*<Original Article>*

# Effects of Introducing the Emergency Severity Index in a Japanese Emergency Department: A Controlled Before-After Study

Koichi TAKAOKA[1], Masaya ONO[2], and Chiharu AKAZAWA[3]

*1 Graduate School of Nursing, Osaka Medical and Pharmaceutical University, Takatsuki, Osaka, Japan*
*2 Social Welfare Organization Saiseikai Imperial Gift Foundation, Inc. Saiseikai Hyogo-ken Hospital, Kobe, Hyogo, Japan*
*3 Department of Nursing, Faculty of Nursing, Osaka Medical and Pharmaceutical University, Takatsuki, Osaka, Japan*

Key words: triage system, emergency department triage, clinical decision, education

**ABSTRACT**

**Introduction:** The objective of this study was to examine whether the Emergency Severity Index (ESI) can be introduced into the emergency department (ED).

**Methods:** This study used a pre-post intervention design. Data were collected from patients who visited the ED after hours and had their urgency assessed using the facility's existing five-level triage system during a three-month period before the intervention. Over a one-month period, the participants underwent an ESI educational program. The ESI was then used for triage, and data were collected for another three months. The primary outcome was the inter-rater agreement between the triage nurse and the emergency physician on the urgency of patients visiting the ED. Secondary outcomes included the rates of under-triage (UT) and over-triage (OT).

**Results:** Two nurses and one emergency physician participated. The total number of triaged cases analyzed was 75 before and 77 after the intervention. Overall inter-rater agreement between the emergency physician and nurse was $\kappa$ (95 % confidence interval) = 0.35 (0.23–0.46 Fair) pre-intervention and 0.66 (0.48–0.83 Substantial) post-intervention. All severity levels showed higher inter-rater agreement after the intervention. There was a significant reduction and large effect size for UT post-intervention.

**Conclusion:** The ESI allows for more accurate triage decisions.

## INTRODUCTION

The yearly increase in patient visits to emergency departments (EDs) has raised concerns about the deterioration of patients' conditions while waiting for a physician's assessment. As a countermeasure, triage, assessing the urgency of a patient's condition before a physician's examination and identifying patients who require immediate care, has been promoted [1]. Triage involves categorizing patients based on degree of severity to prioritize those with

Address correspondence to:
Koichi Takaoka, Graduate School of Nursing, Osaka Medical and Pharmaceutical University, 2-7 Daigaku-machi, Takatsuki, Osaka 569-8686, Japan
Email: verydislikecarrotshiitake@yahoo.co.jp

the highest urgency and ensure timely physician assessment. Initially, a three-level triage system was standard, but following recommendations from the American College of Emergency Physicians in 2010, a more reliable and valid five-level triage system became the standard [2]. This triage system is typically performed by triage nurses, who are the first to interact with patients in the ED. To enhance the standardization and consistency of triage and improve patient safety, triage tools have been introduced to assist in triage decision-making, contributing to improved reliability and validity [3]. In line with international trends, Japan has seen the widespread adoption of triage tools, starting in 2012 with the Japan Triage and Acuity Scale (JTAS), developed based on the Canadian Triage and Acuity Scale (CTAS) [4]. However, despite the progress in adopting triage systems in Japan, there remain challenges in fully implementing the five-level system and ensuring the thorough integration of post-education programs [5]. Furthermore, a survey of Japanese emergency centers indicated that the lack of triage education systems, insufficient manpower (both physicians and nurses), and the cost of implementing triage tools were cited as reasons for not adopting triage tools [6]. In other words, one of the challenges in the ED in Japan is the lack of implementation of triage tools in many facilities. Instead, these facilities rely on their own criteria to determine the level of urgency, which has exposed issues regarding the reliability and validity of the triage outcomes. However, the same survey found that, even in facilities without triage tools, triage nurses expressed a willingness to learn if given the opportunity. Considering these issues in Japan, the present study focused on the Emergency Severity Index (ESI) Version 4, an internationally recognized system.

