The Reliability and Accuracy of International Triage Scale in the Emergency Department (ED): A Literature Review

T. Abdur Rasyid, Cecep Eli Kosasih, dan Ristina Mirwanti
Fakultas Keperawatan, Universitas Padjadjaran
Email: abdur18002@mail.unpad.ac.id

Abstract

A large number of patient visits to the Emergency Department (ED) will influence the outcome of the services provided. The triage scale is one method designed to manage patient screening for quality service improvement. Several triage scales are employed internationally in the EDs including the Australasian Triage Scale (ATS), the Manchester Triage System (MTS), the Canadian Triage and Accuracy Scale (CTAS), and the Emergency Severity Index (ESI). Several studies have a concern to identify the reliability of the triage scale, but only a few of them identified the accuracy of the triage scale. The purpose of this literature review was to identify the best reliability and accuracy among ATS, MTS, CTAS, and ESI based on the literature. The literature search was conducted on electronic databases EBSCO and PubMed with keywords including (triage OR emergency department triage) AND reliability AND ((the Canadian Triage and Accuracy Scale OR CTAS) AND (the Australasian Triage Scale OR ATS) AND (the Manchester Triage System OR MTS) AND the Emergency Severity Index OR ESI)). Assessment of articles was composed based on the PRISMA format with criteria including primary research articles containing the reliability and accuracy of the triage scale in English and published between 2009 – 2019. A total of 271 publications were identified and only 10 studies were included in this literature review. The results reveal that ATS has a moderate level of reliability (k = 04 – 0.57) with an accuracy of 46.2% – 58.3%, CTAS has a good level reliability (k = 0.770) with accuracy of 49%, MTS have good to excellent level of reliability (k = 0.61 – 0.95) with accuracy of 49%, and ESI have moderate to excellent level reliability (k = 0.45 – 0.94) with accuracy of 59.6 % – 72.5%. Based on this review, MTS and ESI are the triage scale with the highest reliability and accuracy. Therefore, MTS and ESI are highly recommended in the ED. However, each EDs need to pay attention to the characteristics, culture, and available resources before choosing and implementing an appropriate triage scale.

Keywords: Accuracy, emergency department, reliability, triage scale

Introduction

The emergency department (ED) is one of the most essential components of the health care system in the hospital. Research has affirmed that about 50% of patient visits to the ED were non-urgent visits (Qureshi, 2010). This has an impact on overcrowding in the ED which leads to an increase in the length of stay in the ED and the hospital, delay in treatment and reduces patient satisfaction (Brouns, Mignot-Evers, Derkx, Lambooij, & Dieleman, et al., 2019). A comprehensive triage system has been designed to screen patients in the ED (Esmailian, Zamani, Azadi, & Ghasemi, 2014; Hinson, Martinez, Schmitz, Toerper, & Radu, et al., 2018).

Triage is defined as a rapid process of categorizing patients upon arrival to determine priorities for further evaluation and treatment (Gräff, Goldschmidt, Glien, Bogdanow, Fimmers, & Hoeft, et al., 2014; Saeed, Al-Fayyadh, Alshomar, Zekry, & Alamiri, et al, 2017). Most of the triage systems applied in the ED are based on a five-level triage where this triage system has been conferred to improve the patient care outcomes historically (Andrade-Silva, Takemura, Bellato, Leonhardt, & Kojima, et al., 2019; Ekins & Morphet, 2015). There are variations in the application of triage systems that are tailored to the availability of resources, the economic situation, and the capacity of patients in each EDs in the wide world (Saeed, Al-Fayyadh, Alshomar, Zekry, & Alamiri, et al., 2017). A triage system with five levels of urgency categories is the most widely adopted triage and published in various literature. The triage includes the Canadian Triage and Accuracy Scale (CTAS), the Australasian Triage Scale (ATS), the Manchester Triage System (MTS), and the Emergency Severity Index (ESI) (Andrade-Silva, Takemura, Bellato, Leonhardt, & Kojima, et al. , 2019; Hinson, Martinez, Schmitz, Toerper, & Radu, et al., 2018).

Inaccuracies or inconsistencies in triage decisions have serious consequences for patients and it has the potential to increase morbidity, mortality, and can also negatively affect resources in the ED (Ekins & Morphet, 2015). Consequently, many studies give special attention to the validity and reliability of the triage scale in the ED (Mirhaghi, Kooshiar, & Ebrahimi, 2015). ED requires a valid and reliable triage system in prioritizing patients based on clinical urgency so that ED staff can plan resources and time for appropriate and efficient treatment (Brouns, Mignot-Evers, Derkx, Lambooij, & Dieleman, et al., 2019). Hinson, Martinez, Schmitz, Toerper, and Radu, et al (2018) stated the accuracy and reliability of triage in the ED was the most essential thing.

