	74
Figure 6 Distribution of 5-level TPTS Scores Across Different Groups: Bland-Altman Plot.....	75
Figure 7 Framework of the Two-Level Multilevel Analysis Model	81
Figure 8 Framework of the Three-Level Multilevel Analysis Model	87
Figure 9 Flow chart of the studied patients.....	95
Figure 10 CIF for different event types - Discharged	103
Figure 11 CIF for different event types - Admission	104
Figure 12 CIF for different event types - Death.....	105
Figure 13 Stratified CIF for Discharge Events by 5-level TPTS	106
Figure 14 Stratified CIF for Admission Events by 5-level TPTS	107
Figure 15 Stratified CIF for Death Events by 5-level TPTS.....	108
Figure 16 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Age Groups: <18 Years	111
Figure 17 Stratified CIF Analysis of Admission Events by 5-level TPTS in Age Groups: <18 Years	112
Figure 18 Stratified CIF Analysis of Death Events by 5-level TPTS in Age Groups: <18 Years	113
Figure 19 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Age Groups: 18-65 Years	114
Figure 20 Stratified CIF Analysis of Admission Events by 5-level TPTS in Age Groups: 18-65 Years	115
Figure 21 Stratified CIF Analysis of Death Events by 5-level TPTS in Age Groups: 18-65 Years	116
Figure 22 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Age Groups: >65 Years	117
Figure 23 Stratified CIF Analysis of Admission Events by 5-level TPTS in Age Groups: >65 Years	118
Figure 24 Stratified CIF Analysis of Death Events by 5-level TPTS in Age Groups: >65 Years	119
Figure 25 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Trauma Cases.....	120
Figure 26 Stratified CIF Analysis of Admission Events by 5-level TPTS in Trauma	

Cases	121
Figure 27 Stratified CIF Analysis of Death Events by 5-level TPTS in Trauma Cases	122
Figure 28 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Non-trauma Cases	123
Figure 29 Stratified CIF Analysis of Admission Events by 5-level TPTS in Non-trauma Cases	124
Figure 30 Stratified CIF Analysis of Death Events by 5-level TPTS in Non-trauma Cases	125
Figure 31 ESI Score distribution by 5-level TPTS	132
Figure 32 ESI Score distribution by TTAS	133
Figure 33 ESI Score distribution by reason for 119 calls	134
Figure 34 ESI Score distribution by age group	135

List of Tables

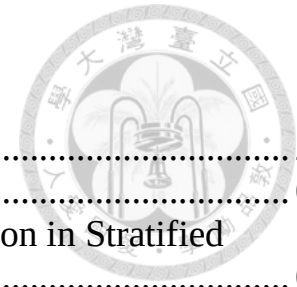


Table 1 Triage system characteristics.....	26
Table 2 Raters demographic data	67
Table 3 Patient Demographics Data Using Proportional Allocation in Stratified Sampling.....	68
Table 4 Inter-rater reliability and the 5-level TPTS score.....	70
Table 5 Correlation coefficients for the 5-level TPTS score measures.....	73
Table 6 Variance Components Models of 5-level TPTS Scores: Distribution of Total Score Variation by Rater, Department, and Patient.....	78
Table 7 Two-Level Variance Components Model of 5-level TPTS Scores: Distribution of Total Score Variation Across Rater, Department, and Patient within the EMT Group.....	79
Table 8 Two-Level Variance Components Model of 5-level TPTS Scores: Distribution of Total Score Variation Across Rater, Department, and Patient within the RN Group	79
Table 9 Two-Level Multilevel analysis of 5-level TPTS score.....	85
Table 10 Three-Level Multilevel analysis of 5-level TPTS score	92
Table 11 Patient demographics data, presented with count (percentage)	98
Table 12 Primary and Secondary Outcome.....	101
Table 13 Subgroup Analysis of Age groups and Reason for calling 119.....	109
Table 14 Multivariable Analysis of Cause-Specific Hazard Models for Discharge Events	126
Table 15 Multivariable Analysis of Cause-Specific Hazard Models for Admission Events	128
Table 16 Multivariable Analysis of Cause-Specific Hazard Models for Death Events	130

List of Appendices

Figure S 1 Franco-German vs Anglo-American Models	154
Figure S 2 Taiwan EMS benchmarks	155
Figure S 3 The traige Algorithm.....	156
Table S 1 Franco-German vs Anglo-American Models	157
Table S 2 Comparison of the TTAS, 2-Level TPTS, 5-Level TPTS.....	158

Chapter 1. Introduction

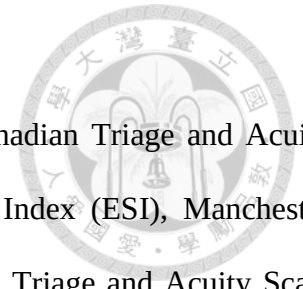


1.1 Background

Emergency Medical Services (EMS) play a crucial role in delivering urgent medical care in acute illness situations[1]. Triage, a critical component of EMS, involves assessing, categorizing, and prioritizing patients based on their presenting symptoms and the urgency of their clinical condition[2]. This process ensures that patients receive timely emergency care. The primary aim of triage in EMS is to efficiently allocate resources to ensure that patients are transported to the appropriate facility at the right time, based on a thorough and accurate assessment of their needs[3].

The unpredictable arrival of patients in an emergency department (ED) can sometimes overwhelm the available resources, leading to a situation known as "overcrowding." This occurs when the number of patients exceeds the ED's ability to handle them efficiently, creating significant risks to patient safety[4]. Five-level triage systems used in emergency care have been shown to be both reliable and valid[5-8]. However, there are inconsistencies in their performance, particularly when assessing critically ill patients, highlighting the need for further research and improvement to ensure high-quality care in emergency settings[7, 8].

In prehospital emergency care, accurately triaging patients with unclear conditions is crucial. Most research from high-income countries focuses on specific adult groups, such as those with stroke or trauma[9]. The goal has been to simplify triage tools to make them easier to use while maintaining accuracy. However, there is still no universal "gold standard" for triaging various acute conditions, highlighting a gap in current methods and the need for a more widely applicable approach in triage practices[1, 10].



Recent studies show that five-level triage systems, such as the Canadian Triage and Acuity Scale (CTAS), Australasian Triage Scale (ATS), Emergency Severity Index (ESI), Manchester Triage System (MTS), South African Triage Scale (SATS), and Taiwan Triage and Acuity Scale (TTAS), are effective in both prehospital and emergency department settings[7, 11, 12]. However, differences in how these systems are applied and their effectiveness across regions like Europe, North America, and Asia suggest a need for continued evaluation and possible standardization[13-25]. These variations highlight the need for a globally effective triage system that can adapt to the diverse and changing demands of emergency healthcare worldwide.

In the last decade, Taiwan's EDs have struggled with overcrowding and increased demand for EMS[26, 27]. This highlights the urgent need for an efficient and reliable triage system. In response, Taiwan's EDs introduced the TTAS, a computerized five-level system that has proven effective in prioritizing care, preventing overtreatment, and optimizing the use of resources[14, 28, 29]. In contrast, two-level Taiwan Prehospital Triage System (TPTS), used by Emergency Medical Technicians (EMTs) in prehospital settings, only classifies patients as emergent or non-emergent. A study comparing the two systems showed TTAS's superior ability to predict hospitalization rates, making it more effective in emergency care.[18].

For more than 20 years, the 2-level TPTS, which categorizes patients as either 'emergent' or 'non-emergent,' has been noted for its simplicity and limited capacity to accommodate complex patient conditions. Ng et al. highlights its limitations in providing effective prehospital care and recommended a more detailed system[18]. In response, the Taoyuan City Fire Department introduced a five-level TPTS in January 2022, combining elements of both the two-level and

five-level systems to improve the accuracy of prehospital triage.



1.2 Definitions

1.2.1 Emergency department (ED)

ED is a specialized medical unit within a hospital that provides immediate care for patients with acute injuries or illnesses. The primary function of the ED is to assess, stabilize, and treat individuals in need of urgent attention. Patients can present with a wide range of conditions, from minor ailments to life-threatening emergencies. EDs operate 24/7, meaning they are open 24 hours a day, 7 days a week, ensuring that medical care is always available. This continuous operation allows patients to receive prompt treatment at any time, day or night, based on the severity of their condition through a triage system[30].

1.2.2 Emergency Department Length of Stay (ED LOS)

ED LOS refers to the total time a patient spends in the emergency department, from the moment they register or are triaged until they are discharged, admitted, or transferred. ED LOS is a critical metric for assessing emergency department efficiency and quality of care. Factors that influence ED LOS include the severity of the patient's condition, diagnostic testing, imaging, treatments, and hospital resources. Long ED stays are associated with overcrowding, delayed treatments, and increased risks for negative patient outcomes[31-33].

1.2.3 Triage Registered Nurses (TRNs)

TRNs are highly trained professionals who assess and prioritize patients based on the severity of their conditions upon arrival in the ED. They use standardized triage protocols, such as the ESI



or other triage scales, to categorize patients into different urgency levels, ensuring that critical patients receive immediate care. Their decision-making process involves quickly assessing vital signs, symptoms, and medical history to allocate resources efficiently and improve patient outcomes[34-36].

1.2.4 Emergency Medical Services (EMS)

EMS are systems designed to provide immediate medical care and transportation to patients experiencing acute medical conditions or trauma. EMS plays a crucial role in delivering prehospital care, stabilizing patients, and facilitating their transfer to healthcare facilities. It involves trained professionals like paramedics, EMTs, and specialized equipment. EMS operates through a network of communication systems to ensure timely response, coordination with hospitals, and effective management of emergencies to minimize mortality and morbidity. Its main objective is to provide rapid and appropriate care in critical situations[37].

1.2.5 Emergency Medical Technicians (EMTs)

EMTs are key healthcare professionals in prehospital emergency care. Originating in the 1960s, EMTs were initially trained to transport patients to hospitals, but their role has significantly expanded over time[38]. In the 1990s, formal EMT training programs began in Taiwan, and EMT certification levels were established (EMT-1, EMT-2, and EMT-P for paramedics). The 2000s saw further professionalization, with EMT-Paramedics gaining skills in advanced life support. Today, EMTs handle a wide range of emergencies, including trauma and cardiac arrest, often acting as the first point of care[26, 27].

Chapter 2. Literature Review



2.1 A Global Overview of EMS System

The global development of EMS traces back to the innovations of Baron Dominique Jean Larrey during Napoleon's Italian campaign in 1794. Larrey introduced the concept of trained medical personnel providing rapid battlefield care and transport, shaping pre-hospital emergency services[39].

From this origin, two key EMS models have evolved (Table S1):

2.1.1 Franco-German Model ("stay and stabilize")

This approach, widely adopted in Europe, involves physicians administering advanced treatment at the scene before transporting patients. This model aims to bring hospital-level care to the patient, stabilizing them before transport. EMS personnel, often physicians, provide immediate on-scene interventions, reducing the number of patients transported to the hospital and minimizing unnecessary emergency department visits. Countries like France, Germany and Austria rely heavily on this model, integrating it closely with the health care system. For example, in France, physicians are part of the SAMU service (Service d'Aide Médicale Urgente, or "Urgent Medical Aid Service"), and ALS ambulances respond to critical cases, ensuring advanced pre-hospital care[38, 40].

2.1.2 Anglo-American Model ("scoop and run")

In contrast, the Anglo-American model, which is prevalent in the US, UK, Canada, and Australia, emphasizes rapid transport to hospitals with limited pre-hospital intervention. Paramedics or EMTs provide basic and advanced life support, stabilizing patients just enough to transport them



to a facility where further treatment can be administered. This system is linked more with public safety services like fire and police departments rather than hospital systems, focusing on quick evacuation to a well-developed emergency department[38, 40].

2.1.3 Development in Different Countries

France and Germany

France's EMS system, known as SAMU service, is centralized and physician-staffed. SAMU operates two tiers of service: Basic Life Support (BLS) provided by fire department ambulances, and Advanced Life Support (ALS) staffed by physicians. The model emphasizes high-quality pre-hospital care, with medical teams often reaching patients in critical conditions, delivering immediate advanced treatment, and then deciding whether to transport them to a hospital[41].

Germany follows a similar physician-led approach. Doctors provide on-scene care, and ambulances are often staffed with paramedics who are highly trained to administer advanced procedures under physician supervision. Medical helicopters are also commonly used to transport patients in rural areas[38].

U.S. and U.K.

In the U.S., the EMS system relies heavily on paramedics and EMTs, who are trained in basic and advanced life support techniques. EMS operates as a decentralized system, with emergency calls routed through 911 dispatch. Many EMS services are fire-based, meaning they are integrated into fire departments. The U.S. emphasizes quick transport to hospital-based care, which contrasts with the physician-led models in Europe[42].



Similarly, the U.K. has a paramedic-based system. The National Health Service (NHS) provides EMS services that focus on stabilizing patients and transporting them to specialized centers for treatment[43].

Australia and New Zealand

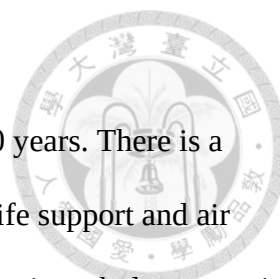
Both Australia and New Zealand also follow an Anglo-American model, with a focus on rapid transport and paramedic interventions. In Australia, paramedics provide a high level of care, including drug administration and advanced life support procedures[44].

Nordic Countries

Nordic countries like Sweden, Norway, and Finland blend aspects of both models. In Norway, for example, anaesthesiologists are involved in pre-hospital care, with paramedics and nurses providing essential support in transporting critically ill patients. Finland also utilizes highly trained EMS personnel, including specialized doctors who contribute to pre-hospital services, although local governance over EMS leads to variability in services[45].

Asia

EMS in many Asian countries has developed relatively recently compared to Europe and the U.S. Japan and Korea have adopted hybrid systems that incorporate both rapid transport and on-scene medical interventions. For instance, Japan's EMS uses physician-staffed ambulances for critical cases, while paramedics provide initial stabilization. Similarly, South Korea's system employs physicians in some cases but primarily relies on paramedics[46].



South Africa

In South Africa, EMS systems have rapidly developed over the last 20 years. There is a significant urban-rural divide, with urban areas benefiting from advanced life support and air response systems. Meanwhile, rural areas remain underserved, often with basic ambulance services. However, the introduction of advanced EMS training programs has improved care quality, particularly in trauma cases[47].

Iran and Turkey

In both Iran and Turkey, a mixed model of emergency medical services is employed, blending aspects of both systems. In Iran, pre-hospital care involves basic life-saving interventions on-site, with rapid patient transportation to hospitals where specialized services are provided. Similarly, in Turkey, paramedics and emergency physicians are dispatched to stabilize patients before swift transport to hospitals for further care[48].

Summary

The development of EMS in low- and middle-income countries (LMICs) faces significant challenges such as a lack of trained personnel, poor infrastructure, and limited transportation. These regions often rely on untrained responders, leading to delays and fragmented care[49]. By contrast, high-income regions follow either the Franco-German model (physician-led on-site care) or the Anglo-American model (rapid transport to hospitals) (Figure S2). Each system adapts to its local needs, but all share the goal of improving patient outcomes through better training, technology, and organizational development worldwide.

2.2 Evolution and Development of Taiwan's EMS



Taiwan's EMS system has evolved significantly since its inception in the 1960s. Initially focused on simple transportation of patients to hospitals, the system underwent major reforms starting in the 1990s. This included the establishment of formal EMT training programs and the implementation of the Emergency Medical Service Act. EMTs in Taiwan are trained at three levels—EMT-1, EMT-2, and EMT-Paramedic—with progressively more advanced skills required at each stage. EMT-Ps (paramedics) undergo extensive training to provide ALS services. Over the past 30 years, Taiwan has developed a robust EMS system that incorporates public defibrillator programs, dispatcher-assisted CPR, and an advanced triage system for prehospital care[26].

The EMS system operates mainly through the National Fire Agency, which oversees both the response and dispatch of emergency services. Recently, the system has been optimized through the use of digital technologies, such as prehospital ECG for cardiac patients, electronic run sheets for ambulances, and comprehensive registries for out-of-hospital cardiac arrests. Moreover, regional variations in the EMS system reflect local needs, with specific benchmarks developed for cities like Taipei, Taichung, and Kaohsiung. These cities have implemented programs (Figure S2) such as the hearing AED program, prehospital triage systems, and crowd-sourcing CPR applications to improve patient outcomes[50-52].

Medical oversight and training have become more formalized since 2007, with the establishment of medical director roles and continuing medical education (CME) programs. Although Taiwan's EMS system has made significant strides, areas like inter-hospital transfers and integration of private EMS services are still under development[27].



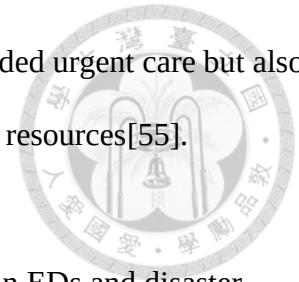
During the COVID-19 pandemic, Taiwan's EMS system played a crucial role in managing the crisis by collaborating with various agencies and developing innovative protective measures for EMTs[53]. Moving forward, the focus will be on integrating prehospital and in-hospital data to further improve the efficiency and responsiveness of the EMS system.

2.3 The Evolution of Triage

Triage, derived from the French word "trier," meaning to sort, has its origins in military medicine, dating back to the Napoleonic Wars. The system was introduced by Baron Dominique-Jean Larrey, Napoleon's chief military surgeon, who recognized the need to prioritize wounded soldiers based on the severity of their injuries, ensuring that those requiring urgent care were treated first, regardless of rank or status. Larrey's system of battlefield triage laid the groundwork for modern triage principles, focusing on immediate care for those with life-threatening injuries and delaying treatment for those who could wait. His innovations, including the use of "flying ambulances," revolutionized battlefield medicine, enabling faster transportation and care for the wounded[54].

Triage evolved through subsequent conflicts, with the American Civil War and World War I marking significant developments. During the Civil War, Union Medical Director Jonathan Letterman implemented triage and ambulance systems that greatly reduced mortality rates[2]. By World War I, the term "triage" became widely used, especially with the advent of new weapons, such as machine guns and poison gas, which increased the number of casualties requiring

prioritization. Triage in these early conflicts emphasized not only who needed urgent care but also who was likely to survive with treatment, guiding the allocation of limited resources[55].



Over time, triage systems adapted to civilian healthcare, particularly in EDs and disaster settings. In modern healthcare, triage systems vary, but they all serve the primary function of categorizing patients based on the urgency of their medical needs. For example, in civilian EDs, patients are typically categorized into five levels, from life-threatening conditions (Level 1) to non-urgent cases (Level 5). The goal is to ensure that patients with the most critical needs receive immediate attention, while those with less severe conditions may wait longer.

Triage has been further adapted in mass casualty incidents (MCIs) and disaster management, where the challenge is often the scarcity of medical resources. In these settings, triage officers rapidly assess the condition of multiple casualties and assign treatment priorities based on the likelihood of survival and the availability of resources. This form of triage differs from routine ED triage, as it often involves making difficult decisions about who can be saved with limited interventions.

In the modern era, several structured triage systems are in place worldwide. The most commonly used systems include the Simple Triage and Rapid Treatment (START) system[56], which is used in many countries for MCI response, and the ESI, which is widely implemented in the U.S. and other countries for ED triage[2]. Each of these systems incorporates the fundamental principles of triage: assessing the patient's condition, categorizing the urgency of care, and allocating resources accordingly.

In conclusion, triage has evolved from its military origins during the Napoleonic Wars to become an essential tool in both military and civilian medicine[57, 58]. It ensures that patients receive care based on the severity of their condition, helping to maximize survival rates in situations where resources are limited. Whether on the battlefield, in an emergency department, or during a disaster, triage remains a critical component of effective medical care[59].

2.4 Global Emergency Department Triage Systems

ED triage plays a critical role in modern healthcare systems by ensuring that patients receive the right level of care at the right time. Triage systems allow healthcare professionals, primarily nurses, to assess, sort, and prioritize patients based on the urgency of their conditions[60]. The goal is to ensure that the most critically ill or injured patients receive immediate care, while those with less severe conditions are treated in an appropriate timeframe. This is particularly important in busy or resource-constrained EDs where demand often exceeds the available resources.

Triage systems have evolved significantly over the past few decades, with many modern approaches incorporating "trace and trigger" tools that assess vital signs and use systematic questionnaires to evaluate each patient's chief complaint and clinical findings. The most widely used triage systems were developed during the 1990s and early 2000s[1]. These include the Canadian Triage and Acuity Scale (CTAS), Australasian Triage Scale (ATS), Emergency Severity Index (ESI), Manchester Triage System (MTS), and South African Triage Scale (SATS). Each system is designed to categorize patients into different levels of urgency, typically using a five-level structure, to guide medical teams in prioritizing care.

In the United States, triage nurses are usually the first point of contact for patients arriving at an ED. They assess and categorize all incoming patients to ensure that those in the most urgent need are treated first, while those with less pressing concerns may wait longer. This system is designed to ensure that care is delivered efficiently and appropriately based on clinical need. Triage not only helps allocate resources effectively but also upholds the principle of "clinical justice," which aims to match the level of care provided to the patient's clinical condition[61].

When EDs experience a surge in patient volume that exceeds the available resources, such as during a mass casualty event or a natural disaster, triage becomes even more crucial. In such cases, triage officers must make decisions on how to distribute limited resources, ensuring that the most critical patients are treated first, while acknowledging that not all needs may be met immediately[2]. This principle of resource allocation underscores the importance of well-functioning triage systems, especially in times of crisis.

The five-level triage system has become the gold standard in many countries, offering a reliable and validated method for assessing patient severity. Five-level systems such as ESI, CTAS, ATS, and MTS have been widely studied and have shown reasonable validity in predicting the severity of patients' conditions in EDs[4]. These systems are designed to optimize the triage process, minimize waiting times for patients, and ensure the efficient allocation of resources.

Despite the global adoption of these systems, their performance varies considerably depending on the context. Research has shown that each triage system has its strengths and weaknesses, with performance influenced by factors such as patient demographics, healthcare settings, and regional resource availability. A systematic review of 33 different triage systems found that ESI, MTS, and

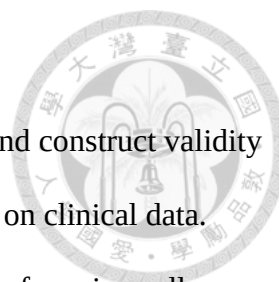
CTAS were the most frequently studied. Other systems, such as the SATS and the Taiwan Triage System, have also been evaluated but have received less global attention[8].



While these established triage systems generally show good validity for assessing patient urgency in the ED, there are still areas for improvement. Previous Studies have examined the reliability, sensitivity, and specificity of these systems in predicting key outcomes such as ED mortality, critical illness, and hospitalization[7, 8]. Common issues identified across all triage systems include variability in triage decisions, particularly in cases of complex or atypical presentations. Furthermore, systems often struggle with accurately predicting certain conditions like sepsis or non-ST elevation myocardial infarction (NSTEMI), where subtle signs may be easily overlooked.

Advancements in healthcare technology, such as the use of electronic health records (EHRs) and decision support tools, offer new opportunities to improve the accuracy and consistency of triage systems[7]. These technologies can help reduce human error and variability, ultimately enhancing the overall performance of triage in EDs.

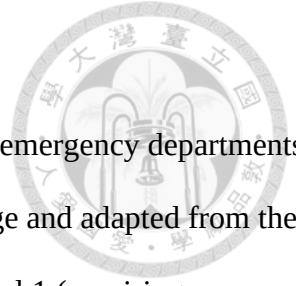
Professional organizations like the American College of Emergency Physicians (ACEP) and the Emergency Nurses Association (ENA) have recognized the importance of standardizing triage practices. In 2003, ACEP and ENA critically reviewed the use of five-level triage systems in the United States, recommending that EDs adopt either CTAS or ESI as their primary triage tool[5]. Both systems are evidence-based and have demonstrated high levels of accuracy in predicting patient outcomes. However, the task force also emphasized the need for ongoing research to further refine these systems and address any weaknesses.



Validation of triage systems is typically conducted through criterion and construct validity studies, which assess how well the system predicts patient outcomes based on clinical data. However, one of the key challenges in validating triage systems is the lack of a universally accepted reference standard for urgency. Most studies either create their own reference standards, often using expert panels, or rely on a combination of severity and resource use measures to approximate urgency. While these methods provide valuable insights, they may not fully capture the complexities of triage in real-world settings.

In conclusion, while modern ED triage systems like ESI, CTAS, ATS, MTS, and SATS have been widely adopted and validated in many healthcare settings, there is still room for improvement. Future research should focus on refining these systems to address their limitations and improve their accuracy, particularly in challenging clinical situations. As healthcare continues to evolve, the integration of technology and data-driven decision-making holds great promise for enhancing the performance of triage systems and ensuring that patients receive the care they need in a timely and efficient manner.

To describe global emergency department triage systems, let's start with a brief overview of some widely recognized systems, including CTAS, ATS, ESI, MTS, and SATS (Table 1). Each of these systems has been adopted to improve patient flow, prioritize care, and manage resources in EDs. Performance between these systems can vary across different regions, situations, and patient populations, and ongoing research assesses their validity and reliability.



2.4.1 Canadian Triage and Acuity Scale (CTAS)

The CTAS is a five-level system designed to prioritize patient care in emergency departments based on the severity of their condition. Developed by Dr. Robert Beveridge and adapted from the Australian National Triage Scale(NTS), CTAS classifies patients from Level 1 (requiring immediate care) to Level 5 (least urgent). It ensures efficient resource allocation and timely medical intervention. CTAS is widely used in Canada and has been adapted by several Asian countries. Taiwan introduced the TTAS in 2010, Japan implemented the Japanese Triage and Acuity Scale (JTAS) in 2012, and Korea developed the Korean Triage and Acuity Scale (KTAS) in 2012[7, 13, 62-64].

CTAS in Adults

For adult patients, CTAS has shown strong correlations between triage levels and critical outcomes like hospitalization, imaging needs, and ICU admissions. Its accuracy helps manage overcrowded EDs, ensuring that urgent cases are treated promptly. In multicenter studies, CTAS has been highly effective in differentiating the urgency of cases, allowing for better resource management in busy ED environments[65, 66].

CTAS in Elderly Patients

Studies show that CTAS effectively predicts resource needs and life-saving interventions for elderly patients. In a study of patients aged 65 and older, CTAS was highly sensitive (97.9%) and specific (89.2%) in identifying those requiring immediate care, such as ICU admission or interventions[66]. This demonstrates CTAS's effectiveness in handling the complex health issues common among elderly patients.



CTAS in Pediatric Patients

In pediatric populations, CTAS has shown good inter-rater reliability, with a kappa score of 0.74. This reliability ensures that triage decisions for children are consistent, allowing for accurate care prioritization. The system's ability to predict hospitalization and resource use in children further demonstrates its utility in pediatric EDs[67].

Nurse Inter-Rater Reliability

Research consistently shows high inter-rater reliability among nurses using CTAS across different patient groups. This consistency is crucial for ensuring standardized care in EDs. In both adult and pediatric settings, nurses' agreement in assigning triage levels has been good, which improves overall patient care quality[63].

In conclusion, CTAS is a validated, reliable tool for assessing and triaging patients of all ages, ensuring timely and appropriate care. Its high sensitivity and nurse reliability make it a cornerstone in modern emergency medicine.

2.4.2 Australasian Triage Scale (ATS)

The ATS is a five-level system designed to prioritize patient care in emergency departments across Australia and New Zealand. It categorizes patients from critical cases needing immediate attention (Category 1) to non-urgent cases (Category 5), ensuring that medical assessments are completed within specific time frames based on the urgency of the condition. Developed in 2000, ATS replaced the older Australian NTS and has since improved efficiency, resource allocation, and

patient outcomes in emergency care settings. It is recognized and used internationally for triage standardization.[7, 12, 68].



Reliability studies, especially those assessing inter-rater reliability, consistently show moderate to substantial agreement among healthcare professionals using ATS. A meta-analysis indicated that while the ATS showed fair to moderate reliability overall, it performed better in adult populations than in pediatric populations[12]. In terms of severe sepsis, the ATS demonstrated a sensitivity of 71%, with the highest validity observed in ATS category 2[69].

ATS has been extensively validated in both urban and rural Australian hospitals, providing reliable benchmarks for maximum waiting times. However, challenges persist in rural vs. urban applications, as the system has shown difficulty in achieving uniform implementation across diverse geographic settings[3]. Despite these issues, ATS has remained an important tool for resource management in emergency care, and it has influenced the development of triage systems in Canada and the United Kingdom.

2.4.3 Emergency Severity Index (ESI)

The ESI is a five-level triage system developed by emergency physicians Richard Wuerz and David Eital. It prioritizes patients from Level 1 (requiring immediate life-saving intervention) to Level 5 (least urgent), helping emergency departments manage patient flow efficiently. ESI is based on the resources needed for patient care, except for Levels 1 and 2, which focus on the urgency of the condition. Initially used for patients over 14, ESI has evolved to include pediatric criteria and

has undergone multiple updates. By 2019, 94% of U.S. emergency departments had adopted the ESI, making it the most widely used triage tool in the U.S.[7, 70-72].



Validation of the ESI

Validation focuses on how accurately the ESI identifies patient severity, predicting outcomes like hospital admission or the need for critical care. Studies have demonstrated that ESI is a valid tool for both pediatric and adult populations. In pediatric settings, ESI has shown good validity with a sensitivity of around 52% and a specificity of 81%, indicating that it moderately predicts which patients require urgent care or hospitalization. The accuracy of ESI can be measured through its Area Under the Curve (AUC) values from ROC analysis. Studies report an AUC of approximately 0.78, which supports ESI's effectiveness in differentiating between high- and low-acuity patients[73].

Reliability of the ESI

Meta-analyses indicate that ESI has substantial inter-rater reliability, with pooled kappa values typically around 0.79[74]. This shows that healthcare professionals tend to agree on patient triage levels when using ESI, which is vital for consistent care delivery. Notably, reliability improves when the ESI is applied by experts or when using its latest versions.

In summary, the ESI is a reliable and valid triage tool widely used in emergency departments. It has shown consistent performance in predicting patient outcomes, though it requires further refinement to minimize under-triage and ensure full reliability across different healthcare contexts. Continued training and the use of the latest versions are essential for maintaining its high reliability and improving patient care.



2.4.4 Manchester Triage System (MTS)

The MTS, developed between 1994 and 1996 by nurses and physicians in Manchester, UK, is a triage tool that categorizes patients into five urgency levels: red (immediate), orange (very urgent), yellow (urgent), green (standard), and blue (non-urgent). It uses clinical algorithms based on the patient's symptoms and history to systematically assess their condition and determine the appropriate triage category. MTS is widely used across Europe, including Germany, where it has been adapted to fit local healthcare needs. It helps ensure timely care for critical patients while efficiently managing less urgent cases.[7, 75].

Validation of the MTS

Several studies have examined the validity of the MTS, typically using outcomes like hospital admissions, mortality, and ICU needs as benchmarks. A prospective study involving over 288,000 patients found that the sensitivity of the MTS in detecting highly urgent patients ranged from 0.47 to 0.87 in adults and from 0.65 to 0.83 in children, depending on the hospital setting. This means that while the MTS correctly identifies high-urgency patients most of the time, there is still a significant proportion of patients who might be under-triaged. Specificity was consistently higher, ranging from 0.84 to 0.94 in adults and from 0.83 to 0.89 in children, indicating that the MTS is effective in identifying lower urgency cases[76].

The system's validity, however, shows variability between different hospitals and patient groups. It performs best in adult populations but tends to under-triage children and the elderly. For instance, up to 14% of adults and 34% of children needing ICU care were misclassified as

low-urgency[77]. This variability highlights the need for continuous adjustments to the MTS to ensure its accuracy across different settings and patient demographics.



Reliability of the MTS

Reliability measures focus on the consistency of triage decisions across different users. Studies using the kappa statistic to measure inter-rater reliability report a wide range of agreement levels, from moderate to very good (kappa values between 0.40 to 0.82). For example, the inter-rater reliability in some settings was as high as 0.82 with quadratic weighting, showing very good agreement, while unweighted kappa values tended to be lower (around 0.49 to 0.76)[77]. These differences suggest that the reliability of MTS is influenced by the use of specific weighting methods and the experience level of the triage staff.

MTS has been validated for both adult and pediatric populations, and studies suggest it performs well in high-resource settings. However, there are concerns about its subjectivity, with some studies highlighting variability in interrater reliability between different clinical staff.

2.4.5 South African Triage Scale (SATS)

The SATS was developed in 2006 to improve emergency care by quickly assessing and prioritizing patients based on the severity of their condition. Initially designed for South African hospitals, SATS uses a color-coded system ranging from red (emergency) to green (routine) and includes a Triage Early Warning Score (TEWS), which is calculated based on vital signs and clinical discriminators. SATS has been widely implemented in low-resource settings and has proven effective in improving patient outcomes and managing emergency department workflow[78-82].



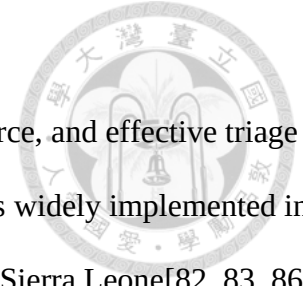
Validation of SATS

SATS has been validated in several low-resource settings, demonstrating good performance in both trauma-only and mixed medical settings. Studies in Afghanistan and Haiti showed that SATS could effectively predict patient outcomes such as hospital admission and mortality, with a sensitivity ranging from 70% to 75% and a specificity of 91% to 97%. In a study at Kenyatta National Hospital in Kenya, the sensitivity of SATS in predicting admission or discharge was found to be 92.2%, though the specificity was lower at 37.7%. Additionally, SATS was effective in improving patient flow and prioritizing critical cases for immediate care, especially in trauma settings.

However, challenges remain, especially in pediatric and non-trauma settings, where SATS has demonstrated under-triage rates as high as 76% for non-trauma cases[83, 84]. This highlights the need for adaptations, such as incorporating hemoglobin levels in malaria-endemic areas to improve its predictive power.

Reliability of SATS

Reliability studies, especially in low-resource settings like Haiti and Afghanistan, show that SATS has moderate inter-rater and intra-rater reliability, with intraclass correlation coefficients (ICC) ranging from 0.50 to 0.60. These findings suggest moderate agreement between healthcare providers using SATS, particularly in settings with limited resources and training[85]. In prehospital settings, SATS has demonstrated an under-triage rate of around 30% and an over-triage rate of 13%, indicating the tool's reliability can vary depending on the context and training of providers[81].



SATS is primarily used in LMICs where healthcare resources are scarce, and effective triage is essential for managing the overwhelming demand for emergency care. It is widely implemented in countries like South Africa, Mozambique, Kenya, Haiti, Afghanistan, and Sierra Leone[82, 83, 86] [84]. Despite its effectiveness, SATS is not universally recognized as an official triage method in some regions due to variability in its performance, especially outside trauma settings[79]. Nevertheless, SATS remains a valuable tool for improving emergency care delivery in resource-constrained environments.

2.4.6 Performance across Different Settings and Groups

A systematic review by Zachariasse et al. evaluated 33 different triage systems, focusing on the ESI, CTAS, and MTS, which were the most studied systems. The review highlighted that while these systems show moderate to good validity in identifying high-urgency patients, performance is highly variable. The ESI, for example, incorporates resource use into its decision-making, making it particularly useful in high-resource settings like the United States. In contrast, SATS, designed for resource-limited settings like South Africa, focuses on matching demand with available capacity.

In terms of identifying high-urgency cases, studies have shown that while most triage systems perform well in detecting critically ill patients requiring ICU admission, there is significant variability. Sensitivity rates for ICU admissions ranged from 58% to 88% for adults and 71% to 93% for children[8]. However, no single system consistently outperforms others, with each showing areas of both strength and weakness. For instance, ESI and MTS have been noted for over-triaging some patients, while under-triaging remains a concern across all systems.



2.4.7 Evidence-Based Validation and Real-World Performance

In real-world settings, the reliability of triage systems depends heavily on contextual factors like patient demographics, healthcare resources, and the experience of healthcare providers. Studies assessing the validity of these systems often use outcomes like ED mortality, ICU admission, and hospital discharge as benchmarks. For example, the ESI has demonstrated good performance in high-resource environments but is less effective in settings with fewer resources. Similarly, the CTAS has shown variability in performance depending on the region, patient volume, and case severity[7].

Zachariasse et al.'s systematic review and meta-analysis revealed that while these triage systems show reasonable validity, performance varies significantly across studies due to differences in methodology and reference standards. ICU admission rates, for example, varied across systems, with the MTS showing ICU admission rates of 21% to 79% for high-urgency cases[8].

Overall, these studies highlight the importance of continuously evaluating and adjusting triage systems to suit local healthcare environments. More research is needed to find ways to optimize these systems to enhance patient outcomes in different settings. Despite the challenges, well-established triage systems like ESI, CTAS, MTS, ATS, and SATS serve as essential tools for managing emergency care globally. However, even though these systems have been validated in various studies, there remains room for improvement, particularly in creating global standards and reducing regional differences in triage accuracy and performance.



Figure 1 Global Distribution of Triage System

This map illustrates the global distribution of triage systems. Countries with no available data or where information is not publicly disclosed are shaded in grey.

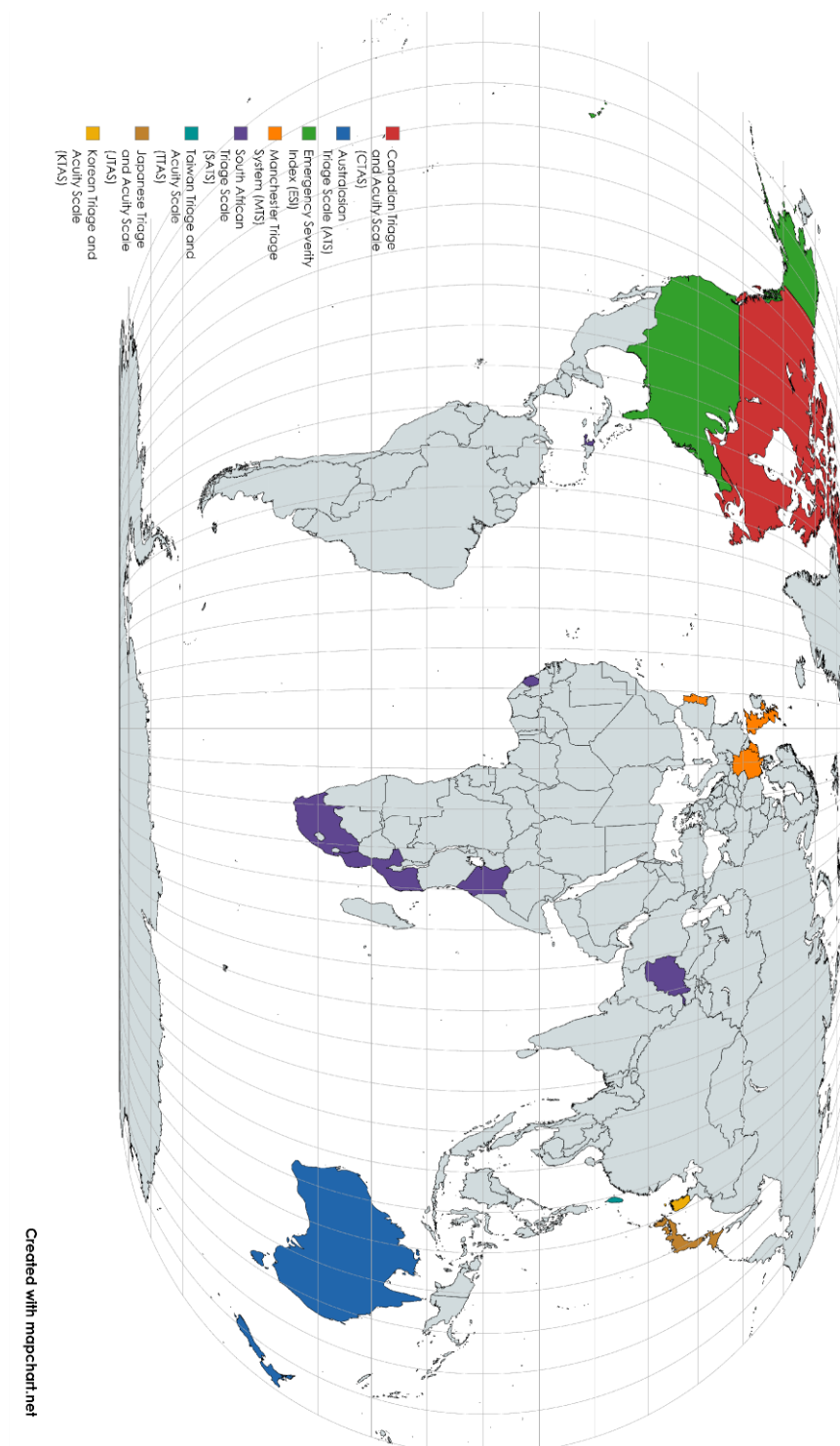




Table 1 Triage system characteristics

Triage System	CTAS	ESI	MTS	ATS	SATS
Stated Objective	Provide timely care based on urgency	Prioritize based on immediacy and resource use	Assess quickly and assign based on clinical need	Treat in order of urgency and allocate to the right area	Prioritize based on urgency, especially in high-demand settings
Recommended time to physician contact, min	1: Immediate 2: ≤ 15 3: ≤ 30 4: ≤ 60 5: ≤ 120	1: Immediate 2: ≤ 15 3: None 4: None 5: None	Red: Immediate Orange: ≤ 10 Yellow: ≤ 60 Green: ≤ 120 Blue: ≤ 240	1: Immediate 2: ≤ 15 3: ≤ 30 4: ≤ 60 5: ≤ 120	Red: Immediate Orange: ≤ 10 Yellow: ≤ 60 Green: ≤ 240 Blue: dead on arrival, certified within ≤ 120
Discriminators					
Clinical	Yes	No	Yes	Yes	Yes
Vital signs	Yes	Yes	Yes	Yes	Yes
Pain score	Yes (10-point)	Yes (visual scale)	Yes (3-point)	Yes (10-point)	Yes (4-point)
Resource use	No	Yes	No	No	No
Pediatrics	Separate version	Different vital sign rules	Considered in algorithm	Considered in algorithm	Separate flowchart



2.5 Global Prehospital Triage Systems

Global prehospital triage systems are essential in ensuring that emergency care begins in the field, facilitating appropriate treatment and resource utilization in a timely manner. These systems, like the Canadian Triage and Acuity Scale (CTAS), Emergency Severity Index (ESI), and various Early Warning Scores (EWS), serve as a critical bridge between prehospital settings and hospitals.

2.5.1 Overview of Global Prehospital Triage Systems

Prehospital triage systems aim to classify patients based on the severity of their condition, guiding paramedics and healthcare workers to prioritize care effectively. Several systems are widely used:

Canadian Triage and Acuity Scale (CTAS)

Developed for both hospital and prehospital use, CTAS is a five-level scale that allows healthcare professionals to assign urgency scores to patients based on their symptoms and vital signs[9, 17].

Emergency Severity Index (ESI)

A similar five-level triage system, ESI is predominantly used in emergency departments but has also been tested in prehospital settings. It categorizes patients based on their need for immediate resources rather than specific diagnoses[87].

Early Warning Scores (EWS)

Early Warning Scores (EWS) are clinical tools designed to detect early signs of patient deterioration by monitoring vital signs like heart rate, respiratory rate, temperature, blood pressure, and consciousness. Systems like NEWS (National Early Warning Score) and its updated version, NEWS2, have gained popularity in prehospital and emergency settings. These scores help healthcare providers, including paramedics and emergency physicians, identify high-risk patients who may require urgent care or hospitalization[88-90].

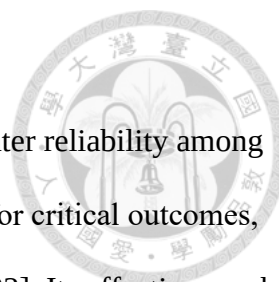
Each system operates under specific frameworks, with slight modifications depending on the region and healthcare infrastructure. Their validation and real-world applicability depend on their ability to maintain consistency and accuracy across diverse healthcare settings.

2.5.2 Comparison Between Different Settings and Groups

Different triage systems perform variably across settings, such as urban versus rural or high-income versus low-income regions. Studies show significant variations in accuracy and reliability between EMS personnel and ED triage nurses. For example:

CTAS in EMS vs. ED Settings

A study comparing EMS-assigned CTAS scores to ED-assigned scores found moderate agreement between paramedics and triage nurses ($\kappa = 0.437$). The system had an 84.3% rate of exact or near-exact agreement between paramedics' prehospital CTAS and ED nurses' initial CTAS[17, 91]. The primary challenges faced by EMS personnel include differences in access to diagnostic tools and time constraints, affecting decision-making.



ESI in EMS vs. ED Settings

ESI has been well-validated in hospital settings, showing high inter-rater reliability among nurses (κ values ranging from 0.61 to 0.89) and strong predictive validity for critical outcomes, such as the need for life-saving interventions and hospital admissions[15, 92]. Its effectiveness has been demonstrated in a wide range of patient populations, including pediatric, adult, and geriatric groups, making it a robust tool for clinical decision-making in the ED.