The ESI is a reliable and valid international triage tool that, in addition to the traditional clinical reasoning-based patient evaluation, assesses the urgency of the patient's condition using its own algorithm [7]. The ESI enables triage nurses to leverage their expertise, potentially reducing waiting times for physician assessments, and creating new opportunities for nurses to play a more active role. Moreover, the ESI can be introduced without the need for specialized equipment or devices, reducing the cost of implementation.

A previous study comparing the accuracy of the ESI with mock patients found that it more accurately determined patient urgency than the JTAS, which is the most widely used tool in Japan, suggesting its potential applicability in the ED [8]. The ESI's ability to provide more precise assessments of patient urgency suggests it could contribute to the widespread adoption of triage tools, particularly in facilities that have hesitated to implement them.

The objective of this study was to introduce the internationally recognized ESI to facilities that have not previously implemented any triage tool and to evaluate its effectiveness based on changes in triage accuracy, waiting time for medical consultation, and length of stay in the ED. This initiative aims to promote the adoption of the internationally standardized five-level triage tool in emergency medical practice. By implementing a reliable and valid triage system, it seeks to contribute to the standardization of emergency patient care.

## MATERIALS AND METHODS

### *Ethics approval*

This study was planned in accordance with the Declaration of Helsinki and Japan's "Ethical Guidelines for Medical and Health Research Involving Human Subjects," and approval was obtained from the Research Ethics Committee of Osaka Medical and Pharmaceutical University (Approval No.: 2023-017). In addition, the study was registered with the University Hospital Medical Information Network (UMIN) (Trial ID: UMIN000052596).

Informed consent was obtained from the study participants after explaining the study content both orally and in writing. For patients visiting the ED where the study took place, an opt-out notice regarding the study was posted in the ED and on the hospital's website.

### *Participants*

The participants in this study were triage nurses and emergency physicians working in the ED. Initially, four nurses from the 24 employed in the ED, who were involved in triage, and one physician from the 44 ED physicians were selected based on selection criteria. Then, the objectives of the study were explained to them verbally. Subsequently, consent was obtained from two nurses and one emergency physician.

Selection criteria included more than three but fewer than 20 years of professional experience, and routine handling of emergency cases at night or on holidays. Exclusion criteria included those who had no prior experience with triage using tools or other reasons deemed unsuitable by the research investigators.

### *Trial design*

This study used a pre-post intervention design. Data were collected during a specific period before the intervention from patients who visited the ED facility at night or on holidays. In the workflow of this study, patients visiting the emergency department were first assessed by a triage nurse, who determined the triage level, including information gathered from the patient history. The triage nurse then communicated findings, excluding the triage level, to the emergency physician. The emergency physician independently determined the triage level based on the history-taking results, following a medical examination. The facility had no systematic triage tool such as JTAS, but the own five-level triage tool assessed by participants. In this triage

system, Level 1 was assigned to patients requiring immediate resuscitation, Level 2 to those with severe conditions such as myocardial infarction or suspected pneumonia with sepsis, Level 3 to those with gastrointestinal symptoms or fractures, Level 4 to those with minor injuries or conditions like a laceration, and Level 5 to non-urgent cases. Prior to the intervention, neither the two triage nurses nor the emergency physician conducted retrospective reviews of the triage results, and no opportunities for reflection were provided.

During the one-month period following the end of the pre-intervention, the participants underwent the intervention, an ESI educational program. The program was led by certified ESI instructors, including a nurse educator who served as the principal investigator and a critical care certified nurse. It lasted six hours and covered topics including an overview of the ESI, guidelines for each triage level, and pediatric triage. Following the educational intervention, the ESI was used for triage, and data were collected for another three months. The participants independently determined the urgency levels without consulting with other medical staff, both before and after the intervention.