Reliability is the main foundation for examining the quality of measuring instruments. Reliability confers the ability of the instrument to measure consistently and accurately what it wants to measure and can provide consistent results in space, time, and with different observers (Souza, Chianca, Juniper, Rausch, & Nascimento, 2018). The reliability of the instrument must be as high as possible so that the measurement method has adequate strength (Gräff, Goldschmidt, Glien, Bogdanow, Fimmers, & Hoeft, et al., 2014). Inter-rater reliability (IRR) is the common widely used method of examining the reliability of a triage scale (Jordi, Grossmann, Gaddis, Cignacco, & Denhaerynck, et al., 2015).

The accuracy of triage is defined as choosing the right level of urgency according to the actual level of urgency. There are three possible triage assessment decisions including undertriage, expected triage, and over triage (Health Policy Priorities Principal Committee, 2011). Undertriage is a failure to recognize and differentiate patients where they are placed in less urgent conditions. This contributes to the delay in time-sensitive interventions, clinical deterioration, morbidity, and mortality. Meanwhile, over triage is the placement of patients who less urgent to a higher level of urgency. Under and over triage are equally dangerous for patients (Hinson, Martinez, Schmitz, Toerper, & Radu, et al., 2018). This literature review aims to identify the reliability and accuracy of the best triage scales among CTAS, ATS, MTS, and ESI based on the literature.

Research methodology

This study uses a literature review method that aims to identify the reliability and

accuracy of the best triage scale based on the literature. Two databases EBSCO and PUBMED were used in searching relevant articles. The selection of literature used the PRISMA format. Some keywords employed include (triage OR emergency department triage) AND reliability AND ((Canadian Triage and Accuracy Scale OR CTAS) AND (Australasian Triage Scale OR ATS) AND (Manchester Triage System OR MTS) AND Emergency Severity Index OR ESI)).

The inclusion criteria in this literature review were: (1) Primary research article; (2) Research on the reliability and accuracy of the CTAS, ATS, MTS, and ESI; (3) Studies containing the determination of triage categories based on case scenarios, direct triage or triage audits in adult patients or mixed populations (children and adults); (4) Studies published in journals; (5) Studies published between 2009 and 2019 and; (6) Studies published in English.

Result

The literature searches generated 137 studies through searches of databases, excluding duplicates. After obtaining 58 full-text articles, we excluded 15 articles specifically included triage reliability in the pediatric population, 31 articles were review and meta-analyses studies, and 2 articles were non-English language. Eventually, 10 articles with a strong design decided to involve in this literature review. The research design on the articles obtained including 1 article with a retrospective study, 3 prospective studies, and 6 cross-sectional studies. The number of studies about international triage system reliability including the Australasian Triage Scale (ATS) (n = 2), the Manchester Triage System (MTS) (n = 3), the Canadian Triage and Accuracy Scale (CTAS) (n = 2) and the Emergency Severity Index (ESI) (n = 3). A total 7 from 10 articles that contained the accuracy of various triage methods including ATS (n = 2), MTS (n = 2), ESI (n = 2), and (CTAS = 1).

The Triage reliability in all articles was analyzed using inter-reliability reliability (IRR). It aims to measure the agreement between rater in allocating the level of triage. IRR analysis in 10 studies was carried out in 3 forms, particularly (1) Agreement between several raters (nurses versus nurses) (n = 6); (2) Agreement between the two raters (between nurse versus doctor) (n = 3); (3) Agreement between the two raters (between doctor versus doctor) (n = 1). There are 3 forms of triage implementation contained in the study include (1) Triage based on case scenarios that have been prepared and validated by experts previously (n = 6); (2) Triage directly to patients (n = 3) and; (3) Audit the implementation of triage based on the patient's medical record (n = 1).