However, when applied in prehospital settings by EMS, ESI faces challenges. A study evaluating the use of ESI by EMS personnel showed only moderate agreement with ED triage nurses, with a kappa value of 0.409. EMS providers, especially those with more experience, were less accurate in assigning triage levels, which may be due to the lack of diagnostic resources and different training compared to ED staff. Specificity for correctly not assigning certain levels was higher than sensitivity, indicating EMS personnel's struggle with under-triage. Despite these challenges, incorporating ESI into prehospital care could streamline transitions from EMS to ED, though further training and adaptations are required for its effective use[92].

ESI and CTAS Comparisons

The use of ESI and CTAS in different settings reveals disparities in agreement rates. A study comparing ESI and CTAS showed weak correlations ($\kappa = -0.003$), emphasizing a need for further alignment between prehospital and ED settings[16]. ESI's resource-based categorization may lead to different interpretations in prehospital environments where access to diagnostics is limited.

EWS

Early Warning Scores (EWS), particularly NEWS2, are effective in predicting short-term outcomes such as early mortality and the need for intensive care. With a sensitivity of up to 81%, NEWS2 has proven to be a reliable triage tool in both emergency departments and prehospital settings. However, its ability to predict longer-term outcomes, like 30-day mortality, is limited. While NEWS2 performs well in identifying immediate clinical deterioration, its predictive accuracy decreases over time, and its effectiveness can vary depending on the patient population and the healthcare context[88, 90].

2.5.3 Validation and Evidence-Based Performance

Despite their widespread adoption, the evidence supporting the reliability and validity of these systems is mixed. Several systematic reviews have highlighted both successes and gaps in the research:

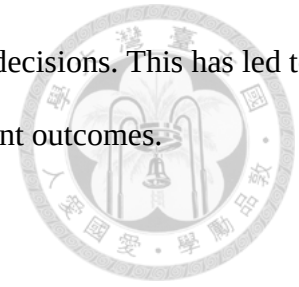
CTAS

The Canadian Triage and Acuity Scale has been validated extensively in both ED and prehospital settings, with moderate reliability ($\kappa = 0.43$ to 0.45)[17]. However, concerns about under- and over-triage remain, particularly in more complex cases where patients' conditions fluctuate between the time of first contact and arrival at the hospital[10, 91]. Studies suggest that real-time clinical decision support tools may improve accuracy[91].

ESI

Evidence supporting ESI in prehospital settings is more limited. Its resource-based categorization method may not translate well outside of hospitals, where EMS personnel lack access

to many of the diagnostic tools necessary for making resource-dependent decisions. This has led to inconsistencies in score assignment, with significant implications for patient outcomes.



EWS

Studies indicate that NEWS2 performs well in predicting short-term mortality but has significant limitations in predicting long-term outcomes, particularly when used in prehospital settings. A systematic review of EWS scores concluded that while they are useful for detecting early signs of deterioration, their effectiveness diminishes when applied outside of hospital environments[88, 90].

Global Scope and Limitations

A scoping review of global prehospital triage systems revealed that most studies are concentrated in high-income countries. There is a gap in research on the implementation and validation of these systems in low- and middle-income countries[10]. Prehospital triage systems like CTAS and ESI may not be fully adaptable to environments with fewer resources and different health burdens[87].

In Summary, Global prehospital triage systems play a critical role in emergency care, but their real-world performance is influenced by a variety of factors, including setting, available resources, and the training of healthcare providers. While systems like CTAS and ESI have shown promise in certain contexts, their application in prehospital settings requires further validation, particularly in low- and middle-income countries. The future of prehospital triage lies in refining these tools to improve accuracy and usability across diverse healthcare environments. Real-time decision-support technologies and enhanced training programs could significantly enhance the reliability of triage in

the field. However, more high-quality research is needed to bridge the gap between hospital-based systems and prehospital realities.



2.6 Triage System in EDs in Asia

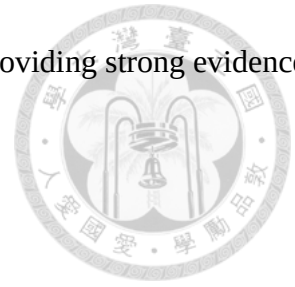
Triage systems in Asian EDs prioritize patients based on urgency to streamline care. Taiwan, South Korea, and Japan have adapted the CTAS into their own five-level systems: Taiwan Triage and Acuity Scale (TTAS), Korean Triage and Acuity Scale (KTAS), and Japan Triage and Acuity Scale (JTAS). These systems help manage patient flow, allocate resources efficiently, and deliver timely care.

2.6.1 Taiwan Triage and Acuity Scale (TTAS)

The TTAS was introduced in 2010 as a five-level system designed to prioritize emergency patients based on the urgency of their medical needs. It was adapted from the CTAS but modified to better fit Taiwan's healthcare context. This computerized system assesses patients' conditions through 163 chief complaints across 14 categories, ranging from respiratory and cardiovascular issues to traumatic injuries. TTAS incorporates both trauma and non-trauma categories, allowing for a more precise assessment of patient severity. It includes shorter reassessment intervals for higher acuity levels and specific criteria for vital signs and pain management. TTAS has since improved patient flow and resource allocation in Taiwan's emergency departments, making it an essential tool in enhancing patient care[14, 28, 93, 94].

Validation and Real-World Evidence

Several studies have validated TTAS in different medical contexts, providing strong evidence for its reliability and practical use in real-world emergency settings.



General Validation

TTAS was validated in a multicenter study involving more than 10,000 patients across academic medical centers and hospitals. The results demonstrated high reliability, with a weighted kappa statistic of 0.87, indicating excellent inter-rater agreement among triage nurses. This shows that TTAS can consistently prioritize patients based on their acuity[14]. Furthermore, TTAS was found to better stratify patients compared to the older four-level system, reducing over-triage and under-triage, thus improving resource utilization[14, 95].

Specific Use Cases

In stroke care, TTAS has been shown to correlate well with patient outcomes, particularly in identifying patients who need urgent interventions like thrombolysis. In one study, 98.8% of thrombolysis candidates were correctly categorized into TTAS level 1 or 2, showing that the system is effective in recognizing high-acuity cases[95]. Similarly, TTAS has been validated for pediatric trauma care, where it demonstrated strong predictive power for key outcomes such as hospitalization rates and the need for emergency surgery[94].

Impact on Emergency Department Operations

TTAS has had a significant impact on managing ED overcrowding, particularly in distinguishing non-urgent from urgent cases. In a study examining non-urgent patients (TTAS levels 4 and 5), predictors for hospitalization were identified, helping refine the triage process.

Although these patients typically do not need immediate care, a subset required further hospital treatment, emphasizing the importance of a nuanced triage system like TTAS[96].



Electronic Decision Support

A key feature of TTAS is its integration with a computerized decision support tool, which enhances its accuracy by reducing the subjectivity of triage decisions. This system is designed to assess patient conditions systematically and suggest appropriate triage levels based on clinical data. The introduction of electronic decision support has further improved the consistency and reliability of the system, as evidenced by studies showing improved triage performance[97].

TTAS has proven to be a reliable and valid triage system in Taiwan's emergency departments. It has improved patient stratification, resource allocation, and overall ED efficiency. With its computerized decision support tool, TTAS represents a significant advancement over older triage systems, offering better accuracy and consistency in patient care. However, ongoing refinements, particularly for special populations like the elderly, are needed to maximize its effectiveness.

2.6.2 Korean Triage and Acuity Scale (KTAS)

The KTAS was developed in 2012 based on the CTAS and has been implemented in all Eds across South Korea since January 2016. KTAS is a five-level system that categorizes patients based on the urgency of their condition, with Level 1 representing resuscitation and Level 5 representing non-urgent cases. The system incorporates both vital signs and chief complaints to prioritize patients for treatment[98, 99].

Validation and Real-World Evidence of KTAS



General Validation

Multiple studies have demonstrated the effectiveness and reliability of KTAS in real-world emergency settings. A prospective study assessed the interrater reliability of KTAS among emergency nurses and found a weighted kappa value of 0.772, indicating good reliability. The study identified that consistency in selecting complaints and modifiers improved the reliability of the triage process[100]. Another study comparing KTAS outcomes with those predicted by CTAS found significant correlations in hospitalization rates, demonstrating KTAS's predictive validity[101].

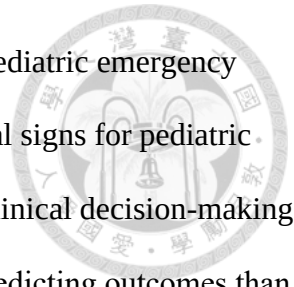
Predictive Validity for Severe Trauma and Mortality

KTAS has been validated in trauma care, particularly in predicting outcomes for severely injured patients. A study of patients with an injury severity score ≥ 16 showed that lower KTAS levels were significantly associated with higher 30-day mortality. Specifically, the mortality rate for patients categorized as Level 1 was 52.5%, compared to 15.5% for Level 2 and 4.1% for Level 3. These findings underscore KTAS's ability to accurately predict mortality in trauma patients, proving its utility in high-acuity settings[98, 102].

Pediatric Triage

The pediatric version of KTAS (PedKTAS) has been in use since 2016 for patients under 15 years of age. Studies evaluating PedKTAS have demonstrated its effectiveness in predicting key outcomes such as hospitalization and ICU admission rates. In a nationwide study involving over 5 million pediatric patients, hospitalization rates were highest for Level 1 patients (63.5%) and lowest

for Level 5 patients (3.7%), showing that PedKTAS effectively stratifies pediatric emergency cases[103]. However, some limitations have been noted in using initial vital signs for pediatric triage, particularly due to anxiety-related abnormal readings in children. Clinical decision-making based on the child's general condition was found to be more accurate in predicting outcomes than relying solely on vital signs[101].



Impact on Emergency Department Efficiency

The implementation of KTAS has been associated with improvements in emergency department operations, including reduced LOS and better resource allocation. A quasi-experimental study conducted at a tertiary hospital showed that after KTAS implementation, the mean LOS in the ED decreased from 343 minutes to 289 minutes, and mortality rates also declined significantly (from 0.36% to 0.28%)[102]. This demonstrates KTAS's positive impact on improving ED efficiency while ensuring that critical patients receive timely care.

Challenges and Limitations

Despite its general reliability, KTAS faces challenges in accurately triaging certain populations, particularly older adults. Studies have shown that KTAS tends to under-triage elderly patients, leading to delayed care for this vulnerable group. One study found that elderly patients had a higher likelihood of up-triage compared to adults, indicating that KTAS might underestimate the severity of conditions in older patients. This has led to calls for the modification of KTAS criteria to account for the complexities of aging[98, 104]. Another study introduced a Clinical Frailty Scale-adjusted KTAS for elderly patients, which improved the system's prognostic accuracy, particularly for predicting ICU admissions and mortality[104].



Mistriage and Triage Accuracy

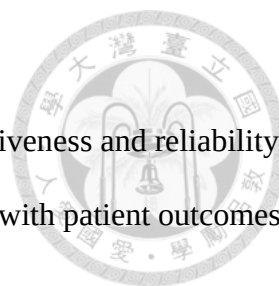
A retrospective study analyzing KTAS mistriage rates in two emergency departments found an overall triage accuracy of 85.3%, with under-triage occurring in 10.3% of cases and over-triage in 4.4%[105]. The primary causes of mistriage were identified as errors in applying the pain scale and misjudgments of the severity of physical symptoms. These findings highlight the need for continuous training to improve the accuracy of KTAS triage.

KTAS has proven to be a reliable and valid tool for prioritizing patients in emergency departments. Its implementation has led to improvements in patient flow, reduced length of stay, and better mortality outcomes. However, certain challenges remain, particularly in the accurate triage of older adults and pediatric patients.

2.6.3 Japan Acuity and Triage Scale (JTAS)

The JTAS was introduced in 2012 as a five-level system derived from the CTAS, tailored to Japan's unique healthcare needs. It was developed by the Japanese Society for Emergency Medicine, the Japanese Association for Acute Medicine, the Japanese Society for Emergency Pediatrics, and the Japanese Association for Emergency Nursing. Since its introduction, JTAS has been widely implemented in Japanese EDs[106, 107]. The JTAS includes modifications specific to Japan, such as the inclusion of conditions like heat stroke and post-vaccination syndrome, which are more common in the country. These additions enhance the system's relevance in Japan's healthcare environment

Validation of JTAS



The JTAS has undergone several validation studies to assess its effectiveness and reliability in real-world settings. These studies have demonstrated its strong correlation with patient outcomes, such as hospital admission rates and LOS in the ED.

General Validation

A study conducted at Kurashiki Central Hospital involving over 27,000 patients validated the JTAS by examining its ability to predict overall hospital and ICU admissions. The study showed that higher acuity levels on the JTAS were strongly associated with greater odds of admission, particularly to the ICU. The area under the receiver operating characteristic curve (AUROC) for overall admission was 0.726, and for ICU admission, it was 0.792, demonstrating good predictive validity. Additionally, ED length of stay was significantly longer for patients with higher JTAS levels, confirming its usefulness in prioritizing care[108].

Impact on Emergency Department Operations

The introduction of JTAS significantly improved ED efficiency by reducing triage times and improving patient flow. A study comparing pre- and post-implementation data revealed that the time from registration to triage was reduced by 3.8 minutes, and the time from registration to physician attention was reduced by 11.2 minutes after the JTAS was implemented[109]. This highlights JTAS's ability to streamline operations in busy ED settings, ensuring that high-acuity patients receive timely care.

Pediatric Triage

JTAS has also been validated for use in pediatric emergency settings. In pediatric triage, a major challenge is over-triage due to abnormal vital signs caused by anxiety, fever, or crying, rather than clinical deterioration. JTAS addresses this by allowing experienced triage nurses to "down-triage" children based on clinical judgment. A study at a pediatric emergency department in Japan showed that down-triaging well-appearing children with abnormal vital signs improved the prediction of hospital admissions, making JTAS more effective in pediatric care[110].

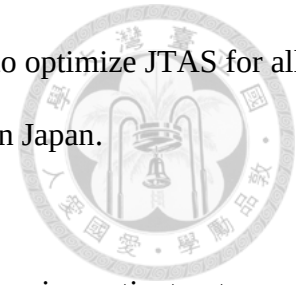
Limitations and Challenges

While JTAS has proven to be an effective triage tool, it faces certain limitations. One major challenge is the system's reliance on web-based tools for real-time decision-making, which may not be practical in all healthcare settings due to limitations in internet connectivity and concerns over patient data privacy[111].

Moreover, studies have shown that JTAS may not perform as well in elderly populations. A study involving over 11,000 elderly patients found that while JTAS effectively predicted hospital and ICU admissions, it tended to under-triage elderly patients with non-specific symptoms or chronic conditions like malignancy[106]. This highlights the need for further refinement of JTAS, particularly for geriatric populations, to avoid delays in care for this vulnerable group.

JTAS has been validated as a reliable and effective tool for emergency triage in Japan. Its implementation has led to improved patient outcomes, reduced triage times, and better resource allocation in EDs. While the system shows strong predictive validity for hospital and ICU admissions, there is room for improvement, especially in addressing the unique challenges posed by

pediatric and elderly patients. Further studies and refinements are needed to optimize JTAS for all populations, ensuring its continued success in enhancing emergency care in Japan.



In conclusion, TTAS, KTAS, and JTAS have proven effective in improving patient outcomes, reducing wait times, and optimizing resource use in Asian EDs. However, challenges persist in accurately triaging pediatric and elderly patients. Continued research and refinement are necessary to further enhance these systems.

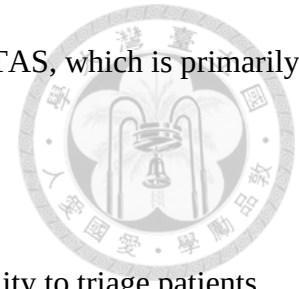
2.7 Prehospital Triage System in Asia

The TTAS, KTAS, JTAS, and 2-level TPTS are essential tools for managing emergency care by assessing patient acuity. In Taiwan, 2-level TPTS has been used for over a decade in prehospital settings, categorizing patients as either emergent or non-emergent. In Japan, a five-level triage system for EMS has been in place since 2014, allowing for more detailed prehospital triage. However, TTAS and KTAS are primarily used in emergency departments, and their prehospital application is still under research and not fully adopted by EMS teams. This section reviews the research validating these systems in prehospital settings and explores their real-world potential.

2.7.1 Two-Level Taiwan Prehospital Triage System (TPTS)

The 2-level TPTS system designed to categorize patients as either emergent or non-emergent based on their initial vital signs and chief complaints. This system has been widely used in Taiwan for over a decade to prioritize EMS responses and ensure that patients who need urgent care are swiftly identified and transported to appropriate medical facilities. However, despite its widespread use, 2-level TPTS has faced criticism for its lack of precision in classifying patients accurately,

especially when compared to more advanced triage systems such as the TTAS, which is primarily used in EDs[18].

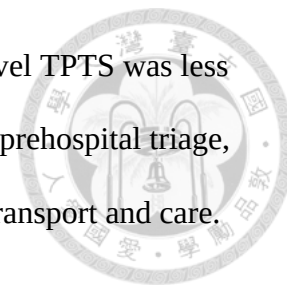


Studies have highlighted significant limitations in 2-level TPTS's ability to triage patients effectively, leading to misclassifying patient severity. A notable study comparing 2-level TPTS with TTAS found that 2-level TPTS frequently misclassified patients, which raised concerns about the system's accuracy in predicting the true acuity of patients. Specifically, the study revealed that 15.2% of patients who were classified as emergent by 2-level TPTS were downgraded to lower-acuity categories (levels 3-5) when reassessed by TTAS upon arrival at the ED. On the other hand, 30.4% of patients initially deemed non-emergent by 2-level TPTS were upgraded to higher-acuity levels (levels 1-2) by TTAS at the ED[18]. This discrepancy demonstrates the limitations of the two-tiered TPTS in providing an accurate reflection of patient needs, especially in cases where more nuanced clinical judgment is required.

Given the limitations of 2-level TPTS, there is increasing support for transitioning to the five-level TTAS in prehospital settings. Unlike 2-level TPTS, which categorizes patients as emergent or non-emergent, TTAS provides a more detailed stratification of patient acuity. This allows for better prioritization of patients who need immediate intervention, such as hospitalization or intensive care, and ensures shorter reassessment intervals for higher-acuity cases, improving care for critical patients.

A key difference between 2-level TPTS and TTAS lies in their ability to predict outcomes. A study found that TTAS had a sensitivity of 70.3% for predicting hospitalizations, significantly higher than 2-level TPTS's 41.1%[18]. TTAS also demonstrated greater specificity in ruling out the

need for hospitalization in lower-acuity patients (levels 3-5), whereas 2-level TPTS was less reliable. These findings suggest that TTAS could enhance the accuracy of prehospital triage, allowing EMS personnel to make more informed decisions about patient transport and care.

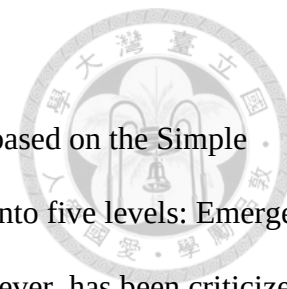


Although TTAS has not been fully adopted in prehospital care, studies show its potential. In one study, EMTs trained in TTAS achieved strong inter-rater reliability ($\kappa = 0.825$) with TRNs[21], suggesting that with proper training, TTAS could be effectively applied in prehospital environments, improving communication between EMS and ED teams. Ng et al. suggest that a more detailed five-level triage system like TTAS might offer better accuracy and resource allocation[18, 21].

TTAS also outperforms 2-level TPTS in resource allocation by better predicting which patients require intensive care, helping reduce unnecessary hospital admissions and improving patient flow through the ED. While TTAS offers many advantages, further validation and training are needed before it can be fully integrated into Taiwan's EMS operations.

2.7.2 Korean Triage and Acuity Scale (KTAS) in Prehospital Settings

KTAS was introduced in 2012 and widely implemented in South Korean hospitals to streamline the triage process. However, its application in prehospital care has been relatively limited. A study aimed at assessing the correlation between the prehospital triage system used by 119 paramedics and KTAS in the hospital setting found a low agreement between the two systems. The study showed that the kappa coefficient—a measure of agreement—was 0.211, indicating poor correlation between prehospital and in-hospital assessments[23].



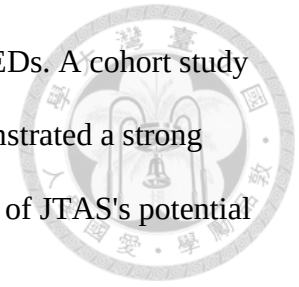
The prehospital triage system used by paramedics in South Korea is based on the Simple Triage and Rapid Treatment (START) model, which categorizes patients into five levels: Emergent, Less Emergent, Potentially Emergent, Extra, and Death. This system, however, has been criticized for its limited ability to accurately stratify patient acuity compared to KTAS. A study involving 1114 patients transported by ambulance showed that many patients were over-triaged or under-triaged during prehospital care compared to their KTAS classification upon arrival at the ED. For instance, among patients categorized as Level 3 in the prehospital setting, 27.3% were reclassified as Level 1 (Resuscitation) upon arriving at the ED[23]. This discrepancy highlights the need for a more integrated and standardized triage system that can be consistently applied across both prehospital and hospital settings.

Despite these challenges, KTAS has demonstrated some potential for improving prehospital care. Studies suggest that training paramedics in KTAS and implementing a unified triage system across both prehospital and in-hospital settings could improve the accuracy of triage decisions in the field. This would ensure that critically ill patients are identified and transported to the appropriate facilities without unnecessary delays. However, additional research is needed to validate KTAS specifically for prehospital use and to address the discrepancies between prehospital and hospital triage assessments.

2.7.3 Japan Triage and Acuity Scale (JTAS) in Prehospital Settings

JTAS, modeled after the CTAS, was introduced in Japan in 2012 and has been widely implemented in hospital EDs. In 2014, Japan began testing JTAS for use in prehospital settings to

reduce the growing demand for ambulance services and overcrowding in EDs. A cohort study aimed at validating the five-level prehospital triage system in Japan demonstrated a strong correlation between triage acuity and ICU admissions, providing evidence of JTAS's potential effectiveness in the field.



The study, conducted in Kurashiki Central Hospital, included 1261 patients transported by ambulance. The researchers found that patients categorized as Level 1 (Resuscitation) were 9.62 times more likely to be admitted to the ICU compared to patients in Levels 4 and 5. This finding supports the use of JTAS in identifying critically ill patients who require immediate intensive care. Furthermore, the study showed that JTAS could help prioritize patients more effectively, with higher acuity levels being strongly associated with ICU admissions and overall hospital admissions. The AUC for predicting ICU admissions was 0.738, indicating good predictive validity[25].

Despite these promising results, challenges remain in applying JTAS in prehospital settings. The system relies heavily on EMS staff selecting the correct modifiers (e.g., vital signs, chief complaints), and any errors in these selections could lead to misclassification. Additionally, while JTAS has demonstrated strong predictive validity for ICU admissions, it has been less effective at identifying certain conditions, such as cerebral infarctions, which were often under-triaged as low-acuity in the prehospital setting. This suggests that further refinement of the triage algorithm, particularly for neurological conditions, is needed to improve JTAS's performance in the field.

The TTAS, KTAS, JTAS, and 2-level TPTS systems have been validated to varying degrees for prehospital use, showing effectiveness in improving patient care and resource allocation. TTAS and JTAS have proven especially useful in predicting hospital admissions and ICU needs,

benefiting EMS in Taiwan and Japan. While KTAS has been validated for trauma care, it struggles with consistency between prehospital and hospital triage. 2-level TPTS, though widely used, is less effective than TTAS and would benefit from being replaced by the five-level system. Ongoing research is needed to improve these triage systems, especially for pediatric and elderly patients.

2.8 Impact of ED LOS and Resource Use on Patient Outcomes

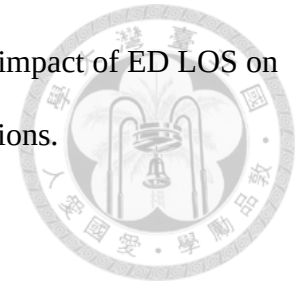
ED LOS and resource utilization have significant impacts on patient outcomes, particularly in critical care admissions. Numerous studies have investigated these effects, highlighting the challenges of prolonged ED stays due to crowding and limited hospital bed availability.

Impact of ED LOS on Mortality

Prolonged ED LOS has been associated with increased mortality, particularly for patients requiring ICU admission. One study showed that critically ill patients who experienced longer ED stays faced a greater risk of hospital mortality. Specifically, patients who stayed in the ED for over six hours had a statistically significant higher mortality rate compared to those who were admitted sooner[112]. Additionally, a study conducted in Western Australia found that overcrowded hospitals with high ED occupancy experienced increased patient mortality, particularly on days with severe overcrowding[113]. This relationship was found to be linear, with the risk of death increasing as the ED overcrowding hazard scale rose.

In contrast, not all studies have found a direct correlation between ED LOS and mortality. A study conducted in Ontario, Canada, noted that although there was a significant increase in mortality after prolonged ED stays, particularly beyond five hours, this increase did not apply

universally across all patient groups[114]. These findings suggest that the impact of ED LOS on mortality can vary depending on patient demographics and hospital conditions.



Resource Utilization and Hospital Crowding

ED overcrowding also leads to increased resource consumption and longer hospital stays for admitted patients. A study in California indicated that days with high ED crowding resulted in longer hospital stays and greater healthcare costs. Specifically, each additional hour of ED crowding added significant costs to the healthcare system and increased patient length of stay in the hospital[115]. Furthermore, studies have shown that crowding often leads to delays in care, with patients waiting longer for admission to the appropriate inpatient units, exacerbating the strain on already limited healthcare resources[116, 117].

The delay in patient transfer from the ED to inpatient beds is particularly concerning because it prolongs exposure to a chaotic ED environment, which can delay treatments like mechanical ventilation, increase the risk of infections, and compromise the continuity of care. In some cases, ED crowding has been linked to increased re-admission rates, as patients may be discharged prematurely or without adequate treatment due to resource constraints[115, 116].

Strategies to Mitigate ED LOS and Resource Strain

Efforts to address ED overcrowding focus on improving hospital flow and minimizing bottlenecks between the ED and inpatient units. Some hospitals have introduced electronic triage systems to expedite decision-making and prioritize patient care more effectively. These systems allow emergency medical teams to assess patient needs in real time and allocate resources more

efficiently, reducing the risk of prolonged stays and ensuring quicker admission for critical patients[118].



Moreover, governments and healthcare institutions are developing policies to optimize bed management and reduce crowding. For example, the implementation of the 4-hour ED discharge rule in the UK aimed to limit patient stays in the ED, requiring either discharge, admission, or transfer within this timeframe. Although not always successful in every instance, this policy has been credited with reducing ED wait times and improving patient flow[118].

The relationship between ED LOS, resource utilization, and patient outcomes is multifaceted. Prolonged ED stays, particularly during periods of crowding, are associated with higher mortality rates and increased healthcare costs. While some variability exists across different healthcare systems, the general consensus is that reducing ED LOS is critical for improving patient outcomes, especially for those requiring intensive care. Moving forward, healthcare systems must continue to innovate and invest in solutions that alleviate crowding and optimize resource allocation to minimize the negative impacts of prolonged ED stays.

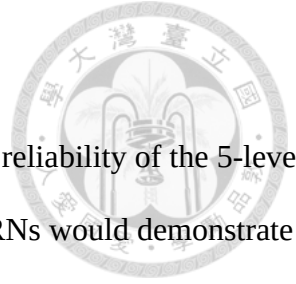
2.9 Research Gap

The limited adoption of the 5-level triage system in Taiwan, mainly in Taoyuan and Taichung, highlights challenges in prehospital emergency care. While effective internationally, Taiwan's reliance on the simpler 2-level system, though time-efficient, often fails to accurately assess patient urgency, leading to over-triage, under-triage, and inefficient resource allocation in EMS.

The 5-level system provides a more detailed, standardized approach aligned with global protocols and could help address rising ambulance demand and misuse of EMS by better managing non-urgent cases. However, implementation faces challenges such as complexity, limited prehospital validation, and greater EMT training requirements.

A key issue is the misalignment between prehospital 5-level triage and hospital-based protocols, which disrupts care continuity and EMS efficiency. Taiwan's diverse geography and infrastructure further demand a flexible, regionally adaptable triage system. Aligning prehospital and hospital protocols could improve communication, reduce delays, and minimize resource misuse.

Research is needed to validate the 5-level system in Taiwan's EMS, focusing on its effects on treatment times, patient outcomes, and workload reduction. Identifying adoption barriers and developing supportive policies, including public education, are crucial. Nationwide implementation can create a more efficient, equitable emergency care framework, improving EMT capabilities and patient safety.



2.10 Research Aims

The primary objective of our initial study was to assess the inter-rater reliability of the 5-level TPTS between EMTs and TRNs. We hypothesized that both EMTs and TRNs would demonstrate a high level of inter-rater reliability, indicating consistent and accurate triage assessments across different professional groups within emergency care.

In our second study, we aimed to evaluate the ability of the 5-level TPTS scores to discriminate patient outcomes following transport to ED by EMTs. These outcomes included discharge, hospital admission, and mortality, as well as their associated ED LOS and medical resource utilization, as measured by the ESI resources. The objective was to investigate which 5-level TPTS scores impacted ED LOS based on different endpoint outcomes.

Chapter 3. Materials and Methods



3.1 Methods of Study 1

3.1.1 Study Design and Methodology

This cross-sectional study was conducted from July 1, 2023, to September 31, 2023, in Taoyuan, a metropolitan city in Taiwan. Informed consent for this study was obtained from the institutional review board of Linkou Chang Gung Memorial Hospital (LCGMH) (Reference: 202202276B0). Taoyuan covers an area of about 1,220 square kilometers and has a population of approximately 2.25 million people. The city's emergency medical services are supported by 948 EMTs, including 112 EMT-P, the highest certification level for EMTs in Taiwan.

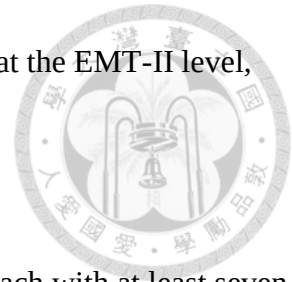
3.1.2 Participants and Setting

This study, conducted from July to December 2023, involved a randomized selection of 20 professionals: 10 EMTs from the TYFD and 10 TRNs from the ED of LCGMH. The goal was to assess the inter-rater reliability of the 5-level TPTS among different emergency care providers in Taoyuan, Taiwan.

LCGMH, a major teaching hospital, tertiary care center, and level I trauma center, has 3,600 beds and handles about 180,000 emergency department visits each year. The hospital employs 102 RNs in its emergency department.

Among the EMTs, five were certified EMT-P, the highest level of EMT training in Taiwan. To become EMT-P certified, candidates must complete EMT-II training, have at least four years of experience or a college degree, and undergo an additional 1,280 hours of training. EMT-Ps are

qualified to perform ALS procedures. The other five EMTs were certified at the EMT-II level, which requires an EMT-I license and an additional 280 hours of training.



The TRNs were divided equally into two groups: five Senior TRNs, each with at least seven years of clinical experience, and five Junior TRNs. This setup allowed for a balanced comparison of experience levels among both EMTs and TRNs.

3.1.3 The Two-Level and Five-Level TPTS

The two-level TPTS categorizes patients as either emergent or non-emergent based on specific criteria that reflect critical health conditions or severe physiological distress[18]. This system uses vital signs, including the Glasgow Coma Scale (GCS), heart rate (HR), respiratory rate (RR), blood pressure (BP), body temperature, oxygen saturation, and blood glucose levels. Chief complaints such as stroke, cardiac issues, seizures, or life-threatening injuries also help determine if a patient is classified as emergent. Patients not meeting these criteria are considered non-emergent, allowing for a structured way to prioritize care in the prehospital setting.

The 5-level TPTS was introduced in August 2022 by Taoyuan medical directors and senior paramedics. Inspired by the TTAS, the 5-level TPTS organizes patient urgency into five categories: Level 1 (Resuscitation), Level 2 (Emergency), Level 3 (Urgent), Level 4 (Less Urgent), and Level 5 (Non-Urgent). This system considers primary variables such as vital signs, pain levels. Special adjustments are made for specific populations, including patients with coagulopathies or stable vital signs following high-energy trauma. The 5-level TPTS also evaluates secondary variables to

address 13 critical non-trauma conditions and 12 critical trauma conditions, ensuring comprehensive patient assessment during prehospital care.



Table S2 outlines the key differences between the two-level and five-level TPTS systems, highlighting their criteria and usage in prehospital care.

3.1.4 Data Collection and Participant Procedure

In this study, patient selection was performed through a randomized process using de-identified electronic health records, with patients' personal information securely stored on Google Cloud Drive. Stratified sampling was utilized to ensure a balanced representation among participants. From the 40 fire corps distributed under 5 fire brigades in Taoyuan City, each brigade randomly selected one EMT-2 and one EMT-P professional, resulting in 2 EMTs per brigade and 10 EMTs enrolled in the study. Additionally, 10 TTAS-certified TRNs participated in the study. These nurses were categorized into two groups based on their clinical experience: the Senior TRN group, comprising 5 nurses with at least 7 years of clinical experience, including a minimum of 3 years of frequent ED triage experience; and the Junior TRN group, consisting of 5 nurses with a minimum of 3 years of frequent ED triage experience. They also received electronic consent forms and personal information forms (collecting data on gender, age, seniority, and job grade) via Google Mail. Participants were required to submit signed consent forms within 14 days to confirm their involvement, resulting in a total of 20 participants for the study.

EMTs from TYFD and TRNs from LCGMH took part in a comprehensive 4-hour training session on the five-level TPTS. The training covered essential aspects of the five-level system,

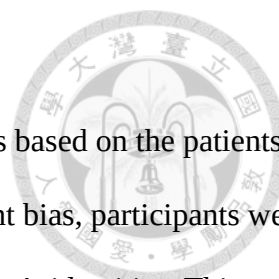
including practical exercises based on real emergency cases. These exercises demonstrated how to implement the five-level triage process, apply triage modifiers, and use a computer-assisted decision-making system.



Following the instruction, participants were tasked with applying their newly acquired skills by assigning 5-level TPTS scores to 20 written scenarios, simulating a range of emergency situations. This hands-on practice helped EMTs and TRNs gain experience in applying the triage methods and modifiers, while also building proficiency in using the computer-assisted triage system for accurate data recording and transmission in real-world emergency situations.

We used a stratified random sampling method to account for factors such as time to hospital arrival and treatment, which can affect outcomes. A representative sample of 50 patients was selected from the Taoyuan City Government Fire Department's patient records, based on a 1/1000 ratio, covering the period from January 1, 2023, to June 30, 2023. The sample was categorized using triage criteria and stratified to account for potential confounding variables, such as transport time to hospital arrival and whether patients received treatment. This process, relying on electronic health records from the Taoyuan Fire Department (TYFD), ensured a carefully selected, unbiased sample for the study.

For each of the 50 randomly selected patients, the participants used a Google Forms online survey to evaluate the five-level TPTS. These evaluations took place during Google Meet sessions, where each session was video recorded and lasted about 40 to 60 minutes. Despite being in different locations, participants were required to keep their video cameras on during the session to verify their identity and participation.



During the study, participants were tasked with assigning triage scores based on the patients' chief complaints and pre-hospital vital signs. To ensure fairness and prevent bias, participants were not allowed to communicate with each other and were unaware of each other's identities. This approach helped maintain the study's integrity and objectivity.

3.1.5 Outcomes

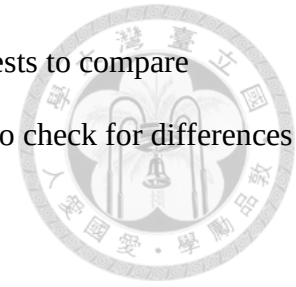
The main aim of this study was to evaluate how consistently different professional groups applied the five-level TPTS when assessing a sample of 50 randomly selected patients. To measure inter-rater reliability, we used the intraclass correlation coefficient (ICC). An ICC value below 0.5 indicates poor reliability, values between 0.5 and 0.75 indicate moderate reliability, between 0.75 and 0.9 indicate good reliability, and values above 0.90 indicate excellent reliability. For this study, we aimed to achieve an ICC range between 0.50 and 0.75.

Additionally, we performed multilevel analysis to examine how different professional backgrounds and patient characteristics impacted the 5-level TPTS evaluations. This approach helped us determine whether the evaluators' professions or the patients' specific traits affected the consistency and accuracy of the five-level TPTS scoring.

3.1.6 Statistical Analysis

To assess how consistently different groups of evaluators applied the five-level TPTS and how patient characteristics were distributed, we analyzed both categorical and continuous variables. Categorical data were presented as absolute numbers and percentages, while continuous data were

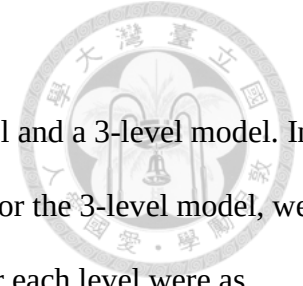
shown as medians with interquartile ranges (IQRs). We used Chi-square tests to compare categorical variables and Mann-Whitney U tests for continuous variables to check for differences across the evaluator and patient groups.



For reliability testing of the triage assessments made by the 20 participants, we calculated the ICC using a two-way random-effects model. This model was chosen because all 20 raters evaluated the same 50 randomly selected patients. The ICC values, along with their 95% confidence intervals (CIs), were calculated using the 'irr' package in R software (version 4.2.2). ICC values were interpreted as follows: values below 0.5 indicated poor reliability, values between 0.5 and 0.75 indicated moderate reliability, between 0.75 and 0.9 indicated good reliability, and values above 0.90 indicated excellent reliability. Our goal was to achieve an ICC between 0.50 and 0.75 for this study.

Before conducting the multilevel analysis, we calculated the average five-level TPTS scores for the 50 patients, as rated by four groups: EMT-II, EMT-P, Junior TRNs, and Senior TRNs. We generated pairwise scatterplots to visualize the relationships between these groups and identify any potential correlations. This was a crucial step in understanding the scoring patterns across the groups before proceeding with a more in-depth multilevel analysis.

We also plotted the scatterplots of the total 5-level TPTS scores for all 20 raters evaluating the 50 patients to examine the distribution patterns of the scores across the groups. Correlation coefficients were calculated for the five-level TPTS scores to assess the relationships. Additionally, we used Bland-Altman plots to check for any systematic bias in how different groups evaluated the 5-level TPTS scores.



For the multilevel analysis, we examined two models: a 2-level model and a 3-level model. In the 2-level model, we considered Level 1 (raters) and Level 2 (patients). For the 3-level model, we added Level 2 (departments) between raters and patients. The variables for each level were as follows:

Level 1 (raters): rater age, gender, seniority, and job grade.

Level 2 (departments): no covariates.

Level 3 (patients): reason for calling 911, vital signs, patient age, gender, treatment, and transport time.

First, we compared a null model (also known as the Variance Component model), the simplest hierarchical model with no explanatory variables, to establish the mean outcome and determine how variance is distributed across the levels. This initial step helped partition the total variance (the square of the standard deviation) between the raters, departments, and patients.

After evaluating the null model, we further developed two of the three models by adding covariates to form random intercept models. These models assume that each higher-level unit (department or patient) shares the same relationship between the outcome and covariates but shows random variation around a collective mean intercept.

This approach allowed us to estimate fixed effects, such as the intercept and effect size, while also calculating how much of the variance is explained by random effects. To assess the fit of each model, we used metrics such as the -2 log likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). We reported all estimates with their 95% confidence

intervals, and a p-value of less than 0.05 was considered statistically significant. All data analysis was performed using R software version 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria) and SAS statistical software version 9.4 (SAS Institute Inc., Cary, NC, USA).



3.2 Methods of Study 2

3.2.1 Study Design

This prospective cohort study was carried out over a one-year period, from January 1, 2023, to December 31, 2023, in Taoyuan City, Taiwan. The research protocol, including all variables of interest and statistical methods, was determined before the study began. Approval for this study was granted by the Institutional Review Board of LCGMH, with the reference number 202301667B0.

3.2.2 Participants and Setting

This study was conducted in Taoyuan, a metropolitan area in northern Taiwan that spans approximately 1,220 square kilometers with a population of around 2.25 million people. The region's EMS are provided by 948 EMTs, including 112 EMT-P, the highest level of EMT certification in Taiwan.

Patients included in the study were those transported by EMS to the ED at LCGMH between January 1, 2023, and December 31, 2023. A total of 125,279 patients utilized the EMS system during this period, and 10,713 of these patients, who were transported to LCGMH, were enrolled in the study.

LCGMH is a large tertiary care facility, academic teaching hospital, and level I trauma center, with a capacity of 3,600 beds. The ED at LCGMH handles approximately 180,000 visits each year and is equipped with one computerized tomography (CT) scan room and two X-ray rooms located within the ED. The hospital provides comprehensive 24-hour surgical and medical subspecialty services. LCGMH serves as a referral center for patients from northern Taiwan who present with a wide variety of emergency complaints.

All EMTs from the TYFD who participated in this study underwent a specialized training program to familiarize themselves with the five-level TPTS. This training included a four-hour session covering both theoretical aspects and practical exercises. EMTs were trained on how to apply the five-level triage process, use triage modifiers, and operate the computer-assisted decision-making system.

As part of their training, EMTs were required to assign 5-level TPTS scores to 20 different written emergency scenarios, simulating a variety of clinical situations. This hands-on practice helped EMTs gain confidence in using the triage system and improved their ability to record and transmit data accurately in real-life emergencies.

3.2.3 Five-Level Taiwan Prehospital Triage System (TPTS)

The five-level TPTS is designed to categorize patients based on the urgency of their condition, helping medical teams prioritize treatment effectively. This system uses a combination of primary and secondary clinical variables to guide triage decisions.

The five urgency levels are as follows:

Level 1 (Resuscitation): Immediate intervention required to prevent death.

Level 2 (Emergency): Potential life-threatening conditions needing urgent treatment.

Level 3 (Urgent): Conditions that require prompt attention but are not immediately life-threatening.

Level 4 (Less Urgent): Non-critical issues that require medical attention but can safely wait.

Level 5 (Non-Urgent): Minor issues that do not need immediate care.

Primary Variables

The 5-level TPTS primarily relies on vital signs to assess patient urgency. These include:

Glasgow Coma Scale (GCS): Evaluates consciousness levels.

Alert, Voice, Pain, Unresponsive (AVPU): Assesses consciousness more quickly than GCS.

Respiratory Rate (RR): Monitors breathing frequency.

Heart Rate (HR): Tracks the heart's speed and rhythm.

Systolic Blood Pressure (SBP): Measures the pressure in arteries during heart contractions.

Body Temperature(BT): Assesses for hypothermia or hyperthermia.

Oxygen Saturation (O2): Monitors blood oxygen levels.

Pain Index: A 10-point scale to rate the intensity of a patient's pain.

Capillary Refill Time (CRT): Assesses peripheral circulation.

EMT Interventions: Tracks any treatments already administered by the EMTs.

Secondary Variables

In addition to the primary variables, the 5-level TPTS also considers specific symptoms and conditions that may require immediate attention:

Out-of-Hospital Cardiac Arrest (OHCA): Sudden cardiac arrest outside the hospital.

Stroke Symptoms: Particularly those occurring within a 4.5-hour window for potential thrombolysis.

Ischemic Heart Disease: Symptoms of coronary artery disease, such as chest pain.

Seizures: Active or recent seizure activity.

Emergent Labor: Imminent childbirth during transport.

Behavioral/Psychiatric Emergencies: Conditions like acute agitation or confusion.

Hypo/Hyperglycemia: Extreme blood sugar levels needing immediate correction.

For trauma-related cases, additional criteria such as major burns or injuries from high-risk mechanisms (e.g., falls from great heights or vehicle collisions) are also considered, ensuring that even patients with stable vital signs but significant injury mechanisms are appropriately triaged.

Process of Triage

Upon EMS arrival, EMTs use these variables to quickly assess the patient's condition. They start by evaluating vital signs and consciousness levels using the GCS or AVPU scale. EMTs also consider any symptoms from the secondary list, such as stroke or trauma, to refine the urgency level.

Once the initial assessment is complete, EMTs use triage modifiers like the pain index or CRT to adjust the patient's urgency level if necessary. If patients exhibit special conditions, such as high-risk trauma mechanisms, EMTs modify the triage decision to ensure proper categorization. For example, an elderly patient with stable vital signs but suffering from a significant fall may be escalated in priority due to the high-energy trauma mechanism.

Finally, EMTs assign a 5-level TPTS score and record the data using a computer-assisted decision-making system, ensuring consistent and accurate transmission to the receiving hospital.