***Research period***

The pre-intervention period was from November 2023 to January 2024, and the post-intervention period was from March to May 2024. The analysis focused on triage during the nine pre-intervention days and 12 post-intervention days when one of the two triage nurses and one emergency physician were actively participating in the triage process. A total of 140 emergency patients pre-intervention and 129 post-intervention visited the hospital specifically during these periods of the study.

***Emergency department situation***

The study was conducted at an emergency medical facility in Japan that handles both emergency transport and walk-in patients on evenings and holidays. The study was conducted in a region with a population of approximately 210,000, where the number of emergency transports in 2023 was 10,129 cases per year. The study facility employs a total of 24 ED nurses, with an average of two nurses on duty during nighttime and holiday shifts responsible for triage and patient management. Medical examinations are conducted by on-call physicians from internal medicine and surgery departments. The annual number of emergency outpatient visits is 6,437, including 2,304 transported by ambulance, and 2,327 hospital admissions originate from the ED. As a secondary emergency facility hospital, the facility plays a central role in the region's emergency medical services.

This study has been done in only hospital with its own approach to triage.

***Outcomes***

The primary outcome of this study was the inter-rater agreement between the triage nurse and emergency physician on the urgency of patients visiting the ED, measured using the weighted kappa coefficient before and after the intervention. Secondary outcomes included the rates of under-triage (UT) and over-triage (OT) and the time from triage initiation to physician assessment (examination response time) and from patient check-in to discharge (ED length of stay). As the evidence for UT and OT, all triage decisions made during the research period were retrospectively reviewed and finalized by a panel of three experts: two individuals certified as ESI providers (one emergency nursing expert and one triage educator) and one emergency physician.

***Statistical methods***

Descriptive statistics were calculated for basic information such as chief complaints, number of patient visits, examination response time, and ED length of stay. Differences between groups were compared using the Mann-Whitney U test (significance level: $p < 0.05$). In addition to significance testing, effect sizes ($r$), which represent the standardized group difference unaffected by sample size, were also calculated. The effect sizes were classified as $r \geq 0.10$ (small effect), $r \geq 0.30$ (medium effect), and $r \geq 0.50$ (large effect) [9].

Inter-rater agreement was assessed using weighted kappa ($\kappa$) coefficients to account for the ordering of categories and provide a measure of reliability beyond chance. Kappa values were interpreted as follows, based on previous literature: 0.81–1.00 Almost perfect; 0.61–0.80 Substantial; 0.41–0.60 Moderate; 0.21–0.40 Fair; 0.00–0.20 Slight; and less than 0.00 No agreement.

Next, OT and UT rates were calculated and compared before and after the intervention using odds ratios (ORs). Data were analyzed using IBM SPSS Statistics ver. 27.0.0.

***Sample size***

The sample size of 150 was calculated based on previous studies to ensure the minimum number of triage case required, with a significance level of 5 %, power of 80 %, and a non-inferiority margin of 0.2 or higher.

## RESULTS

***Participants***

The study was conducted in the ED of a hospital that receives emergency patients during nights and holidays. Two nurses who regularly perform triage and one emergency physician, all of whom met the inclusion criteria and provided consent, participated in the study (**Table 1**).

The number of patient visits including transported patients by ambulance (in the parenthesis) during the study

**Table 1 Overview of study participants**

| | Triage Nurse A | Triage Nurse B | Physician |
|---|---|---|---|
| Age | 48 | 36 | 55 |
| Years of experience in the profession | 19 | 13 | 19 |
| Years of experience in emergency department | 9 | 2 | 19 |
| Years of experience in triage | 0.5 | 2 | 0.5 |
| Experience using triage tools | none | none | none |