The Reliability and Accuracy International Triage Scale in the

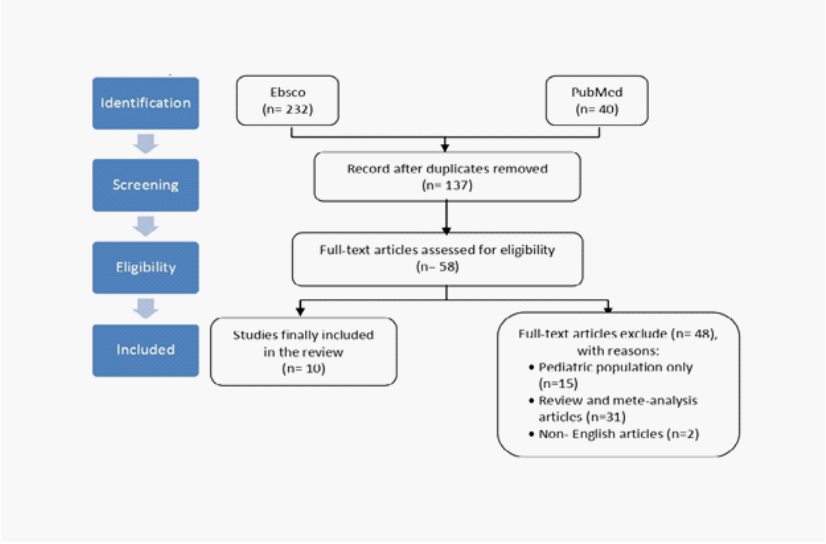

Diagram 1. PRISMA Format for Literature Search

Emergency Department

The triage system was first introduced by Dominique Jean Larrey, a French doctor in 1792 who tried to prioritize the medical needs of military victims during the first World War (Yuksen, Sawatmongkornkul, Suttabuth, Sawanyawisuth, & Sittichanbuncha, 2016; Fry & Burr, 2002; Dippenaar & Bruijns, 2016). This system eventually became the concern of civil health care providers to be employed in civil hospital EDs. During the late 1970s and early 1980s, EDs began developing, implementing and reviewing their triage systems (FitzGerald, Jelinek, Scott, & Gerdtz, 2010). The various triage instruments currently available are designed in various methods such as using color codes, scales with three, four, or five levels of urgency (Christ, Grossmann, Winter, Bingisser, & Platz, 2010). Triage with 5 levels of urgency received more favorable acceptance because it proves to be more reliable (Iserson & Moskop, 2007). This triage system can also significantly predict resource utilization, hospital admission rates, ED length of stay (LOS), intensive unit LOS, and mortality rates. Countries such as Canada, Spain, England, Australia, and the United States (US) have implemented this 5-level triage system in the ED (Christ, Grossmann, Winter, Bingisser, & Platz, 2010).

Triage scale reliability is very essential to determine whether the scale can be used repeatedly and can measure what is intended (Parenti, Reggiani, Iannone, Percudani, & Dowding, 2014). The reliability of the triage system addresses the consistency of its performance, where the same patient assignment categories must appear in the same answer despite who is doing the triage (Dippenaar, 2016; Christ, Grossmann, Winter, Bingisser, & Platz, 2010). Measurement of triage instrument reliability can be done by IRR assessment which aims to assess the agreement between two or more rater about the features of a set of assessed subjects (Hallgren, 2012; Chong, & Romkey, 2017). IRR takes several forms including the joint probability of agreement, kappa statistics (Cohen kappa and Fleiss kappa), correlation coefficients, limit agreements, alpha Krippendorff and inter-class correlation coefficients (ICC). All types of measurements above have various kinds of differences (Dippenaar, 2016).

Most triage scale reliability studies use the Cohen kappa or Fleiss kappa coefficients. This analysis has two types of weighting, particularly unweighted and weighted (linear or quadratic). However, most of these studies did not reveal the type of weighting used. The kappa coefficient (k) has a range of values between -1 to 1 while in ICC and alpha Krippendorff has a range of 0 to 1. The most popular reference level reference from kappa statistics was presented by Landis and Kock in 1977 where they ranked kappa coefficients from poor to the excellent agreement levels (Dippenaar, 2016; Twomey, 2011).

The accuracy of triage is a match between the assessment of primary complaints and the allocation of patient urgency (Twomey, 2011). Accuracy is achieved by comparing triage categories assessed by rater with gold standard values based on expert validation. Triage accuracy reveals three possible triage decisions including undertriage, expected triage, and over triage (Health Policy Priorities Principal Committee, 2011). Under triage is a condition when the patient's urgency level is recognized lower than actual urgency (Twomey, 2011; Hinson, Martinez, Schmitz, Toerper, & Radu, et al., 2018). Over triage is a determination of the level of urgency the patient is recognized higher than actual urgency. Both over and undertriage have an impact on the decrease in patient care outcomes (Twomey, 2011; Lampi, 2017). The American College of Surgeons Committee on Trauma sets under triage tolerated <5% and 25-50% for over triage (Lampi, 2017).

The Australian Triage Scale (ATS)

The Australasian Triage Scale (ATS) is a triage that was formalized and standardized by the Australasian College for Emergency Medicine (ACEM) in 2000. The ATS is a revision of the National Triage Scale (NTS) developed in 1993. ATS is a scale with 5 levels of urgency employed and implemented in all EDs in Australia and New Zealand. This has also provided a foundation for other countries to develop their national triage scale (Health Policy Priorities Principal Committee, 2011; Monash Institute of Health

Services Research, 2001).