3.2.4 Data Collection

In August 2022, Taoyuan medical directors and senior paramedics introduced the five-level TPTS. The implementation of this triage system in prehospital settings began in January 2023. EMTs assigned the 5-level TPTS score during on-scene triage using a tablet. For emergency scenarios where immediate transport was necessary, the triage score was also confirmed upon arrival at the ED. The triage algorithm used by the EMTs is illustrated in Figure S3. Briefly, this system allows EMTs to select from predefined options on a tablet screen, automatically determining the triage level based on the patient's condition. EMTs can modify the score to account for real-time assessments and specific considerations regarding the patient's condition.

Patient data was stored in the TYFD EMS cloud database. This data included details such as the date, time of dispatch, day of the week, patient demographics, and the reason for calling 119 (categorized as Trauma, Adult Non-Trauma, and Pediatric Non-Trauma [≤ 18 years]). Both the two-level and five-level TPTS scores were recorded.

Additional data was collected from the Linkou Chang Gung Memorial Hospital (LCGMH) administrative database. This included the recommended patient disposition, such as:

- Admission to the intensive care unit (ICU),
- Admission to a general ward,
- Death in the ED,
- Transfer to another hospital,

Discharge from the ED,
Discharge against medical advice, or
Leaving the ED without being seen by a physician.



Patient LOS in the ED was defined as the time from registration to departure from the ED. We also collected the five-level TTAS score assigned by RNs and the medical resource consumption during the first six hours of emergency care. To assess resource usage, we utilized the ESI resource categories, with each physician order counted as one resource unit. A six-hour window was chosen for analysis to prevent overestimating resource usage associated with extended ED stays.

The ESI resource categories included labs (blood, urine), electrocardiograms, radiographs, advanced imaging such as computed tomography, magnetic resonance imaging, ultrasound, and angiography, intravenous fluids for hydration, administration of intravenous, intramuscular, or nebulized medications, specialty consultations, and procedures. Simple procedures, such as laceration repair or urinary catheter placement, were assigned a value of one resource unit, while complex procedures, including procedural sedation, were assigned two resource units.

We merged the TYFD EMS database and LCGMH hospital data by matching key identifiers, such as patient age, sex, de-identified number, and EMS dispatch time. This allowed us to integrate prehospital and hospital data for a comprehensive analysis of patient outcomes.

3.2.5 Outcomes

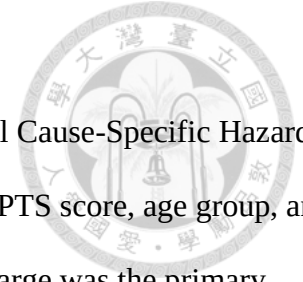
The primary objective of this study was to evaluate the discrimination ability of the five-level TPTS in discriminating patient outcomes following transportation by TYFD EMTs to the ED at LCGMH. The outcomes of interest included hospital discharge, admission, and mortality. Additionally, we analyzed the LOS in the ED and the utilization of medical resources, measured by the ESI resources.

To achieve this, we examined how each 5-level TPTS score category, patient gender, patient age group, and reason for calling 119 influenced ED LOS and the corresponding outcomes of hospital discharge, admission, or death. We also assessed medical resource consumption for each score level to understand the relationship between triage scores and the intensity of care needed.

3.2.6 Statistical Analysis

In this study, we used a variety of statistical methods to analyze patient outcomes, focusing on ED LOS, discharge status (discharged, admitted, or death), and ESI resource. Categorical variables like gender and triage levels were compared between groups using chi-square tests, while continuous variables like age and ED LOS, which did not follow a normal distribution, were analyzed using the Kruskal-Wallis test. For these continuous variables, we reported medians and interquartile ranges (IQR).

We also used the Cumulative Incidence Function (CIF) to visualize the probability of different outcomes (discharge, admission, or death) over time, stratified by 5-level TPTS score category, patient gender, patient age group, and the reason for calling 119. This helped us understand how ED LOS varied between different groups.



Since these outcomes are competing risks , we employed Proportional Cause-Specific Hazards Models. These models allowed us to estimate how variables like 5-level TPTS score, age group, and reason for calling 119 affected the likelihood of each outcome. Since discharge was the primary event of interest, admissions and deaths were treated as competing events and censored in the analysis, along with other usual censored observations like incomplete data. This method helped us focus on the risk of the primary event of interest while accounting for the influence of other potential outcomes.

To explore the distribution of ESI scores, stratified analyses were performed using key demographic and clinical variables. The ESI scores were categorized based on stratification by the 5-level TPTS levels, TTAS levels, reason for calling 119 (trauma, pediatric non-trauma, adult non-trauma), gender, and age groups. Bar plots were generated to visualize the distribution across these strata, enabling a comparative assessment of ESI score variations among different patient subgroups.

We reported all estimates with their 95% confidence intervals, and a p-value of less than 0.05 was considered statistically significant. All data analysis was performed using R software version 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria) and SAS statistical software version 9.4 (SAS Institute Inc., Cary, NC, USA).

Chapter 4. Results



4.1 Results of Study 1

4.1.1 Patient Flow and Enrollment

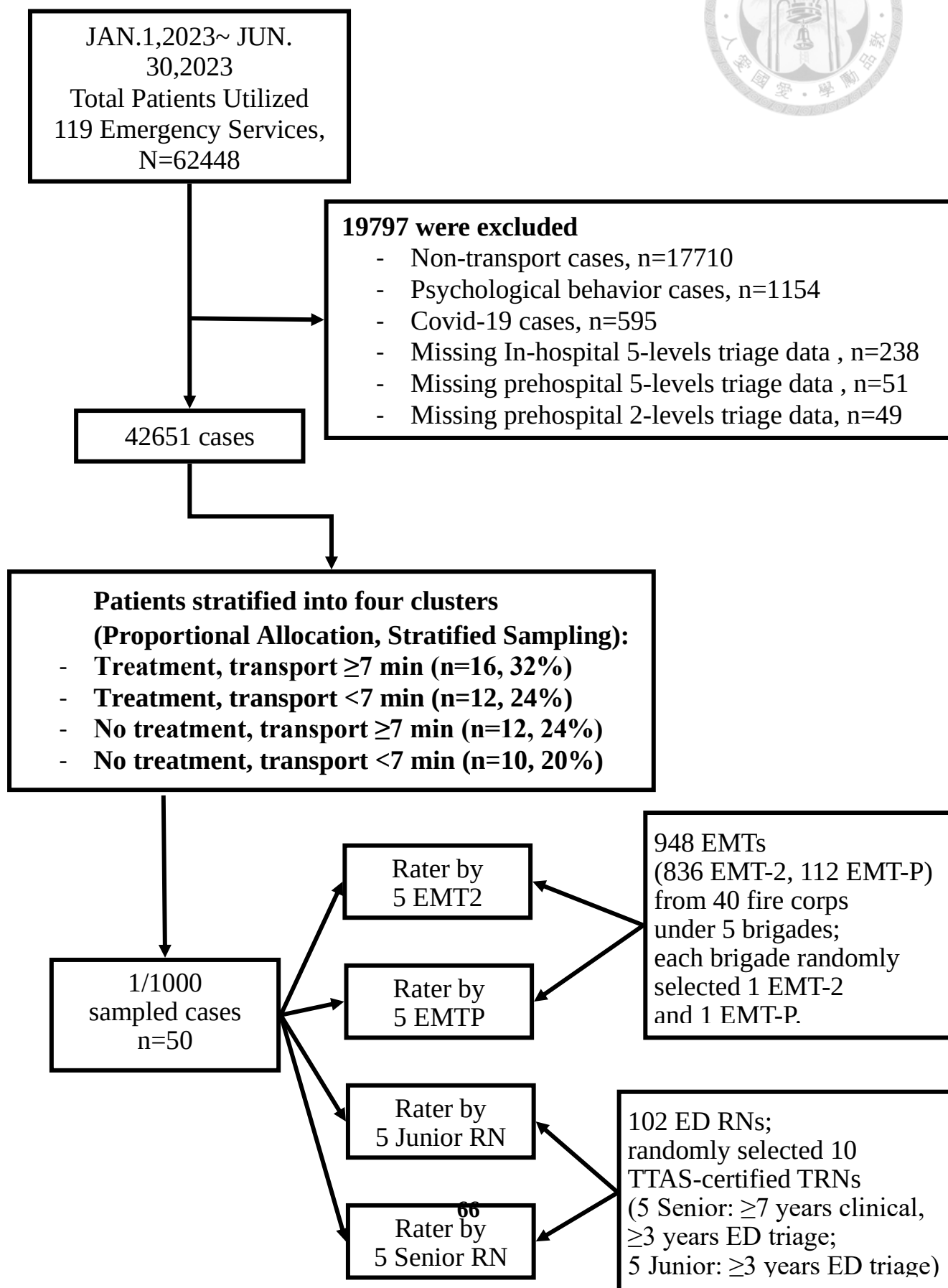
From January 1, 2023, to June 30, 2023, a total of 62,448 patients utilized 119 emergency services (Figure 2). Some patients were excluded, including non-transport cases ($n=17,710$), psychological or behavioral cases ($n=1,154$), COVID-19 cases ($n=595$), and cases missing triage data ($n=338$). This left a total of 42,651 eligible cases.

To account for potential confounding factors such as transport time and treatment received (like oxygen therapy or IV fluids), patients were stratified using 7 minutes as the median transport time. A stratified random sampling method was used to select 1/1000 of the patients from the TYFD's database, resulting in 50 patients being chosen for the five-level TPTS evaluation.

The evaluation was conducted by a group of 20 trained evaluators, which included 5 EMT-IIs, 5 EMT-Ps, 5 Senior TRNs, and 5 Junior TRNs. The triage assessments were recorded during a Google Meet session to ensure transparency and accuracy in the process. The sample was divided into four groups based on transport time and treatment received, ensuring proportional representation in the analysis.



Figure 2 Flow chart of the studied patients



4.1.2 Characteristics of the study population

The characteristics of the evaluators are listed in Table 2. The study involved 20 evaluators, including EMTs and TRNs, divided into four groups: 5 EMT-IIs, 5 EMT-Ps, 5 Junior TRNs, and 5 Senior TRNs. The median age of the evaluators was 30.5 years (IQR: 25.5-33.5 years). There was a significant difference in gender distribution ($p < 0.05$), with a higher proportion of males in the EMT-II and EMT-P groups. The evaluators' seniority also varied, with Senior TRNs having the highest median seniority of 12 years (IQR: 11-13 years).

Table 2 Raters demographic data

Job Grade	Overall N=20	EMT2 N=5	EMTP N=5	Junior RN N=5	Senior RN N=5	p-value
Age(IQR)	30.5 (25.5-33.5)	30 (29-33)	31 (28-33)	25 (24-25)	34 (32-35)	<0.05*
Gender						<0.05*
Male(%)	12(60)	5(100)	4(80)	2(40)	1(20)	
Female(%)	8(40)	0(0)	1(20)	3(60)	4(80)	
Seniority(IQR)	7(4-8.5)	7(3-7)	7(5-7)	4(4-5)	12(11-13)	<0.05*

Abbreviations: EMT= Emergency Medical Technician, RN= Registered Nurse.* $p < 0.05$

The demographics of the patients selected through random sampling are shown in Table 3. A total of 50 patients were included in the study, with a median age of 61.5 years (IQR: 44-72 years). Of these, 58% were male and 42% were female. The reasons for calling 911 were divided into trauma (40%) and non-trauma (60%) cases. The majority of patients had stable vital signs, with 78% classified as stable and 22% as unstable. The median transport time to the hospital was 7 minutes (IQR: 4-11 minutes), with 54% of transports taking 7 minutes or more.

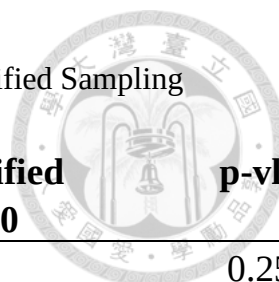


Table 3 Patient Demographics Data Using Proportional Allocation in Stratified Sampling

Patients	Overall N=42651	Proportional Stratified Sampling N=50	p-value
Age(IQR)	56(34-73)	61.5(44-72)	0.2539
Age-n(%)			0.8809
<40	12986(30.45)	11(20)	
40-49	4952(11.61)	7(14)	
50-59	5182(12.15)	5(10)	
60-69	6440(15.10)	8(16)	
70-79	5913(13.86)	10(20)	
80-89	4766(11.17)	6(12)	
>89	2401(5.63)	3(6)	
unknown	11(0.03)	—	
Gender-n(%)			0.6288
Male	23285(54.59)	29(58)	
Female	19366(45.41)	21(42)	
Reason for calling 911-n(%)			0.0618
Trauma	19965(46.81)	20(40)	
Non-Trauma	22686(53.19)	30(60)	
vital sign-n(%)			0.9318
stable	31576(74.03)	39(78)	
unstable	11075(25.97)	11(22)	
Treatment-n(%)			0.8400
Yes	24487(57.41)	28(56)	
No	18164(42.59)	22(44)	
Transport time(IQR)	7(4-11)	7(4-11)	0.4648
Transport time-n(%)			0.9935
≥7 minutes	23860(55.94)	27(54)	
<7 minutes	18791(44.06)	23(46)	

4.1.3 Main results

As shown in Table 4, the inter-rater reliability of the five-level TPTS assessments was analyzed using two models: absolute agreement for single raters and consistency across multiple raters.

Model a: For absolute agreement (single rater), the overall ICC was 0.57 (95% CI: 0.47-0.68), indicating moderate reliability. In particular, the EMT-II group had an ICC of 0.60 (95% CI: 0.48-0.72), and the EMT-P group had an ICC of 0.62 (95% CI: 0.49-0.74). For the nurses, the Junior TRNs showed an ICC of 0.71 (95% CI: 0.60-0.80), while the Senior TRNs had an ICC of 0.69 (95% CI: 0.58-0.79), indicating moderate to good reliability.

Model b: For consistency across multiple raters, the overall ICC was 0.96 (95% CI: 0.95-0.98), showing excellent reliability. The EMT-II group had an ICC of 0.88 (95% CI: 0.82-0.93), while the EMT-P group demonstrated even higher consistency with an ICC of 0.96 (95% CI: 0.95-0.98). Both the Junior TRNs and Senior TRNs had ICCs of 0.92 (95% CI: 0.88-0.95 and 0.87-0.95, respectively), showing high agreement among raters.

The average 5-level TPTS scores assigned by the evaluators were 3.04 for EMT-II, 3.10 for EMT-P, 2.76 for Junior TRNs, and 2.73 for Senior TRNs, with an overall mean score of 2.90 and a standard deviation of 0.74, indicating strong consensus in scoring across all groups.

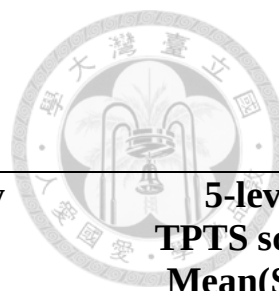


Table 4 Inter-rater reliability and the 5-level TPTS score

Job Grade	Inter-rater reliability ICC(95%CI) ^a	Inter-rater reliability ICC(95%CI) ^b	5-level TPTS score Mean(SE)
EMT-2	0.60(0.48-0.72)	0.88(0.82-0.93)	3.04(0.79)
EMT-P	0.62(0.49-0.74)	0.96(0.95-0.98)	3.10(0.76)
JRN	0.71(0.60-0.80)	0.92(0.88-0.95)	2.76(0.73)
SRN	0.69(0.58-0.79)	0.92(0.87-0.95)	2.73(0.59)
Overall	0.57(0.47-0.68)	0.96(0.95-0.98)	2.90(0.74)

To measure the reliability in the triage assessment among the 20 raters, we calculated ICC with its

95% confidence interval (CI) using a two-way random-effect model

to take into account that the 50 scenarios were rated by the same set of 20 independent raters.

^a Two-way random effects, absolute agreement, single rater/measurement(ICC (2,1)).

$$ICC(2,1) = (MSR - MSE) / [MSR + (k - 1) * MSE + (k / n) * (MSC - MSE)]$$

^b Two-way random effects, consistency, multiple raters/measurements(ICC (2,k)).

$$ICC(2,k) = (MSR - MSE) / [MSR + MSC - MSE / n]$$

Definitions:

MSR: Mean Square for Rows, representing the variance among subjects.

MSE: Mean Square for Error, representing random error variance.

MSC: Mean Square for Columns, representing the variance among raters.

n: Number of subjects.

k: Number of raters or measurements.

Values less than 0.5 are indicative of poor reliability.

Values between 0.5 and 0.75 indicate moderate reliability.

Values between 0.75 and 0.9 indicate good reliability.

Values greater than 0.90 indicate excellent reliability.

Figure 3 Distribution of 5-level TPTS Scores between EMT vs. RN: Scatterplot

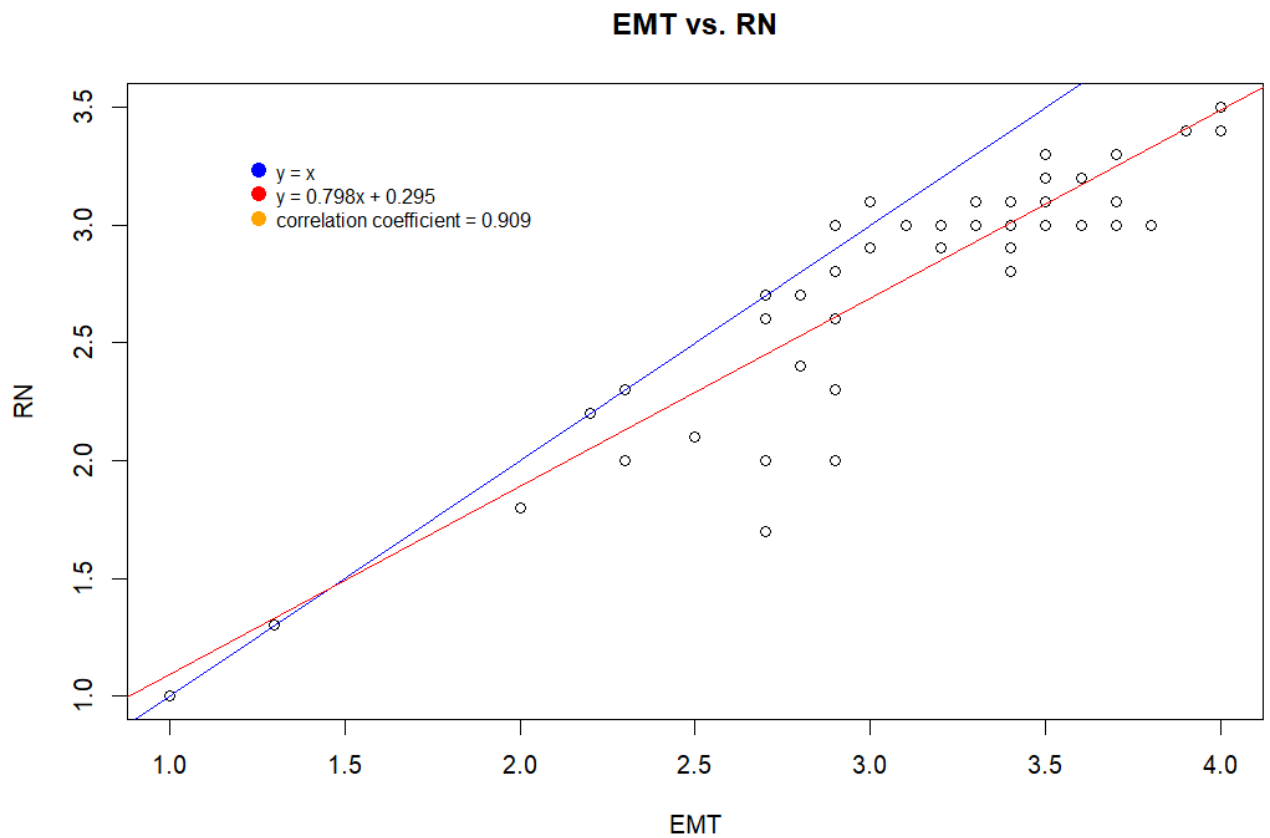


Figure 3 illustrates the strong positive correlation ($r = 0.909$) between 5-level TPTS scores assigned by EMTs and RNs. The blue line indicates a perfect correlation (mean coefficient = 1), while the red regression line shows that each unit increase in EMT scores corresponds to a 0.798 increase in RN scores. EMTs generally assign higher scores, whereas RNs give lower scores, reflecting systematic biases between the two groups in triage scoring.



Figure 4 Distribution of 5-level TPTS Scores Across Different Groups: Scatterplot

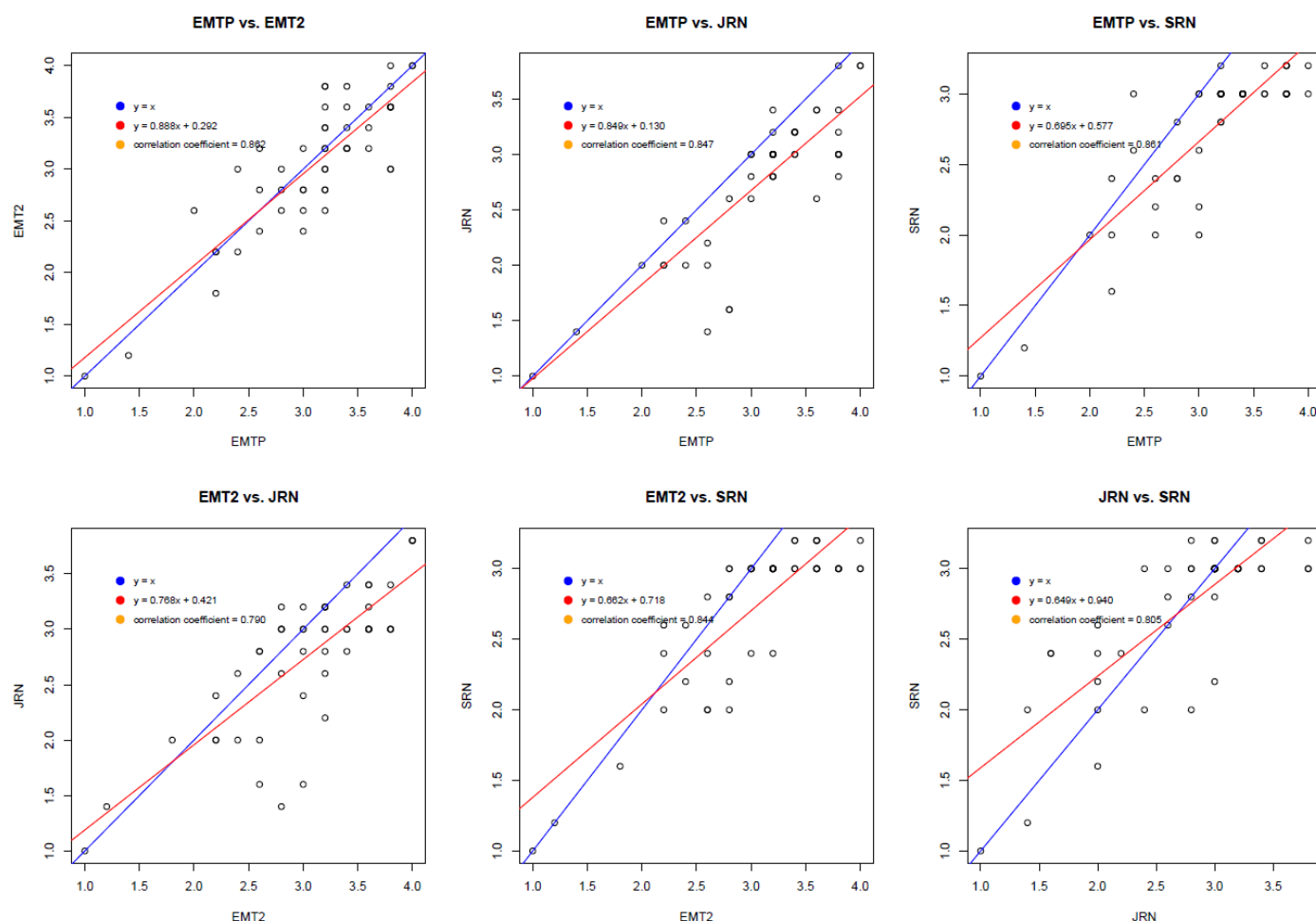


Figure 4 highlights a strong positive correlation between 5-level TPTS scores across different groups. The blue line represents a perfect correlation (mean coefficient = 1), while the red regression line shows that an increase of one unit in the x-axis scores consistently corresponds to a one-unit increase in the y-axis scores. All groups show a high degree of linear correlation based on their coefficients. However, there is noticeable variability in scores between 3 and 4 points within each group, which may reflect differences in individual case or variations in scoring by different raters.



Table 5 Correlation coefficients for the 5-level TPTS score measures

	EMT-2	EMT-P	JRN	SRN
EMT-2	1.00	0.86***	0.79***	0.84***
EMT-P	—	1.00	0.84***	0.86***
JRN	—	—	1.00	0.80***
SRN	—	—	—	1.00

***p<0.001 Correlations using average score for each groups

Table 5 demonstrates strong positive correlations for 5-level TPTS scores across different groups, all statistically significant ($p < 0.001$). EMT-2 and EMT-P show a high correlation ($r = 0.86$), as do EMT-P and both JRN ($r = 0.84$) and SRN ($r = 0.86$). JRN also correlate strongly with SRNs ($r = 0.80$) and EMT-2 ($r = 0.79$). These results highlight the consistency of 5-level TPTS scoring among diverse groups.



Figure 5 Distribution of 5-level TPTS Scores between EMT vs. RN: Bland-Altman Plot

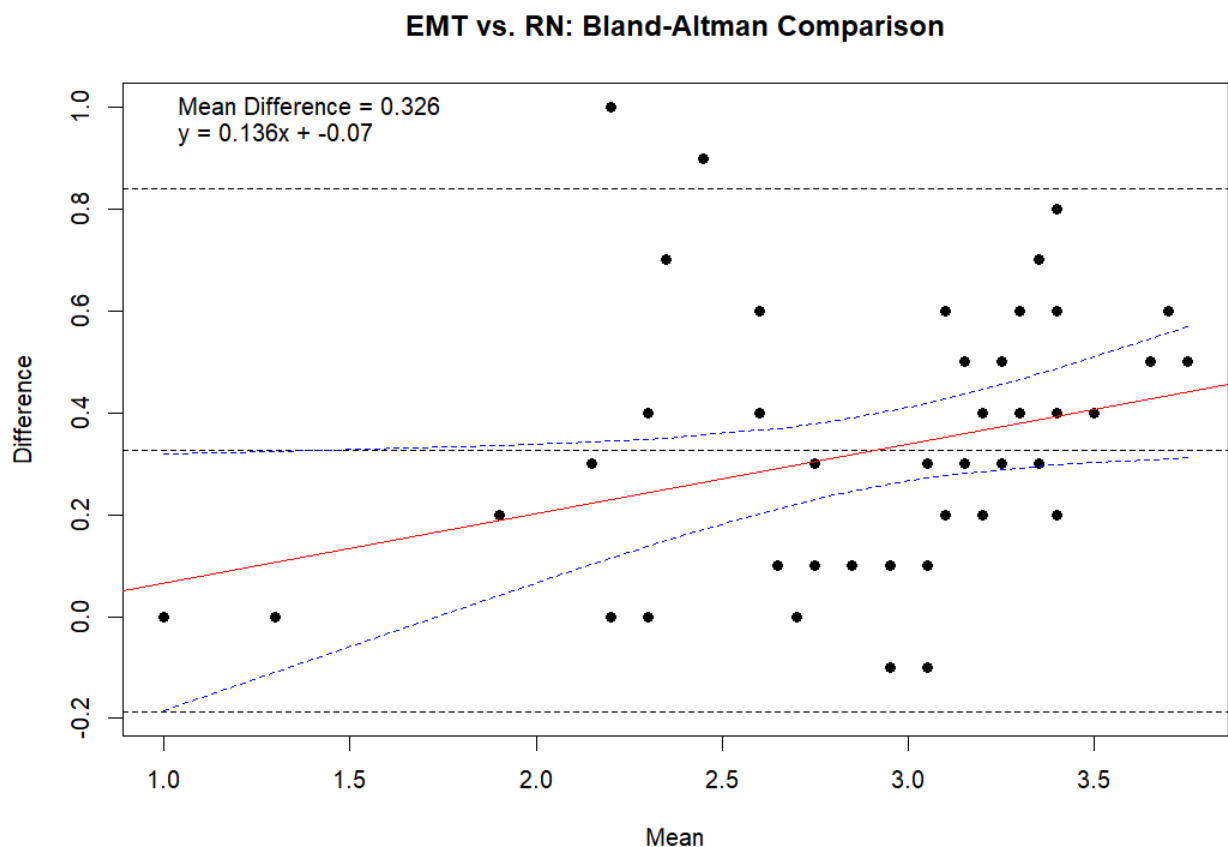


Figure 5 illustrates a Bland-Altman plot comparing 5-level TPTS scores assigned by EMTs and RNs. The mean difference between the two groups is 0.326, indicating that EMTs generally score slightly higher than RNs on average. This mean difference is depicted as a horizontal line at 0.326, reflecting the overall agreement with a small systematic bias.

The regression equation ($Y=0.136X-0.07$) describes the relationship between the differences and the mean scores. The positive slope (0.136) suggests that the difference between EMT and RN scores increases slightly as the mean score rises. The intercept of -0.07 indicates the predicted difference when the mean score is zero. Continuous lines representing the confidence intervals show the variability around the regression line.

These findings suggest a modest bias between EMT and RN assessments, with differences slightly amplifying at higher mean scores. While overall agreement is strong, further analysis may help understand the factors contributing to this trend and guide improvements in scoring alignment.

Figure 6 Distribution of 5-level TPTS Scores Across Different Groups: Bland-Altman Plot

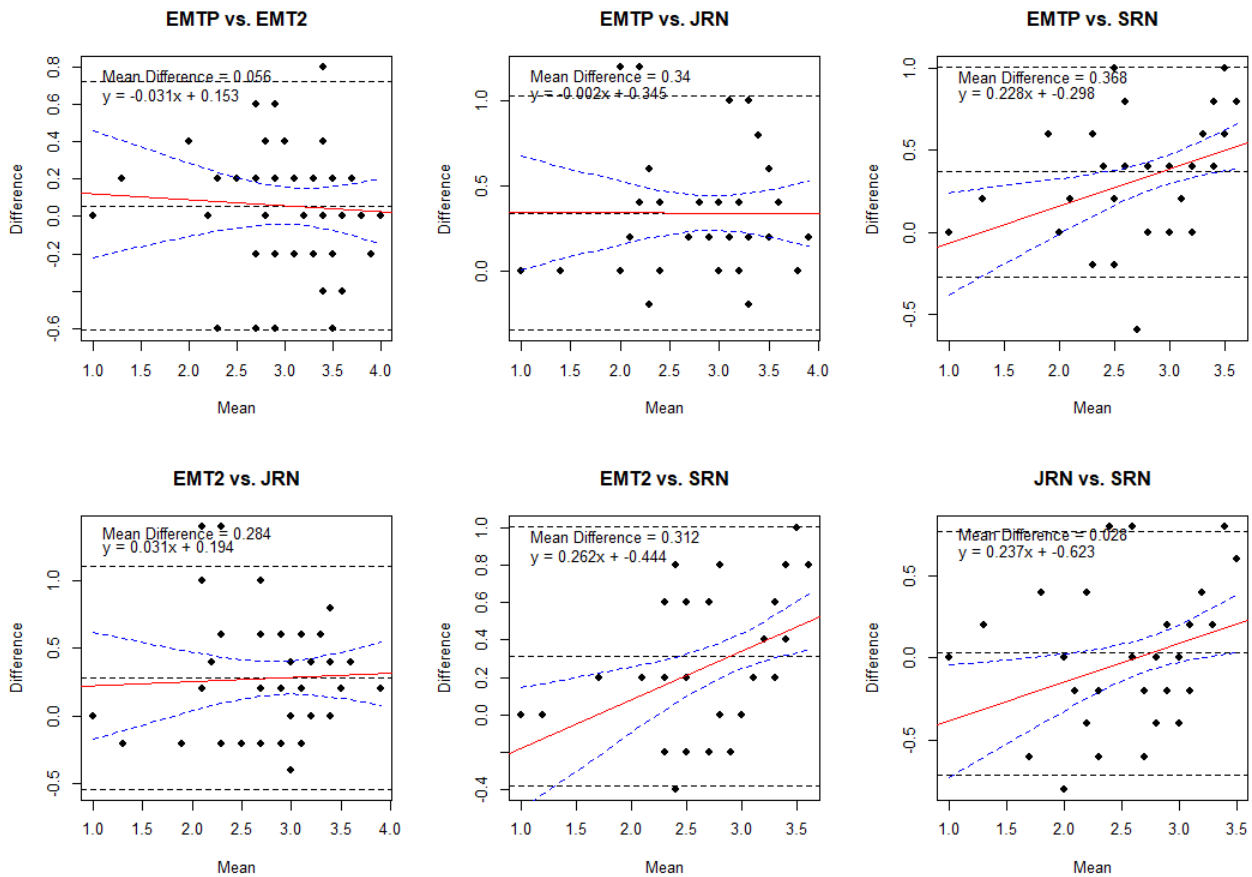


Figure 6 illustrates the agreement in 5-level TPTS scores among EMT-P, EMT-2, JRN, and SRN groups, focusing on mean differences and systematic biases. The mean differences range from 0.028 (JRN vs. SRN), reflecting near-equal scoring, to 0.368 (EMT-P vs. SRN), indicating a moderate positive bias with EMT-P scoring higher. EMT-2 generally shows slight positive biases compared to both JRN (mean difference = 0.284) and SRN (mean difference = 0.312). Regression equations further explore systematic biases, with slopes ranging from -0.092 (EMT-P vs. JRN, $Y = -0.092X - 0.345$) to 0.262 (EMT-2 vs. SRN, $Y = 0.262X - 0.444$). Negative slopes, such as in

EMT-P vs. JRN, suggest decreasing differences as mean scores rise, while positive slopes, such as in EMT-2 vs. SRN, indicate increasing discrepancies at higher mean scores. EMT-P vs. SRN ($Y=0.228X-0.298$) and EMT-2 vs. SRN ($Y=0.262X-0.444$) exhibit notable increases in bias at higher mean scores. These findings reveal strong overall agreement between groups but also identify varying levels of bias, particularly between EMT-P and SRN as well as EMT-2 and SRN.

4.1.4 Variance Components Model



Variance Components Models

2-levels Model

$$y_{ij} = \beta_{0j} + e_{ij} \quad (\text{equation 1})$$

$$\beta_{0j} = \beta_0 + u_{0j} \quad (\text{equation 2})$$

y_{ij} : the total 5-level TPTS score assessed for for each patient j by the i th rater

β_{0j} : unique intercept for each patient j .

$e_{ij} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for each patient j by the i th rater)

β_0 : the average intercept across patients

$u_{0j} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for each patient)

3-levels Model

$$y_{ijk} = \beta_{0jk} + e_{ijk} \quad (\text{equation 3})$$

$$\beta_{0jk} = b_{0(j)k} + u_{0jk} \quad (\text{equation 4})$$

$$b_{0(j)k} = b_{0(jk)} + v_{0(j)k} \quad (\text{equation 5})$$

y_{ijk} : the total 5-level TPTS score assessed for each patient k and department j by rater i

β_{0jk} : unique intercept for each patient k and department j

$e_{ijk} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for for each patient k and department j by rater i)

$b_{0(j)k}$: the average intercept at patient k

$u_{0jk} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for department j)

$b_{0(jk)}$: the average intercept across patients

$v_{0(j)k} \sim N(0, \sigma_v^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for patient k)

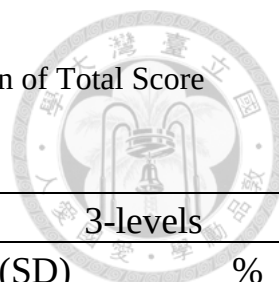


Table 6 Variance Components Models of 5-level TPTS Scores: Distribution of Total Score Variation by Rater, Department, and Patient

Overall	2-levels		3-levels	
	Estimate (SD)	%	Estimate (SD)	%
Variances*			0.567	
Patient level	0.310(0.557)	56.6	0.319(0.565)	56.2
Department level	—	—	0.059(0.242)	10.4
Rater level	0.237(0.487)	43.4	0.189(0.435)	33.3
Mean(SE)	2.90(0.74)	(<i>P</i> <0.001)†	2.97(0.08)	(<i>P</i> <0.001)†

SD = standard deviation. SE = standard error. *All variance estimates were significantly greater than zero (*P* < 0.001), emphasizing the necessity of accounting for the clustered hierarchical structure of the data. †the mean outcome was anticipated to be non-zero, the formal significance of the intercept coefficient was considered negligible.

Table 6 presents the distribution of total score variation in 5-level TPTS scores across three hierarchical levels: patient, department, and rater. Variance estimates and their corresponding percentages are reported for both a 2-level and a 3-level variance model, providing insights into how much variation each level contributes.

In the variance components analysis of 5-level TPTS scores, patient-level differences emerged as the main source of variation, accounting for 56.6% in the 2-level model (0.310±0.557) and 56.2% in the 3-level model (0.319±0.565). Rater-level variation contributed 43.4% in the 2-level model (0.237±0.487) and decreased to 33.3% in the 3-level model (0.189±0.435) when department-level variance was included. Notably, department-level differences accounted for 10.4% of the total variation in the 3-level model (0.059±0.242), highlighting a modest but significant influence of inter-department variability. All variance components were significantly greater than zero (*P*<0.001), underscoring the importance of considering the hierarchical clustering in the data. These findings emphasize that while patient-level factors are the primary contributors to score



variability, rater and department-level influences are also significant and must be accounted for in triage assessments.

Table 7 Two-Level Variance Components Model of 5-level TPTS Scores: Distribution of Total Score Variation Across Rater, Department, and Patient within the EMT Group

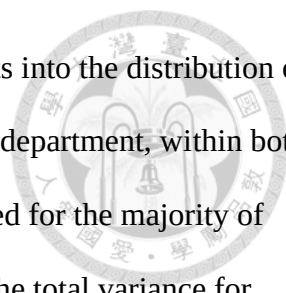
EMT Group	EMT-2		EMT-P	
	Estimate (SD)	%	Estimate (SD)	%
Variances*				
Patient level	0.372(0.610)	59.4	0.354(0.595)	61.2
Rater level	0.254(0.504)	40.6	0.224(0.473)	38.8
Mean(SE)	3.04(0.79)	($P < 0.001$)†	3.10(0.76)	($P < 0.001$)†

SD = standard deviation. SE = standard error. *All variance estimates were significantly greater than zero ($P < 0.001$), emphasizing the necessity of accounting for the clustered hierarchical structure of the data. †the mean outcome was anticipated to be non-zero, the formal significance of the intercept coefficient was considered negligible.

Table 8 Two-Level Variance Components Model of 5-level TPTS Scores: Distribution of Total Score Variation Across Rater, Department, and Patient within the RN Group

RN Group	JRN		SRN	
	Estimate (SD)	%	Estimate (SD)	%
Variances*			0.348	
Patient level	0.368(0.607)	70.0	0.238(0.489)	68.4
Rater level	0.158(0.398)	30.0	0.110(0.332)	31.6
Mean(SE)	2.76(0.73)	($P < 0.001$)†	2.73(0.59)	($P < 0.001$)†

SD = standard deviation. SE = standard error. *All variance estimates were significantly greater than zero ($P < 0.001$), emphasizing the necessity of accounting for the clustered hierarchical structure of the data. †the mean outcome was anticipated to be non-zero, the formal significance of the intercept coefficient was considered negligible.



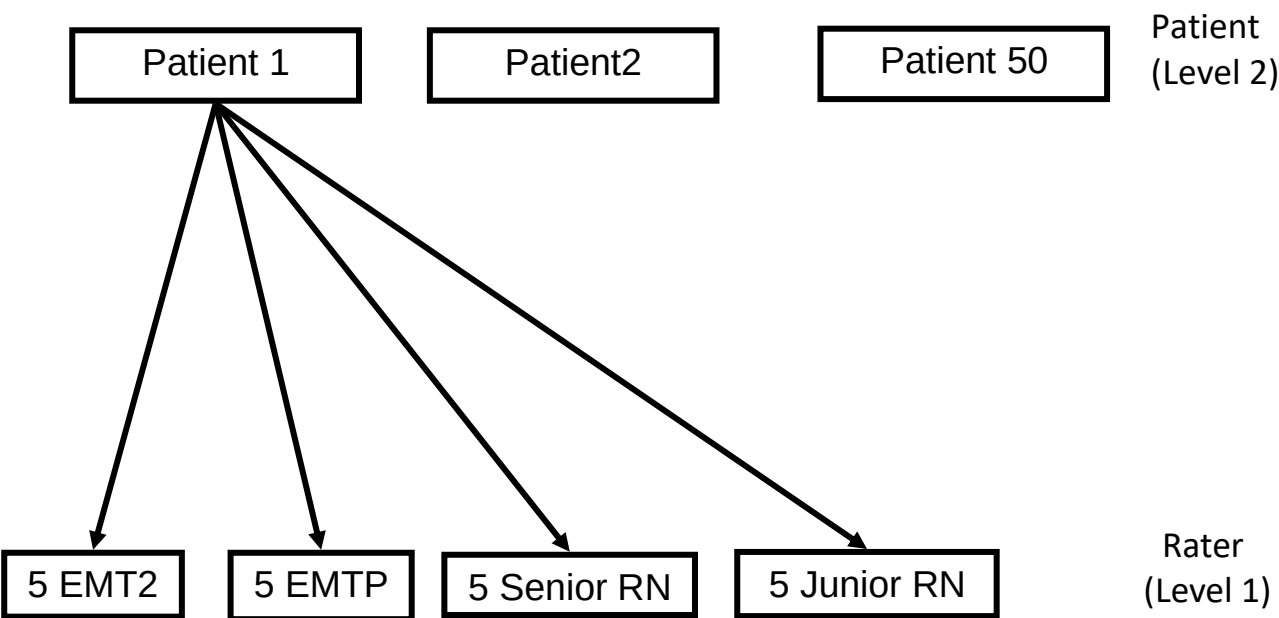
Tables 7 and 8 present the variance components models, offering insights into the distribution of total score variation across hierarchical levels, including patient, rater, and department, within both the EMT and RN groups. Among EMTs, patient-level differences accounted for the majority of variation, contributing 59.4% (0.372 ± 0.610) and 61.2% (0.354 ± 0.595) of the total variance for EMT-2 and EMT-P, respectively. Rater-level variation explained the remaining 40.6% (0.254 ± 0.504) and 38.8% (0.224 ± 0.473) in these groups.

In contrast, the RN group demonstrated higher consistency in individual scoring, with patient-level variation accounting for 70.0% (0.368 ± 0.607) among JRN and 68.4% (0.238 ± 0.489) among SRN. Rater-level variation in this group was substantially lower, contributing only 30.0% (0.158 ± 0.398) for JRN and 31.6% (0.110 ± 0.332) for SRN. These findings highlight that nurses (RNs) exhibit greater consistency in individual scoring, as reflected by the reduced rater-level variability compared to EMTs.

All variance estimates were significantly greater than zero ($P < 0.001$), affirming the need to account for hierarchical clustering in the analysis. The mean 5-level TPTS scores were significantly different from zero across both groups ($P < 0.001$), though the intercept's formal significance was negligible.

4.1.5 Two-Level Multilevel analysis

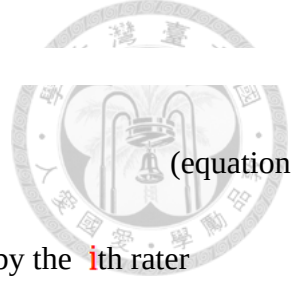
Figure 7 Framework of the Two-Level Multilevel Analysis Model



covariates

level-1(Rater): rater age, rater gender, rater seniority, job grade

level-2(Patient):reason for calling 911, vital signs, patient age, patient gender, treatment, transport time



Single level Model(Simple Linear Model)

- M1(Single level model):

$$y_i = \beta_0 + e_i \quad (\text{equation 6})$$

y_i : the total 5-level TPTS score assessed for for each patient by the i th rater

β_0 : the average intercept across patients

$e_{ij} \sim N(0, \sigma_e^2)$,: residual error term

Variance Components Model

- M2(Variance Components Model):

$$y_{ij} = \beta_{0j} + e_{ij} \quad (\text{equation 7})$$

$$\beta_{0j} = \beta_0 + u_{0j} \quad (\text{equation 8})$$

y_{ij} : the total 5-level TPTS score assessed for for each patient j by the i th rater

β_{0j} : unique intercept for each patient j .

$e_{ij} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for each patient j by the i th rater)

β_0 : the average intercept across patients

$u_{0j} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for each patient)

Random Intercept Models

Two-Level Multilevel Models

level-1(Rater): rater age, rater gender, rater seniority, job grade

level-2(Patient): reason for calling 911, vital signs, patient age, patient gender, treatment, transport time

- M3(Level-1 Predictors+ job grade):

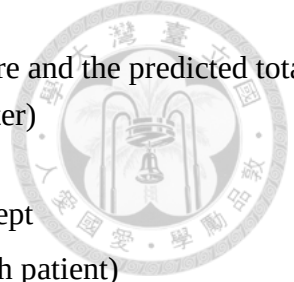
$$y_{ij} = \beta_{0j} + \beta_{1ij} + e_{ij} \quad (\text{equation 9})$$

$$\beta_{0j} = \beta_0 + u_{0j} \quad (\text{equation 10})$$

y_{ij} : the total 5-level TPTS score assessed for for each patient j by the i th rater

β_{0j} : unique intercept for each patient j .