period number was 140 (20) before the intervention and 129 (18) after ($p$ = 0.94). Of them, the number of triaged cases was 75 (15) before and 80 (16) after the intervention ($p$ = 1.00). In terms of triage severity, there were 0 (0 %) cases at Level 1, 10 (13.3 %) at Level 2 (3 admitted and 7 discharged; hospital admission rate: 30.0 %), 24 (32.0 %) at Level 3 (11 admitted and 13 discharged; hospital admission rate: 45.8 %), 39 (52.0 %) at Level 4 (3 admitted and 36 discharged; hospital admission rate: 7.7 %), and 2 (2.7 %) at Level 5 (0 admitted and 2 discharged; hospital admission rate: 0.0 %) before the intervention. After the intervention, there were 3 (3.7 %) cases at Level 1 (3 admitted and 0 discharged; hospital admission rate: 100.0 %), 18 (22.5 %) at Level 2 (7 admitted and 11 discharged; hospital admission rate: 63.6 %), 25 (31.3 %) at Level 3 (12 admitted and 13 discharged; hospital admission rate: 48.0 %), 22 (27.5 %) at Level 4 (0 admitted and 22 discharged; hospital admission rate: 0.0 %), and 12 (15.0 %) at Level 5 (0 admitted and 12 discharged; hospital admission rate: 0.0 %). Given the clear severity of Level 1 cases, they were excluded from the analysis.

Chief complaints included altered consciousness (9 cases pre-intervention, 4 post-intervention), gastrointestinal symptoms (12 pre, 12 post), respiratory and circulatory symptoms (42 pre, 45 post), orthopedic symptoms (8 pre, 6 post), and minor symptoms such as dermatological or ophthalmological issues (4 pre, 10 post).

***Interrater agreement***

The total number of triaged cases analyzed was 75 before and 77 after the intervention. The overall inter-rater agreement between the emergency physician and nurse was $\kappa$ (95 % CI) = 0.35 (0.23–0.46, Fair) pre-intervention and 0.66 (0.48–0.83, Substantial) post-intervention, showing higher agreement after the ESI educational intervention (**Table 2**).

Regarding severity-specific inter-rater agreement, for Level 2, $\kappa$ was 0.12 (0.13–0.32, Slight) pre-intervention and 0.54 (0.23–0.67, Moderate) post-intervention. For Level 3, $\kappa$ was 0.26 (0.16–0.29, Fair) pre-intervention and 0.87 (0.59–1.03, Almost perfect) post-intervention. For Level 4, $\kappa$ was 0.24 (0.17–0.26, Fair) pre-intervention and 0.87 (0.60–1.04, Almost perfect) post-intervention. For Level 5, $\kappa$ was 0.05 (−0.10–0.12, Slight) pre-intervention and 0.86 (0.61–1.05, Almost perfect) post-intervention. All severity levels showed higher inter-rater agreement after the intervention (**Table 3**).

**Table 2 Cohen's Kappa coefficient of the inter-rater agreement between physician and triage nurses before and after intervention**

| inteivention | Cohen's $\kappa$ (95 % CI) | category |
|---|---|---|
| Before (n = 75) | 0.35 (0.23–0.46) | Fair |
| After (n = 77) | 0.66 (0.48–0.83) | Substantial |

Note. 95 % CI = 95 % Confidence Interval.

***Under-triage and over-triage***

The UT rate was 75.6 % (59 cases) pre-intervention and 11.5 % (9 cases) post-intervention. The OT rate was 2.7 % (2 cases) pre-intervention and 9.0 % (7 cases) post-intervention. When comparing pre-intervention and post-intervention rates using ORs, the OR (95 % CI) for UT was 0.34 (0.14–0.83; $p < 0.01$, $r$ = 0.68), showing a significant reduction and large effect size for UT after the ESI intervention.

OT increased post-intervention, with an ORs of 3.50 (0.70–17.40; $p$ = 0.17, $r$ = 0.13), but the difference was not significant, with a small effect size.

***Examination response time and length of stay***

The median (IQR) examination response time was 10.0 (5.0–20.0) minutes pre-intervention and 12.0 (7.0–32.0) minutes post-intervention, with no significant difference between groups and a small effect size ($p$ = 0.46, $r$ = 0.04). The ED length of stay was 50.0 (25.0–78.0) minutes pre-intervention and 52.0 (31.0–77.0) minutes post-intervention, again with no significant difference and a small effect size ($p$ = 0.80, $r$ = 0.01).