The Australasian Triage Scale (ATS) consists of 5 levels of urgency categories that contain an ideal time limit for patients to first evaluate and get treatment by a doctor. The urgency categories include (1) Category 1- immediately life-threatening (evaluation is carried out immediately after the patient arrives at the emergency room); (2) Category 2- imminently life-threatening (evaluation in 10 minutes); (3) Category 3- potentially life-threatening (evaluation carried out in 30 minutes), category 4- potentially serious (evaluation in 60 minutes), and category 5- less urgent (evaluation in 120 minutes). The ATS also includes a performance threshold percentage that forms the basis for reporting triage implementation in IGD in Australia (Christ, Grossmann, Winter, Bingisser, & Platz, 2010; Hodge, Hugman, Varndell, & Howes, 2013).

The ATS contains physiological parameters according to the primary survey including airway patency, respiratory status, circulation, and disability (Australasian College for Emergency Medicine, 2016). Patients with airway obstruction, breathing arrest, severe respiratory distress, cardiac arrest, severe hemodynamic disorders, uncontrolled bleeding, and patients with a Glasgow coma scale (GCS) less than 9 were allocated into ATS 1. Patients with airway patency and or have moderate respiratory distress, moderate hemodynamic disorders and GCS 9 - 12 are allocated into ATS 2. Patients without airway patency, and/or have mild respiratory distress, mild hemodynamic disorders and GCS> 12 are allocated into ATS 3. Patients without airway patency and/or respiratory distress, patients without hemodynamic disorders are allocated into categories 4 and 5 (Health Policy Priorities Principal Committee, 2011; Australasian College For Emergency Medicine, 2016).

The results of a review of 2 studies employing ATS with sample size (n = 3979) divided into 3952 direct triaged patients and 27 paper-based triage case scenarios presented that reliability ATS was at a moderate level (k = 04 - 0.57) (Ekins & Morphet, 2015; Varndell, Hodge, Ryan, & Fry, 2019). Two studies with an equal number of samples presented ATS accuracy was 46.2% - 58.3%. One study with sample size (n = 3952) presented over triage and undertriage on ATS ware 22.9% and 12.4% respectively (Ekins & Morphet, 2015; Varndell, Hodge, Ryan, & Fry, 2019).

The Manchester Triage System (MTS)

This five-level triage system was developed by a consensus group in the UK. This scale is designed into 52 flowcharts that represent several patient symptoms when visiting the emergency room. These symptoms include shortness of breath, abdominal pain, gastrointestinal bleeding, head injury, and other symptoms. Every flowchart consists of six key discriminators designed according to the primary ABCDE survey format including life-threatening hazards, level of consciousness, bleeding, temperature, pain, and acute conditions arranged to distinguish each category of urgency (Christ, Grossmann, Winter, Bingisser, & Platz, 2010; Manchester Triage Group, 2014).

MTS also includes the maximum waiting time for the first evaluation by doctors that have been designed according to each category including (1) Category 1-immediately (red); (2) Category 2- very urgent (orange), evaluation in 10 minutes; (3) Category 3 - urgent (yellow), evaluation in 60 minutes; (4) Category 4-standard (green), evaluation in 120 minutes and; (5) Category 5- non-urgent (blue), evaluation in 240 minutes. A flow chart of sick adult patients on MTS shows patients with airway obstruction, and/or inadequate breathing, severe bleeding with shock, and not responsive are allocated into red category. Patients with uncontrolled heavy bleeding and/or pulse abnormalities, changes in the level of consciousness, extreme heat, cold, and severe pain are allocated into the orange category. Patients with mild uncontrolled bleeding and/or history of unconsciousness, heat, and pain are allocated into the yellow category. Patients who exhibited warm and mild acute pain are allocated into green category and patients without the problems stated previously are allocated into blue category (Manchester Triage Group, 2014).

The results of the review in 3 studies employing MTS with sample size (n = 412) which were divided into 398 direct triaged

patients and 14 case scenarios presented the reliability of MTS was at good to excellent levels (k= 0.61 - 0.95) (Olofsson, Gellerstedt, & Carlström, 2009; Gräff, Goldschmidt, Glien, Bogdanow, Fimmers, & Hoeft, et al., 2014; Andrade-Silva, Takemura, Bellato, Leonhardt, & Kojima, et al., 2019). While in 2 studies with sample size (n = 245) which were divided into 231 direct triaged patients and 14 triage case scenarios presented MTS accuracy was 64% - 73%. One study with sample size (n = 14 case scenarios) determines the percentage of under and over triage on MTS were 13% and 14% respectively (Olofsson, Gellerstedt, & Carlström, 2009; Andrade-Silva, Takemura, Bellato, Leonhardt, & Kojima, et al., 2019). While 1 other study with a sample size (n = 231) presented over triage in MTS occurred in 38% of traumatic patients and 69% in non-traumatic patients. While the percentage of undertriage in trauma and non-traumatic patients is 9% and 13% respectively (Andrade-Silva, Takemura, Bellato, Leonhardt, & Kojima, et al., 2019).