$e_{ij} \sim N(0, \sigma_e^2)$,: residual error term



(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for each patient j by the i th rater)

β_0 : the average intercept across patients

$u_{0j} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for each patient)

β_{1ij} : the regression coefficients for level-1(Rater): job grade

- M4(Level-1 Predictors+ rater age, rater gender, rater seniority, job grade):

$$y_{ij} = \beta_{0j} + \beta_{1ij} + \beta_{2ij} + \beta_{3ij} + \beta_{4ij} + e_{ij} \quad (\text{equation 11})$$

$$\beta_{0j} = \beta_0 + u_{0j} \quad (\text{equation 12})$$

y_{ij} : the total 5-level TPTS score assessed for for each patient j by the i th rater

β_{0j} : unique intercept for each patient j .

$e_{ij} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for each patient j by the i th rater)

β_0 : the average intercept across patients

$u_{0j} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for each patient)

$\beta_{1,2,3,4ij}$: the regression coefficients for level-1(Rater): rater age, rater gender, rater seniority, job grade

- M5(Level-1 Predictors+ rater age, rater gender, rater seniority, job grade,

Level-2 Predictors+ reason for calling 911, vital signs):

$$y_{ij} = \beta_{0j} + \beta_{1ij} + \beta_{2ij} + \beta_{3ij} + \beta_{4ij} + \beta_{5ij} + \beta_{6ij} + e_{ij} \quad (\text{equation 13})$$

$$\beta_{0j} = \beta_0 + u_{0j} \quad (\text{equation 14})$$

y_{ij} : the total 5-level TPTS score assessed for for each patient j by the i th rater

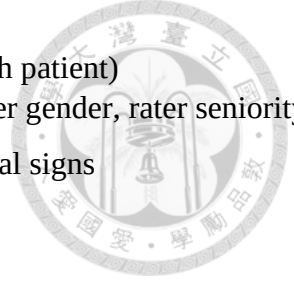
β_{0j} : unique intercept for each patient j .

$e_{ij} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for each patient j by the i th rater)

β_0 : the average intercept across patients

$u_{0j} \sim N(0, \sigma_u^2)$: the random effects for the intercept



(the variation in estimated total 5-level TPTS score for each patient)

$\beta_{1,2,3,4,5,6ij}$: the regression coefficients for level-1(Rater): rater age, rater gender, rater seniority, job grade and level-2(Patient): reason for calling 911, vital signs

• M6(Level-1 Predictors+ rater age, rater gender, rater seniority, job grade, Level-2 Predictors+ reason for calling 911, vital signs, patient age, patient gender, treatment, transport time):

$$y_{ij} = \beta_{0j} + \beta_{1ij} + \beta_{2ij} + \beta_{3ij} + \beta_{4ij} + \beta_{5ij} + \beta_{6ij} + \beta_{7ij} + \beta_{8ij} + \beta_{9ij} + \beta_{10ij} + e_{ij}$$

(equation 15)

$$\beta_{0j} = \beta_0 + u_{0j} \quad (\text{equation 16})$$

y_{ij} : the total 5-level TPTS score assessed for for each patient j by the i th rater

β_{0j} : unique intercept for each patient j .

$e_{ij} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for each patient j by the i th rater)

β_0 : the average intercept across patients

$u_{0j} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for each patient)

$\beta_{1,2,3,4,5,6,7,8,9,10ij}$: the regression coefficients for level-1(Rater): rater age, rater gender, rater seniority, job grade and level-2(Patient):

reason for calling 911, vital signs, patient age, patient gender, treatment, transport time

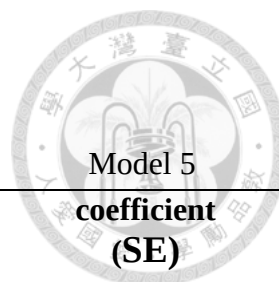


Table 9 Two-Level Multilevel analysis of 5-level TPTS score

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	coefficient (SE)	coefficient (SE)	coefficient (SE)	coefficient (SE)	coefficient (SE)	coefficient (SE)
Fixed effects						
Intercept†	2.91(0.23)**	2.91(0.08)**	3.04(0.08)**	2.49(0.19)**	2.49(0.20)**	2.59(0.28)**
Raters-level variables						
EMT-P(Ref: EMT-2)	—	—	0.06(0.04)	0.08(0.04)*	0.08(0.04)*	0.08(0.04)*
JRN(Ref: EMT-2)	—	—	-0.28(0.04)**	-0.14(0.06)**	-0.14(0.06)**	-0.14(0.06)**
SRN(Ref: EMT-2)	—	—	-0.31(0.04)**	-0.24(0.07)**	-0.24(0.07)**	-0.24(0.07)**
Rater Age	—	—	—	0.02(0.01)**	0.02(0.01)**	0.02(0.01)**
Rater Gender (Ref: female)	—	—	—	0.08(0.04)*	0.08(0.04)*	0.08(0.04)*
Rater Seniority	—	—	—	-0.01(0.01)	-0.01(0.01)	-0.01(0.01)
Patients-level variables						
Trauma (Ref:non-trauma)	—	—	—	—	0.30(0.12)**	0.41(0.13)**
unstable V/S (Ref:normal V/S)	—	—	—	—	-0.83(0.14)**	-0.61(0.16)**
Patient Age	—	—	—	—	—	0.00(0.00)
Patient Gender (Ref: female)	—	—	—	—	—	-0.20(0.12)
Treatment(Ref:No)	—	—	—	—	—	-0.21(0.13)
Transport time	—	—	—	—	—	0.01(0.01)
Random effects	Variance (SD)	Variance (SD)	Variance (SD)	Variance (SD)	Variance (SD)	Variance (SD)
Patients-level	—	0.31(0.56)	0.31(0.56)	0.31(0.56)	0.14(0.38)	0.13(0.36)
Raters-level	0.55(0.74)	0.24(0.49)	0.21(0.46)	0.20(0.45)	0.20(0.45)	0.20(0.45)
-2loglikelihood	2234.05	1565.20	1457.76	1467.79	1433.57	1447.79
AIC	2238.05	1571.20	1469.76	1485.79	1455.57	1477.79
BIC	2247.87	1585.92	1499.18	1529.90	1509.46	1551.21

Abbreviations: SE= Standard Error, SD= Standard Deviation, EMT= Emergency Medical

Technician, JRN=Junior Registered Nurse, SRN=Senior Registered Nurse.

V/S=Vital signs,AIC, Akaike Information Criteria; BIC, Bayesian Information Criteria.

* borderline significance $p \approx 0.05$, ** $p < 0.05$ † The intercept coefficient estimate holds little formal significance when covariates are included in the model.

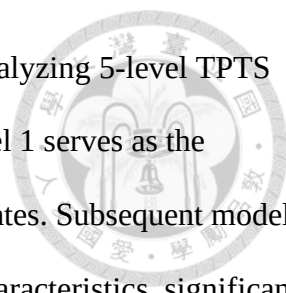
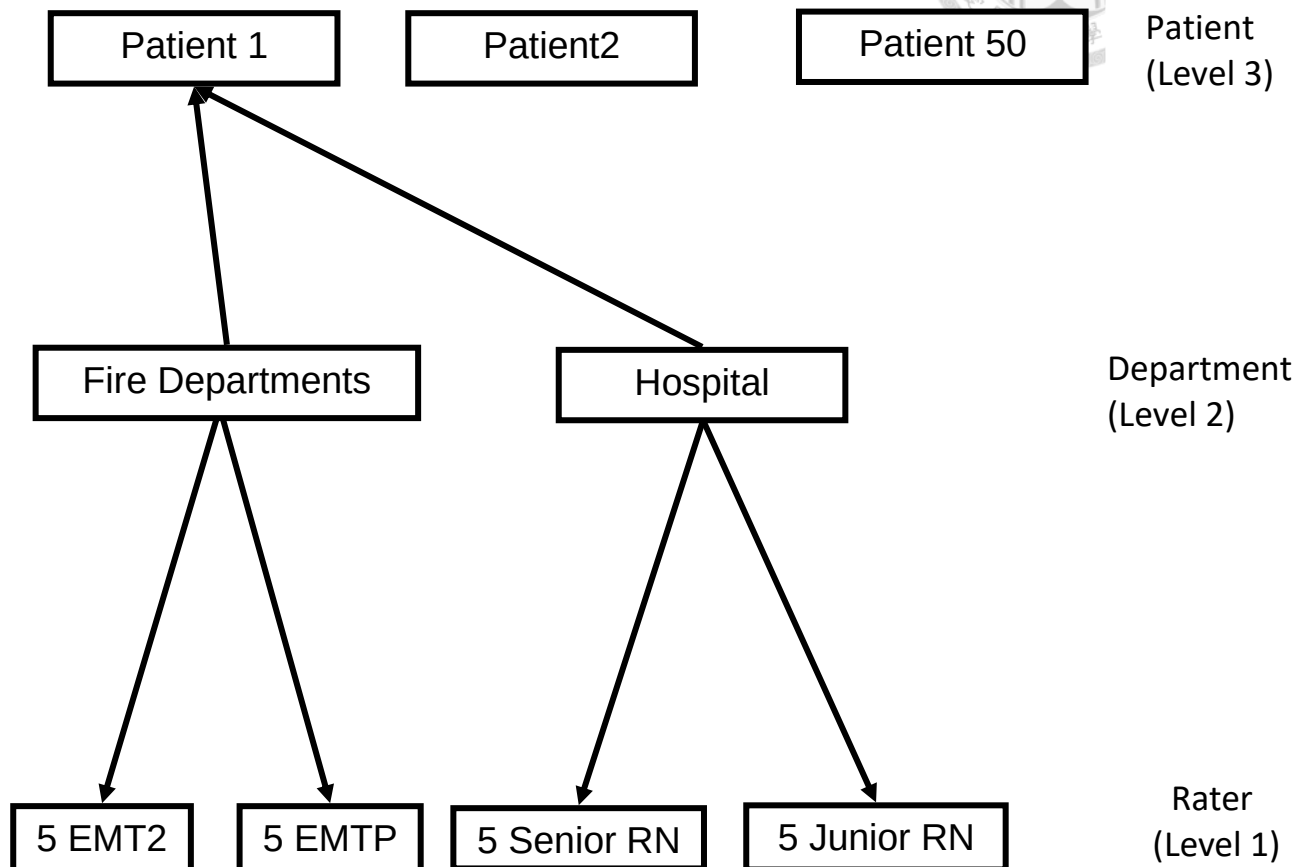


Table 9 summarizes the results of six multilevel regression models analyzing 5-level TPTS scores, progressively incorporating rater- and patient-level variables. Model 1 serves as the baseline, showing a mean 5-level TPTS score of 2.91 ± 0.23 without covariates. Subsequent models reveal that rater-level factors, including qualifications and demographic characteristics, significantly influenced scoring. For instance, EMT-Ps scored slightly higher than EMT-2s ($0.08 \pm 0.04, p < 0.05$), while JRN and SRN assigned significantly lower scores compared to EMT-2s (-0.28 ± 0.04 and -0.31 ± 0.04 , respectively, $p < 0.05$). Among patient-level variables, trauma cases were associated with higher scores ($0.41 \pm 0.13, p < 0.05$), while patients with unstable vital signs had lower scores ($-0.61 \pm 0.16, p < 0.05$). Model fit improved across iterations, with reductions in -2loglikelihood, AIC and BIC values. Notably, the unexplained variance at the patient and rater levels decreased substantially as more covariates were added, stabilizing at 0.13 ± 0.36 and 0.20 ± 0.45 , respectively, in Model 6.

Although rater-level variability was statistically significant, its contribution to the overall mean score was minimal, with patient-level variability emerging as the primary determinant of score differences.

4.1.6 Three-Level Multilevel analysis

Figure 8 Framework of the Three-Level Multilevel Analysis Model

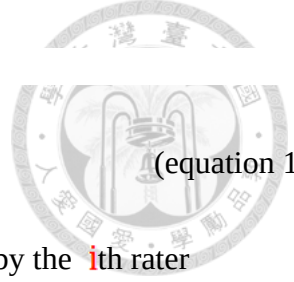


covariates

level-1(Rater): rater age, rater gender, rater seniority, job grade

level-2(Department): no covariates

level-3(Patient): reason for calling 911, vital signs, patient age, patient gender, treatment, transport time



Single level Model(Simple Linear Model)

- M1(Single level model):

$$y_i = \beta_0 + e_i \quad (\text{equation 17})$$

y_i : the total 5-level TPTS score assessed for for each patient by the i th rater

β_0 : the average intercept across patients

$e_{ij} \sim N(0, \sigma_e^2)$,: residual error term

Variance Components Model

- M2(Variance Components Model):

$$y_{ijk} = \beta_{0jk} + e_{ijk} \quad (\text{equation 18})$$

$$\beta_{0jk} = b_{0(j)k} + u_{0jk} \quad (\text{equation 19})$$

$$b_{0(j)k} = b_{0(jk)} + v_{0(j)k} \quad (\text{equation 20})$$

y_{ijk} : the total 5-level TPTS score assessed for each patient k and department j by rater i

β_{0jk} : unique intercept for each patient k and department j

$e_{ijk} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for for each patient k and department j by rater i)

$b_{0(j)k}$: the average intercept at patient k

$u_{0jk} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for department j)

$b_{0(jk)}$: the average intercept across patients

$v_{0(j)k} \sim N(0, \sigma_v^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for patient k)

Random Intercept Models

Three-Level Multilevel Models

level-1(Rater): rater age, rater gender, rater seniority, job grade

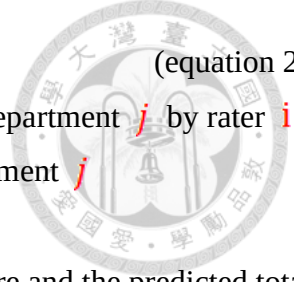
level-2(Department): no covariates

level-3(Patient): reason for calling 911, vital signs, patient age, patient gender, treatment, transport time

- M3(Level-1 Predictors+ job grade):

$$y_{ijk} = \beta_{0jk} + \beta_{1ijk} + e_{ijk} \quad (\text{equation 21})$$

$$\beta_{0jk} = b_{0(j)k} + u_{0jk} \quad (\text{equation 22})$$



$$b_{0(j)k} = b_{0(jk)} + u_{0(j)k} \quad (\text{equation 23})$$

y_{ijk} : the total 5-level TPTS score assessed for each patient k and department j by rater i

β_{0jk} : unique intercept for each patient k and department j

$e_{ijk} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for for each patient k and department j by rater i)

$b_{0(j)k}$: the average intercept at patient k

$u_{0jk} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for department j)

$b_{0(jk)}$: the average intercept across patients

$v_{0(j)k} \sim N(0, \sigma_v^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for patient k)

β_{1ijk} : the regression coefficients for level-1(Rater): job grade

- M4(Level-1 Predictors+ job grade, rater age, rater gender, rater seniority):

$$y_{ijk} = \beta_{0jk} + \beta_{1ijk} + \beta_{2ijk} + \beta_{3ijk} + \beta_{4ijk} + e_{ijk} \quad (\text{equation 24})$$

$$\beta_{0jk} = b_{0(j)k} + u_{0jk} \quad (\text{equation 25})$$

$$b_{0(j)k} = b_{0(jk)} + v_{0(j)k} \quad (\text{equation 26})$$

y_{ijk} : the total 5-level TPTS score assessed for each patient k and department j by rater i

β_{0jk} : unique intercept for each patient k and department j

$e_{ijk} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for for each patient k and department j by rater i)

$b_{0(j)k}$: the average intercept at patient k

$u_{0jk} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for department j)

$b_{0(jk)}$: the average intercept across patients

$v_{0(j)k} \sim N(0, \sigma_v^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for patient k)

$\beta_{1,2,3,4ijk}$: the regression coefficients for level-1(Rater): job grade, rater age, rater gender, rater seniority



- M5(Level-1 Predictors+ job grade, rater age, rater gender, rater seniority

Level-3 Predictors+ reason for calling 911, vital signs):

$$y_{ijk} = \beta_{0jk} + \beta_{1ijk} + \beta_{2ijk} + \beta_{3ijk} + \beta_{4ijk} + \beta_{5k} + \beta_{6k} + e_{ijk} \quad (\text{equation 27})$$

$$\beta_{0jk} = b_{0(j)k} + u_{0jk} \quad (\text{equation 28})$$

$$b_{0(j)k} = b_{0(jk)} + v_{0(j)k} \quad (\text{equation 29})$$

y_{ijk} : the total 5-level TPTS score assessed for each patient k and department j by rater i

β_{0jk} : unique intercept for each patient k and department j

$e_{ijk} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for for each patient k and department j by rater i)

$b_{0(j)k}$: the average intercept at patient k

$u_{0jk} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for department j)

$b_{0(jk)}$: the average intercept across patients

$v_{0(j)k} \sim N(0, \sigma_v^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for patient k)

$\beta_{1,2,3,4ijk}$: the regression coefficients for level-1(Rater): job grade, rater age, rater gender, rater seniority

$\beta_{5,6k}$: the regression coefficients for level-3(Patient): reason for calling 911, vital signs

- M6(Level-1 Predictors+ job grade, rater age, rater gender, rater seniority

Level-3 Predictors+ reason for calling 911, vital signs, patient age, patient gender, treatment, transport time):

$$y_{ijk} = \beta_{0jk} + \beta_{1ijk} + \beta_{2ijk} + \beta_{3ijk} + \beta_{4ijk} + \beta_{5k} + \beta_{6k} + \beta_{7k} + \beta_{8k} + \beta_{9k} + \beta_{10k} + e_{ijk}$$

(equation 30)

$$\beta_{0jk} = b_{0(j)k} + u_{0jk} \quad (\text{equation 31})$$

$$b_{0(j)k} = b_{0(jk)} + v_{0(j)k} \quad (\text{equation 32})$$

y_{ijk} : the total 5-level TPTS score assessed for each patient k and department j by rater i

β_{0jk} : unique intercept for each patient k and department j

$e_{ijk} \sim N(0, \sigma_e^2)$,: residual error term

(which shows the difference between the observed total 5-level TPTS score and the predicted total 5-level TPTS score for for each patient k and department j by rater i)



$b_{0(j)k}$: the average intercept at patient k

$u_{0jk} \sim N(0, \sigma_u^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for department j)

$b_{0(jk)}$: the average intercept across patients

$v_{0(j)k} \sim N(0, \sigma_v^2)$: the random effects for the intercept

(the variation in estimated total 5-level TPTS score for patient k)

$\beta_{1,2,3,4ijk}$: the regression coefficients for level-1(Rater): job grade, rater age, rater gender, rater seniority:

$\beta_{6,7,8,9,10k}$: the regression coefficients for level-3(Patient): reason for calling 911, vital signs, patient age, patient gender, treatment, transport time

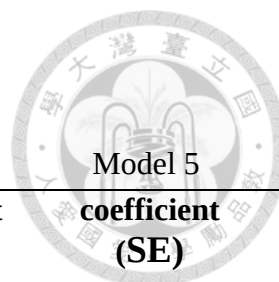


Table 10 Three-Level Multilevel analysis of 5-level TPTS score

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	coefficient (SE)	coefficient (SE)	coefficient (SE)	coefficient (SE)	coefficient (SE)	coefficient (SE)
Fixed effects						
Intercept†	2.91(0.23)**	2.97(0.08)**	3.04(0.09)**	2.45(0.19)**	2.45(0.20)**	2.57(0.29)**
Raters-level variables						
EMT-P(Ref: EMT-2)	—	—	0.06(0.04)	0.09(0.04)**	0.08(0.04)**	0.08(0.04)**
JRN(Ref: EMT-2)	—	—	-0.28(0.04)**	-0.13(0.06)**	-0.13(0.06)**	-0.13(0.06)**
SRN(Ref: EMT-2)	—	—	-0.31(0.04)**	-0.23(0.07)**	-0.23(0.07)**	-0.23(0.07)**
Rater Age	—	—	—	0.02(0.01)**	0.02(0.01)**	0.02(0.01)**
Rater Gender (Ref: female)	—	—	—	0.08(0.04)**	0.08(0.04)**	0.08(0.04)**
Rater Seniority	—	—	—	-0.01(0.01)	-0.01(0.01)	-0.01(0.01)
Patients-level variables						
Trauma (Ref:non-trauma)	—	—	—	—	0.30(0.12)**	0.40(0.13)**
unstable V/S (Ref:normal V/S)	—	—	—	—	-0.84(0.15)**	-0.63(0.17)**
Patient Age	—	—	—	—	—	0.00(0.00)
Patient Gender (Ref: female)	—	—	—	—	—	-0.20(0.12)
Treatment(Ref:No)	—	—	—	—	—	-0.21(0.13)
Transport time	—	—	—	—	—	0.01(0.01)
Random effects	Variance (SD)	Variance (SD)	Variance (SD)	Variance (SD)	Variance (SD)	Variance (SD)
Patients-level	—	0.32(0.57)	0.32(0.57)	0.32(0.57)	0.14(0.38)	0.14(0.37)
Department level	—	0.06(0.24)	0.02(0.14)	0.02(0.15)	0.02(0.15)	0.02(0.15)
Raters-level	0.55(0.74)	0.19(0.44)	0.19(0.44)	0.19(0.44)	0.19(0.44)	0.19(0.44)
-2loglikelihood	2234.05	1494.99	1448.73	1457.17	1423.52	1438.31
AIC	2238.05	1502.99	1462.73	1477.17	1447.52	1470.31
BIC	2247.87	1522.62	1497.06	1526.18	1506.31	1548.62

Abbreviations: SE= Standard Error, SD= Standard Deviation, EMT= Emergency Medical

Technician, JRN=Junior Registered Nurse, SRN=Senior Registered Nurse.

V/S=Vital signs,AIC, Akaike Information Criteria; BIC, Bayesian Information Criteria.

* borderline significance $p \approx 0.05$, ** $p < 0.05$ † The intercept coefficient estimate holds little formal significance when covariates are included in the model.

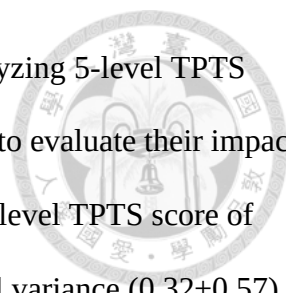


Table 10 presents the results of six three-level regression models analyzing 5-level TPTS scores, progressively incorporating variables at the rater and patient levels to evaluate their impact on score variability. Model 1 serves as the baseline, showing an average 5-level TPTS score of 2.91 ± 0.23 without considering covariates. Model 2 introduces patient-level variance (0.32 ± 0.57) and department-level variance (0.06 ± 0.24), significantly reducing unexplained rater-level variance from 0.55 ± 0.74 to 0.19 ± 0.44 . Model 3 introduces rater-level variables, revealing that EMT-Ps scored slightly higher than EMT-2s (0.06 ± 0.04), while JRN and SRNs scored significantly lower (-0.28 ± 0.04 and -0.31 ± 0.04 , respectively, $p < 0.05$). Models 4 through 6 progressively incorporate additional rater characteristics, such as age ($0.02 \pm 0.01, p < 0.05$) and gender ($0.08 \pm 0.04, p < 0.05$), as well as patient-level factors, including trauma status ($0.40 \pm 0.13, p < 0.05$) and unstable vital signs ($-0.63 \pm 0.17, p < 0.05$). The addition of patient-level variables substantially reduced unexplained variance at the patient level from 0.32 ± 0.57 in Model 2 to 0.14 ± 0.37 in Model 6, indicating that these factors were the primary contributors to score differences.

Although individual-level variability, such as rater characteristics, reached statistical significance, its contribution to the overall mean score was minimal. Patient-level variability remained the dominant factor influencing the average 5-level TPTS scores.

4.2 Results of Study 2



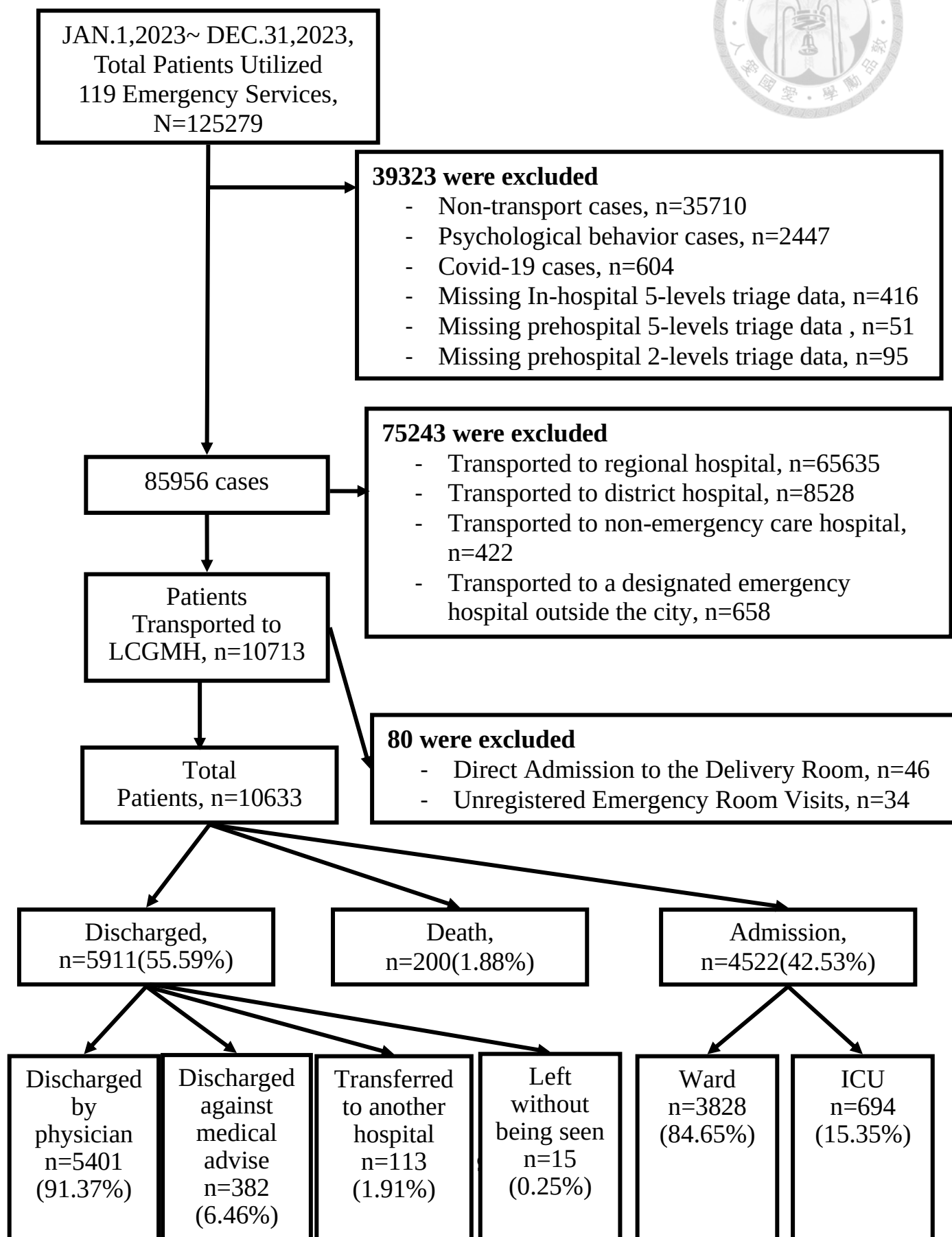
4.2.1 Patient Flow and Enrollment

From January 1, 2023, to December 31, 2023, a total of 125,279 patients utilized the 119 emergency services (Figure 9). Of these, 39,323 cases were excluded due to non-transport incidents (n=35,710), psychological or behavioral cases (n=2,447), COVID-19-related cases (n=604), or missing prehospital and in-hospital triage data (n=562). After these exclusions, 85,956 cases involved patients transported to healthcare facilities, with 10,713 arriving at LCGMH. Following further exclusions, including direct admissions to the delivery room (n=46) and unregistered emergency visits (n=34), the final study population consisted of 10,633 patients.

In the emergency department, 5,911 patients (55.59%) were discharged, including 5,401 (91.37%) discharged by physicians, 382 (6.46%) who left against medical advice, and 15 (0.25%) who left without being seen. Additionally, 113 patients (1.91%) were transferred to other hospitals. Admissions accounted for 4,522 patients (42.53%), with 3,828 (84.65%) admitted to general wards and 694 (15.35%) to the ICU. A total of 200 patients (1.88%) died during their emergency department stay.



Figure 9 Flow chart of the studied patients



4.2.2 Characteristics of the study population

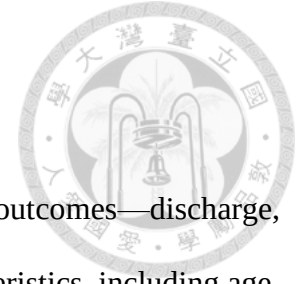


Table 11 presents an analysis of 10,633 patients categorized by their outcomes—discharge, admission, or death. The data highlight demographics and clinical characteristics, including age distribution, gender differences, reasons for utilizing emergency services, triage classifications, and dispatch timing, offering insights into factors influencing patient outcomes.

The median age of the population demonstrated significant differences across outcomes ($p < 0.0001$), increasing from discharged patients (47 years) to those admitted (65 years) and deceased (68.5 years). Younger patients (< 20 years) were predominantly discharged (12.21%), whereas older patients (60–79 years and > 79 years) accounted for higher admission rates (38.37% and 22.71%, respectively) and deaths (41% and 24.50%, respectively).

Males constituted 57.29% of the cases and were more likely to experience adverse outcomes, representing 65% of deaths. Females, who made up 42.71% of the cohort, exhibited a higher discharge rate compared to males (46% vs. 54%, $p < 0.0001$).

Trauma cases were a frequent reason for emergency service utilization, with 45.53% of such cases resulting in discharge. However, non-trauma cases, particularly adult non-trauma emergencies, dominated admissions (67.91%) and deaths (82.00%), indicating greater severity compared to trauma cases ($p < 0.0001$).

Triage level significantly correlated with outcomes. Using the 2-level TPTS, 93% of deaths occurred in the "Emergency" category ($p < 0.0001$). Within the 5-level TPTS, Level 1 (critical) cases exhibited the highest mortality rate, accounting for 80% of deaths, while most discharges occurred at Level 3 (56.05%). A similar trend was observed with the 5-level TTAS, where Level 1 comprised 85% of deaths.

Dispatch timing also influenced outcomes. The middle shift (16:00–24:00) had the highest mortality rate, with 45.50% of deaths occurring during this period, while the day shift (08:00–16:00) accounted for the largest proportion of discharges (43.18%, $p = 0.0044$). There were no significant differences in outcomes by day of the week, as the distribution remained relatively even across all weekdays ($p = 0.9117$).

Table 11 Patient demographics data, presented with count (percentage)



Variables	Total N=10633	Discharged n=5911 (55.59%)	Admission n=4522 (42.53%)	Death n=200 (1.88%)	p-value
Age^a	57 (34-72)	47 (28-66)	65 (47-78)	68.5 (53-79)	<.0001
Age group					<.0001
<20	1055(9.92)	722(12.21)	328(7.25)	5(2.50)	
20-39	2106(19.81)	1643(27.80)	447(9.89)	16(8.00)	
40-59	2535(23.84)	1502(25.41)	985(21.78)	48(24.00)	
60-79	3292(30.96)	1475(24.95)	1735(38.37)	82(41.00)	
>79	1645(15.47)	569(9.63)	1027(22.71)	49(24.50)	
Gender					<.0001
Male	6092(57.29)	3192(54.00)	2770(61.26)	130(65.00)	
Female	4541(42.71)	2719(46.00)	1752(38.74)	70(35.00)	
Reason for calling 119					<.0001
Trauma	3935(37.01)	2691(45.53)	1211(26.78)	33(16.50)	
Adult non-trauma	6096(57.33)	2861(48.40)	3071(67.91)	164(82.00)	
Pediatric non-trauma (≤18 y)	602(5.66)	359(6.07)	240(5.31)	3(1.50)	
2-level TPTS (Emergency)	3008(28.29)	938(15.87)	1884(41.66)	186(93.00)	<.0001
5-level TPTS Triage level					<.0001
Level 1	1174(11.04)	254(4.30)	760(16.81)	160(80.00)	
Level 2	3664(34.46)	1751(29.62)	1884(41.66)	29(14.50)	
Level 3	4897(46.05)	3313(56.05)	1575(34.83)	9(4.50)	
Level 4	463(4.35)	295(4.99)	168(3.72)	0(0)	
Level 5	435(4.09)	298(5.04)	135(2.99)	2(1.00)	
5-level TTAS Triage level					<.0001

Level 1	997(9.38)	161(2.72)	666(14.73)	170(85.00)		0.0044
Level 2	2993(28.15)	1168(19.76)	1803(39.87)	22(11.00)		
Level 3	6461(60.76)	4414(74.67)	2039(45.09)	8(4.00)		
Level 4	180(1.69)	166(2.81)	14(0.31)	0(0)		
Level 5	2(0.02)	2(0.03)	0(0)	0(0)		
Dispatch time						
Day shift(08-16)	4591(43.18)	1208(20.44)	890(19.68)	46(23.00)		
Middle shift(16-24)	3898(36.66)	2462(41.65)	2038(45.07)	91(45.50)		
Night shift(24-08)	2144(20.16)	2241(37.91)	1594(35.25)	63(31.50)		
Weekday						0.9117
Monday	1646(15.48)	913(15.45)	703(15.55)	30(15.00)		
Tuesday	1563(14.70)	855(14.46)	675(14.93)	33(16.50)		
Wednesday	1533(14.42)	862(14.58)	639(14.13)	32(16.00)		
Thursday	1515(14.25)	852(14.41)	644(14.24)	19(9.50)		
Friday	1548(14.56)	863(14.60)	657(14.53)	28(14.00)		
Saturday	1416(13.32)	786(13.30)	604(13.36)	26(13.00)		
Sunday	1412(13.28)	780(13.20)	600(13.27)	32(16.00)		

a Presented with median (Interquartile range).

Abbreviations: TPTS=Taiwan Prehospital Triage System, TTAS=Taiwan Triage and Acuity

Scale

4.2.3 Main results



Table 12 analyzes 10,633 patients, highlighting significant differences in ED LOS and ESI resource utilization across discharge, admission, and death outcomes.

The median ED LOS varied notably among the groups ($p < 0.0001$). Discharged patients had the shortest median LOS at 2.73 hours, compared to 13.90 hours for admitted patients and 2.29 hours for those who died. In terms of LOS distribution, 78.70% of discharged patients stayed less than 6 hours, while admitted patients were predominantly in the 6–24 hours category (39.10%). A considerable portion of admitted patients (35.27%) stayed beyond 24 hours. Patients who died showed a more balanced distribution, with 70% staying less than 6 hours and 30% staying longer.

ESI resource utilization patterns also differed significantly among the groups. Most patients (88.95%) required many resources, with this proportion rising to 99.80% for admitted patients and 96.0% for those who died. In contrast, 15.23% of discharged patients utilized only one resource, while nearly none of the admitted (0.18%) or deceased (0.50%) patients required just one. Additionally, 2.50% of patients required no resources, a group mainly composed of discharged patients (4.36%). Notably, 3.50% of patients who died used no resources, possibly due to rapid clinical deterioration.

These results reveal a strong relationship between ED LOS, ESI resource utilization, and patient outcomes, emphasizing the importance of optimizing triage processes and resource allocation in emergency care settings.

Table 12 Primary and Secondary Outcome



Outcome	Total N=10633	Discharged n=5911 (55.59%)	Admission n=4522 (42.53%)	Death n=200 (1.88%)	p-value
ED LOS(hr)^a	4.94 (2.29-15.24)	2.73 (1.59-5.31)	13.90 (5.87-34.42)	2.29 (1.22-7.05)	<.0001
ED LOS group(hr)					<.0001
<6	5951(55.97)	4652(78.70)	1159(25.63)	140(70.00)	
6-24	2768(26.03)	964(16.31)	1768(39.10)	36(18.00)	
24-48	984(9.25)	161(2.72)	805(17.80)	18(9.00)	
>48	930(8.75)	134(2.27)	790(17.47)	6(3.00)	
ESI Resources					
Many	9458(88.95)	4753(80.41)	4513(99.80)	192(96.0)	<.0001
One	909(8.55)	900(15.23)	8(0.18)	1(0.50)	<.0001
None	266(2.50)	258(4.36)	1(0.02)	7(3.50)	<.0001

a Presented with median (Interquartile range).

Abbreviations: ED LOS= Emergency department length of stay

4.2.4 Cumulative Incidence for Different Events



Cumulative Incidence Function $F_j(t)$

The cumulative incidence function describes the probability of a specific event j occurring by time t , accounting for the presence of competing risks.

It is mathematically defined as:

$$F_j(t) = \int_0^t S(u)h_j(u)du$$

$F_j(t)$: The cumulative incidence function for event j , representing the probability that event j has occurred by time t , considering the presence of other competing events.

$S(t)$: The overall survival function, which represents the probability of surviving free of any event (of any type) up to time t .

It is further defined as:

$$S(t) = \exp\left(-\sum_j H_j(t)\right)$$

where $H_j(t) = \int_0^t h_j(u)du$ is the cumulative cause-specific hazard for event j .

$H_j(t)$: The cause-specific hazard rate for event j , representing the instantaneous rate at which event j occurs at time t in individuals who are still at risk (i.e., have not experienced any event yet).

Figure 10 CIF for different event types - Discharged

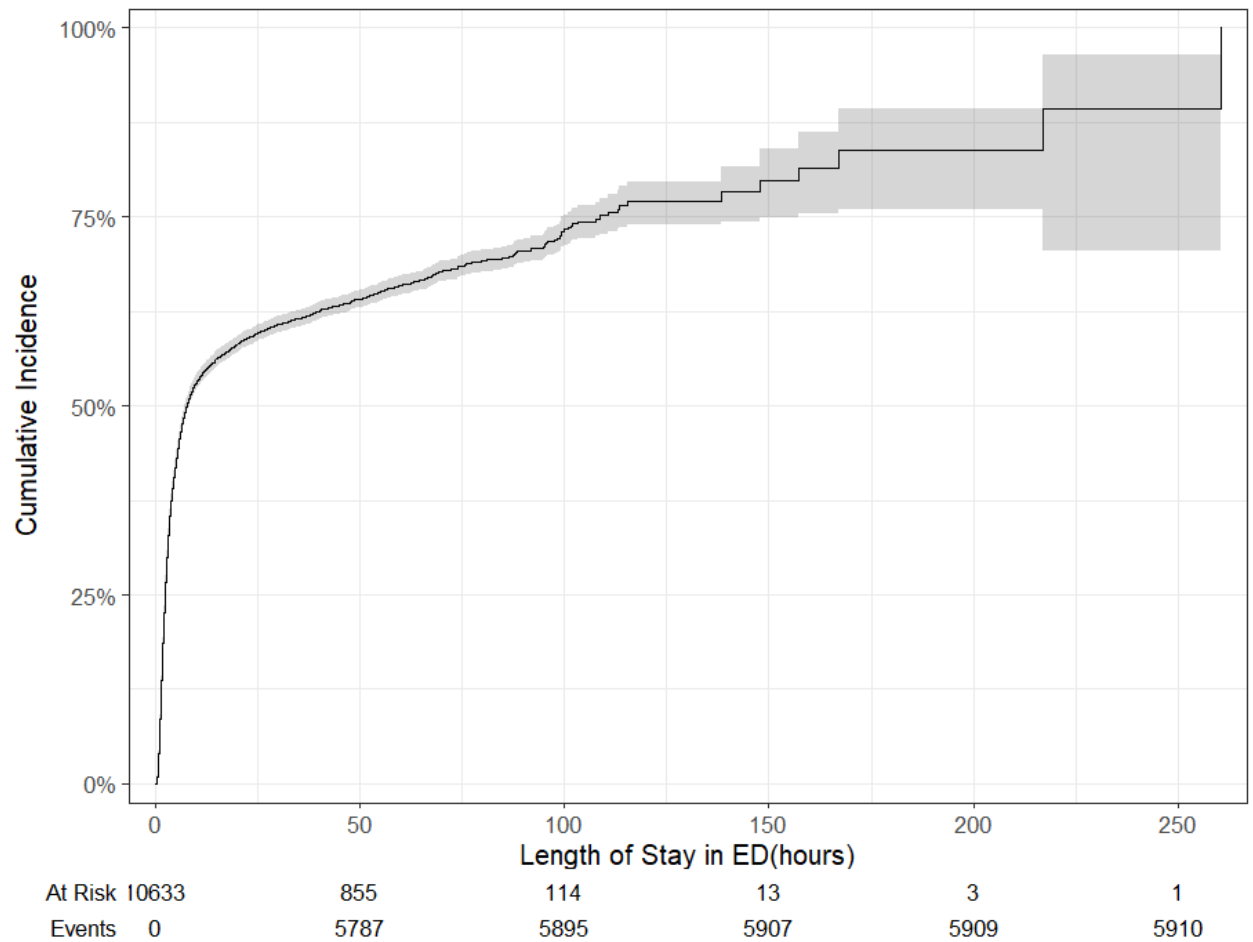


Figure 10 presents the non-parametric CIF estimates for discharge events (status == 1) in comparison to other outcomes, such as admission and death, which are treated as censored (status == 0). The CIF for discharge events, demonstrating the proportion of patients discharged over time and the corresponding number of patients at risk. The CIF rises steeply within the first few hours, with the median discharge time occurring at approximately 4–5 hours, indicating that 50% of all discharges are completed within this timeframe. Initially, all 10,633 patients are at risk, but this number declines rapidly as discharges occur, with only 855 patients remaining at risk by 50 hours and 114 by 100 hours.

Figure 11 CIF for different event types - Admission

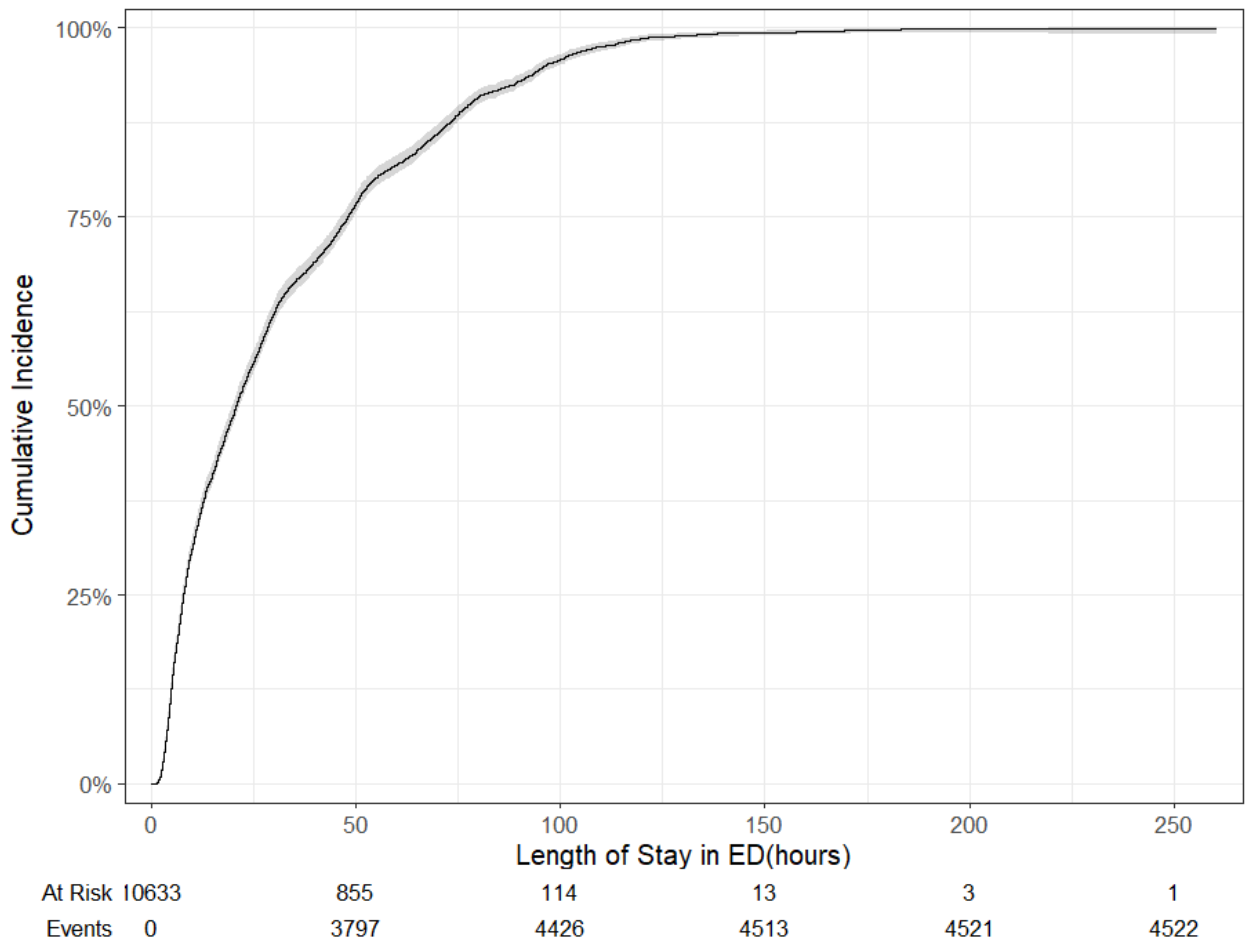


Figure 11 presents the non-parametric CIF estimates for admission events (status == 1) in comparison to other outcomes, such as discharge and death, which are treated as censored (status == 0). The CIF illustrates the proportion of patients admitted over time while accounting for competing risks. The CIF rises steadily during the first 24 hours, with the cumulative incidence reaching approximately 50% around 13–14 hours, indicating that half of all admissions occur within this timeframe. Initially, all 10,633 patients are at risk for admission, but this number declines significantly as events occur. By 50 hours, only 855 patients remain at risk, and by 100 hours, the at-risk population decreases further to 114.