## DISCUSSION

This study evaluated the effects of implementing the

**Table 3 Fleiss's $\kappa$ comparison of the inter-rater agreement between physician and triage nurses before and after intervention**

| | Before (n = 75) | | After (n = 77) | |
|---|---|---|---|---|
| Rating Category | $\kappa$ (95 % CI) | Category | $\kappa$ (95 % CI) | Category |
| 2 | 0.12 (0.13, 0.32) | Slight | 0.54 (0.23, 0.67) | Moderate |
| 3 | 0.26 (0.16, 0.29) | Fair | 0.87 (0.59, 1.03) | Almost perfect |
| 4 | 0.24 (0.17, 0.26) | Fair | 0.87 (0.60, 1.24) | Almost perfect |
| 5 | 0.05 (−0.10, 0.12) | Slight | 0.86 (0.61, 1.05) | Almost perfect |

Note. 95 % CI = 95 % Confidence Interval.

ESI, an international five-level triage system, in an ED where triage tools had not previously been introduced. By comparing triage decision accuracy between the pre-intervention and post-intervention phases, the inter-rater agreement between emergency physician and triage nurses was assessed.

The overall inter-rater agreement between emergency physician and triage nurses was $\kappa$ = 0.35 pre-intervention and $\kappa$ = 0.66 post-intervention, showing improved agreement after the intervention. Previous meta-analyses of 260 cases using the ESI reported an overall inter-rater agreement of $\kappa$ = 0.79, which is consistent with the post-intervention results of the present study [10]. Another study of ESI use in a tertiary emergency center reported inter-rater agreement between triage nurses and physician of $\kappa$ = 0.60 [11]. These results suggest that ESI use in Japan can achieve high inter-rater agreement, comparable to that of other countries already using the ESI.

Next, when examining agreement by severity level, high levels of agreement were observed at all levels post-intervention. Level 2 showed significantly higher agreement post-intervention. This criterion is often the most critical and difficult to assess [12]. Levels 3 to 5, however, showed nearly perfect agreement. These demonstrate indicating that triage nurses could accurately predict the level of medical resources a patient would require based on the ESI's methodology, which factors in the number of medical resources likely to be used [12].

UT rates decreased significantly after the intervention, with a high effect size, further supporting the reliability of the ESI for accurate triage decisions. Though OT rates increased, there was no significant difference, and the effect size was small. Previous studies reported UT rates of 12.2 % [13], indicating that the post-intervention results of this study were comparable to international benchmarks. One possible factor contributing to the reduction in UT is that the ESI system enables decisions to be made not solely based on individual clinical experience but rather using a standardized algorithm and the number of required medical resources. This approach likely helped prevent the inappropriate assignment of actual emergency patients to lower urgency levels, thereby contributing to a decrease in UT.

The time metrics, such as examination response time and ED length of stay, showed no significant differences between the pre-intervention and post-intervention periods, suggesting that ESI implementation did not add to the time burden of triage. This aligns with the ESI manual, which emphasizes that the tool is designed to enhance reliability and validity without increasing time delays [11].

### *Limitations*

This study used a pre-post intervention design, and the possibility cannot be ruled out that participants' decisions and behaviors after the intervention were influenced by learning effects based on their prior clinical experience. Furthermore, since this study was conducted in a single facility and included only patients seen in the ED, its generalizability is limited. Biases may have arisen due to individual triage abilities and differences in patient demographics. In addition, the inability to fully match patient chief complaints before and after the intervention and seasonal factors may have affected the results. This study collected data from an actual emergency department; therefore, not all triage cases for emergency patients were captured, which may have influenced the triage outcomes and the results regarding emergency department length of stay. The small number of Level 1 cases also limits the ability to generalize the results to the most critical patients. Future studies should replicate this research in different facilities to further validate the accuracy of triage tools.