The Canadian Triage and Accuity Scale (CTAS)

CTAS is a 5-level triage scale developed in the late 1990s by the Canadian Association of Emergency Physicians and the National Emergency Nurse Affiliation. CTAS has revised several revisions particularly in 2004 and 2008 (Dallaire, Poitras, Aubin, Lavoie, & Moore, 2012). CTAS also has a time guideline for the doctor's first evaluation. CTAS contains an extensive list that presents patient complaints, clinical signs, vital parameters, respiratory problems, and pain as predictors of the level of urgency. CTAS also recommends triage re-triage after a specified waiting time or when changes occur in a patient's clinical symptoms (Christ, Grossmann, Winter, Bingisser, & Platz, 2010).

The CTAS categories include: (1) category 1- resuscitation (blue) immediately evaluated; (2) category 2- emergencies (red), evaluation in 15 minutes; (3) category 3-urgent (yellow), evaluation in 30 minutes; (4) category 4- less urgent (green), evaluation in 60 minutes and; (5) category 5- non-urgent (white), evaluation in 120 minutes. Patients with cardiac arrest and/or respiratory arrest, severe trauma (shock), severe respiratory distress, and unconscious patients with GCS 3 - 9 are allocated into the resuscitation category. Patients with moderate respiratory distress and/or vomiting of blood with dizziness when seated, severe hypertension, decreased consciousness with GCS 10-13, fever (> 38oC), sepsis patients with 3 criteria for systemic inflammatory response syndrome (SIRS), cardiac chest pain, severe chest pain non-cardiac disease, severe and sudden headaches, and severe trauma are allocated into emergency category. Patients with mild respiratory distress and/or moderate hypertension, vomiting and/or nausea with mild dehydration, acute moderate pain of stomach, and head pain, and diarrhea with uncontrolled blood are allocated into the urgent category. Patients with confusion and/or urinary tract infections with mild dysuria and constipation with mild pain fall into the less urgent category. Patients with diarrhea with or no mild dehydration, small bites with or without mild acute peripheral pain, patients with wound dressing replacement, or medication requests allocated into the non-urgent category (Canadian Association of Emergency Physicians, 2013).

The results of the review in 1 study employing CTAS with sample size (n = 160 case scenarios) presented the reliability of CTAS was at a good level (substantial) with k = 0.770 (Alquraini, Awad, & Hijazi, 2015). While 1 study with sample size (n = 49 case scenarios) presented the accuracy of CTAS was 49%. The percentage of over triage and undertriage of CTAS were 55.93% and 44.07% respectively (Saeed, Al-Fayyadh, Alshomar, Zekry, & Alamiri, et al, 2017).

The Emergency Severity Index (ESI)

Emergency Severity Index (ESI) is a five-level triage algorithm developed in the US in the late 1990s. Treatment priority on this triage scale is determined based on the severity of the disease and the estimated amount of resources needed by patients in the ED (Christ, Grossmann, Winter, Bingisser, & Platz, 2010). Nowadays, ESI has been revised and refined as ESI version 4. The ESI algorithm contains four key questions in

allocating patient urgency levels, including: (1) Does this patient need immediate life-saving intervention ?; (2) Does the patient not have to wait ?; (3) How many resources does the patient need ?, and; (4) What is the patient's vital signs?. First, the nurse observes the patient's urgency level to determine whether the patient falls into the high urgency criteria (ESI level 1 or 2). If the patient needs to get an immediate life-saving intervention, then the patient is allocated to ESI level 1. If the patient has potentially life-threatening symptoms, then the patient is allocated to ESI level 2 (Agency for Healthcare Research and Quality, 2012).

Patients who did not allocate into the high level of acuity (levels 3, 4, and 5) were assessed based on predictions of the number of resources needed during treatment in the ED. Patients who do not need resources are allocated into ESI level 5, patients who need 1 resource are allocated into ESI level 4, and patients who need some resources are allocated into ESI 3. Before determining ESI level 3, nurses require to consider the patient's vital signs wherever if the patient has vital signs in the danger zone so the patient is allocated into ESI 2 (Agency for Healthcare Research and Quality, 2012).

The results of the review in 3 studies employing ESI with sample size (n = 651) divided into 601 direct triaged patients and 50 paper-based scans presented the reliability of ESI was at moderate to excellent levels with k = 0.45 - 0.94 (Jordi, Grossmann, Gaddis, Cignacco, & Denhaerynck, et al., 2015; Esmailian, Zamani, Azadi, & Ghasemi, 2014; Mirhaghi, Kooshiar, Esmaeili, & Ebrahimi, 2015). While ESI in 2 studies with sample size (n = 50 case scenarios) presented ESI accuracy was 59.6% - 72.5%. One study with sample size (n = 30 case scenarios) presented the percentage of under and over triage were 26.8% and 13.6% respectively (Jordi, Grossmann, Gaddis, Cignacco, & Denhaerynck, et al., 2015; Mirhaghi, Kooshiar, Esmaeili, & Ebrahimi, 2015).