Figure 12 CIF for different event types - Death

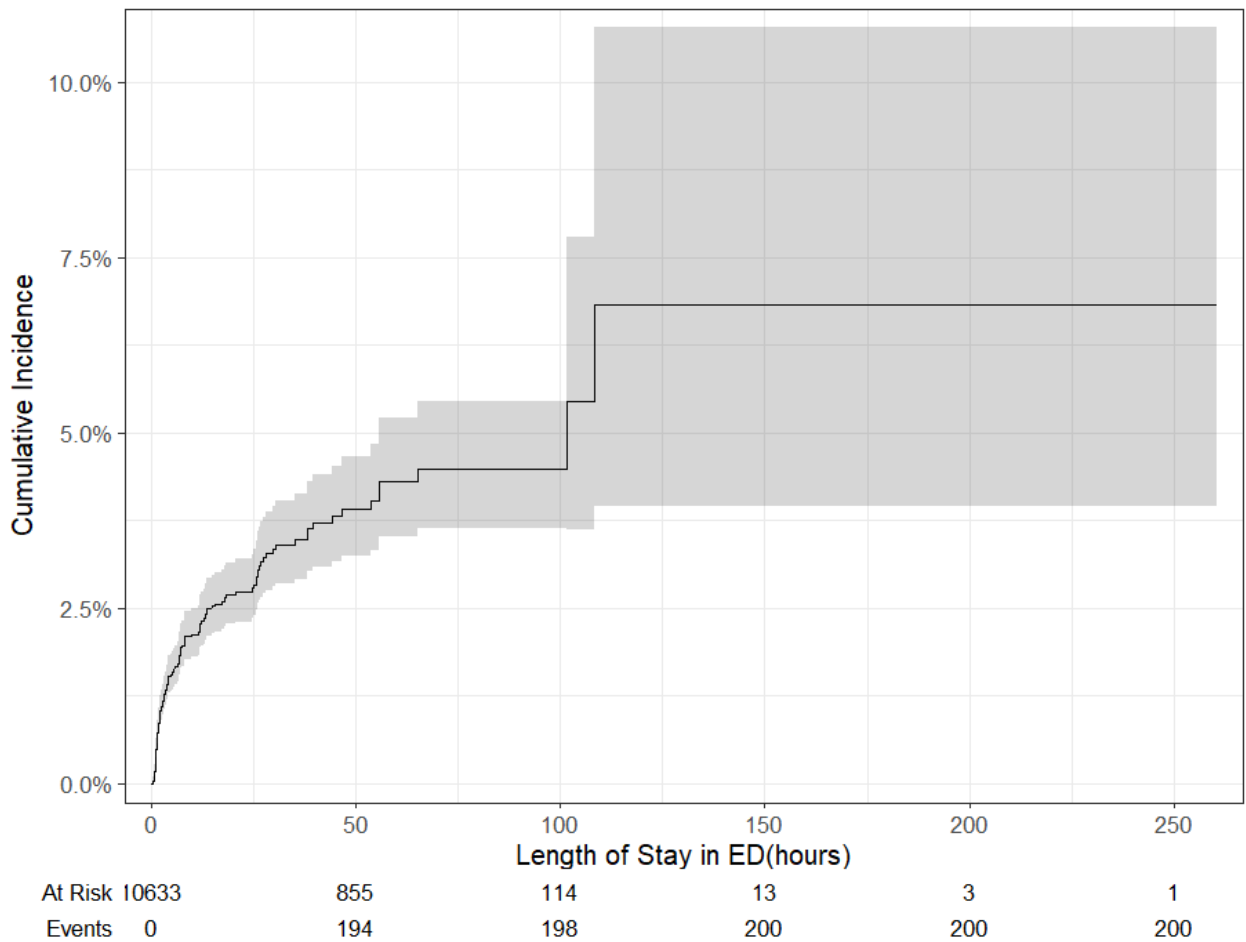


Figure 12 presents the non-parametric CIF estimates for death events (status == 1) in comparison to other outcomes, such as discharge and admission, which are treated as censored (status == 0). The CIF illustrates the cumulative proportion of deaths in the emergency department (ED) over time, measured in hours, while accounting for competing risks. The cumulative incidence of death rises gradually over time, with no sharp increase observed. The CIF does not reach 50%, reflecting that less than half of the patients experience death as an outcome. At-risk patients decline steadily from the initial cohort of 10,633 to only 194 by 50 hours, 114 by 100 hours, and minimal numbers at later time points. This progressive decline corresponds to the occurrence of competing events, such as discharge or admission, which reduce the at-risk population.

Figure 13 Stratified CIF for Discharge Events by 5-level TPTS

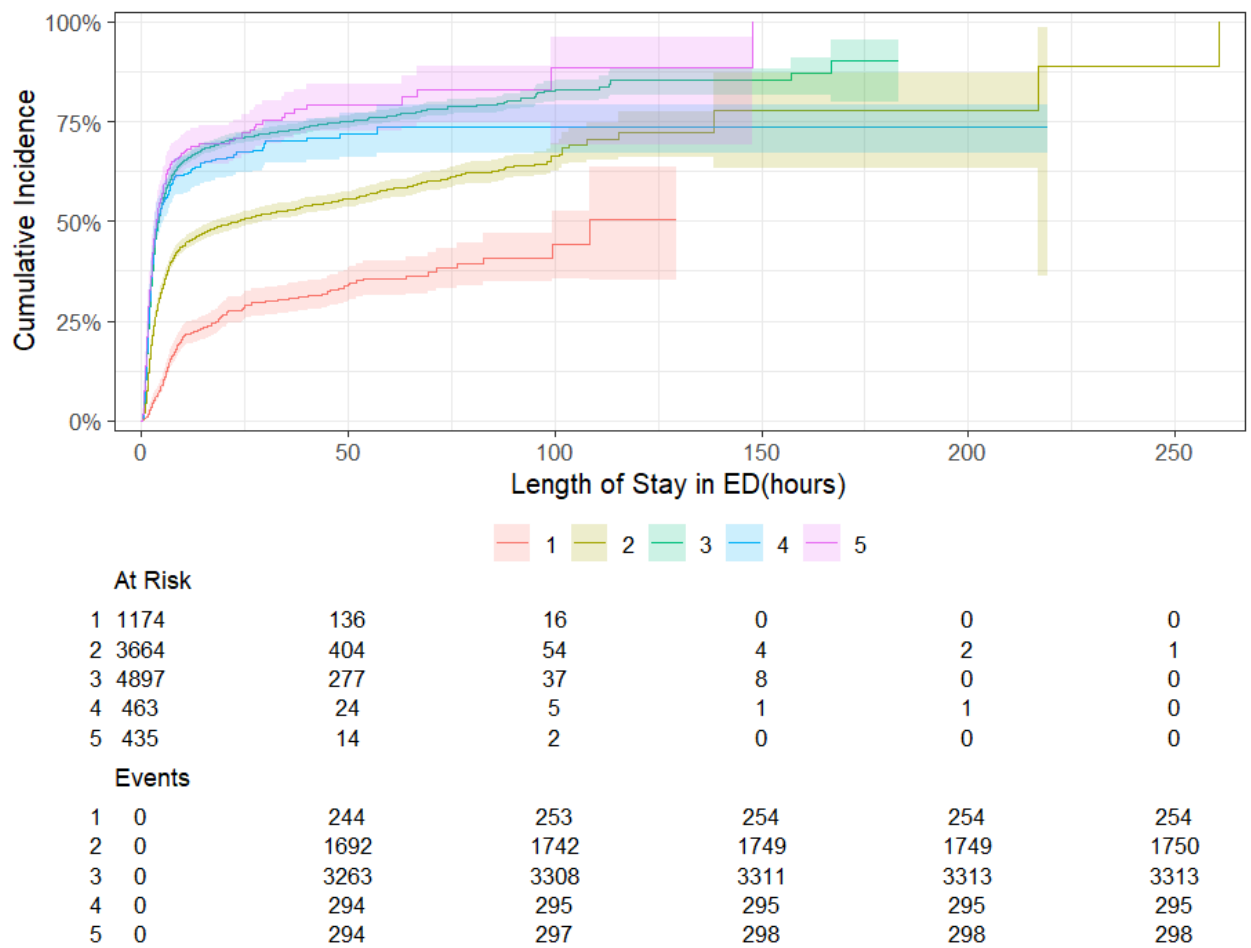


Figure 13 illustrates the 5-level TPTS's strong discrimination ability for discharge events during the early LOS (<6 hours), with lower-severity levels (e.g., Level 5) showing rapid discharge rates and higher-severity levels (e.g., Level 1) lagging significantly behind. This distinction persists moderately in the 6–24 hour interval but diminishes for LOS 24–48 hours, where Levels 3–5 show overlapping cumulative incidences. Beyond 48 hours, the curves for all levels converge, indicating minimal discrimination. Overall, 5-level TPTS effectively stratifies patients for early discharges but shows weaker differentiation for longer ED stays, particularly among lower-severity levels.

Figure 14 Stratified CIF for Admission Events by 5-level TPTS

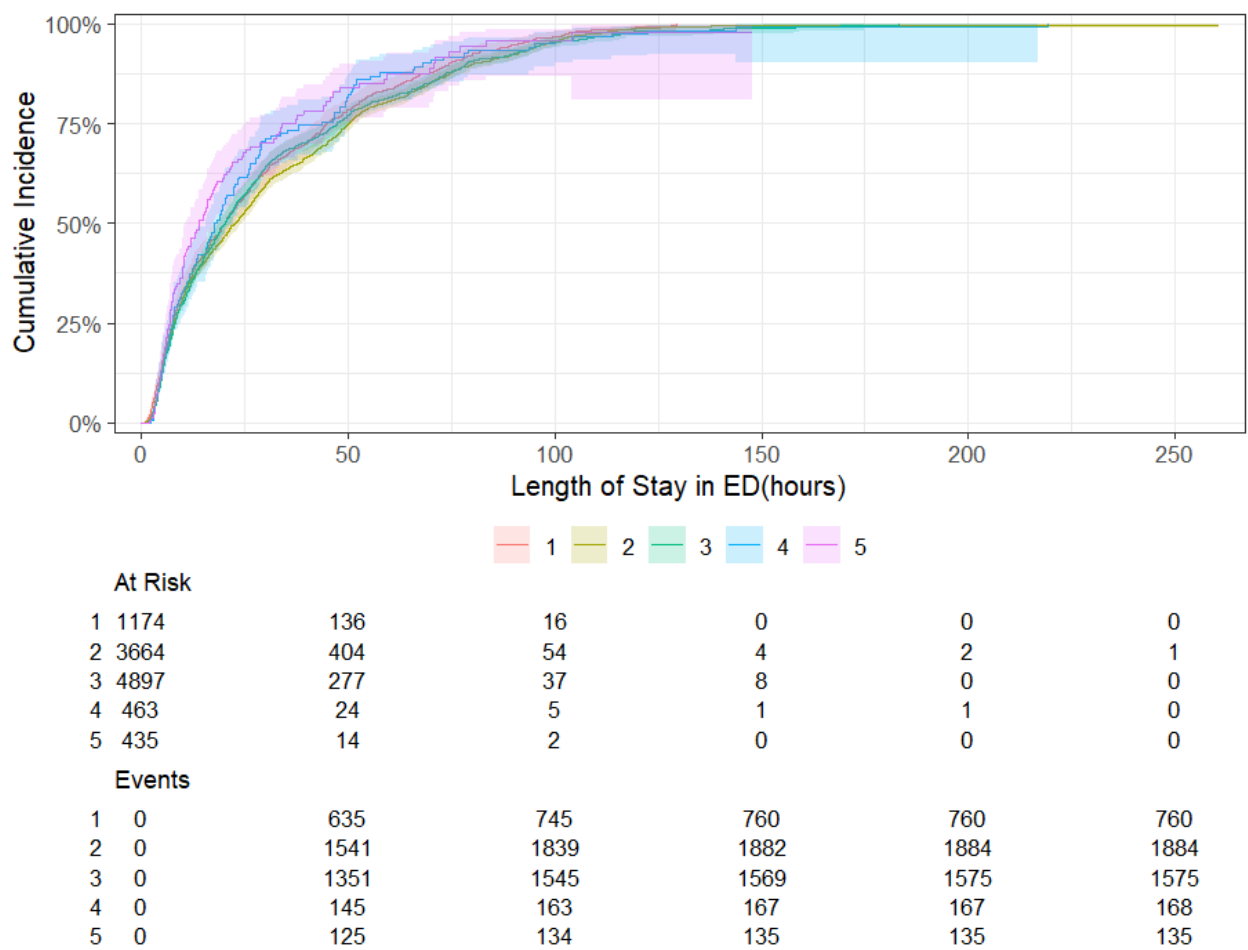


Figure 14 demonstrates the 5-level TPTS's overall discrimination ability for admission events across different LOS intervals. For LOS <6 hours, Levels 1–3 overlap significantly, showing similar rapid admission rates, while Levels 4 and 5 progress faster, with Level 5 achieving the highest cumulative incidence. Between 6–24 hours, Levels 4 and 5 maintain faster trajectories, clearly separating from Levels 1–3, which continue to overlap. During the 24–48 hour interval, the distinction between Levels 4 and 5 and the other levels narrows, and beyond 48 hours, all levels converge, reflecting minimal discrimination. Overall, the 5-level TPTS effectively differentiates Levels 4 and 5 in the early and moderate LOS but struggles to separate Levels 1–3.

Figure 15 Stratified CIF for Death Events by 5-level TPTS

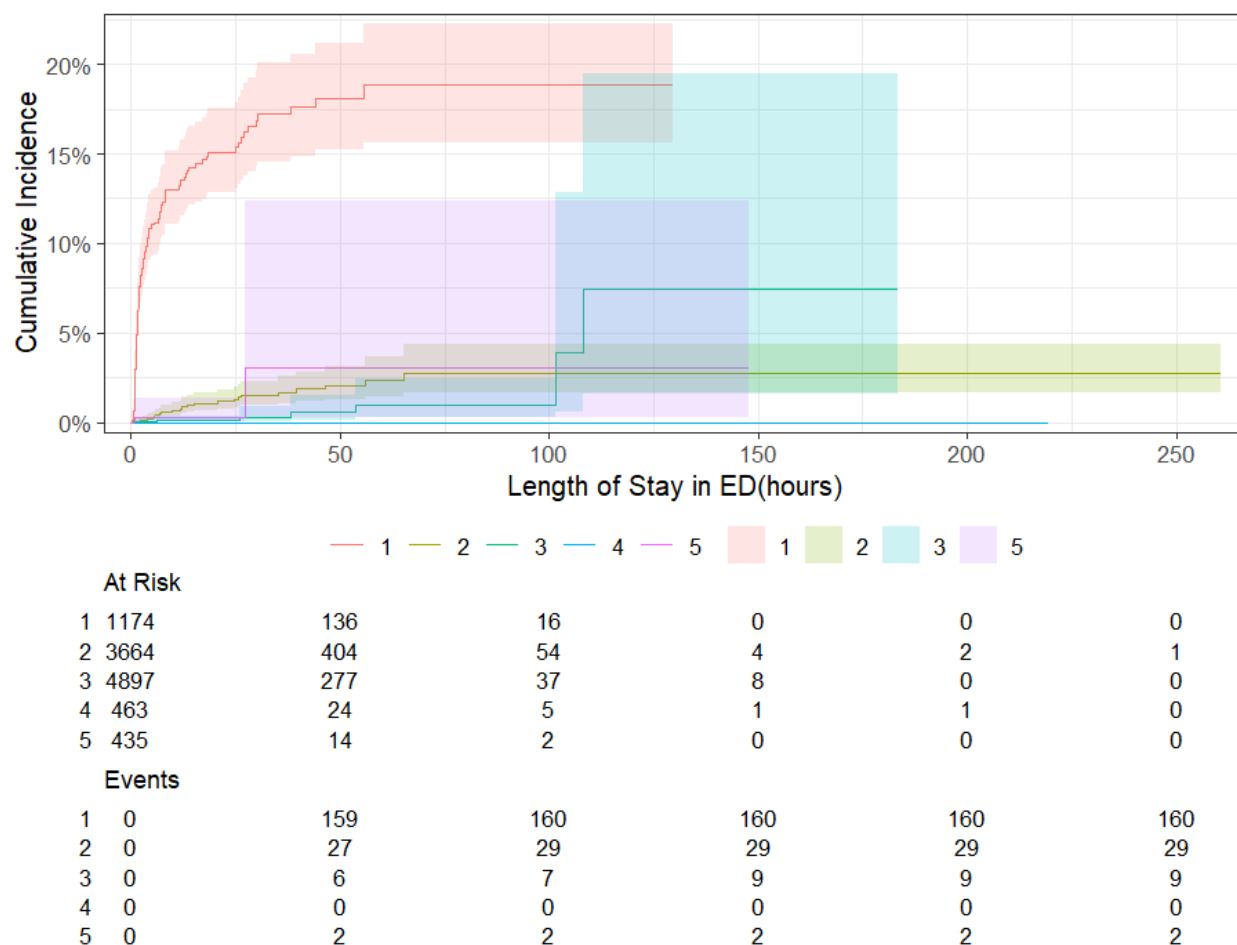


Figure 15 demonstrates the 5-level TPTS's discrimination ability for death events across different LOS intervals. For LOS <6 hours, Level 1 clearly separates from the other levels, with rapid progression to nearly 15% cumulative incidence, while Levels 2–5 show minimal events. Between 6–24 hours, Level 1 continues to dominate with the highest cumulative incidence, though Levels 2–3 begin to diverge slightly from Levels 4–5. During the 24–48 hour interval, Levels 2 and 3 show gradual progression but overlap significantly, while Levels 4 and 5 remain relatively flat. Beyond 48 hours, the cumulative incidence for Levels 1–5 stabilizes, with minimal additional events and decreasing separation. Overall, 5-level TPTS effectively distinguishes Level 1 from other levels in the early LOS but shows weaker stratification among Levels 2–5 over time.



Table 13 Subgroup Analysis of Age groups and Reason for calling 119

Outcome	Age Groups			Reason for calling 119	
	<18 n=847	18-65 n=5890	>65 n=3896	Non-trauma n=6698	Trauma n=3935
5-level TPTS					
Triage level					
Level 1	132(15.58)	446(7.57)	596(15.30)	1022(15.26)	152(3.86)
Level 2	355(41.91)	1854(31.48)	1455(37.35)	2632(39.30)	1032(26.23)
Level 3	294(34.71)	3030(51.44)	1573(40.37)	2581(38.53)	2316(58.86)
Level 4	39(4.60)	272(4.62)	152(3.90)	268(4.00)	195(4.96)
Level 5	27(3.19)	288(4.89)	120(3.08)	195(2.91)	240(6.10)
ED LOS(hr)^a	3.82 (2.02-7.64)	3.87 (1.92-10.30)	7.80 (3.40-29.02)	7.40 (3.00-26.08)	3.11 (1.68-5.70)
ED LOS group(hr)					
<6	535(63.16)	3735(63.41)	1681(43.15)	2934(43.80)	3017(76.67)
6-24	288(34.00)	1419(24.09)	1061(27.23)	1978(29.53)	790(20.08)
24-48	22(2.60)	390(6.62)	572(14.68)	897(13.39)	87(2.21)
>48	2(0.24)	346(5.87)	582(14.94)	889(13.27)	41(1.04)
ESI Resources					
Many	637(75.21)	5092(86.45)	3729(95.71)	6221(92.88)	3237(82.26)
One	143(16.88)	634(10.76)	132(3.39)	372(5.55)	537(13.65)
None	67(7.91)	164(2.78)	35(0.90)	105(1.57)	161(4.09)

A subgroup analysis was performed to evaluate the distribution of the five-level TPTS, ED

LOS, and required ESI resources across different age groups and reasons for calling 119. Among

patients aged <18 years (n=847), the majority were triaged as Level 2 (41.91%), followed by Level 3 (34.71%). In contrast, patients aged 18–65 years (n=5,890) predominantly fell under Level 3 (51.44%), with similar trends observed in those aged >65 years (n=3,896), where 40.37% were triaged as Level 3. When comparing non-trauma cases (n=6,698) to trauma cases (n=3,935), Level 2 triage was most common in the non-trauma group (39.30%), while trauma cases were more frequently categorized as Level 3 (58.86%).

The median ED LOS varied significantly between groups. Patients aged >65 years experienced the longest LOS, with a median of 7.80 hours (interquartile range [IQR], 3.40–29.02 hours), compared to 3.82 hours (IQR, 2.02–7.64 hours) for those aged <18 years. Non-trauma cases had a median LOS of 7.40 hours (IQR, 3.00–26.08 hours), which was notably longer than the 3.11 hours (IQR, 1.68–5.70 hours) observed in trauma cases. Additionally, most patients aged <18 years (63.16%) and 18–65 years (63.41%) had an ED LOS of <6 hours, whereas only 43.15% of those aged >65 years fell into this category.

Finally, analysis of ESI resource utilization revealed that older patients (>65 years) required significantly more resources, with 95.71% categorized as needing multiple resources. Non-trauma cases also demonstrated a higher proportion of multi-resource use (92.88%) compared to trauma cases (82.26%).

Figure 16 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Age Groups: <18 Years

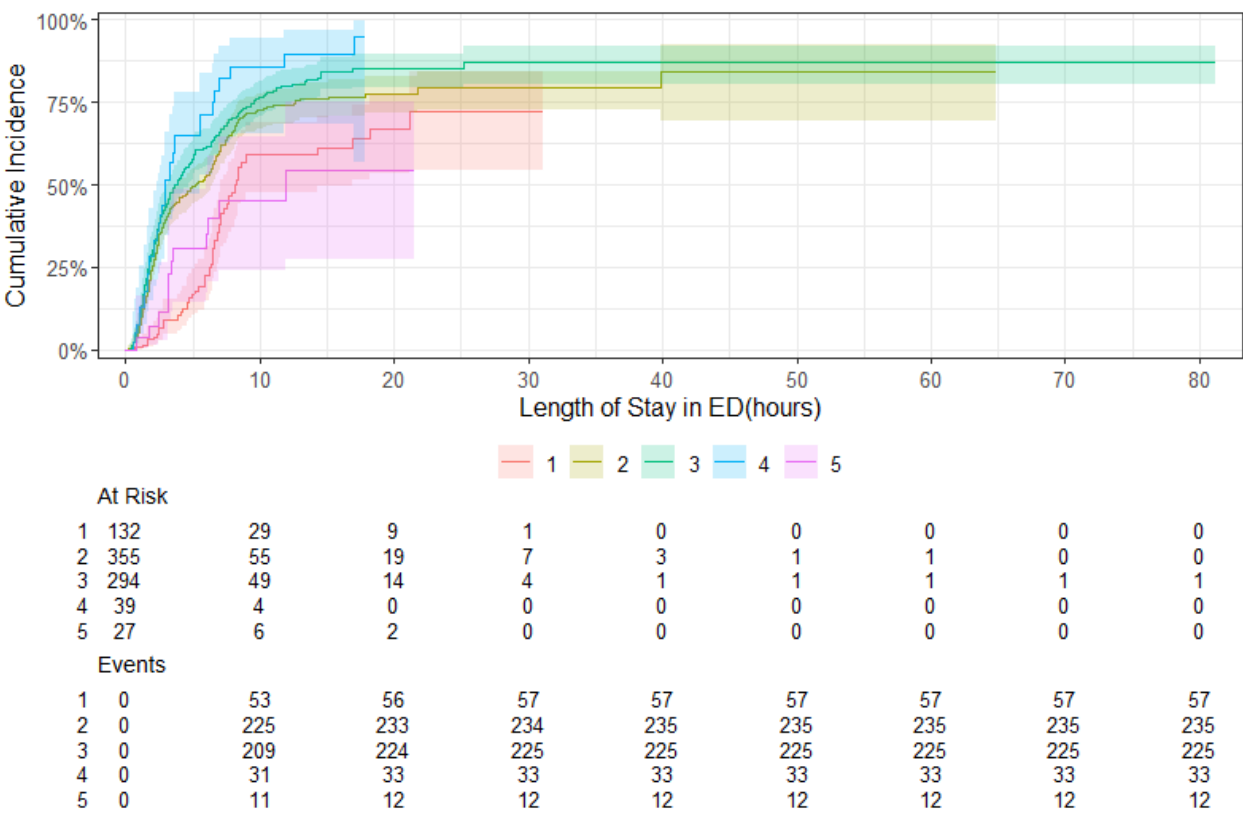


Figure 16 illustrates the 5-level TPTS's discrimination ability for discharge events in patients under 18 years of age across different LOS intervals. For LOS <6 hours, Level 1 shows the slowest progression, with minimal discharges, while Level 5 demonstrates rapid cumulative incidences, reaching near 50%, clearly separating it from the other levels. Between 6–24 hours, Level 5 continues to progress faster and maintains separation, while Levels 2–4 gradually converge with moderate discharge rates. During the 24–48 hour interval, Levels 2–4 show overlapping trajectories, while Level 1 remains distinct with the slowest progression. Beyond 48 hours, the cumulative incidences for all levels converge, reflecting limited discrimination. Overall, the 5-level TPTS effectively distinguishes Level 1 (highest severity) and Level 5 (lowest severity) in early LOS but shows reduced differentiation among intermediate levels over time.

Figure 17 Stratified CIF Analysis of Admission Events by 5-level TPTS in Age Groups: <18 Years

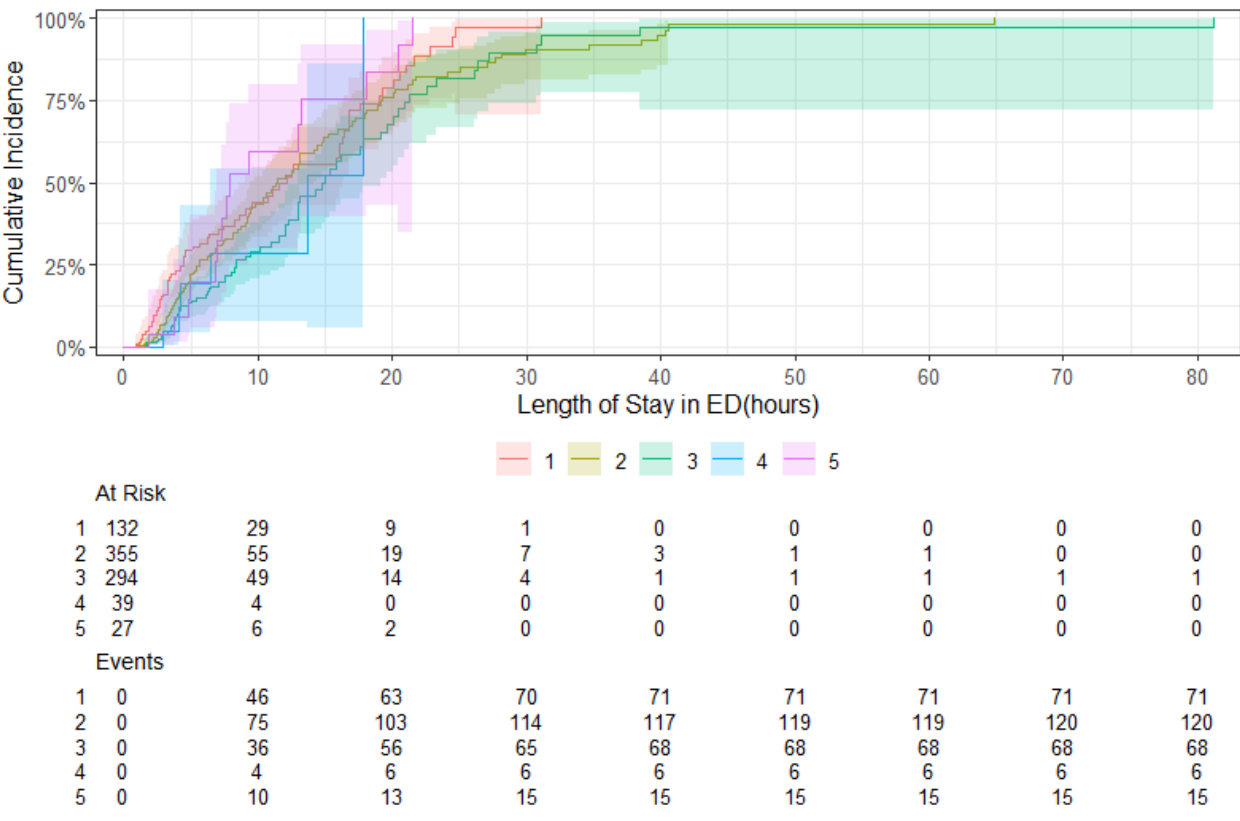


Figure 17 illustrates the 5-level TPTS’s discrimination ability for admission events in patients under 18 years of age across different LOS intervals. For LOS <6 hours, Level 1 shows the highest cumulative incidence, with rapid admission rates, while Level 3 demonstrates the lowest progression among all levels. After 6 hours, the curves for most levels, including Levels 1, 2, and 3, start to overlap significantly, reducing the distinction between them. Levels 4 and 5 maintain slight separation but converge gradually with the other levels. By the 24–48 hour and >48 hour intervals, cumulative incidences for all levels align closely, showing limited stratification. Overall, 5-level TPTS effectively distinguishes Level 1 in the first 6 hours but demonstrates weaker discrimination for Levels 2–5 over longer LOS intervals.

Figure 18 Stratified CIF Analysis of Death Events by 5-level TPTS in Age Groups: <18 Years

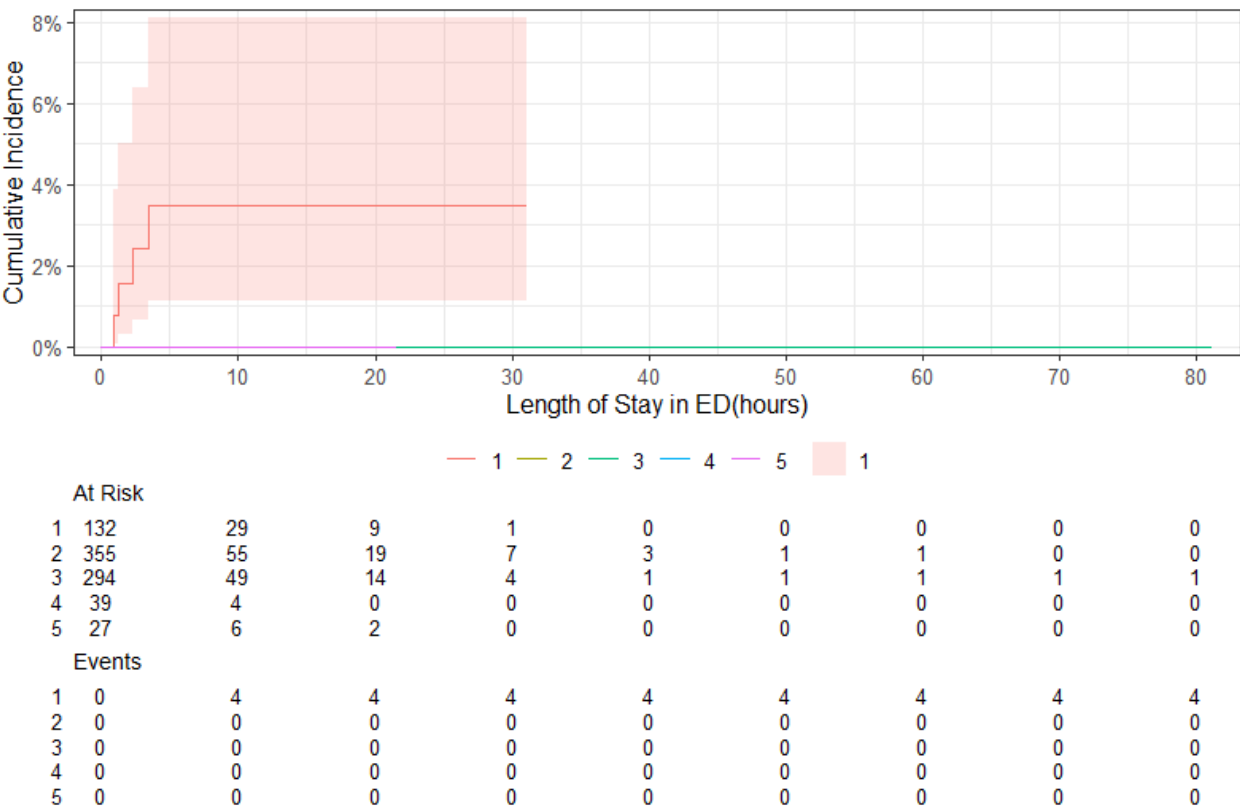


Figure 18 illustrates the 5-level TPTS’s discrimination ability for death events in patients under 18 years of age across different LOS intervals. For LOS <6 hours, Level 1 shows a distinct trend, with a cumulative incidence reaching approximately 4%, while no events occur for Levels 2–5, resulting in flat cumulative incidence curves. Between 6–24 hours, as well as during the 24–48 hour and >48 hour intervals, the absence of events for Levels 2–5 continues, leaving only the trajectory of Level 1 to be observed. Overall, this analysis highlights that death events are concentrated exclusively in Level 1, limiting the ability to evaluate stratification for the other levels due to the lack of events.

Figure 19 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Age Groups: 18-65 Years

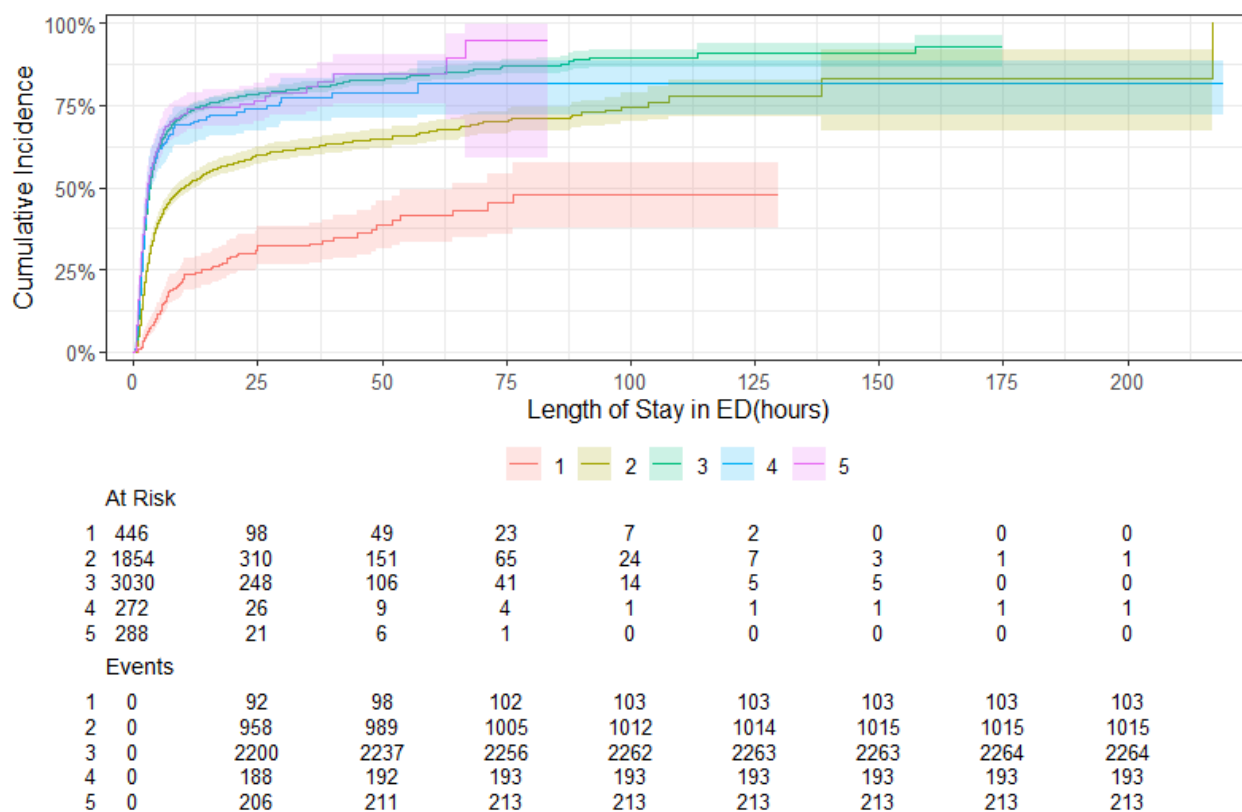


Figure 19 demonstrates the 5-level TPTS's discrimination ability for discharge events in patients aged 18–65 years across different LOS intervals. For LOS <6 hours, Level 5 shows the fastest cumulative incidence, indicating rapid discharges, while Level 1 progresses much slower, with distinct separation. Between 6–24 hours, Levels 4 and 5 maintain faster discharge rates, while Levels 1–3 converge and show overlapping trends. During the 24–48 hour interval, the separation between Levels 4 and 5 and the other levels diminishes as the curves for Levels 1–3 remain closely aligned. Beyond 48 hours, cumulative incidences for all levels approach similar values, reflecting minimal discrimination. Overall, the TPTS effectively distinguishes Level 1 and Level 5 in the early LOS but struggles to maintain clear stratification for intermediate levels (2–4) over longer stays.

Figure 20 Stratified CIF Analysis of Admission Events by 5-level TPTS in Age Groups: 18-65 Years

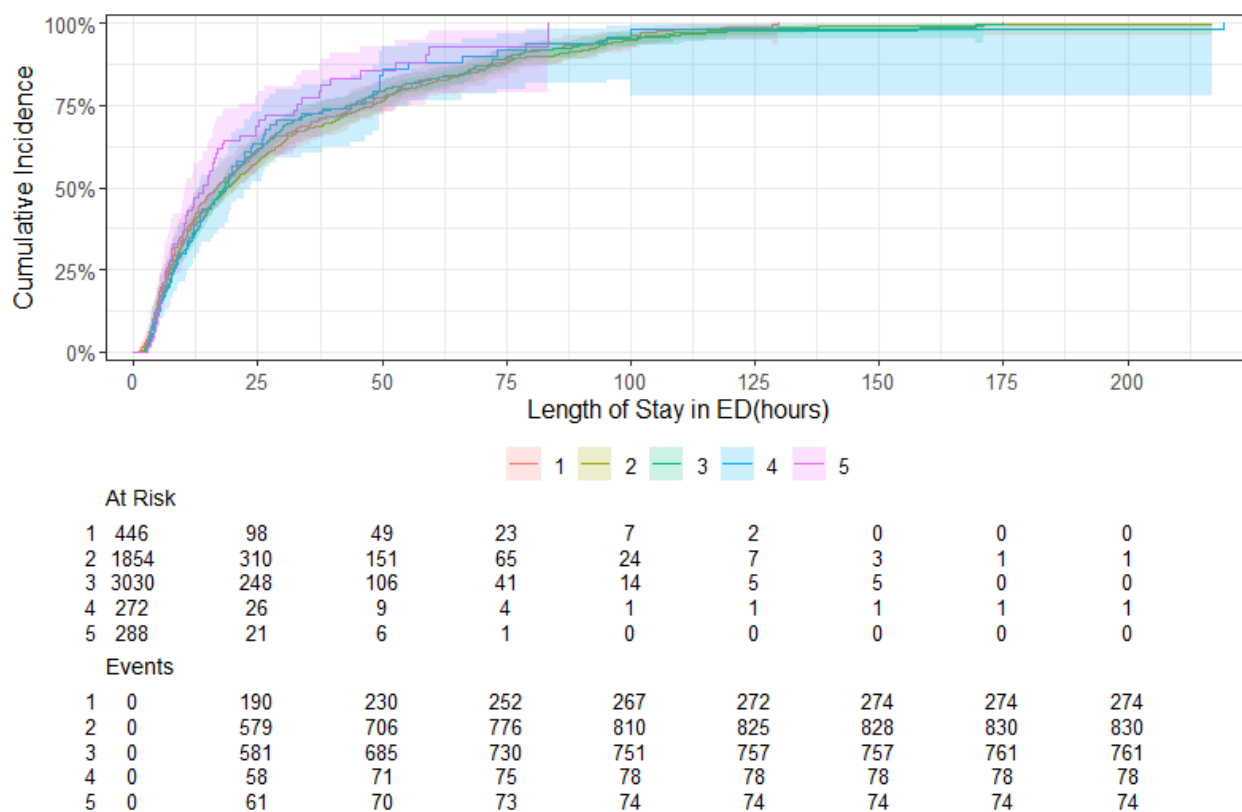


Figure 20 illustrates the 5-level TPTS's ability to stratify admission events in patients aged 18–65 years across LOS intervals. While Level 5 shows distinct separation with the fastest admission rates, Levels 1–4 overlap significantly throughout all intervals, indicating minimal differentiation among them. For LOS <6 hours, Level 5 achieves a high cumulative incidence, but Levels 1–4 progress more slowly with overlapping trends. Between 6–24 hours, Level 5 maintains separation, while Levels 1–4 converge further, reducing distinction. During the 24–48 hour interval and beyond 48 hours, cumulative incidences for all levels align closely. This suggests that, apart from Level 5, the 5-level TPTS lacks strong discrimination ability for admission events in this age group, as it fails to clearly stratify Levels 1–4.

Figure 21 Stratified CIF Analysis of Death Events by 5-level TPTS in Age Groups: 18-65 Years

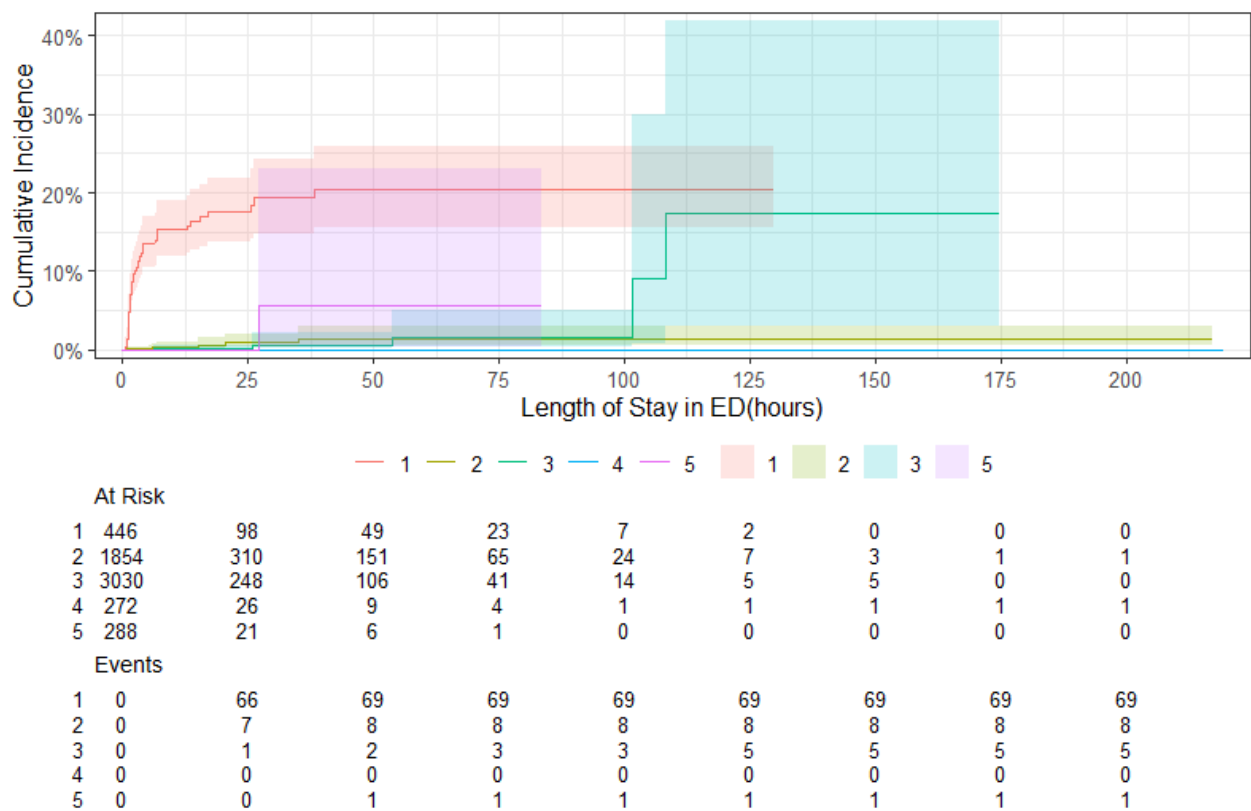


Figure 21 illustrates the 5-level TPTS’s ability to stratify death events in patients aged 18–65 years across LOS intervals. For LOS <6 hours, Level 1 demonstrates the highest cumulative incidence, distinctly separating from other levels, which show minimal or no progression. Between 6–24 hours, Level 1 continues to rise steadily, while Levels 2–5 remain nearly flat, indicating limited events. During the 24–48 hour interval and beyond 48 hours, Level 1 stabilizes, and the cumulative incidence for the other levels shows no substantial change. These patterns suggest that death events are primarily concentrated in Level 1, with 5-level TPTS showing minimal stratification ability for Levels 2–5 across all LOS intervals due to the lack of events in these groups.