### *Conclusions*

This study compared the effects of implementing the ESI in a Japanese ED before and after an educational intervention. The results showed improved inter-rater agreement in all categories post-intervention. This suggests that ESI allows for more accurate triage decisions and could contribute to the dissemination of a reliable and valid five-level triage tool, meeting the increasing demand for accurate triage in the ED.

**FUNDING**

The authors would like to express their deep gratitude to Professor Tsuda and Professor Teraguchi for their helpful advice on how to conduct the research and the framework. This work was supported by JSPS KAKENHI [Grant No. JP 20K19121].

**DISCLOSURE STATEMENT**

The authors declare no conflicts of interest associated with this manuscript.

**REFERENCES**

1. Christ M, Grossmann F, Winter D, Bingisser R, Platz E. Modern triage in the emergency department. *Dtsch Arztebl Int.* 2010;107(50):892–898. doi: 10.3238/arztebl.2010.0892
2. The American College of Emergency Physicians. ACEP policy statements: Triage Scale Standardization: Joint Statement by the American College of Emergency Physicians and the Emergency Nurses Association. 2010. *Ann Emerg Med.* 2010;56:451. doi: 10.1016/j.annemergmed.2010.08.008
3. Welch SJ, Asplin BR, Stone-Griffith S, Davidson SJ, Augustine J, Schuur J; Emergency Department Benchmarking Alliance. Emergency department operational metrics, measures and definitions: results of the Second Performance Measures and Benchmarking Summit. *Ann Emerg Med.* 2011;58(1):33–40. doi: 10.1016/j.annemergmed.2010.08.040
4. Japanese Society for Emergency Medicine. *Japan Triage and Acuity Scale 2023 guidebook (3rd edition)*. Tokyo: Herusu Shuppan, Co. Inc; 2023.
5. Takaoka K, Kakeda T. Comparative study of international triage tools in emergency departments. *Emergency Care.* 2017;30(11):1142–1148.
6. Takaoka K, Tsuda Y, Akazawa C. The current situation and issues of triage tools in Emergency Department in Japan: survey by questionnaire. *Journal of Japanese Association for Emergency Nursing.* in press.
7. Tanabe P, Gimbel R, Yarnold PR, Kyriacou DN, Adams JG. Reliability and validity of scores on The Emergency Severity Index version 3. *Acad Emerg Med.* 2004;11(1):59–65. doi: 10.1197/j.aem.2003.06.013
8. Takaoka K, Ooya K, Ono M, Kakeda T. Utility of the emergency severity index by accuracy of interrater agreement by expert triage nurses in a simulated scenario in Japan: a randomized controlled trial. *J Emerg Nurs.* 2021;47(4):669–674. doi: 10.1016/j.jen.2021.03.009
9. Cohen J. *Statistical Power Analysis for the Behavioral sciences*. 2nd ed., Lawrence Erlbaum, Hillsdale; 1988: 273–406.
10. Mirhaghi A, Heydari A, Mazlom R, Hasanzadeh F. Reliability of the emergency severity index: meta-analysis. *Sultan Qaboos Univ Med J.* 2015; 15(1):e71–e77.
11. Busch JM, Arnold I, Kellett J, Brabrand M, Bingisser R, Nickel CH. Validation of a simple score for mortality prediction in a cohort of unselected emergency patients. *Int J Clin Pract.* 2022;2022:7281693. doi: 10.1155/2022/7281693
12. Gilboy N, Tanabe P, Travers D, Alexander M. Eds: Emergency Severity Index (ESI) (Implementation Handbook 2012 Edition). https://www.sgnor.ch/fileadmin/user_upload/Dokumente/Downloads/Esi_Handbook.pdf. Accessed: February 20, 2023
13. Razavian A, Mohamadirizi S, Heydari F, Nasr-Esfahani M. Evaluation of Emergency Severity Index (ESI) triage quality by nurses and associated factors in Iran. *J Educ Health Promot.* 2024;13:165. doi: 10.4103/jehp.jehp_1142_22

Received October 18, 2024
Accepted January 8, 2025