Table 1.1. Research Characteristics Table

Author (Year)	Study Design	Triage Method	Partisipant	Research Method	Result
Ekins & Morphet (2015)	**Descriptive Correlational design with cross-sectional**	**ATS**	**65 triage nurses were selected using the purposive sampling method**	**27 paper-based scenarios (14 adult cases and 13 pediatric cases) were validated. Triage accuracy was assessed using descriptive statistics and triage accuracy was analyzed using Kappa coefficients.**	**Triage accuracy: 46.2%. The most accurate triage in category 2 (78.3%), and the least accurate triage in category 5 (40.7%) (no under and over triage data).** **Inter-rater reliability: Fleiss coefficient (unspecified weight or unweight kappa) of 0.4 (moderate level)**

Vandell et al (2019)	**Descriptive, explorative audit and retrospective surveys**	**ATS**	**1 triage data set by nurses in 3952 adult and pediatric patients**	**39 audits conducted in 12 months during 2018 on 3952 triage records. Triage agreements were tested between nurses versus The Triage Quality Assessment Software system.**	**Triage accuracy: 58.3% with over triage (22.9%) and under triage (12.4%).** **Inter-rater reliability: Weighted Kappa (unspecified linear or quadratic) of 0.57 at a moderate level**
Jordi et al (2015)	**Cross-sectional, multicenter study conducted on 4 EDs**	**ESI versi 4**	**69 nurses who have been trained using ESI version 4.**	**Triage conducted by nurses uses 30 case scenarios where 6 of them are the pediatric case. All cases have been validated by experts.**	**Triage accuracy: 59.6% with under triage of 26.8% and over triage of 13.6%. There was no difference in accuracy for the 4 hospitals (X2 = 3.88; df = 3; p = 0.27)** **Inter-rater reliability: Krippendorff's alpha value of 0.78 was at the level of substantial / good agreement.**
Esmailian et al (2014)	**Prospective cross-sectional study**	**ESI versi 4**	**Trias was conducted separately by doctors and nurses (the number of triage officers is not stated) and the triage results are compared.**	**Triage was conducted on 601 patients who came to the ED.**	**Accuracy: No data available** **Inter-rater reliability: Weighted kappa (unspecified linear or quadratic specific) of 0.94 was in the excellent level.**

Mirhaghi et al (2015)	**Single-center study with cross-sectional**	**ESI versi 4**	**20 emergency room nurses were included in the study.**	**20 paper-based scenarios were compiled from ESI Handbook version 4. Nurses were asked to do triage based on cases in 2 separate periods. The nurse had ESI version 4 training previously.**	**Accuracy: The accuracy of the first period (60%) and the second period (72.5%). There is no under and over triage data.** **Inter-rater reliability: -first period: Percent of agreement (79%) with Fleiss un-weighted kappa (unspecified linear or quadratic) (κ = 0.548; 95% CI 0.531 - 0.564) at moderate level -In the second period: Percent of agreement (71%), Fleiss un-weighted kappa (unspecified linear or quadratic) in the moderate level (κ = 0.455; 95% CI 0.413 - 0.497).**
Olofsson et al (2009)	**Descriptive prospective research**	**MTS in Sweden**	**79 nurses were included in the study**	**The nurse triaged 9 case scenarios extracted from previous studies plus 5 self-composed cases (a total of 14 cases).**	**Accuracy: 73% with under triage (13%) and over triage (14%).** **Inter-rater reliability: unweighted k = 0.61 (SD 95%, CI 0.57–0.65), linear weighted k = 0.71, and quadratic weighted k = 0.81 (all at good to very good levels).**

Graff et al (2014)	**Research on prospective construct validity and inter-rater reliability assessment**	**MTS translated to the German version**	**1 triage data set was performed on 167 patients by 10 nurses within 24 hours for 2 months. The number of experts is not stated.**	**167 patients (within 2 months) were triaged by triage nurses and experts separately to assess the consistency of triage. Cohen's weighted Kappa and Spearman's rank correlation were chosen to determine the agreement between raters.**	**Accuracy: No data available** **Inter-rater reliability: percent of agreement (97.01%). Cohen's weighted kappa (unspecificied linear or quadratic) with k = 0.954; 95% CI 0.912 - 0.996) at a excellent level with Spearman's rank correlation coefficient (p = 0.956; 95% CI 0.910 to 1.0).**
Andrade-Silva et al (2019)	**Prospective study**	**MTS**	**2 orthopedic doctors were included in the study.**	**Triage was conducted on 231 patients aged ≥ 18. Triage reliability between 2 doctors was measured.**	**Accuracy: 64% with over triage of 38% in trauma patients and 69% in non-trauma. The under triage rate was 9% in trauma patients and 13% in non-trauma patients.** **Inter-rater reliability: Percent of agreement at 84%, with a coefficient of kappa (unspecificied weighted or unweighted kappa) of 0.77 (p <0.001) at a good (substantial) level.**