Figure 22 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Age Groups: >65 Years

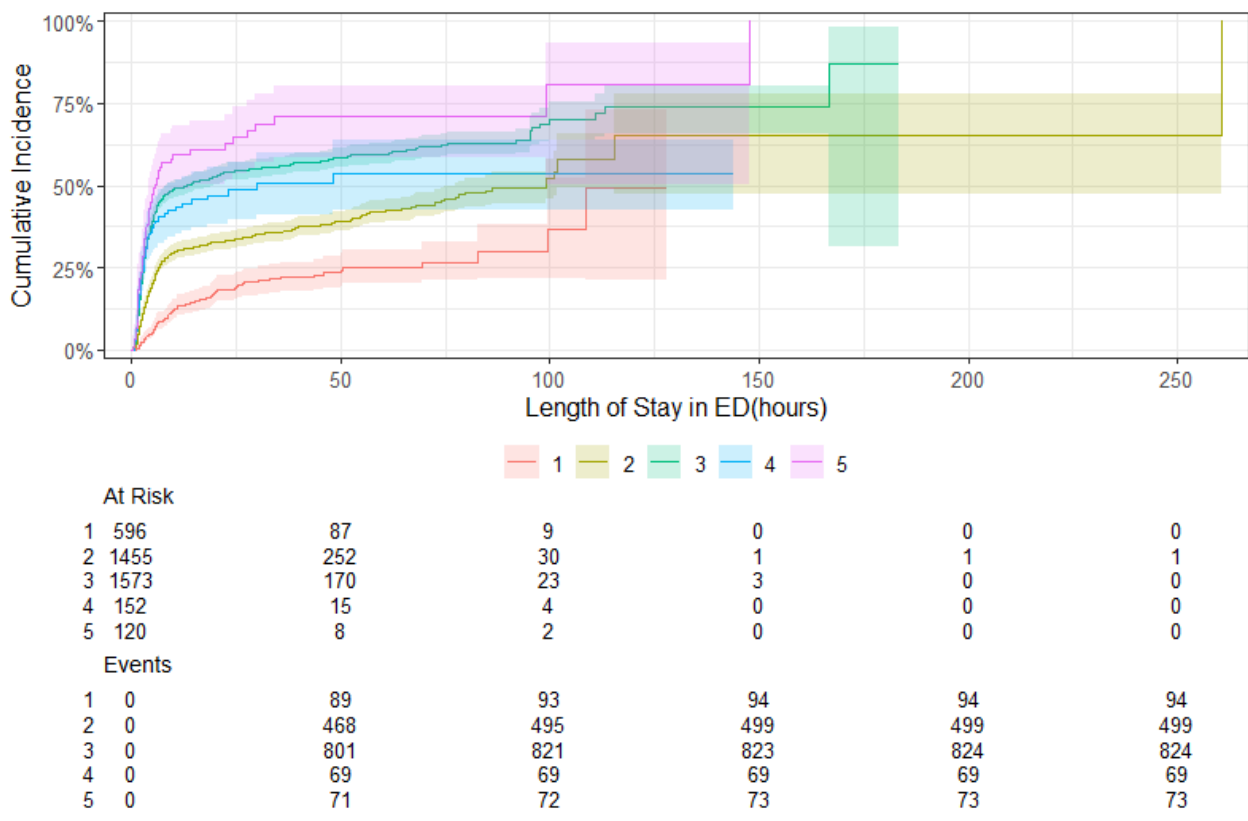


Figure 22 demonstrates the 5-level TPTS's discrimination ability for discharge events in patients aged >65 years across different LOS intervals. For LOS <6 hours, Levels 1 and 2 show slower progression compared to Levels 3–5, which exhibit faster cumulative incidences. This stratification is maintained during the 6–24 hour interval, where Levels 1 and 2 remain distinct from Levels 3–5. During the 24–48 hour interval, the separation among Levels 1, 2, and Levels 3–5 continues to be clear, demonstrating effective stratification. Beyond 48 hours, however, cumulative incidences for all levels converge, indicating minimal differentiation. Overall, 5-level TPTS shows strong discrimination within the first 48 hours, effectively distinguishing Levels 1 and 2 from Levels 3–5, but loses stratification ability over longer LOS intervals.

Figure 23 Stratified CIF Analysis of Admission Events by 5-level TPTS in Age Groups: >65 Years

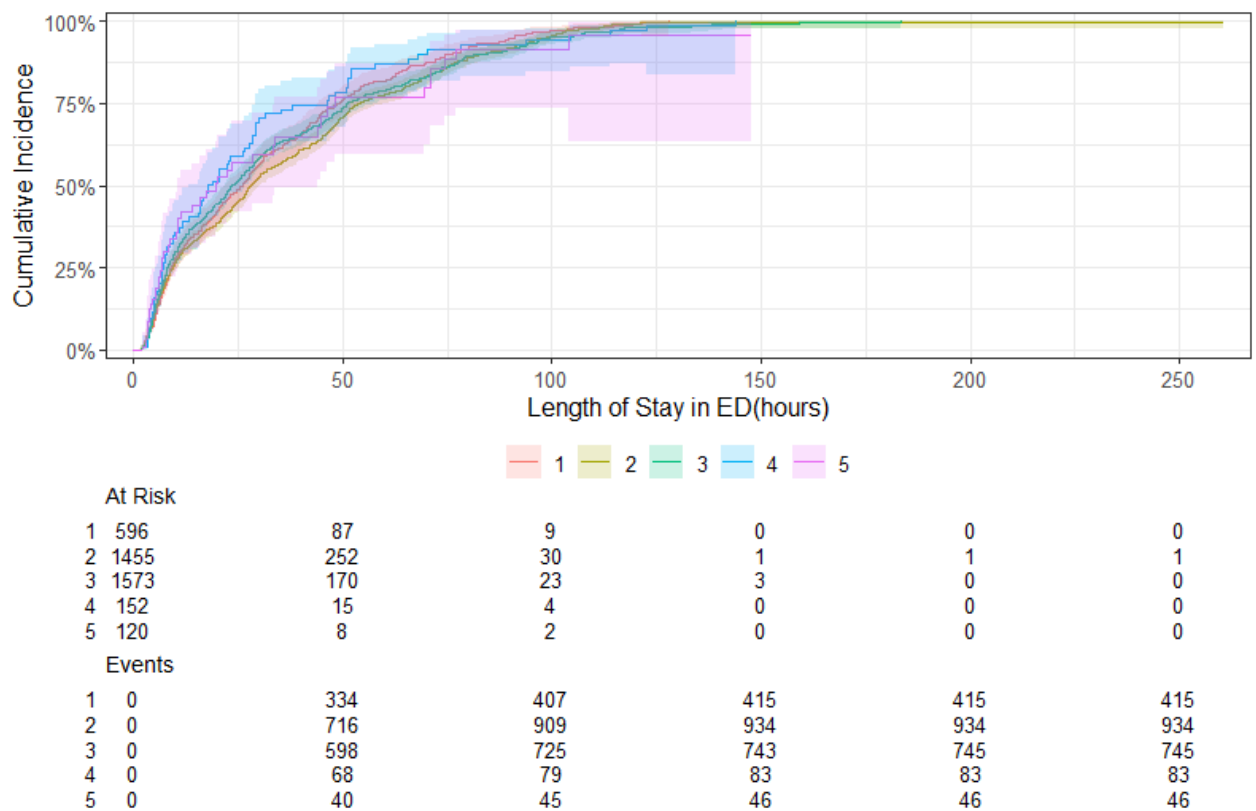


Figure 23 demonstrates the 5-level TPTS's discrimination ability for admission events in patients aged >65 years across LOS intervals. Throughout all intervals—LOS <6 hours, 6–24 hours, 24–48 hours, and beyond 48 hours—the cumulative incidence curves for Levels 1–5 overlap significantly, showing no clear separation between the levels. This lack of stratification suggests that 5-level TPTS does not effectively distinguish admission likelihood among different severity levels in this age group.

Figure 24 Stratified CIF Analysis of Death Events by 5-level TPTS in Age Groups: >65 Years

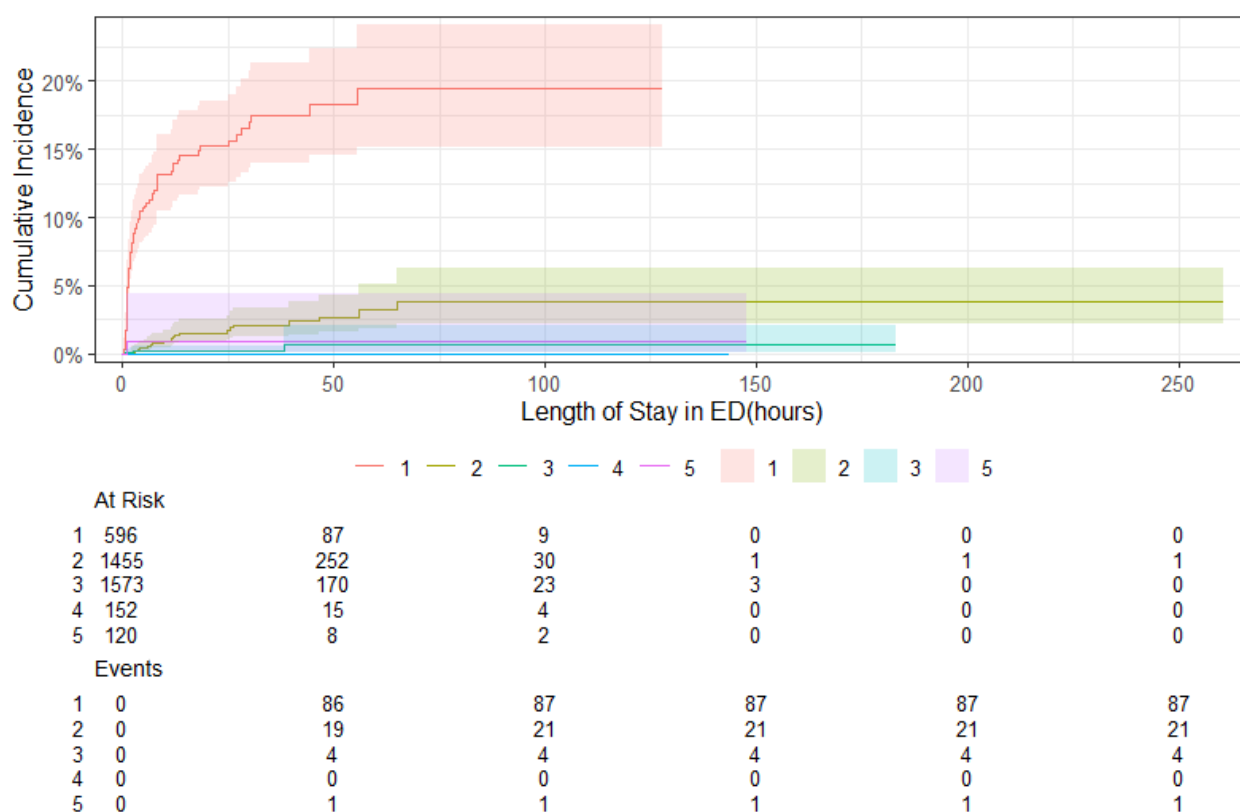


Figure 24 illustrates the 5-level TPTS's discrimination ability for death events in patients aged >65 years across LOS intervals. For LOS <6 hours, Level 1 shows the highest cumulative incidence with a steep progression, distinctly separating from the other levels. In contrast, Levels 2–5 demonstrate minimal or no events, with flat or near-flat curves. Between 6–24 hours, the trend remains consistent, with Level 1 continuing to progress while Levels 2–5 remain nearly unchanged. During the 24–48 hour interval and beyond 48 hours, Level 1 stabilizes, and the other levels show little to no progression, indicating no additional events. Overall, death events are concentrated almost exclusively in Level 1, with 5-level TPTS showing minimal ability to stratify Levels 2–5 due to the lack of events in these groups.

Figure 25 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Trauma Cases

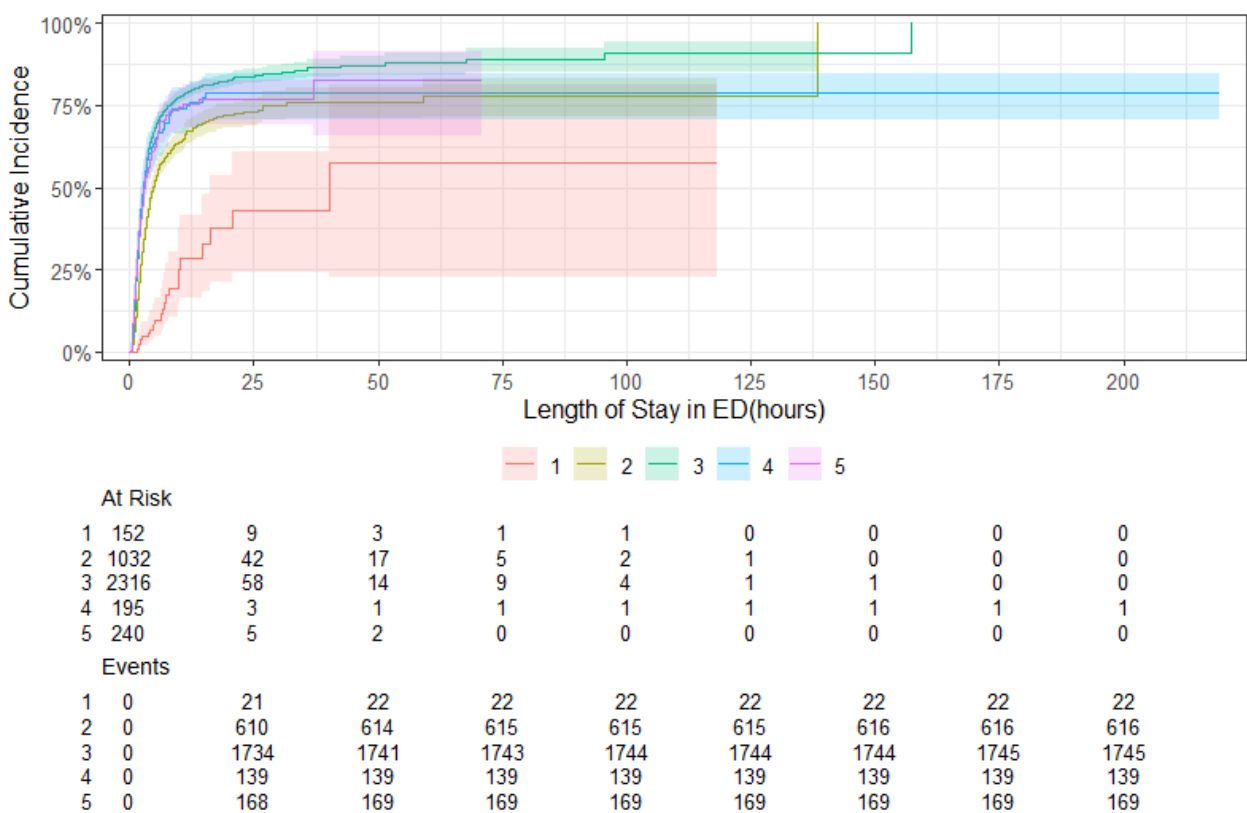


Figure 25 illustrates the 5-level TPTS’s discrimination ability for discharge events in trauma cases across LOS intervals. For LOS <6 hours, Levels 3–5 exhibit the fastest discharge rates, followed by Level 2, while Level 1 progresses the slowest, showing clear separation. Between 6–24 hours, Levels 3–5 overlap significantly, but Levels 1 and 2 remain distinct, with Level 1 progressing slower than Level 2. After 24 hours, Levels 2–5 converge and overlap, while Level 1 remains clearly distinct with slower progression. By 48 hours, Level 1 continues to demonstrate clear separation, while the other levels show little stratification. Overall, 5-level TPTS effectively distinguishes Level 1 throughout all intervals, particularly beyond 24 hours, but struggles to differentiate among Levels 2–5 as LOS increases.

Figure 26 Stratified CIF Analysis of Admission Events by 5-level TPTS in Trauma Cases

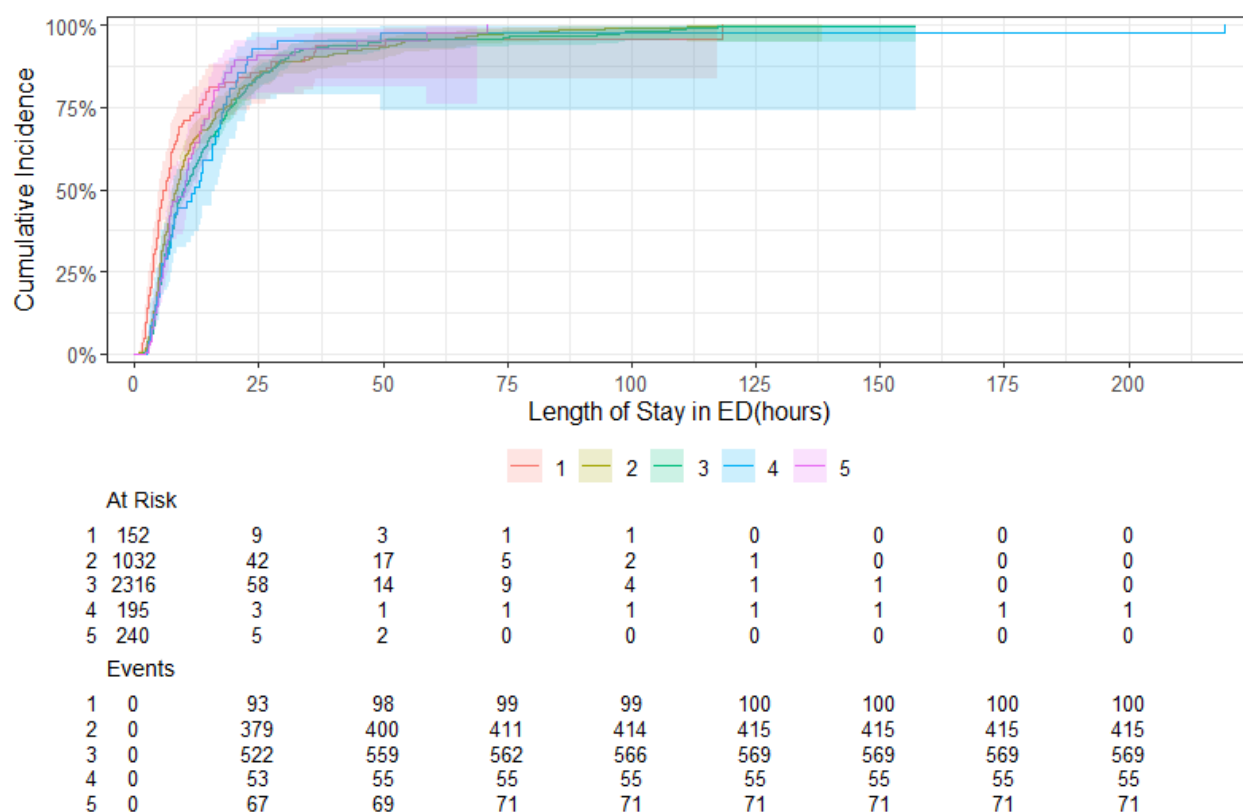


Figure 26 illustrates the 5-level TPTS's discrimination ability for admission events in trauma cases across LOS intervals. For LOS <6 hours, Level 1 shows the fastest progression, with a steep cumulative incidence curve, clearly separating from the other levels. In contrast, Levels 2–5 overlap significantly, showing similar slower admission rates. Between 6–24 hours, this pattern persists, with Level 1 maintaining its distinction, while Levels 2–5 remain closely aligned. After 24 hours and beyond 48 hours, cumulative incidences for Levels 2–5 continue to overlap, with minimal separation, while Level 1 remains distinct throughout. Overall, 5-level TPTS effectively distinguishes Level 1 but fails to differentiate among Levels 2–5.

Figure 27 Stratified CIF Analysis of Death Events by 5-level TPTS in Trauma Cases



Figure 27 illustrates the 5-level TPTS's discrimination ability for death events in trauma cases across LOS intervals. For LOS <6 hours, Level 1 shows the highest cumulative incidence, with a steep rise in deaths, clearly separating it from Levels 2–5, which show minimal or no events. Between 6–24 hours, Level 1 continues to dominate with significantly higher cumulative incidence, while Levels 2–5 remain flat with no additional events. During the 24–48 hour interval and beyond 48 hours, Level 1 stabilizes, and the other levels continue to show no progression. Overall, death events are almost exclusively concentrated in Level 1, with 5-level TPTS showing minimal stratification for Levels 2–5 due to the lack of events in these groups.

Figure 28 Stratified CIF Analysis of Discharge Events by 5-level TPTS in Non-trauma Cases

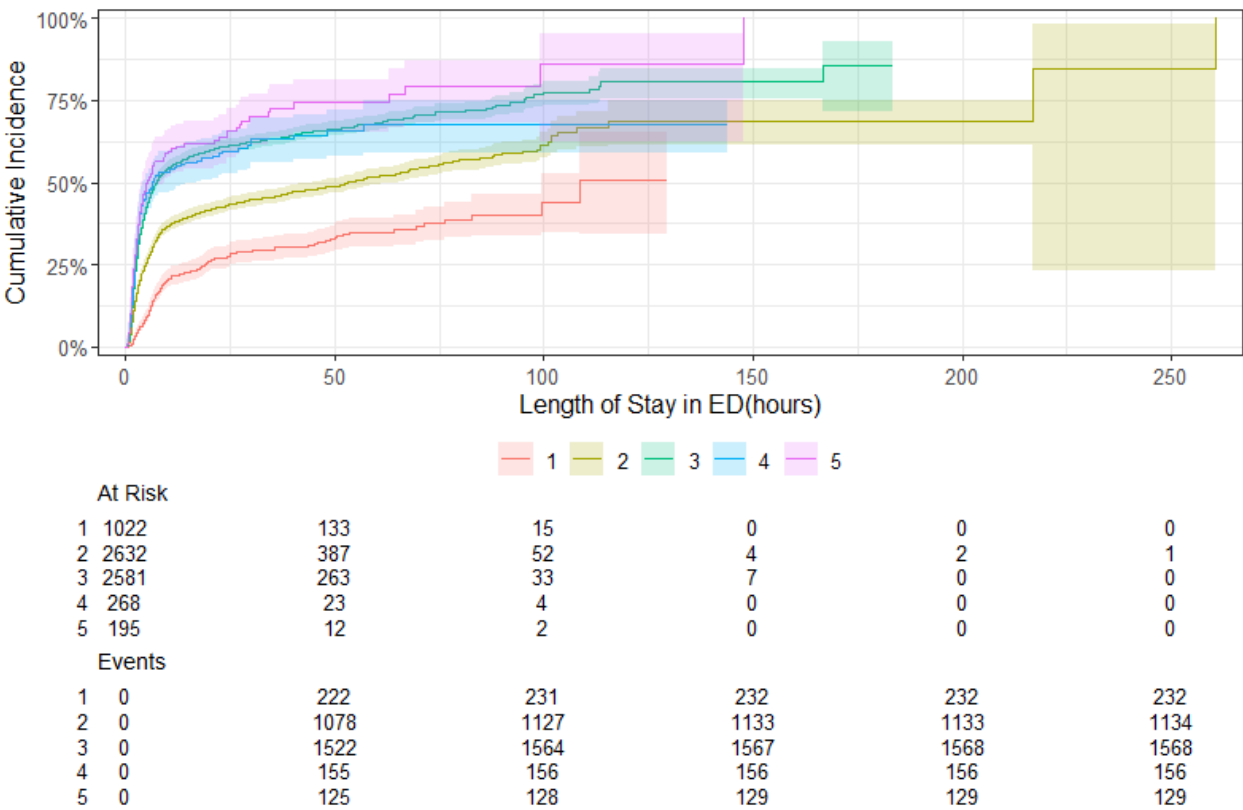


Figure 28 illustrates the 5-level TPTS’s discrimination ability for discharge events in non-trauma cases across LOS intervals. For LOS <6 hours, Levels 1, 2, and 5 show distinct separation, with Level 5 demonstrating the fastest discharge rates and Level 1 the slowest progression. Level 2 progresses at a moderate pace, remaining distinct from Levels 3 and 4, which overlap significantly. Between 6–24 hours, the clear separation among Levels 1, 2, and 5 persists, while Levels 3 and 4 continue to overlap. During the 24–48 hour interval, Levels 1 and 2 maintain distinction, but the overlap between Levels 3 and 4 becomes more pronounced, and Level 5 stabilizes. Beyond 48 hours, cumulative incidences for all levels begin to converge, indicating minimal stratification. Overall, 5-level TPTS effectively distinguishes Levels 1, 2, and 5 early on but struggles to separate Levels 3 and 4 as LOS increases.

Figure 29 Stratified CIF Analysis of Admission Events by 5-level TPTS in Non-trauma Cases

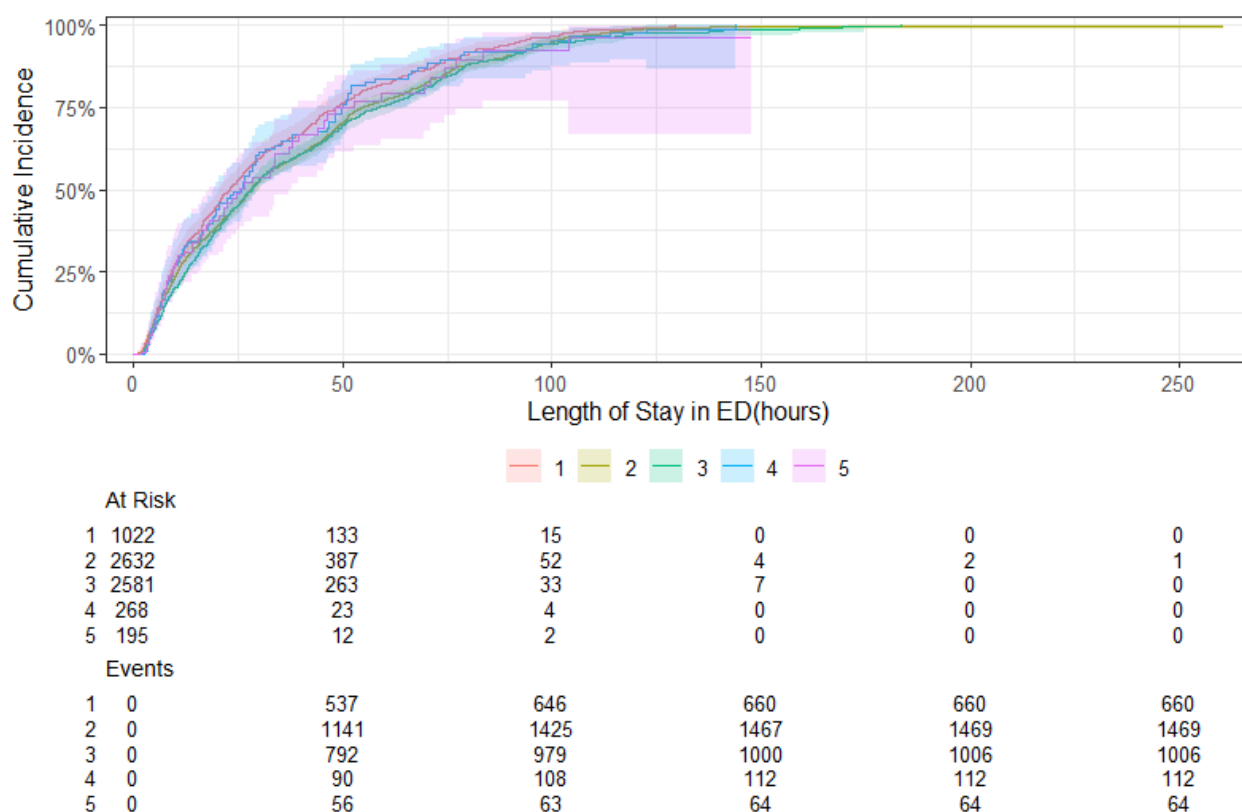


Figure 29 illustrates the 5-level TPTS's discrimination ability for admission events in non-trauma cases across LOS intervals. Throughout all intervals—LOS <6 hours, 6–24 hours, 24–48 hours, and beyond 48 hours—the cumulative incidence curves for Levels 1–5 overlap significantly, with minimal separation observed. This lack of distinction indicates that 5-level TPTS shows weak discrimination ability for stratifying admission events among different severity levels in non-trauma cases.

Figure 30 Stratified CIF Analysis of Death Events by 5-level TPTS in Non-trauma Cases

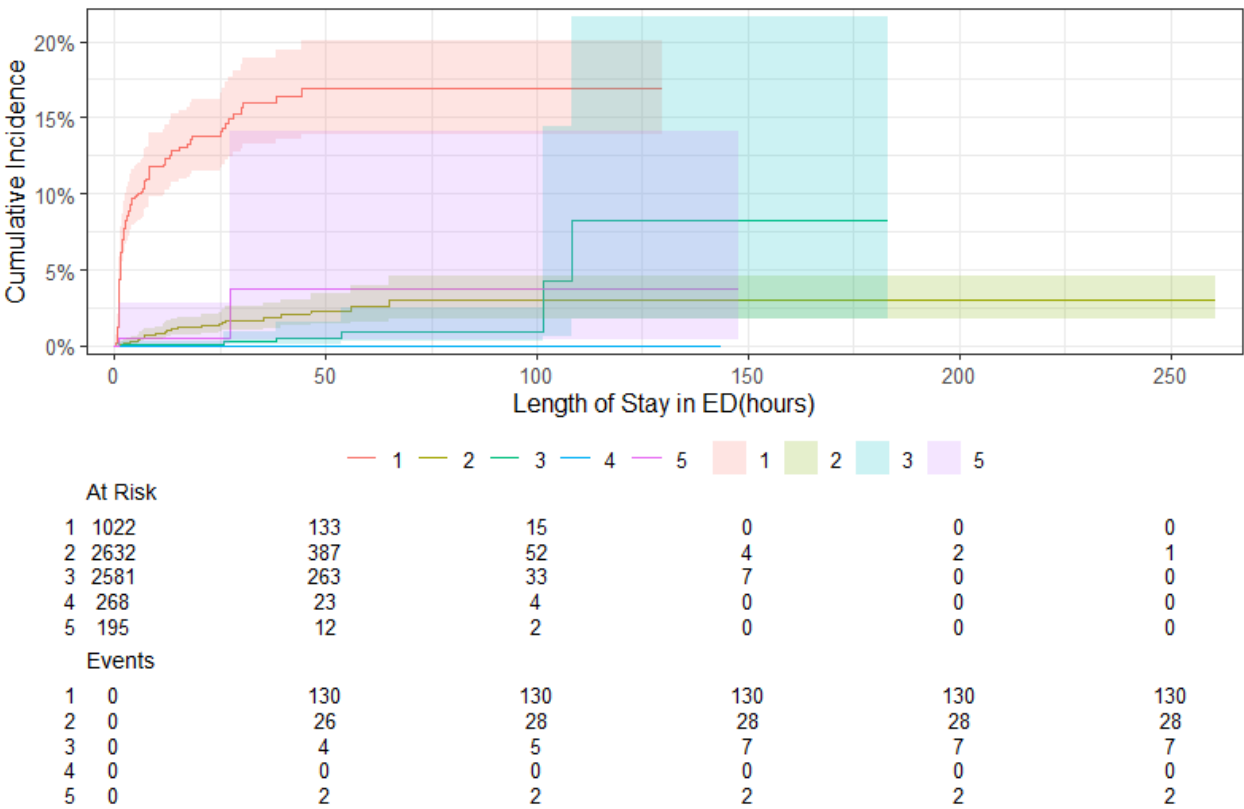


Figure 30 illustrates the 5-level TPTS’s discrimination ability for death events in non-trauma cases across LOS intervals. For LOS <6 hours, Level 1 shows the highest cumulative incidence, clearly separating from Levels 2–5, which remain nearly flat with minimal or no events. Between 6–24 hours, this pattern persists, with Level 1 continuing to rise while Levels 2–5 show no significant progression. During the 24–48 hour interval and beyond 48 hours, Level 1 stabilizes, and Levels 2–5 remain flat, reflecting the absence of additional events. Overall, death events are concentrated exclusively in Level 1, with 5-level TPTS showing minimal stratification ability for Levels 2–5 due to the lack of events.

4.2.5 Proportional Cause-Specific Hazards Models for different events

Proportional Cause-Specific Hazards Model

$$h_j(t|X) = h_{j0}(t)\exp(X\beta_j)$$

$h_j(t|X)$: The cause-specific hazard for the j-th event at time t, given covariates X.

$h_{j0}(t)$: The baseline cause-specific hazard for the j-th event at time t.

X : A vector of covariates.

β_j : A vector of regression coefficients specific to the j-th event.

$\exp(X\beta_j)$: The relative risk for the j-th event as a function of covariates X.

Table 14 Multivariable Analysis of Cause-Specific Hazard Models for Discharge Events
Model 1.-Discharge Events with Others Treated as Censored

Variables	Coeff.	se(coef)	HR (95% CI)	p-value
Gender				
Female	Reference			
Male	-0.23	0.03	0.79(0.75-0.83)	<.0001*
TPTS Levels				
1	-1.24	0.07	0.29(0.25-0.33)	<.0001*
2	-0.48	0.03	0.62(0.58-0.66)	<.0001*
3	Reference			
4	-0.02	0.06	0.98(0.87-1.10)	0.68
5	0.01	0.06	1.00(0.89-1.13)	0.93
Age group				
<18	Reference			
18-65	-0.33	0.05	0.72(0.66-0.79)	<.0001*
>65	-0.92	0.05	0.40(0.36-0.44)	<.0001*
Reason for calling 119				
Trauma	Reference			
Non-trauma	-0.60	0.03	0.55(0.52-0.58)	<.0001*

Abbreviations: HR= hazard ratio *p<0.05

The multivariable analysis identified several significant factors influencing discharge outcomes.

Male patients had a 21% lower likelihood of discharge compared to female patients (HR = 0.79,

95% CI: 0.75–0.83, $p < 0.0001$). TPTS levels were strongly associated with discharge probabilities. Patients with TPTS Level 1 and Level 2 scores had substantially reduced chances of discharge, with hazard ratios of 0.29 (95% CI: 0.25–0.33, $p < 0.0001$) and 0.62 (95% CI: 0.58–0.66, $p < 0.0001$), respectively, relative to those with Level 3 scores. Levels 4 and 5 showed no significant difference in discharge likelihood compared to Level 3 (HR = 0.98, 95% CI: 0.87–1.10, $p = 0.68$, and HR = 1.00, 95% CI: 0.89–1.13, $p = 0.93$, respectively).

Age was another critical determinant. Compared to patients under 18 years, those aged 18–65 had a 28% lower likelihood of discharge (HR = 0.72, 95% CI: 0.66–0.79, $p < 0.0001$), while patients over 65 experienced a 60% reduction in discharge probability (HR = 0.40, 95% CI: 0.36–0.44, $p < 0.0001$). The reason for contacting emergency services was also significant. Non-trauma cases had a 45% lower likelihood of discharge compared to trauma cases (HR = 0.55, 95% CI: 0.52–0.58, $p < 0.0001$).

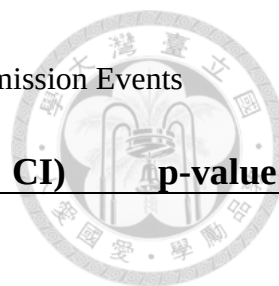


Table 15 Multivariable Analysis of Cause-Specific Hazard Models for Admission Events
Model 2.-Admission Events with Others Treated as Censored

Variables	Coeff.	se(coef)	HR (95% CI)	p-value
Gender				
Female	Reference			
Male	-0.06	0.03	0.94(0.88-0.99)	0.03*
TPTS Levels				
1	0.26	0.05	1.29(1.18-1.41)	<.0001*
2	0.08	0.03	1.09(1.02-1.16)	0.01*
3	Reference			
4	0.09	0.08	1.09(0.93-1.28)	0.27
5	0.09	0.09	1.10(0.92-1.31)	0.31
Age group				
<18	Reference			
18-65	-0.71	0.06	0.49(0.44-0.56)	<.0001*
>65	-0.73	0.06	0.48(0.42-0.55)	<.0001*
Reason for calling 119				
Trauma	Reference			
Non-trauma	-0.97	0.04	0.38(0.35-0.41)	<.0001*

Abbreviations: HR= hazard ratio *p<0.05

The multivariable analysis revealed several factors significantly influencing hospital admission outcomes. Male patients had a 6% lower likelihood of admission compared to females (HR = 0.94, 95% CI: 0.88–0.99, p = 0.03). TPTS levels were strongly associated with admission probabilities. Patients with TPTS Level 1 scores had a 29% higher likelihood of admission (HR = 1.29, 95% CI: 1.18–1.41, p < 0.0001), while those with Level 2 scores exhibited a 9% increase in admission likelihood (HR = 1.09, 95% CI: 1.02–1.16, p = 0.01), relative to Level 3 patients. Levels 4 and 5 did not show significant differences compared to Level 3, with hazard ratios of 1.09 (95% CI: 0.93–1.28, p = 0.27) and 1.10 (95% CI: 0.92–1.31, p = 0.31), respectively.

Age was another critical factor. Compared to patients under 18 years, those aged 18–65 had a 51% lower likelihood of admission (HR = 0.49, 95% CI: 0.44–0.56, $p < 0.0001$), and patients over 65 experienced a 52% reduction in admission likelihood (HR = 0.48, 95% CI: 0.42–0.55, $p < 0.0001$). Furthermore, the reason for contacting emergency services was a significant determinant. Non-trauma cases were associated with a 62% lower likelihood of admission compared to trauma cases (HR = 0.38, 95% CI: 0.35–0.41, $p < 0.0001$).

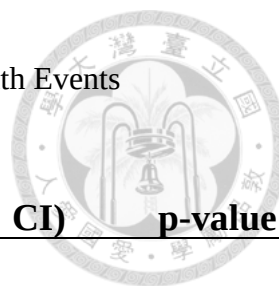
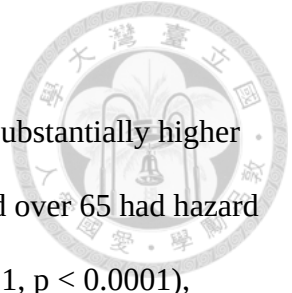


Table 16 Multivariable Analysis of Cause-Specific Hazard Models for Death Events
Model 3.-Death Events with Others Treated as Censored

Variables	Coeff.	se(coef)	HR (95% CI)	p-value
Gender				
Female	Reference			
Male	0.07	0.15	1.08(0.81-1.45)	0.60
TPTS Levels				
1	4.10	0.35	6.03(30.4-119.3)	<.0001*
2	1.29	0.38	3.65(1.72-7.73)	<.0001*
3	Reference			
4	-0.14	0.001	8.1×10^{-7} (0.00- ∞)	0.99
5	1.00	0.78	2.73(0.59-12.7)	0.20
Age group				
<18	Reference			
18-65	1.43	0.51	4.17(1.52-11.5)	<.0001*
>65	1.43	0.51	4.43(1.63-12.1)	<.0001*
Reason for calling 119				
Trauma	Reference			
Non-trauma	-0.29	0.20	0.75(0.50-1.12)	0.16

Abbreviations: HR= hazard ratio *p<0.05

The multivariable analysis identified several factors significantly associated with mortality outcomes. Gender did not significantly affect the likelihood of death, with male patients showing a HR of 1.08 (95% CI: 0.81–1.45, $p = 0.60$) compared to females. TPTS levels were strongly linked to mortality risk. Patients with TPTS Level 1 and Level 2 scores had markedly higher mortality risks compared to Level 3, with hazard ratios of 6.03 (95% CI: 3.04–119.3, $p < 0.0001$) and 3.65 (95% CI: 1.72–7.73, $p < 0.0001$), respectively. Levels 4 and 5 did not exhibit statistically significant differences, with HRs of 8.1×10^{-7} (95% CI: 0.00– ∞ , $p = 0.99$) and 2.73 (95% CI: 0.59–12.7, $p = 0.20$), respectively.



Age was a critical determinant, with older age groups demonstrating substantially higher mortality risks. Compared to patients under 18 years, those aged 18–65 and over 65 had hazard ratios of 4.17 (95% CI: 1.52–11.5, $p < 0.0001$) and 4.43 (95% CI: 1.63–12.1, $p < 0.0001$), respectively. The reason for contacting emergency services was not significantly associated with mortality outcomes, as non-trauma cases had a hazard ratio of 0.75 (95% CI: 0.50–1.12, $p = 0.16$) compared to trauma cases.

4.2.6 Exploring the Distribution of ESI Score through Stratification



Figure 31 ESI Score distribution by 5-level TPTS

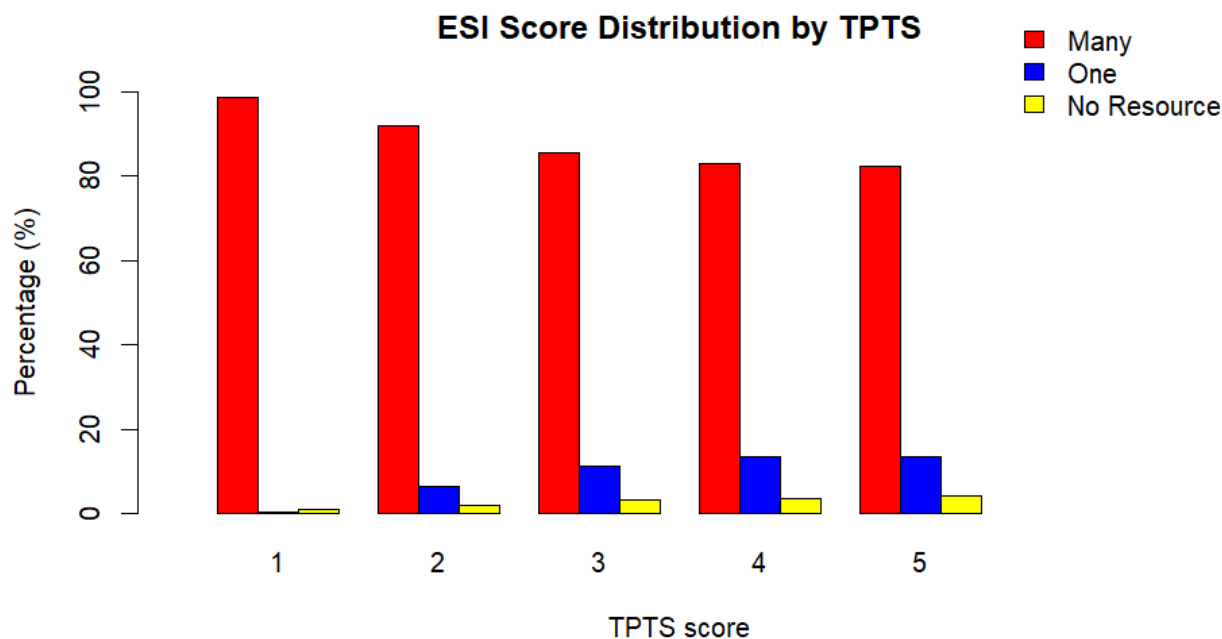


Figure 25 illustrates the distribution of ESI categories by 5-level TPTS scores. Patients with a TPTS score of 1 predominantly fall under the "Many resources" category, with minimal representation in "One resource" or "No resource." Similarly, TPTS 2 patients are mostly classified as requiring "Many resources," with a slight increase in lower resource categories.

For TPTS scores of 3 to 5, a gradual shift is observed, with a higher proportion of patients requiring fewer resources, though "Many resources" remains the largest category. These findings demonstrate a strong correlation between lower 5-level TPTS scores and higher resource needs, validating 5-level TPTS as a discrimination of resource utilization in the emergency department.



Figure 32 ESI Score distribution by TTAS

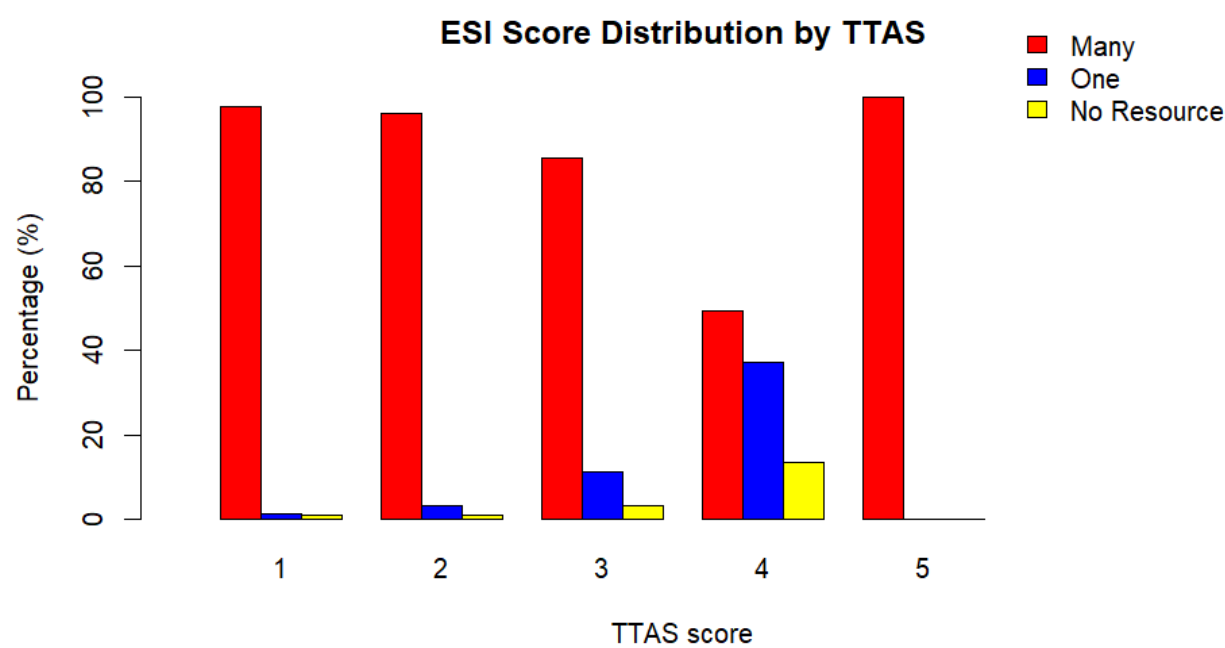
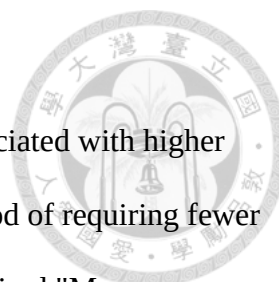


Figure 26 illustrates the distribution of ESI categories by TTAS scores. Patients with a TTAS score of 1 predominantly fall under the "Many resources" category, with negligible representation in the "One resource" or "No resource" categories. Similarly, TTAS 2 patients are largely classified as requiring "Many resources," with a slightly increased proportion falling into the lower resource categories.

For TTAS scores of 3 and 4, a more notable shift is observed. Patients with a TTAS score of 3 show an increase in the "One resource" category, while a smaller proportion falls into the "No resource" category. Patients with a TTAS score of 4 display a more balanced distribution, with significant representation in the "One resource" and "No resource" categories. However, for TTAS score 5, all patients are classified in the "Many resources" category, with no representation in either the "One resource" or "No resource" categories.



These findings demonstrate that lower TTAS scores are strongly associated with higher resource utilization, while higher TTAS scores show an increased likelihood of requiring fewer resources, with the exception of TTAS score 5, where all patients still required "Many resources." This pattern supports the discriminant validity of TTAS in estimating emergency department resource needs.

Figure 33 ESI Score distribution by reason for 119 calls

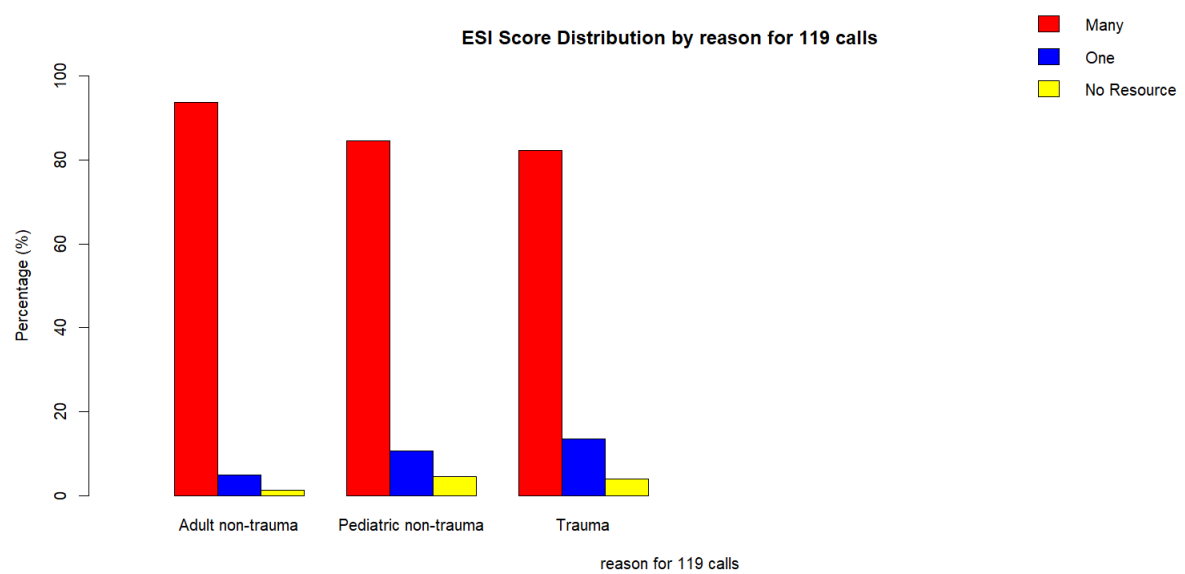


Figure 27 depicts the distribution of ESI categories by the reason for 119 emergency calls, divided into "Adult non-trauma," "Pediatric non-trauma," and "Trauma." Across all categories, the majority of patients are classified under the "Many resources" category, accounting for over 80% of cases in each group.

For "Adult non-trauma" and "Pediatric non-trauma," a small proportion of patients fall into the "One resource" and "No resource" categories, with "One resource" being slightly more common. A similar pattern is observed in the "Trauma" category, where the "One resource" and "No resource" proportions remain low but are slightly higher compared to non-trauma groups.

These findings indicate that most emergency cases, regardless of the reason for the 119 call, require substantial resource utilization. However, trauma cases show a marginally greater likelihood of requiring fewer resources compared to non-trauma cases. This suggests that resource needs can vary slightly depending on the type of emergency, but the overall demand for "Many resources" remains high.

Figure 34 ESI Score distribution by age group

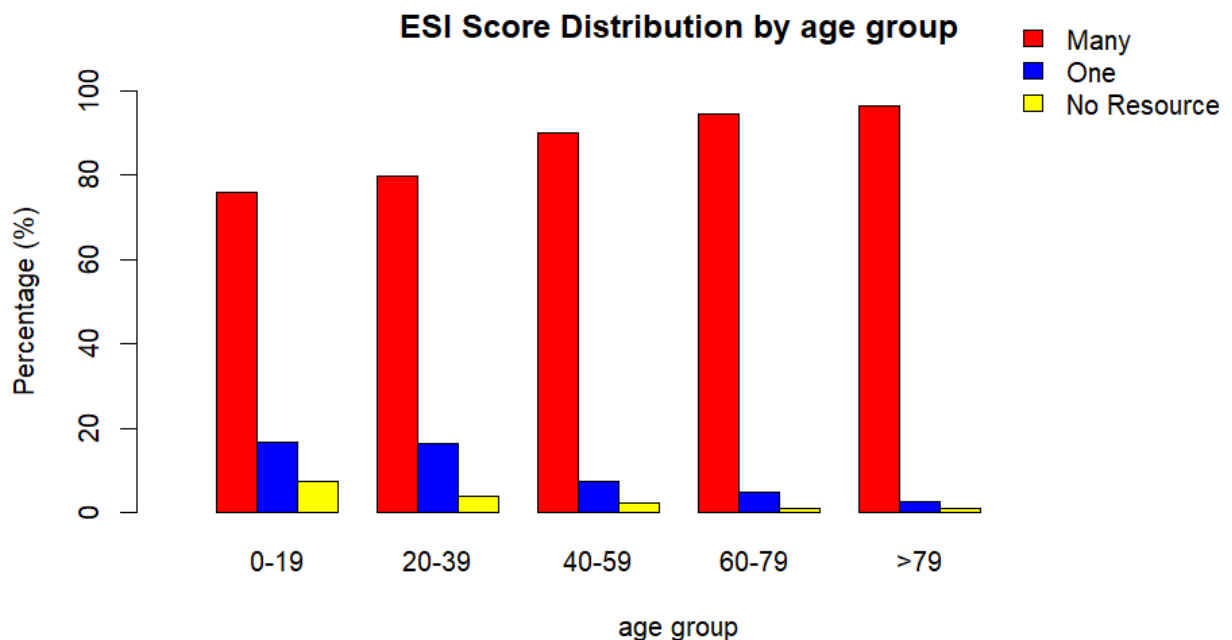
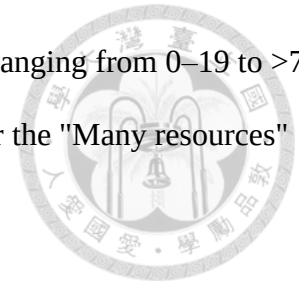


Figure 28 illustrates the distribution of ESI categories by age group, ranging from 0–19 to >79 years. Across all age groups, the majority of patients are categorized under the "Many resources" category, consistently accounting for over 80% of cases.



In the younger age group (0–19), a notable proportion of patients fall into the "One resource" category, with a small representation in the "No resource" category. As the age groups increase (20–39, 40–59, 60–79, and >79), the proportion of patients requiring "One resource" or "No resource" slightly decreases. Patients in older age groups (60–79 and >79) show an even higher predominance in the "Many resources" category, with minimal representation in lower resource categories.

These findings suggest that while resource utilization is consistently high across all age groups, younger patients are more likely to require fewer resources compared to older patients. This pattern highlights the relationship between age and resource needs in the emergency department, emphasizing greater resource demands among older populations.

Chapter 5. Discussion

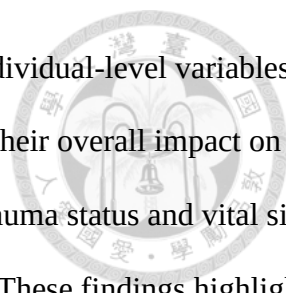
5.1 Discussion of Study 1



5.1.1 Major findings

This study conducted in Taiwan evaluated the reliability and validity of the 5-level TPTS system among different groups, including EMTs and RNs, using various statistical measures. The analysis of ICCs revealed key findings: under Model A, which assessed absolute agreement, moderate reliability was observed across all groups, with ICC values ranging from 0.57 overall to 0.71 among different groups. In contrast, Model B, focusing on consistency across multiple raters, demonstrated excellent reliability, with ICC values exceeding 0.9 in most groups. These results suggest that while absolute agreement shows moderate reliability, the consistency among raters is notably strong.

Scatter plot analyses reinforced these findings, demonstrating strong positive correlations between 5-level TPTS scores across groups, with EMTs generally assigning higher scores compared to RNs. The Bland-Altman plot further highlighted a mean difference of 0.326, reflecting a systematic bias where EMTs tended to rate cases slightly higher than RNs. This bias increased slightly as mean scores rose, as indicated by a positive regression slope. These findings suggest that while 5-level TPTS scoring exhibits robust reliability and agreement overall, some degree of systematic scoring variation exists between EMTs and RNs, which warrants further investigation to ensure optimal inter-rater consistency in triage scenarios.



The multilevel analysis of 5-level TPTS scores revealed that while individual-level variables, such as rater qualifications, age, and gender, were statistically significant, their overall impact on score variation was minimal. In contrast, patient-level factors, including trauma status and vital sign stability, demonstrated a much stronger influence on 5-level TPTS scores. These findings highlight that the primary determinants of 5-level TPTS score variability are rooted in patient characteristics rather than individual rater differences.

This suggests that although systematic biases exist among different professional groups—such as EMT-Ps scoring slightly higher and nurses scoring lower on average—the 5-level TPTS system remains robust and reliable for use in prehospital settings. The relatively minor influence of individual-level variability on scores supports the consistency and generalizability of the 5-level TPTS system, underscoring its applicability in diverse and dynamic emergency care environments. These results reaffirm the utility of 5-level TPTS as a valuable tool for structured triage, even when accounting for the professional diversity of its users.

5.1.2 Comparison with previous studies

Our study contributes valuable insights to the domain of prehospital triage by evaluating the inter-rater reliability of the 5-level TPTS system among diverse professional groups. We observed moderate agreement in absolute terms (ICCs 0.57 to 0.71) and excellent consistency (ICCs >0.9) across rater groups, underscoring the system's robustness in diverse settings. Notably, EMTs tended to assign slightly higher scores than RNs, with a mean difference of 0.326, as demonstrated by Bland-Altman analysis. These findings align with prior research on inter-rater reliability in triage systems and highlight potential differences in scoring tendencies between professional roles,

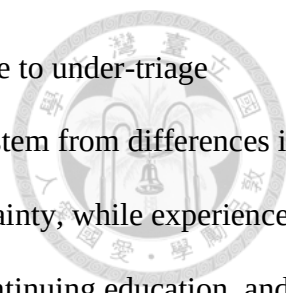
offering a basis for further exploration of calibration strategies to enhance consistency in prehospital triage practices.

For example, Leeies et al. evaluated the CTAS and found moderate inter-rater reliability between paramedics and nurses, with a kappa value of 0.437. Despite some systematic variation, the study reported that paramedic and nurse scores matched exactly or were within one level 84.3% of the time, indicating sufficient concordance for practical use[17]. Similarly, Smith et al. reported moderate agreement ($\kappa = 0.409$) in ESI scores between paramedics and nurses, with experience levels influencing the reliability of scoring[15].

Additionally, some studies have highlighted the potential of targeted training to improve concordance between EMTs and nurses. Ghanbarzehi et al., for instance, showed that pre-training concordance using the ESI was poor ($\kappa = 0.20$), but improved to a good level ($\kappa = 0.62$) following structured training programs. These results underscore the importance of consistent training to enhance inter-rater reliability across professional groups[92].

Supporting evidence includes the Taiwanese study by Ng et al., which validated the five-level TTAS system. After EMT training, inter-rater reliability between EMTs and RNs improved to a weighted kappa of 0.825, indicating very good agreement[21]. In contrast, the Korean study on the KTAS observed only low agreement ($\kappa = 0.211$) between prehospital and hospital triage scores, emphasizing the need for consistent training and systems integration[23].

Levis-Elmelech et al. identified that nurse experience significantly influences triage accuracy, with less-experienced nurses (2–3 years) performing better in high-urgency classifications (49.6%) but frequently over-triaging low-urgency cases (16.8%). In contrast, more experienced nurses (13+



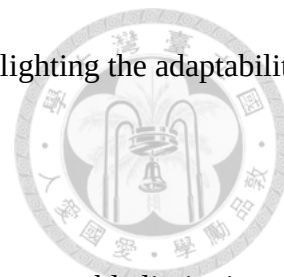
years) excelled in low-urgency classifications (39.6%) but were more prone to under-triage high-urgency patients (26.9%). The authors suggest that these trends may stem from differences in confidence levels, where less-experienced nurses over-triage due to uncertainty, while experienced nurses may underestimate urgency. They recommend targeted training, continuing education, and decision-support tools to address these discrepancies and improve triage accuracy across experience levels[119].

Moreover, the introduction of real-time decision-support tools has significantly improved reliability. For example, the implementation of the electronic Canadian Triage and Acuity Scale (eCTAS) enhanced inter-rater reliability, increasing exact modal agreement from 75.4% to 92.7%, while the quadratic-weighted kappa rose from 0.79 to 0.93. This also led to reduced rates of over-triage and under-triage[120]. These findings emphasize the benefits of integrating electronic systems to standardize triage protocols, improve consistency, and minimize errors, making them valuable additions to existing triage practices.

A key factor in comparing inter-rater reliability lies in the metrics used to evaluate it. Studies have employed various measures such as linear weighted kappa, quadratically weighted kappa, triage-weighted kappa, and ICC, each reflecting distinct assumptions and implications for interpreting reliability[121, 122].

For instance, Dalwai et al. assessed the reliability of the SATS using ICC and weighted kappa, reporting moderate inter-rater agreement (ICC = 0.50–0.60; QWK range: 0.50–0.59) in low-resource settings, demonstrating the challenges of achieving consistency in such environments[85]. Similarly, the Swiss Emergency Triage Scale (SETS) showed high reliability,

with an ICC of 0.84, when paramedics evaluated simulated scenarios, highlighting the adaptability of standardized triage tools across different professional contexts[19].

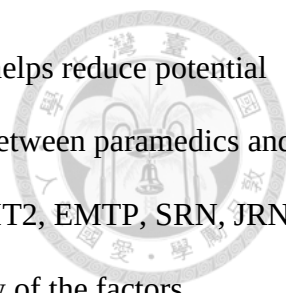


Kappa statistics are widely used to assess inter-rater agreement but have notable limitations. Tu's research highlighted that Kappa values are sensitive to prevalence and category distribution, where imbalanced datasets can result in low Kappa despite high observed agreement. Additionally, Kappa values are influenced by the number of categories, with fewer categories often yielding inflated agreement. Assumptions of independent raters and reliance on null hypothesis testing further complicate interpretation. To address these issues, alternative methods like weighted Kappa for ordinal data or multilevel modeling can provide more reliable insights. Future research should adopt these advanced tools to improve the accuracy and transparency of agreement assessments[123]

Our study adds to the existing research by showing that the 5-level TPTS system is reliable, even though EMTs tend to give slightly higher scores than RNs. Differences in training, experience, and triage protocols across studies may explain some of the variability in results[17, 91]. However, cultural and operational differences, such as the lack of shared triage protocols, further complicate inter-professional concordance[92, 124]. This highlights the need for improved training and standardized protocols to align scoring between different professional groups.

5.1.3 Strengths

Our study has several notable strengths. First, it is one of the few studies to compare the 5-level TPTS across two distinct groups: EMTs and RNs. We used stratified sampling and random



selection, ensuring a fair representation of both raters and patients, which helps reduce potential bias. Second, unlike many previous studies that focus only on agreement between paramedics and nurses, our research also considers differences based on job roles (e.g., EMT2, EMTP, SRN, JRN) and the characteristics of patients themselves. This provides a broader view of the factors influencing triage accuracy. Third, even though some differences were found in scoring between raters, our findings show that patient condition remains the key factor guiding triage decisions. This suggests that the 5-level TPTS is practical and reliable in real-world prehospital settings. Lastly, this study fills a research gap in Asia, where studies on prehospital triage systems are very limited[10]. Among the few available studies from Taiwan, Japan, and Korea[18, 21, 23, 25], our research is the first to validate the reliability and consistency of 5-level TPTS in Taiwan, providing valuable evidence for its use in emergency medical services.

5.1.4 Limitations and future work

This study has several limitations that should be considered. First, while random sampling and electronic medical records (EMRs) were used, the findings may have limited generalizability. The study was conducted in a single region, Taoyuan, and focused on a specific medical center, which may introduce selection bias. Additionally, EMRs might not fully reflect the complexity and variability of real-world patient interactions. The patient population and EMS practices in Taoyuan may differ from those in other regions of Taiwan, limiting broader applicability.

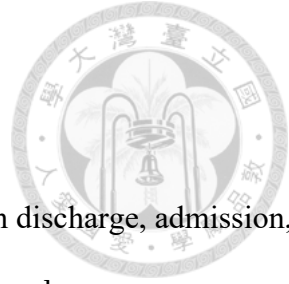
Another limitation is the observed differences between EMTs and triage nurses, which may be influenced by variations in their training, particularly in areas such as pain scale assessment. This study did not include physicians, as prior research indicates they may upgrade triage levels based on clinical judgment. However, in Taiwan, triage is primarily performed by nurses, not physicians.

Inter-rater reliability among experts was not assessed, and there is no established gold standard for triage systems, which could affect the consistency of our findings. Future studies should examine the 5-level TPTS system across diverse patient populations, such as pediatric, geriatric, and pregnant patients, as well as in specific clinical scenarios like head trauma, stroke, or myocardial infarction. Addressing these gaps and exploring targeted training for EMTs and nurses may help improve the reliability and applicability of triage systems.

5.2 Discussion of Study 2

5.2.1 Major findings

This study assessed the 5-level TPTS's ability to discriminate between discharge, admission, and death events across various LOS intervals and subgroups based on age and trauma status.



Overall Discrimination Ability

The 5-level TPTS demonstrated its strongest discrimination ability for discharge events during early LOS (<6 hours), effectively stratifying lower-severity levels (e.g., Level 5) with faster discharges and higher-severity levels (e.g., Level 1) with slower progression. However, this ability diminished over time, with levels converging beyond 48 hours. For admission events, the 5-level TPTS showed limited overall stratification, as Levels 1–5 often overlapped. Death events were primarily concentrated in Level 1, with minimal progression observed in Levels 2–5, highlighting strong stratification for the most severe cases but little differentiation for others.

Subgroup Analysis

Age <18 Years: The 5-level TPTS effectively separated Level 1 (slowest) and Level 5 (fastest) for discharge events during early LOS, but intermediate levels (2–4) showed weaker differentiation. Admission events exhibited significant overlap among Levels 2–5, and death events were exclusive to Level 1.

Age 18–65 Years: Level 5 showed faster discharges and Level 1 slower discharges during early LOS, but stratification weakened after 24 hours. Admission and death events demonstrated substantial overlap for Levels 2–5, while Level 1 remained distinct.



Age >65 Years: Levels 1 and 2 showed slower discharges compared to Levels 3–5, with stratification persisting up to 48 hours. Admission events showed minimal separation across all levels, while death events were concentrated in Level 1.

Trauma and Non-trauma Cases

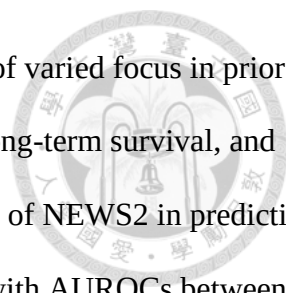
In trauma cases, the 5-level TPTS effectively separated discharge events for Levels 3–5 during early LOS, while Levels 2–5 overlapped for admission events. Death events were primarily concentrated in Level 1.

In non-trauma cases, discharge events showed good stratification for Levels 1, 2, and 5 early on, but Levels 3 and 4 overlapped. Admission and death events exhibited weak stratification, with significant overlap among Levels 1–5.

Summary

The 5-level TPTS performed best for early LOS, particularly in stratifying discharge events and distinguishing Level 1 for death events. However, its ability to differentiate among lower-severity levels weakened over time and in non-trauma and admission scenarios. These findings highlight the need for further refinement of the 5-level TPTS to improve its stratification accuracy across all scenarios.

5.2.2 Comparison with previous studies



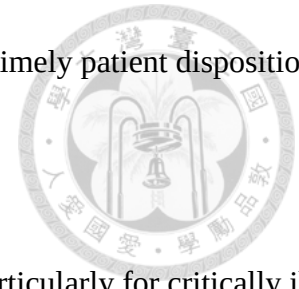
The predictive value of prehospital triage systems has been a subject of varied focus in prior research, examining short-term outcomes like mortality within 48 hours, long-term survival, and resource utilization. For instance, Lindskou et al. assessed the performance of NEWS2 in predicting 30-day mortality and ICU admissions, reporting moderate discrimination with AUROCs between 0.680 (0.675-0.685) and 0.681 (0.673-0.689)[125]. Similarly, Martín-Rodríguez et al. compared seven prehospital EWS and found MREMS to be the most effective for predicting one-year mortality, with an AUROC of 0.77, though overall performance remained moderate[126].

Prehospital triage systems such as RETTS-A and NEWS2 have demonstrated varying degrees of success in identifying high-risk patients. RETTS-A showed superior sensitivity (73%) for detecting time-sensitive conditions compared to NEWS (37%) and NEWS 2 (35%), but at the cost of specificity[24, 90].

CTAS has proven effective in predicting short-term outcomes, such as resource utilization and hospitalization, with higher acuity levels (1–3) strongly linked to increased admissions and resource needs. Reported sensitivities range from 10.77% to 47.44%, and specificities from 60.87% to 93.97%, depending on the setting[16]. However, its ability to predict long-term outcomes, such as 30-day or one-year mortality, remains limited.

Prolonged ED LOS has well-documented adverse effects on patient safety and outcomes. Extended LOS is associated with higher mortality risks, as demonstrated by Jones et al., who observed an 8% increase in 30-day mortality for patients with delays exceeding six to eight hours[114]. Similarly, Sprivulis et al. highlighted that each additional hour in a crowded ED increased the likelihood of adverse outcomes, particularly within seven to 30 days of the ED

visit[113]. These findings underscore the need for interventions targeting timely patient disposition to improve survival rates.



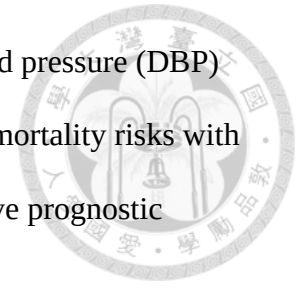
In addition to mortality, prolonged LOS exacerbates clinical risks, particularly for critically ill patients. Rose et al. reported that delays in transferring patients from the ED to ICUs increased the incidence of ventilator-associated pneumonia and prolonged mechanical ventilation, leading to worse prognoses[112]. Efficient transitions to specialized care settings are critical for mitigating such complications.

From a resource perspective, longer ED LOS significantly impacts operational efficiency and economic costs. High ED crowding, as analyzed by Sun et al., not only increased inpatient stays by 0.8% but also raised hospital expenses, illustrating the cascading effects of inefficiencies[115]. Moreover, Derosé et al. found that prolonged ED boarding due to limited inpatient capacity resulted in a six-hour extension of total hospital LOS when boarding times exceeded 14 hours, further straining healthcare resources[116].

Triage systems, such as the ESI, play a pivotal role in optimizing resource utilization by predicting patient acuity and disposition needs. Tanabe et al. demonstrated that patients classified as ESI level 1 required an average of five resources, compared to just 0.2 for those at level 5, highlighting the importance of acuity-driven resource allocation[127]. This stratification not only aids in efficient resource planning but also enhances patient flow and care delivery.

Vital signs also influence triage accuracy and clinical outcomes. Candel et al. showed that deviations in vital signs, such as systolic blood pressure (SBP) and oxygen saturation (O₂), had

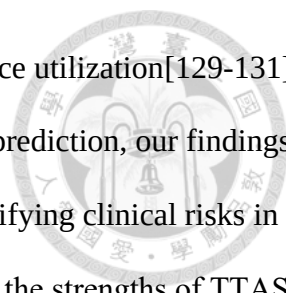
linear associations with mortality, while heart rate (HR) and diastolic blood pressure (DBP) followed U-shaped patterns. Importantly, older patients exhibited greater mortality risks with deviations, underscoring the need for age-adjusted triage criteria to improve prognostic accuracy[128].



In conclusion, addressing prolonged LOS and improving resource utilization require systemic interventions, including optimizing triage systems like ESI and integrating age-adjusted risk stratification tools. Such measures can enhance operational efficiency, reduce adverse outcomes, and ensure the delivery of high-quality emergency care.

Previous studies in Taiwan have assessed the effectiveness of various triage systems in emergency medical settings, focusing predominantly on the two-level TPTS and the five-level TTAS. Tsai et al. compared the two-level TPTS with TTAS upon ED arrival, demonstrating that TTAS outperformed the two-level TPTS in predicting hospitalizations, with an AUROC of 0.75 compared to 0.66. TTAS also demonstrated a higher sensitivity in predicting hospitalization rate (70.3% vs. 41.1%)[18, 21]. Additionally, Ng et al. validated the use of TTAS by EMTs, reporting strong inter-rater reliability and better predictive performance for hospital admissions and resource consumption compared to the two-level TPTS[21].

Our study expands on this body of work by evaluating the discriminant capabilities of the newly implemented five-level TPTS in prehospital settings, marking a significant shift from the older two-level system. Unlike its predecessor, the five-level TPTS demonstrated strong discriminant capacity for a broader range of post-arrival outcomes, including mortality, ED LOS and resource needs. Our findings are supported by previous evidence showing that lower triage

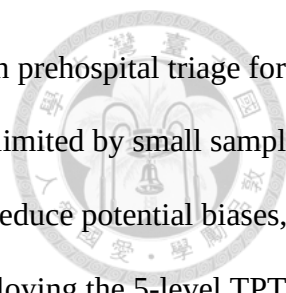


scores are consistently linked to worse clinical outcomes and higher resource utilization[129-131]. While TTAS remained effective for resource prioritization and admission prediction, our findings highlight the complementary role of the five-level TPTS in accurately stratifying clinical risks in prehospital scenarios. These results align with Tsai et al.'s observations on the strengths of TTAS while underscoring the efficiency and granularity of the updated five-level TPTS for urgent risk stratification[18, 21].

The complementary roles of TTAS and the five-level TPTS emphasize their potential for integrated application. TTAS's focus on resource optimization is well-matched with the five-level TPTS's enhanced ability to provide granular, rapid, and reliable assessments in prehospital settings. While earlier studies criticized the two-level TPTS for its limited granularity, the five-level TPTS addresses these concerns, offering a more nuanced approach. Differences in study populations and outcome measures may explain discrepancies in previous findings, as our study broadens the scope to include post-arrival mortality, ED stay duration and resource utilization.

In summary, our findings support the integration of the five-level TPTS and TTAS to optimize prehospital and ED triage in Taiwan. This combined system provides a robust framework for improving patient care, from prehospital risk assessment to in-hospital resource allocation. The hybrid approach has the potential to serve as a scalable model for enhancing emergency care systems globally.

5.2.3 Strengths



Our study possesses several notable strengths. First, previous research on prehospital triage for discriminant patient disposition and resource utilization has often been limited by small sample sizes. In contrast, We analyzed a full year of data to increase sample size, reduce potential biases, and enhance the robustness and reliability of our findings. Second, by employing the 5-level TPTS levels, we demonstrated its strong discriminant capacity for post-arrival outcomes. The 5-level TPTS effectively stratifies patients by severity, aiding in identifying those requiring urgent care and estimating the expected consumption of ED resources. Third, our analysis revealed that patients categorized as 5-level TPTS levels 3–5, indicative of mild conditions, exhibited a low mortality rate of 0.1% (11/5795 patients). This finding is especially relevant for Taiwan, where an aging population and increasing ambulance demand highlight the need for data-driven policies encouraging alternative healthcare-seeking behaviors for non-urgent cases[27, 132]. Lastly, clinical efficacy studies on prehospital triage systems remain scarce globally. While prior studies have primarily focused on inter-rater reliability metrics, such as ICC or kappa statistics, this research uniquely explores the clinical utility of the 5-level TPTS. Our study is the first in Taiwan to evaluate the 5-level TPTS's discriminant capability for patient disposition, ED length of stay, and resource utilization, providing critical insights into the distribution of ED resource consumption.

5.2.4 Limitations and future work

This study has several limitations that should be acknowledged. First, although our analysis reflects real-world triage performance, it was conducted in a single medical center, which may not represent the broader patient population in Taoyuan. This could introduce selection bias.

Second, patients transported to medical centers are often chosen based on paramedics' clinical judgment and considerations such as the need for advanced resources or family preferences. Patient

outcomes may vary between hospitals due to differences in treatment capabilities and available specialties, adding variability to the results.

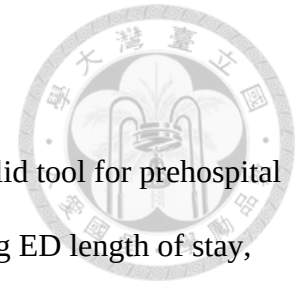
Third, we did not evaluate the total medical costs for each patient or estimate the time and manpower required to provide care. This omission limits our ability to assess the broader resource implications of our findings.

Fourth, differences between EMT and triage nurse assessments could result from variations in their roles, training, and timing of evaluations. EMTs perform triage scoring in the prehospital setting, while nurses evaluate patients upon hospital arrival. Changes in a patient's condition during transport or improvements after initial treatment could naturally cause discrepancies. Additionally, differences in areas such as pain scale assessment may further contribute to these variations. Future research could focus on developing training programs to reduce these inconsistencies and improve alignment between prehospital and in-hospital triage assessments.

Fifth, while triage levels were documented in the electronic medical records, it is unclear whether these scores were auto-generated by the system or adjusted by paramedics based on clinical judgment. Future studies should record both the initial system-generated scores and any paramedic-modified scores to better understand the decision-making process.

Sixth, this study was conducted in Taoyuan, Taiwan, where the EMS system exclusively uses the 5-level TPTS and 2-level TPTS. This limits the applicability of our findings to other regions or countries that utilize different triage systems. Finally, the use of distinct triage systems pre- and post-hospital reflects the specific needs of each setting. However, further research is needed to explore the clinical utility of alternative prehospital triage systems. Future studies should include other regions and populations, such as elderly and pediatric patients, and evaluate triage performance in time-sensitive conditions like stroke, myocardial infarction, and major trauma.

Chapter 6. Conclusion

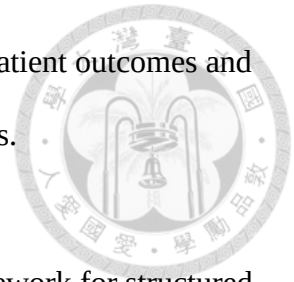


This study confirms that the 5-level TPTS system is a reliable and valid tool for prehospital triage, with strong discriminant capabilities for patient outcomes, including ED length of stay, resource utilization, and in-hospital mortality. Despite some observed differences in scoring between EMTs and RNs, these variations were relatively small compared to the overall high level of reliability and consistency demonstrated across all groups. The study emphasizes that patient characteristics, such as trauma status and vital sign stability, are the primary drivers of 5-level TPTS score variability, underscoring its clinical relevance in guiding triage decisions.

The findings also highlight the complementary roles of 5-level TPTS and TTAS in Taiwan's emergency care framework. 5-level TPTS is highly effective for prehospital triage and discriminating immediate post-arrival outcomes, while TTAS provides additional utility in prioritizing in-hospital resources, particularly for patients with intermediate triage levels. Together, these systems can enhance the accuracy of risk stratification, streamline resource allocation, and improve overall operational efficiency in emergency settings.

Additionally, this study fills a critical research gap in Asia, offering one of the first comprehensive evaluations of 5-level TPTS in Taiwan. By validating its discriminant capacity and clinical utility in a large-scale, real-world dataset, we provide strong evidence for the adoption of 5-level TPTS as a standardized tool in prehospital and emergency care. Future research should build on these findings by examining 5-level TPTS performance in broader populations, including pediatric, geriatric, and high-risk patients, and exploring its application in time-sensitive conditions such as stroke, myocardial infarction, and major trauma. Expanding these insights will further

enhance the effectiveness and adaptability of 5-level TPTS in improving patient outcomes and supporting data-driven policy development for emergency medical services.



In conclusion, the 5-level TPTS system is a robust and practical framework for structured triage in Taiwan's dynamic prehospital care environment. Its integration with other triage systems, like TTAS, offers significant potential to advance emergency care by improving the alignment between prehospital and in-hospital processes. These findings highlight the critical role of standardized triage tools in optimizing patient care, ensuring efficient use of healthcare resources, and addressing the growing demands on emergency medical services in Taiwan and beyond.

Appendix

I. Supplemental figures

Figure S 1 Franco-German vs Anglo-American Models

This map illustrates the global distribution of EMS systems. Countries with no available data or where information is not publicly disclosed are shaded in grey.

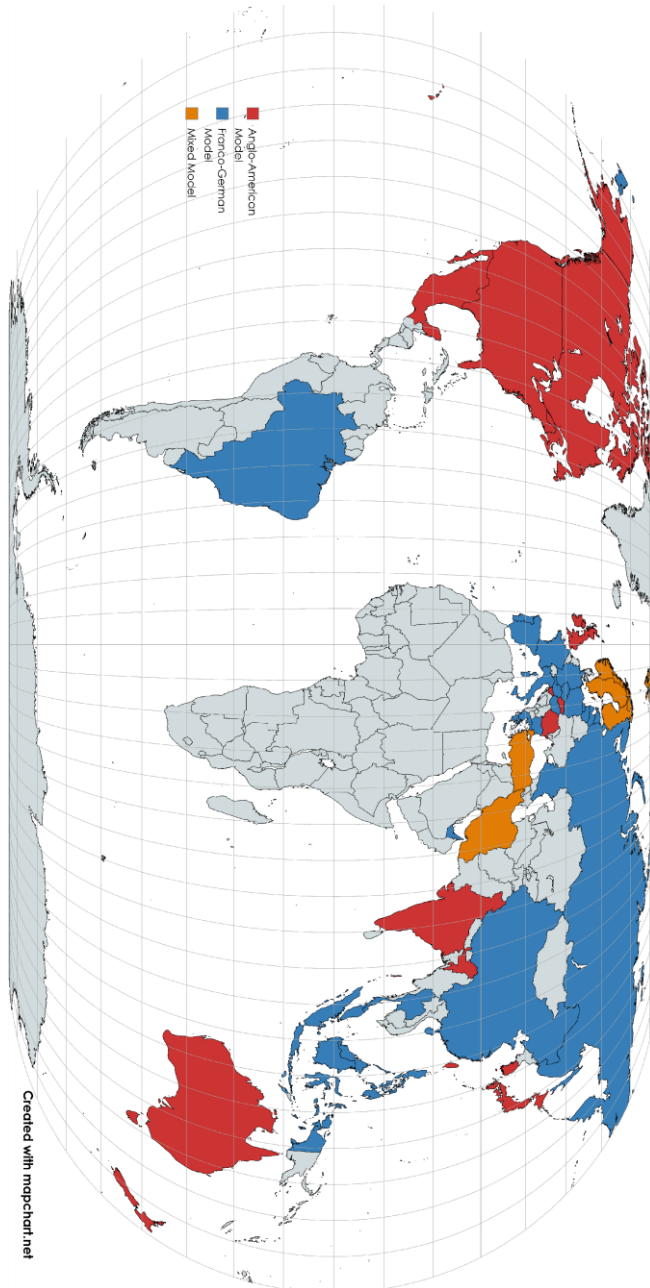


Figure S 2 Taiwan EMS benchmarks

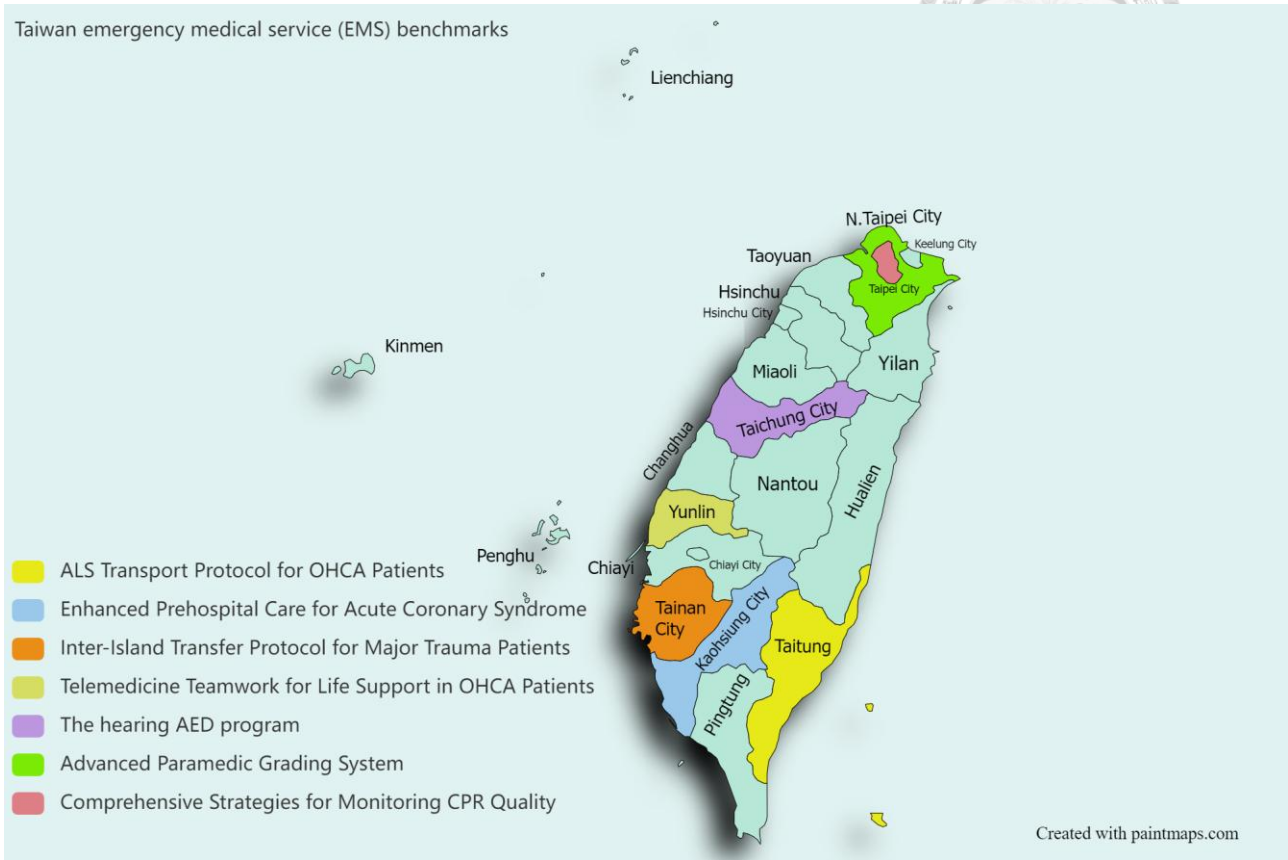
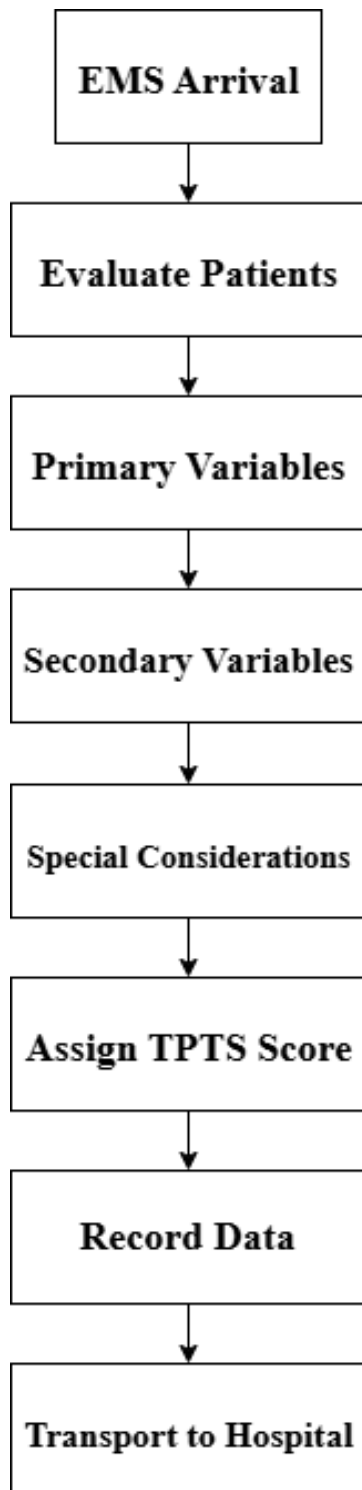


Figure S 3 The traige Algorithm



II. Supplemental tables

Table S 1 Franco-German vs Anglo-American Models

Model	Franco-German Model	Anglo-American Model
Principle	More treated on scene. Few transported to hospitals.	Few treated on scene. More transported to hospitals.
Dispatch Mode	Ambulances respond to severe cases; minor cases self-transport.	Ambulances are dispatched for all cases, regardless of severity.
Care Providers	Medical doctors supported by paramedics	Paramedics with medical oversight
Organizational Structure	Typically part of a broader healthcare system. Some ambulances are hospital-based.	Often linked to public safety services like fire departments. Ambulances are usually fire-based for deployment.
Overarching Organization	EMS is a part of public health organization	EMS is a part of public safety organization
Main Motive	Brings the hospital to the patient	Brings the patient to the hospital
Destination for Transported Patients	Direct transport to hospital wards (bypassing EDs)	Direct transport to emergency departments (EDs)
Countries	Germany, France, Greece, Austria, and similar European countries	U.S., U.K., Canada, New Zealand, Taiwan, Australia, and others

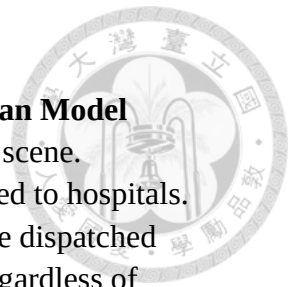


Table S 2 Comparison of the TTAS, 2-Level TPTS, 5-Level TPTS

Criteria	TTAS	2-level TPTS	5-level TPTS
Triage decision criteria	Primary Variables:1. GCS score 3-15 2. HR 50-140 beats/minute 3. RR 10-30/minute 4. Systolic BP 70-220 mm Hg 5. Body temperature 32-41°C 6. Oxygen saturation via pulse oximeter 90-94 % 7. Pain Index score 1-10 8. CRT 2-4 seconds 9. High risk Mechanism of injury	Critical Vital sign: 1. GCS score <14 2. HR >140 or <50 beats/minute 3. RR >30 or <10/minute 4. Systolic BP >220 or <90mm Hg 5. Body temperature >40 or <32°C 6. Oxygen saturation via pulse oximeter <90%	Primary Variables: 1. GCS score 3-15 via AVPU 2. HR 50-140 beats/minute 3. RR 10-30/minute 4. Systolic BP 70-220 mm Hg 5. Body temperature 32-41°C 6. Oxygen saturation via pulse oximeter 90-94 % 7. Pain Index score 1-10 8. CRT 2-4 seconds 9. EMT Interventions. Special Considerations: Adjust for unique vital signs in special populations, history of coagulopathy, and high-energy trauma with stable vitals.
Chief complaint list	Secondary Variables: 14 Adult Non-Trauma Categories: 125 Chief Complaints 14 Trauma Conditions: 42 Chief Complaints 14 Child Non-Trauma Categories: 127 Chief Complaints Environment: 11 Chief Complaints	8 Critical Non-Trauma Conditions: Acute stroke, ischemic heart disease, seizures, life-threatening intoxication, precipitate labor, snake bite, drowning, and hypo-/hyperglycemia. 8 Critical Trauma Conditions: Major burns (second- or third-degree burns >18% TBSA), burns involving the face or perineum, flail chest, penetrating injuries to the trunk, long bone fractures, pelvic fractures, limb amputation, and high-risk trauma mechanisms (e.g., high-speed collision or falls >6 meters).	Secondary Variables: 13 Critical Non-Trauma Conditions: airway and respiratory issues, altered consciousness, chest and abdominal pain, general ailments (like headaches and fever), suspected poisonings, seizures, public collapses, psychiatric emergencies, emergent labor, OHCA, and other unspecified conditions. 12 Critical Trauma Conditions: Major Burn injuries- Adults (18-65) >25% second-degree, Elderly/Children >20% second-degree, All ages >10% third-degree, general trauma (head, chest, abdomen, back, limbs), traffic-related injuries (cars, motorcycles, bicycles, pedestrians), drowning incidents, fall-related injuries, puncture wounds, electric shock injuries, bites and stings, gunshot wounds, OHCA, injuries from falls from height and other unspecified causes.
Pain scale	10-point scale	Not specified	10-point scale
Implementation guidelines	Specified	Not specified	Specified
Educational materials	Yes, Web-based training material	None	Yes, 4-hour training session

III. Supplementary R and SAS Code for Data Analysis

1. Statistical software information

In our study, SAS was used to calculate interquartile ranges (IQR), frequencies, and proportions. It also performed Chi-square tests for categorical variables and Mann-Whitney U tests for continuous variables. R was used for stratified random sampling to select 1/1000 of the patients, calculate the Intraclass Correlation Coefficient (ICC), and generate scatterplots and Bland-Altman plots. These analyses explored relationships between groups and assessed systematic bias in 5-level TPTS evaluations.