Alquraini et al (2015)	**A quantitative observational study with cross-sectional**	**CTAS**	**Triage comparison between 5 senior nurses (SN1) versus juniors (SN2).**	**160 case scenarios were extracted based on the patient's medical record reviewed by the investigator.**	**Accuracy: No data available** **Inter-rater reliability: Nurse agreement between SN1 versus SN2 with weighted kappa (unspecified linear or quadratic) of 0.770; 95% CI (0.742-0.797) was at a good level.**
Saeed et al (2017)	**A quantitative observational study with cross-sectional**	**CTAS**	**69 nurses were included in the study**	**49 paper-based scenarios extracted from real cases of patients in the ED who were randomly selected and validated by an emergency physician consultant (gold standard).**	**Accuracy: 49% and miss triage of 51%. Over and undertriage were 55.93% and 44.07%, respectively.** **Interrater reliability: Only displays percent of agreements without kappa coefficients. Agreements among nurses were highest in category 1 (62.2%) and lowest in category 5 (42.7%).**

The results of this review reveal that ESI and MTS are triage scales with the highest level of reliability, which is equal at an excellent level. While Triage with the highest level accuracy was also found in MTS and ESI with accuracy were 73% and 72.5%, respectively. The lowest undertriage was found in MTS and highest in CTAS. While the lowest over triage was also found in MTS and the highest was also found in CTAS. All triage scales reveal that the percentage of undertriage exceeds the limit tolerated by the American College of Surgeons Committee on Trauma which establishes a target of undertriage <5%. While over triage that exceeds the tolerance limit is found in CTAS which is greater than 50%.

Discussion

The five-level triage scale is the gold standard in the EDs in the wide world. The five-level triage systems that are most widely studied and adopted are ATS, CTAS, MTS, and ESI, all of which have good validity and reliability (Christ, Grossmann, Winter, Bingisser, & Platz, 2010). CTAS, ATS, MTS triage instruments have almost similar indicators in determining the allocation of patient urgency levels. The instrument also contains an ideal time limit for patients to first evaluate and receive treatment by doctors (Christ, Grossmann, Winter, Bingisser, & Platz, 2010; Manchester Triage Group,

2014; Canadian Association of Emergency Physicians, 2013). While ESI uses a different approach, especially in patients with low acuity level (ESI level 3,4, and 5) where the allocation of patient urgency is created based on the estimated amount of resources needed by patients while being treated in the ED (Christ, Grossmann, Winter, Bingisser, & Platz, 2010; Agency for Healthcare Research and Quality, 2012).

The emergency department requires a valid and reliable triage system (Brouns, Mignot-Evers, Derkx, Lambooij, & Dieleman, et al., 2019). Hinson, Martinez, Schmitz, Toerper, and Radu, et al (2018) stated the accuracy and reliability of triage in the ED is the most important thing. In this literature review, ESI and MTS are the triage scale with the highest reliability which are at an excellent level of reliability. This result supported by the literature study by Christ, Grossmann, Winter, Bingisser, and Platz (2010) showing 12 analyzes that discuss ESI has a good to the excellent level of reliability ($\kappa = 0.46 - 0.91$). A meta-analysis study by Mirhaghi, Heydari, Mazlom, and Hasanzadeh (2015) presented the combined coefficients for ESI were at a substantial level of reliability ($k = 0.791$; 95% CI: 0.787-0.795).

The Study of Parenti, Reggiani, Iannone, Percudani, and Dowding (2014) also reveals that MTS reliability was at fair to excellent levels with unweighted, weighted kappa, and quadratic weighted kappa were 0.31-0.76; 0.40–0.80, and 0.81– 0.82 respectively. Most triage reliability analyzes in various research articles in this literature review do not confirm the type of weighting used. Quadratic weights kappa is part of the weighted kappa analysis which is the best analysis in assessing the reliability of the triage scale (Dippenaar, 2016; Twomey, 2011).

The triage accuracy on ESI and MTS also reveals an almost equal percentage. This result supported by a study of Storm-Versloot, Ubbink, Kappelhof, and Luitse (2011) which presented that ESI and MTS in adult patients have an almost equal sensitivity and specificity in hospital admissions and predictions of patient mortality. However, MTS has a lower percentage of under and over triage compared to ESI. This is in line with a systematic study by Parenti, Reggiani, Iannone, Percudani, and Dowding (2014) revealed that MTS has lower undertriage with a range between 11% to 25%. A study by Storm-Versloot, Ubbink, Kappelhof, and Luitse (2011) also presented that the percentage of undertriage in MTS is lower of 11% compared to ESI of 20%. Nevertheless, the percentage of undertriage on both triage scales still exceeds the tolerance limit of undertriage by the American College of Surgeons Committee on Trauma which establishes under triage targets was <5%.