Additionally, R was used to compute two-level and three-level multilevel models, visualize outcome probabilities with the CIF under competing risks, apply Proportional Cause-Specific Hazards Models, and create bar charts to analyze the distribution of ESI scores.

Detailed software information is provided below:

-R software version 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria)

-SAS statistical software version 9.4 (SAS Institute Inc., Cary, NC, USA).

2. Mann-Whitney U tests

We used SAS statistical software (version 9.4) for analysis. Descriptive statistics, including the first quartile (Q1), median, third quartile (Q3), and IQR, were calculated using **Proc Means**. The Wilcoxon rank-sum test (Mann-Whitney U test) was performed with **proc npar1way** to compare differences in continuous variables between groups.

SAS code:

```
Proc Means data=data  Q1 median Q3 qrange;
```

```
class Group;
```

```
var  VAR;
```

run;

```
proc npar1way data=data wilcoxon; /* Mann-Whitney U */
```

```
class Group;
```

```
var VAR;
```

run;



3. Chi-square tests

SAS was employed to analyze categorical variables using the **proc freq** procedure, which calculated frequencies and conducted Chi-square tests to examine associations between the variable of interest and group classifications.

SAS code:

```
proc freq data=data;
```

```
tables VAR*Group/chisq norow nopercent ;
```

run;

4. Stratified random sampling

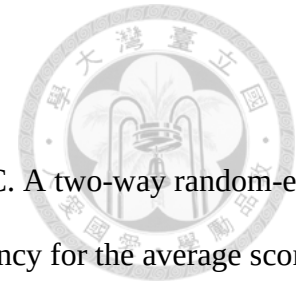
R was used for random sampling with replacement to select participants. A seed value was set using the current date to ensure reproducibility, and the `sample()` function was applied to randomly draw individuals from the target population.

R code:

```
set seed(date) #based on the current date
```

```
smample(1:N,sample number,replace=TRUE) #N: Total population size.
```

#sample number: The size of the sample to be drawn.



5. Intraclass Correlation Coefficient (ICC)

The reliability of 5-level TPTS assessments was evaluated using the ICC. A two-way random-effects model was applied to assess agreement for single measurements and consistency for the average scores of multiple raters. Model A focused on absolute agreement, while Model B examined consistency across raters. These analyses quantified the reliability of 5-level TPTS scoring in a multilevel context.

R code:

```
library("irr")
```

```
# Model A: two-way model based on agreement for single measurements
```

```
icc(  
  data, model = "twoway",  
  type = "agreement", unit = "single"  
)
```

```
# Model B: two-way model based on consistency for the average of multiple raters
```

```
icc(  
  data, model = "twoway",  
  type = "consistency", unit = "average"  
)
```

6. Scatterplots

R was utilized to visualize pairwise relationships between the performance rates of EMTP, EMT2, JRN, and SRN groups. Scatterplots with linear regression lines and $y = x$ reference lines were generated to examine correlations. The linear regression equations and correlation coefficients were displayed to assess the degree of agreement and potential systematic biases between groups.

R code:

```
# Define a custom plotting function
```

```

plot_comparison <- function(x, y, xlabel, ylabel) {
  # Fit a linear model and calculate the correlation coefficient
  fit <- lm(y ~ x)
  r <- cor(x, y)

  # Create equation and correlation coefficient text for the plot
  eq <- sprintf("y = %.3fx + %.3f", coef(fit)[2], coef(fit)[1])
  cor_text <- sprintf("correlation coefficient = %.3f", r)

  # Generate the scatterplot
  plot(x, y, main=paste(xlabel, "vs.", ylabel), xlab=xlabel, ylab=ylabel)
  abline(0, 1, col='blue')
  abline(fit, col='red')

  # Add a legend
  legend("topleft",
        legend = c("y = x", eq, cor_text),
        col = c('blue', 'red', 'orange'),
        pch = c(20, 20),
        bty = "n",
        pt.cex = 2,
        cex = 0.8,
        text.col = "black",
        inset = c(0.1, 0.1))
}

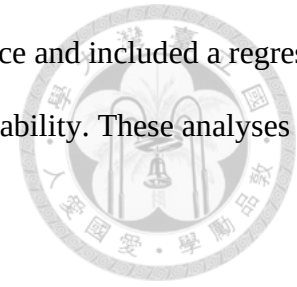
# Generate six comparison plots
plot_comparison(rate$EMTP, rate$EMT2, "EMTP", "EMT2")
plot_comparison(rate$EMTP, rate$JRN, "EMTP", "JRN")
plot_comparison(rate$EMTP, rate$SRN, "EMTP", "SRN")
plot_comparison(rate$EMT2, rate$JRN, "EMT2", "JRN")
plot_comparison(rate$EMT2, rate$SRN, "EMT2", "SRN")
plot_comparison(rate$JRN, rate$SRN, "JRN", "SRN")

```



7. Bland-Altman plots

R was used to create Bland-Altman plots for pairwise comparisons between performance rates of EMTP, EMT2, JRN, and SRN groups. Each plot displayed the mean difference and included a regression line to evaluate systematic bias, with confidence intervals for visualizing variability. These analyses provided insights into agreement and potential bias between groups.



R code:

Function for Bland-Altman plot with regression line and text annotations

```
ba_plot_with_reg <- function(x, y, title) {  
  # Create basic Bland-Altman plot  
  bland.altman.plot(x, y,  
                    xlab="Mean", ylab="Difference",  
                    main=title,  
                    conf.int=FALSE, pch=19)  
  
  # Calculate means and fit regression  
  means <- (x + y)/2  
  fit <- lm(x - y ~ means)  
  pred <- predict(fit, data.frame(means=means), interval="confidence")  
  
  # Add regression line and confidence intervals  
  abline(fit, col="red")  
  lines(means[order(means)], pred[order(means),"lwr"], col="blue", lty=2)  
  lines(means[order(means)], pred[order(means),"upr"], col="blue", lty=2)  
  
  # Calculate mean difference  
  mean_diff <- mean(x - y)  
  
  # Add mean difference text  
  text(min(means), max(x - y)-0.05,  
        paste("Mean Difference =", round(mean_diff, 3)),  
        pos = 4)  
  
  # Add regression equation text  
  eq <- paste0("y = ", round(coef(fit)[2], 3), "x + ", round(coef(fit)[1], 3))  
  text(min(means), max(x - y) - 0.15, eq, pos = 4)  
  
  # Return invisible list of statistical results (optional)
```

```
invisible(list(
  mean_difference = mean_diff,
  regression = fit
))
}
```



Generate Bland-Altman plots for all group comparisons

```
ba_plot_with_reg(rate$EMTP, rate$EMT2, "EMTP vs. EMT2")
ba_plot_with_reg(rate$EMTP, rate$JRN, "EMTP vs. JRN")
ba_plot_with_reg(rate$EMTP, rate$SRN, "EMTP vs. SRN")
ba_plot_with_reg(rate$EMT2, rate$JRN, "EMT2 vs. JRN")
ba_plot_with_reg(rate$EMT2, rate$SRN, "EMT2 vs. SRN")
ba_plot_with_reg(rate$JRN, rate$SRN, "JRN vs. SRN")
```

8. Multilevel models

Multilevel modeling was used to analyze 5-level TPTS scores with both 2-level and 3-level hierarchical structures. For 2-level models, random effects accounted for clustering at the patient level, and predictors included job grade, rater characteristics (age, gender, seniority), and patient-level variables (age, gender, treatment, vital signs, transport time). In 3-level models, site-level clustering was incorporated. Model fit was assessed using AIC, BIC, and log-likelihood values, providing insights into factors influencing 5-level TPTS scoring across hierarchical levels.

R code:

```
library(nlme)
```

#2-level multilevel models

#M1(Simple Linear Model)

```
model1<-lm(Score ~ 1, data=rate)
```

```
summary(model1)
```

```
AIC(model1)
```

BIC(model1)

-2*logLik(model1)

#M2(Variance Components Model)

model2<-lme(Score ~ 1, random=~1|pt, data=rate)

summary(model2)

-2*logLik(model2)

#Random Intercept Models

#M3(Level-1 Predictors+ job grade)

model3<-lme(Score ~ work, random=~1|pt, data=rate)

summary(model3)

-2*logLik(model3)

#M4(Level-1 Predictors+ rater age, rater gender, rater seniority, job grade)

model4<-lme(Score ~ work+ratergender+raterage+Seniority, random=~1|pt, data=rate)

summary(model4)

-2*logLik(model4)

#M5(Level-1 Predictors+ rater age, rater gender, rater seniority, job grade,

Level-2 Predictors+ reason for calling 911, vital signs)

model5<-lme(Score ~ work+ratergender+raterage+Seniority+vs+cause, random=~1|pt,

data=rate)

summary(model5)

-2*logLik(model5)

#M6(Level-1 Predictors+ rater age, rater gender, rater seniority, job grade,

Level-2 Predictors+ reason for calling 911, vital signs, patient age, patient gender, treatment, transport time)

model6<-lme(Score ~



```
work+ratergender+raterage+Seniority+vs+cause+ptage+ptgender+Trt+transtime,
random=~1|pt, data=rate)
summary(model6)
-2*logLik(model6)
```



#3-level multilevel models

#M1(Simple Linear Model)

```
model1<-lm(Score ~ 1, data=rate)
summary(model1)
AIC(model1)
BIC(model1)
-2*logLik(model1)
```

#M2(Variance Components Model)

```
model2<-lme(Score ~ 1, random=~1|pt/site, data=rate)
summary(model2)
-2*logLik(model2)
```

#Random Intercept Models

#M3(Level-1 Predictors+ job grade)

```
model3<-lme(Score ~ work, random=~1|pt/site, data=rate)
summary(model3)
-2*logLik(model3)
```

#M4(Level-1 Predictors+ job grade, rater age, rater gender, rater seniority)

```
model4<-lme(Score ~ work+ratergender+raterage+Seniority, random=~1|pt/site, data=rate)
summary(model4)
-2*logLik(model4)
```

#M5(Level-1 Predictors+ job grade, rater age, rater gender, rater seniority

Level-3 Predictors+reason for calling 911, vital signs)

```
model5<-lme(Score ~ work+ratergender+raterage+Seniority+vs+cause, random=~1|pt/site,
```

```
data=rate)
```

```
summary(model5)
```

```
-2*logLik(model5)
```

#M6(Level-1 Predictors+job grade, rater age, rater gender, rater seniority

Level-3 Predictors+reason for calling 911, vital signs, patient age, patient
gender, treatment, transport time)

```
model6<-lme(Score ~
```

```
work+ratergender+raterage+Seniority+vs+cause+ptage+ptgender+Trt+transtime,
```

```
random=~1|pt/site, data=rate)
```

```
summary(model6)
```

```
-2*logLik(model6)
```



9. Visualize outcome probabilities with the CIF

We used R to analyze and visualize cumulative incidence functions for competing risks in the ED.

The outcomes included discharge, hospital admission, and death. The cuminc function was applied to calculate cumulative incidence, and ggcuminc was used to create Kaplan-Meier-like plots for each outcome, with confidence intervals and risk tables included. The x-axis represents the length of stay in the ED in hours, providing a clear visual representation of event probabilities over time.

R code:

```
library(survival)
```

```
library(tidycmprsk)
```

```
library(gtsummary)
```

```
library(ggsurvfit)
```

```
# For patients discharged from the ED
```

```
cuminc(Surv(time, status2, type="mstate")~ 1, data=tt) %>%  
ggcuminc(outcome =c("1")) +  
add_confidence_interval() +  
add_risktable() +  
scale_ggsurvfit()+  
labs(  
  x = "Length of Stay in ED(hours)",  
)
```

```
# For patients admitted from the ED
```

```
cuminc(Surv(time, status3, type="mstate")~1, data=tt) %>%  
ggcuminc(outcome =c("2")) +  
add_confidence_interval() +  
add_risktable() +  
scale_ggsurvfit()+  
labs(  
  x = "Length of Stay in ED(hours)",  
)
```

```
# For patients who died in the ED
```

```
cuminc(Surv(time, status4, type="mstate")~1, data=tt) %>%  
ggcuminc(outcome =c("3")) +  
add_confidence_interval() +  
add_risktable() +  
scale_ggsurvfit()+  
labs(  
  x = "Length of Stay in ED(hours)",  
)
```



10. Proportional Cause-Specific Hazards Models

Proportional cause-specific hazards models were applied using Cox regression to estimate risks for competing events in the ED. Separate models were constructed for discharge, hospital admission, and death, treating other events as censored. Covariates included gender, 5-level TPTS, age group, and reason for emergency. Hazard ratios and confidence intervals were calculated to identify significant predictors for each outcome, providing insights into the factors influencing ED disposition.

R code:

```
tt$ageG <- factor(tt$ageG)
tt$ageG = relevel(tt$ageG, ref = "0-19")
tt$reason.119 <- factor(tt$reason.119)
tt$reason.119 = relevel(tt$reason.119, ref = "Trauma")
tt$TPTS <- factor(tt$TPTS)
tt$TPTS = relevel(tt$TPTS , ref = "3")

# Discharged-specific hazard with other censored (status==1)
Model.1 <- coxph(Surv(time, status==1) ~ gender+TPTS+ageG+reason.119, data=tt,
ties="breslow")

# Admission-specific hazard with other censored (status==2)
Model.2 <- coxph(Surv(time, status==2) ~ gender+TPTS+ageG+reason.119, data=tt,
ties="breslow")

# Death-specific hazard with other censored (status==3)
Model.3 <- coxph(Surv(time, status==3) ~ gender+TPTS+ageG+reason.119, data=tt,
ties="breslow")
summary(Model.1)
summary(Model.2)
summary(Model.3)
```

11. Bar charts

Bar plots were created using R to visualize the percentage distribution of ESI resources across various grouping variables, including 5-level TPTS scores, TTAS scores, age groups, gender, and reasons for 119 calls. Percentages were calculated for each group, and legends were included for clarity. These visualizations highlight the relationship between ESI resource levels and key clinical or demographic factors, providing insights into triage decision-making patterns.

R code:

```
library(ggplot2)
```

```
library(dplyr)
```

```
library(tidyr)
```

```
# Function for plotting score distributions
```

```
plot_distribution <- function(data, var1, var2, main_title = "Score Distribution",  
                             x_label = "Score") {
```

```
# Calculate percentages
```

```
counts <- table(data[[var1]], data[[var2]])
```

```
percent <- prop.table(counts, 2) * 100
```

```
# Create a bar plot
```

```
barplot(percent,  
        main = main_title,  
        xlab = x_label,  
        ylab = "Percentage (%)",  
        col = c("red", "blue", "yellow"),  
        legend = rownames(percent),  
        beside = TRUE,  
        xlim = c(0, 25),  
        ylim = c(0, 100),  
        args.legend = list(x = 26.5,  
                           y = 120,  
                           bty = "n"))  
}
```

```
# Plot for 5-level TPTS scores
```

```
plot_distribution(esi, "ESI.Resource", "TPTS",
```

```
main_title = "ESI Score Distribution by TPTS",  
x_label = "TPTS score")
```

Plot for TTAS scores

```
plot_distribution(esi, "ESI.Resource", "TTAS",  
main_title = "ESI Score Distribution by TTAS",  
x_label = "TTAS score")
```

Plot for age group

```
plot_distribution(esi, "ESI.Resource", "ageG",  
main_title = "ESI Score Distribution by age group",  
x_label = "age group")
```

Plot for gender

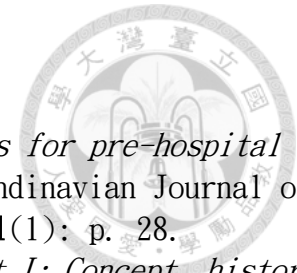
```
plot_distribution(esi, "ESI.Resource", "gender",  
main_title = "ESI Score Distribution by gender",  
x_label = "gender")
```

Plot for reason for 119 calls

```
plot_distribution(esi, "ESI.Resource", "reason.119.eng",  
main_title = "ESI Score Distribution by reason for 119 calls",  
x_label = "reason for 119 calls")
```

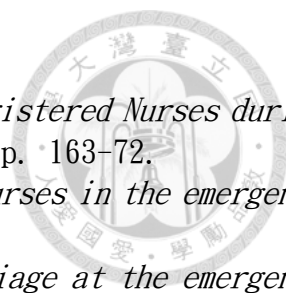


References



1. Lidal, I.B., H.H. Holte, and G.E. Vist, *Triage systems for pre-hospital emergency medical services – a systematic review*. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, 2013. **21**(1): p. 28.
2. Iserson, K.V. and J.C. Moskop, *Triage in medicine, part I: Concept, history, and types*. Ann Emerg Med, 2007. **49**(3): p. 275–81.
3. Peta, D., et al., *Triage: A Global Perspective*. J Emerg Nurs, 2023. **49**(6): p. 814–825.
4. Christ, M., et al., *Modern triage in the emergency department*. Dtsch Arztebl Int, 2010. **107**(50): p. 892–8.
5. Fernandes, C.M., et al., *Five-level triage: a report from the ACEP/ENA Five-level Triage Task Force*. J Emerg Nurs, 2005. **31**(1): p. 39–50; quiz 118.
6. Kuriyama, A., S. Urushidani, and T. Nakayama, *Five-level emergency triage systems: variation in assessment of validity*. Emerg Med J, 2017. **34**(11): p. 703–710.
7. Hinson, J.S., et al., *Triage Performance in Emergency Medicine: A Systematic Review*. Ann Emerg Med, 2019. **74**(1): p. 140–152.
8. Zachariasse, J.M., et al., *Performance of triage systems in emergency care: a systematic review and meta-analysis*. BMJ Open, 2019. **9**(5): p. e026471.
9. Bhaumik, S., et al., *Prehospital triage tools across the world: a scoping review of the published literature*. Scand J Trauma Resusc Emerg Med, 2022. **30**(1): p. 32.
10. Kim, K. and B. Oh, *Prehospital triage in emergency medical services system: A scoping review*. Int Emerg Nurs, 2023. **69**: p. 101293.
11. Abetz, J.W., et al., *Review article: Pre-hospital provider clinical judgement upon arrival to the emergency department: A systematic review and meta-analysis*. Emerg Med Australas, 2020. **32**(6): p. 917–923.
12. Ebrahimi, M., et al., *The reliability of the Australasian Triage Scale: a meta-analysis*. World J Emerg Med, 2015. **6**(2): p. 94–9.
13. Dallaire, C., et al., *Interrater agreement of Canadian Emergency Department Triage and Acuity Scale scores assigned by base hospital and emergency department nurses*. Cjem, 2010. **12**(1): p. 45–9.
14. Ng, C.J., et al., *Validation of the Taiwan triage and acuity scale: a new computerised five-level triage system*. Emerg Med J, 2011. **28**(12): p. 1026–31.
15. Buschhorn, H.M., et al., *Emergency medical services triage using the emergency severity index: is it reliable and valid?* J Emerg Nurs, 2013. **39**(5): p. e55–63.
16. Smith, D.T., et al., *Analyzing the Usability of the 5-Level Canadian Triage and Acuity Scale By Paramedics in the Prehospital Environment*. J Emerg Nurs, 2015. **41**(6): p. 489–95.
17. Leeies, M., et al., *Prehospital Application of the Canadian Triage and Acuity*

- Scale by Emergency Medical Services*. Cjem, 2017. **19**(1): p. 26-31.
18. Tsai, L.H., et al., *Comparison of prehospital triage and five-level triage system at the emergency department*. Emerg Med J, 2017. **34**(11): p. 720-725.
 19. Groscurin, O., et al., *Reliability and performance of the Swiss Emergency Triage Scale used by paramedics*. Eur J Emerg Med, 2019. **26**(3): p. 188-193.
 20. Magnusson, C., et al., *The performance of the EMS triage (RETTs-p) and the agreement between the field assessment and final hospital diagnosis: a prospective observational study among children < 16 years*. BMC Pediatr, 2019. **19**(1): p. 500.
 21. Ng, C.J., et al., *Validation of the five-tier Taiwan Triage and Acuity Scale for prehospital use by Emergency Medical Technicians*. Emerg Med J, 2019. **36**(8): p. 472-478.
 22. Skjøt-Arkil, H., et al., *Do prehospital providers and emergency nurses agree on triage assignment?: an efficacy study*. Eur J Emerg Med, 2019. **26**(1): p. 29-33.
 23. Sung, S., et al., *Correlation between the pre-hospital triage scale and emergency department triage scale*. Hong Kong Journal of Emergency Medicine, 2019. **26**(5): p. 281-287.
 24. Magnusson, C., J. Herlitz, and C. Axelsson, *Pre-hospital triage performance and emergency medical services nurse's field assessment in an unselected patient population attended to by the emergency medical services: a prospective observational study*. Scand J Trauma Resusc Emerg Med, 2020. **28**(1): p. 81.
 25. Yamamoto, A., A. Kuriyama, and T. Ikegami, *Validity of a five-level prehospital triage system in Japan: A cohort study*. Am J Emerg Med, 2021. **45**: p. 329-334.
 26. Chiang, W.C., et al., *EMS in Taiwan: past, present, and future*. Resuscitation, 2009. **80**(1): p. 9-13.
 27. Huan, T.L., et al., *Emergency Medical Services in Taiwan: Past, Present, and Future*. J Acute Med, 2023. **13**(3): p. 91-103.
 28. Ng, C.J., et al., *Comparison between Canadian Triage and Acuity Scale and Taiwan Triage System in emergency departments*. J Formos Med Assoc, 2010. **109**(11): p. 828-37.
 29. Cheng, M.T., et al., *Physician gestalt for emergency department triage: A prospective videotaped study*. Acad Emerg Med, 2022. **29**(9): p. 1050-1056.
 30. Richardson, L.D. and U. Hwang, *Access to care: a review of the emergency medicine literature*. Acad Emerg Med, 2001. **8**(11): p. 1030-6.
 31. Robinson, D.J., *An integrative review: triage protocols and the effect on ED length of stay*. J Emerg Nurs, 2013. **39**(4): p. 398-408.
 32. Stey, A.M., et al., *Emergency Department Length of Stay and Mortality in Critically Injured Patients*. J Intensive Care Med, 2022. **37**(2): p. 278-287.
 33. Habib, H. and M.K. Sudaryo, *Association Between the Emergency Department Length of Stay and in-Hospital Mortality: A Retrospective Cohort Study*. Open Access

- 
- Emerg Med, 2023. **15**: p. 313-323.
34. Göransson, K. E., et al., *Thinking strategies used by Registered Nurses during emergency department triage*. J Adv Nurs, 2008. **61**(2): p. 163-72.
 35. Vatnøy, T.K., et al., *Triage assessment of registered nurses in the emergency department*. Int Emerg Nurs, 2013. **21**(2): p. 89-96.
 36. Yuzeng, S. and L.L. Hui, *Improving the wait time to triage at the emergency department*. BMJ Open Qual, 2020. **9**(1).
 37. Shah, M. N., *The formation of the emergency medical services system*. Am J Public Health, 2006. **96**(3): p. 414-23.
 38. Al-Shaqsi, S., *Models of International Emergency Medical Service (EMS) Systems*. Oman Med J, 2010. **25**(4): p. 320-3.
 39. Brewer, L. A., 3rd, *Baron Dominique Jean Larrey (1766-1842). Father of modern military surgery, innovater, humanist*. J Thorac Cardiovasc Surg, 1986. **92**(6): p. 1096-8.
 40. Dick, W. F., *Anglo-American vs. Franco-German emergency medical services system*. Prehosp Disaster Med, 2003. **18**(1): p. 29-35; discussion 35-7.
 41. Adnet, F. and F. Lapostolle, *International EMS systems: France*. Resuscitation, 2004. **63**(1): p. 7-9.
 42. Pozner, C. N., et al., *International EMS systems: The United States: past, present, and future*. Resuscitation, 2004. **60**(3): p. 239-44.
 43. Black, J. J. and G. D. Davies, *International EMS systems: United Kingdom*. Resuscitation, 2005. **64**(1): p. 21-9.
 44. Trevithick, S., et al., *International EMS systems: New South Wales, Australia*. Resuscitation, 2003. **59**(2): p. 165-70.
 45. Langhelle, A., et al., *International EMS Systems: the Nordic countries*. Resuscitation, 2004. **61**(1): p. 9-21.
 46. Ong, M. E., et al., *Comparison of emergency medical services systems in the pan-Asian resuscitation outcomes study countries: Report from a literature review and survey*. Emerg Med Australas, 2013. **25**(1): p. 55-63.
 47. MacFarlane, C., C. van Loggerenberg, and W. Kloeck, *International EMS systems in South Africa—past, present, and future*. Resuscitation, 2005. **64**(2): p. 145-8.
 48. Babaei, H. A., et al., *A comparative study on specialized services in pre-hospital emergencies in Iran and selected countries*. Journal of Education and Health Promotion, 2023. **12**(1).
 49. Bhattarai, H. K., et al., *Prehospital Emergency Care in Low- and Middle-Income Countries: A Systematic Review*. Prehosp Disaster Med, 2023. **38**(4): p. 495-512.
 50. Chen, C. Y., et al., *"Hearing AED integration program" – A innovative way to activate the emergency service system of community*. Resuscitation, 2022. **181**: p. 10-11.
 51. Lee, A. F., et al., *Effect of Placement of a Supraglottic Airway Device vs*

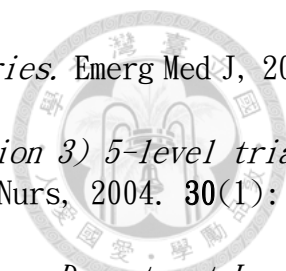
- Endotracheal Intubation on Return of Spontaneous Circulation in Adults With Out-of-Hospital Cardiac Arrest in Taipei, Taiwan: A Cluster Randomized Clinical Trial*. JAMA Netw Open, 2022. **5**(2): p. e2148871.
52. Lin, H.Y., et al., *Outcomes of out-of-hospital cardiac arrests after a decade of system-wide initiatives optimising community chain of survival in Taipei city*. Resuscitation, 2022. **172**: p. 149-158.
 53. Yu, J.H., et al., *Impact of the COVID-19 pandemic on emergency medical service response to out-of-hospital cardiac arrests in Taiwan: a retrospective observational study*. Emerg Med J, 2021. **38**(9): p. 679-684.
 54. Blagg, C.R., *Triage: Napoleon to the present day*. Journal of nephrology, 2004. **17**(4): p. 629-632.
 55. Frykberg, E.R., *Triage: principles and practice*. Scand J Surg, 2005. **94**(4): p. 272-8.
 56. Bhalla, M.C., et al., *Simple Triage Algorithm and Rapid Treatment and Sort, Assess, Lifesaving, Interventions, Treatment, and Transportation mass casualty triage methods for sensitivity, specificity, and predictive values*. Am J Emerg Med, 2015. **33**(11): p. 1687-91.
 57. Burkle, F.M., Jr., et al., *Emergency medicine in the Persian Gulf War--Part 2. Triage methodology and lessons learned*. Ann Emerg Med, 1994. **23**(4): p. 748-54.
 58. Burkle, F.M., Jr., S. Orebaugh, and B.R. Barendse, *Emergency medicine in the Persian Gulf War--Part 1: Preparations for triage and combat casualty care*. Ann Emerg Med, 1994. **23**(4): p. 742-7.
 59. Moskop, J.C. and K.V. Iserson, *Triage in medicine, part II: Underlying values and principles*. Ann Emerg Med, 2007. **49**(3): p. 282-7.
 60. FitzGerald, G., et al., *Emergency department triage revisited*. Emerg Med J, 2010. **27**(2): p. 86-92.
 61. Tam, H.L., S.F. Chung, and C.K. Lou, *A review of triage accuracy and future direction*. BMC Emerg Med, 2018. **18**(1): p. 58.
 62. M, J.M., *The Canadian Triage and Acuity Scale: A Canadian perspective on emergency department triage*. Emerg Med (Fremantle), 2003. **15**(1): p. 6-10.
 63. Mirhaghi, A., et al., *The Reliability of the Canadian Triage and Acuity Scale: Meta-analysis*. N Am J Med Sci, 2015. **7**(7): p. 299-305.
 64. Lee, J.H., et al., *Over-triage occurs when considering the patient's pain in Korean Triage and Acuity Scale (KTAS)*. PLoS One, 2019. **14**(5): p. e0216519.
 65. Jiménez, J.G., et al., *Implementation of the Canadian Emergency Department Triage and Acuity Scale (CTAS) in the Principality of Andorra: Can triage parameters serve as emergency department quality indicators?* Cjem, 2003. **5**(5): p. 315-22.
 66. Lee, J.Y., et al., *The validity of the Canadian Triage and Acuity Scale in predicting resource utilization and the need for immediate life-saving*

- interventions in elderly emergency department patients.* Scand J Trauma Resusc Emerg Med, 2011. **19**: p. 68.
67. Gravel, J., et al., *The Canadian Triage and Acuity Scale for Children: A Prospective Multicenter Evaluation.* Annals of Emergency Medicine, 2012. **60**(1): p. 71-77.e3.
 68. *The Australasian Triage Scale.* Emerg Med (Fremantle), 2002. **14**(3): p. 335-6.
 69. Chamberlain, D.J., et al., *Identification of the severe sepsis patient at triage: a prospective analysis of the Australasian Triage Scale.* Emerg Med J, 2015. **32**(9): p. 690-7.
 70. Wuerz, R.C., et al., *Implementation and refinement of the emergency severity index.* Acad Emerg Med, 2001. **8**(2): p. 170-6.
 71. Sax, D.R., et al., *Evaluation of Version 4 of the Emergency Severity Index in US Emergency Departments for the Rate of Mistriage.* JAMA Network Open, 2023. **6**(3): p. e233404-e233404.
 72. Chmielewski, N. and J. Moretz, *ESI Triage Distribution in U.S. Emergency Departments.* Advanced Emergency Nursing Journal, 2022. **44**(1): p. 46-53.
 73. Aeimchanbanjong, K. and U. Pandee, *Validation of different pediatric triage systems in the emergency department.* World J Emerg Med, 2017. **8**(3): p. 223-227.
 74. Mirhaghi, A., et al., *Reliability of the Emergency Severity Index: Meta-analysis.* Sultan Qaboos Univ Med J, 2015. **15**(1): p. e71-7.
 75. Gräff, I., et al., *The German Version of the Manchester Triage System and its quality criteria--first assessment of validity and reliability.* PLoS One, 2014. **9**(2): p. e88995.
 76. Zachariasse, J.M., et al., *Validity of the Manchester Triage System in emergency care: A prospective observational study.* PLoS One, 2017. **12**(2): p. e0170811.
 77. Parenti, N., et al., *A systematic review on the validity and reliability of an emergency department triage scale, the Manchester Triage System.* Int J Nurs Stud, 2014. **51**(7): p. 1062-9.
 78. Twomey, M., et al., *The South African Triage Scale (adult version) provides reliable acuity ratings.* Int Emerg Nurs, 2012. **20**(3): p. 142-50.
 79. Meyer, G.D., T.N. Meyer, and C.B. Gaunt, *Validity of the South African Triage Scale in a rural district hospital.* Afr J Emerg Med, 2018. **8**(4): p. 145-149.
 80. Dixon, J., et al., *Using the South African Triage Scale for prehospital triage: a qualitative study.* BMC Emerg Med, 2021. **21**(1): p. 125.
 81. Mould-Millman, N.K., et al., *Validity and reliability of the South African Triage Scale in prehospital providers.* BMC Emerg Med, 2021. **21**(1): p. 8.
 82. Markussen, D.L., et al., *Validation of a modified South African triage scale in a high-resource setting: a retrospective cohort study.* Scand J Trauma Resusc Emerg Med, 2023. **31**(1): p. 13.
 83. Wangara, A.A., et al., *Implementation and performance of the South African*

- Triage Scale at Kenyatta National Hospital in Nairobi, Kenya.* Int J Emerg Med, 2019. **12**(1): p. 5.
84. Dalwai, M., et al., *Is the South African Triage Scale valid for use in Afghanistan, Haiti and Sierra Leone?* BMJ Glob Health, 2017. **2**(2): p. e000160.
 85. Dalwai, M., et al., *Inter-rater and intrarater reliability of the South African Triage Scale in low-resource settings of Haiti and Afghanistan.* Emerg Med J, 2018. **35**(6): p. 379–383.
 86. Conti, A., et al., *Implementation of the South African Triage Scale (SATS) in a New Ambulance System in Beira, Mozambique: A Retrospective Observational Study.* Int J Environ Res Public Health, 2022. **19**(16).
 87. Bazzyar, J., et al., *The Principles of Triage in Emergencies and Disasters: A Systematic Review.* Prehosp Disaster Med, 2020. **35**(3): p. 305–313.
 88. Guan, G., et al., *The use of early warning system scores in prehospital and emergency department settings to predict clinical deterioration: A systematic review and meta-analysis.* PLoS One, 2022. **17**(3): p. e0265559.
 89. Saberian, P., et al., *Comparing the prehospital NEWS with in-hospital ESI in predicting 30-day severe outcomes in emergency patients.* BMC Emerg Med, 2022. **22**(1): p. 42.
 90. Wei, S., et al., *The accuracy of the National Early Warning Score 2 in predicting early death in prehospital and emergency department settings: a systematic review and meta-analysis.* Ann Transl Med, 2023. **11**(2): p. 95.
 91. Smith, K.E., K. Lobay, and M. Bullard, *PI22: The use of decision support tools in the implementation of the Prehospital Canadian Triage Acuity Score (Pre-CTAS).* CJEM, 2016. **18**(S1): p. S119–S119.
 92. Ghanbarzehi, N., et al., *Effect of Triage Training on Concordance of Triage Level between Triage Nurses and Emergency Medical Technicians.* J Clin Diagn Res, 2016. **10**(5): p. Ic05–ic07.
 93. Chang, Y.C., et al., *Effectiveness of a five-level Paediatric Triage System: an analysis of resource utilisation in the emergency department in Taiwan.* Emerg Med J, 2013. **30**(9): p. 735–9.
 94. Liu, T.T., et al., *Validation of a five-level triage system in pediatric trauma and the effectiveness of triage nurse modification: A multi-center cohort analysis.* Front Med (Lausanne), 2022. **9**: p. 947501.
 95. Sung, S.F., et al., *Validity of a computerised five-level emergency triage system for patients with acute ischaemic stroke.* Emerg Med J, 2013. **30**(6): p. 454–8.
 96. Ng, C.J., et al., *Predictive factors for hospitalization of nonurgent patients in the emergency department.* Medicine (Baltimore), 2016. **95**(26): p. e4053.
 97. Chang, W., et al., *Using the Five-Level Taiwan Triage and Acuity Scale Computerized System: Factors in Decision Making by Emergency Department Triage Nurses.* Clin Nurs Res, 2017. **26**(5): p. 651–666.

98. Lim, Y.D., et al., *Validity of the Korean Triage and Acuity Scale for predicting 30-day mortality due to severe trauma: a retrospective single-center study.* Eur J Trauma Emerg Surg, 2020. **46**(4): p. 895–901.
99. Chung, H.S., et al., *Validity of the Korean triage and acuity scale in older patients compared to the adult group.* Exp Gerontol, 2023. **175**: p. 112136.
100. Park, J.B., et al., *Reliability of Korean Triage and Acuity Scale: Interrater Agreement between Two Experienced Nurses by Real-Time Triage and Analysis of Influencing Factors to Disagreement of Triage Levels.* J Korean Med Sci, 2019. **34**(28): p. e189.
101. Lee, B., et al., *Clinical Considerations When Applying Vital Signs in Pediatric Korean Triage and Acuity Scale.* J Korean Med Sci, 2017. **32**(10): p. 1702–1707.
102. Kwon, H., et al., *The Korean Triage and Acuity Scale: associations with admission, disposition, mortality and length of stay in the emergency department.* Int J Qual Health Care, 2019. **31**(6): p. 449–455.
103. Bae, W., et al., *Predictive Validity of a New Triage System for Outcomes in Patients Visiting Pediatric Emergency Departments: A Nationwide Study in Korea.* Children (Basel), 2023. **10**(6).
104. Chung, H.S., et al., *Validation of the Korean Version of the Clinical Frailty Scale-Adjusted Korean Triage and Acuity Scale for Older Patients in the Emergency Department.* Medicina (Kaunas), 2024. **60**(6).
105. Moon, S.H., et al., *Triage accuracy and causes of mistriage using the Korean Triage and Acuity Scale.* PLoS One, 2019. **14**(9): p. e0216972.
106. Kuriyama, A., T. Kaihara, and T. Ikegami, *Validity of the Japan Acuity and Triage Scale in elderly patients: A cohort study.* Am J Emerg Med, 2019. **37**(12): p. 2159–2164.
107. Koyama, T., et al., *A study of the effect of introduction of JTAS in the emergency room.* Acute Med Surg, 2017. **4**(3): p. 262–270.
108. Kuriyama, A., et al., *Validity of the Japan Acuity and Triage Scale in adults: a cohort study.* Emerg Med J, 2018. **35**(6): p. 384–388.
109. Hamamoto, J., H. Yamase, and Y. Yamase, *Impacts of the introduction of a triage system in Japan: a time series study.* Int Emerg Nurs, 2014. **22**(3): p. 153–8.
110. Takahashi, T., et al., *'Down-triage' for children with abnormal vital signs: evaluation of a new triage practice at a paediatric emergency department in Japan.* Emerg Med J, 2016. **33**(8): p. 533–7.
111. Funakoshi, H., et al., *Validation of the modified Japanese Triage and Acuity Scale-based triage system emphasizing the physiologic variables or mechanism of injuries.* Int J Emerg Med, 2016. **9**(1): p. 1.
112. Rose, L., et al., *Emergency Department Length of Stay for Critical Care Admissions. A Population-based Study.* Ann Am Thorac Soc, 2016. **13**(8): p. 1324–32.
113. Sprivulis, P.C., et al., *The association between hospital overcrowding and*

- mortality among patients admitted via Western Australian emergency departments.* Med J Aust, 2006. **184**(5): p. 208-12.
114. Jones, S., et al., *Association between delays to patient admission from the emergency department and all-cause 30-day mortality.* Emerg Med J, 2022. **39**(3): p. 168-173.
 115. Sun, B.C., et al., *Effect of emergency department crowding on outcomes of admitted patients.* Ann Emerg Med, 2013. **61**(6): p. 605-611.e6.
 116. Deroose, S.F., et al., *Emergency department crowding predicts admission length-of-stay but not mortality in a large health system.* Med Care, 2014. **52**(7): p. 602-11.
 117. Maynou, L., et al., *Factors associated with longer wait times, admission and reattendances in older patients attending emergency departments: an analysis of linked healthcare data.* Emerg Med J, 2023. **40**(4): p. 248-256.
 118. Guttman, A., et al., *Association between waiting times and short term mortality and hospital admission after departure from emergency department: population based cohort study from Ontario, Canada.* Bmj, 2011. **342**: p. d2983.
 119. Levis-Elmelech, T., D. Schwartz, and Y. Bitan, *The effect of emergency department nurse experience on triage decision making.* Human Factors in Healthcare, 2022. **2**: p. 100015.
 120. McLeod, S.L., et al., *Interrater Reliability, Accuracy, and Triage Time Pre- and Post-implementation of a Real-Time Electronic Triage Decision-Support Tool.* Ann Emerg Med, 2020. **75**(4): p. 524-531.
 121. Koo, T.K. and M.Y. Li, *A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research.* J Chiropr Med, 2016. **15**(2): p. 155-63.
 122. van der Wulp, I. and H.F. van Stel, *Adjusting weighted kappa for severity of mistriage decreases reported reliability of emergency department triage systems: a comparative study.* J Clin Epidemiol, 2009. **62**(11): p. 1196-201.
 123. Tu, Y.K. and M.S. Gilthorpe, *Key statistical and analytical issues for evaluating treatment effects in periodontal research.* Periodontol 2000, 2012. **59**(1): p. 75-88.
 124. Worster, A., et al., *Triage tool inter-rater reliability: a comparison of live versus paper case scenarios.* J Emerg Nurs, 2007. **33**(4): p. 319-23.
 125. Lindskou, T.A., et al., *Prehospital Early Warning Scores to Predict Mortality in Patients Using Ambulances.* JAMA Netw Open, 2023. **6**(8): p. e2328128.
 126. Martin-Rodriguez, F., et al., *Comparison of seven prehospital early warning scores to predict long-term mortality: a prospective, multicenter, ambulance-based study.* Eur J Emerg Med, 2023. **30**(3): p. 193-201.
 127. Tanabe, P., et al., *Reliability and validity of scores on The Emergency Severity Index version 3.* Acad Emerg Med, 2004. **11**(1): p. 59-65.
 128. Candel, B.G., et al., *The association between vital signs and clinical outcomes*

- 
- in emergency department patients of different age categories. Emerg Med J, 2022. 39(12): p. 903-911.*
129. Tanabe, P., et al., *The Emergency Severity Index (version 3) 5-level triage system scores predict ED resource consumption. J Emerg Nurs, 2004. 30(1): p. 22-9.*
130. Chaou, C.H., et al., *Analyzing Factors Affecting Emergency Department Length of Stay-Using a Competing Risk-accelerated Failure Time Model. Medicine (Baltimore), 2016. 95(14): p. e3263.*
131. Chaou, C.H., et al., *Predicting Length of Stay among Patients Discharged from the Emergency Department-Using an Accelerated Failure Time Model. PLoS One, 2017. 12(1): p. e0165756.*
132. Lin, Y.-Y. and C.-S. Huang, *Aging in Taiwan: Building a Society for Active Aging and Aging in Place. The Gerontologist, 2015. 56(2): p. 176-183.*