This literature review reveals that ESI and MTS have the highest reliability and accuracy compared to ATS and CTAS. Accordingly, ESI and MTS can be recommended as a triage scale choice in the ED. However, there are serious concerns, especially in the employment of ESI in the EDs of other countries. This is because there are significant differences in the structure of the health care system and culture of patient care compared to the United States where ESI was developed. Consequently, before choosing a triage method, the ED needs to consider the suitability of the triage process with the conditions of each local EDs. This is because not all triage scales are suitable to use in all areas, such as ESI requires good experience and skills for emergency nurses to allocate levels of patient urgency based on the possible resources needed by patients during treatment in the ED (Mirhaghi, Kooshiar, Esmaeili, & Ebrahimi, 2015). Meanwhile, MTS has 52 different flowcharts that represent several patient symptoms in the ED. As a result, the ED requires many forms of MTS based on the patient's symptoms that must be available.

T he limitation of the study is not all studies that identify the reliability of the international triage scale also contains triage accuracy. As a result, articles on triage accuracy were fewer in number. Besides, some articles relevant to the purpose of this literature review cannot be involved due to the publication year and the language of the article.

Conclusion

Triage is a rapid process in categorizing patients who come to the ED. Most of the ED's triage systems in the wide world are

using the five-level triage scale. Inaccuracies or inconsistencies in triage decisions can have serious consequences for patients and potentially increase poor treatment outcomes. Therefore, the study that gives special attention to the reliability and accuracy of the triage scale in the ED is needed. The results of this literature review reveal that MTS and ESI are the triage scales with the highest reliability and accuracy. Consequently, it is important to recommend both triage scales as the choice of triage in the ED. However, each EDs need to pay attention to the characteristics, culture, and available resources before choosing the appropriate triage scale will be adopted.

References

Agency for Healthcare Research and Quality. (2012). Emergency severity index (ESI): A triage tool for emergency department care version 4. Retrieved from https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/systems/hospital/esi/esihandbk.pdf

Alquraini, M., Awad, E., & Hijazi, R. (2015). Reliability of Canadian Emergency Department Triage and Acuity Scale (CTAS) in Saudi Arabia. International journal of emergency medicine, 8(1), 80. doi:10.1186/s12245-015-0080-5

Andrade-Silva, F. B., Takemura, R. L., Bellato, R. T., Leonhardt, M. D. C., & Kojima, K. E., et al. (2019). Validity and reliability of the Manchester scale used in the orthopedic emergency department. Acta Ortopedica Brasileira, 27(1), 50–54. https://doi.org/10.1590/1413-785220192701191577

Australasian College For Emergency Medicine. (2016). Guidelines on the Implementation of the ATS in Emergency Departments, pp. 1–8. Melbourne: Australasian College For Emergency Medicine.

Brouns, S. H. A., Mignot-Evers, L., Derkx, F., Lambooij, S. L., & Dieleman, J. P., et al. (2019). Performance of the Manchester triage system in older emergency department patients: A retrospective cohort study. BMC Emergency Medicine, 19(1), 1–12. https://doi.org/10.1186/s12873-018-0217-y

Canadian Association of Emergency Physicians. (2013). Applying CTAS to adults. Retrieved from http://www.qehaedbb.org/wp-content/uploads/2018/03/2011_-_CTAS_Module_2_Notes.pdf

Chong, A., & Romkey, L. (2017). Testing Inter-Rater Reliability in Rubrics for Large Scale Undergraduate Independent Projects. Proceedings of the Canadian Engineering Education Association (CEEA), pp. 1–9.

Christ, M., Grossmann, F., Winter, D., Bingisser, R., & Platz, E. (2010). Modern Triage in the Emergency Department. Deutsches Ärzteblatt International, 107(50), 892–898. https://doi.org/10.3238/arztebl.2010.0892

Dippenaar, E. (2016). Standardisation and validation of a triage system in a private hospital group in the United Arab Emirates (Doctoral dissertation). University of Cape Town.

Dippenaar, E., & Bruijns, S. (2016). Triage is easy, said no triage nurse ever. International Emergency Nursing, 29, 1–2. https://doi.org/10.1016/j.ienj.2016.09.005

Ekins, K., & Morphet, J. (2015). The accuracy and consistency of rural, remote and outpost triage nurse decision making in one Western Australia Country Health Service Region. Australasian Emergency Nursing Journal. https://doi.org/10.1016/j.aenj.2015.05.002

Esmailian, M., Zamani, M., Azadi, F., & Ghasemi, F. (2014). Inter-Rater Agreement of Emergency Nurses and Physicians in Emergency Severity Index (ESI) Triage. Emergency (Tehran, Iran), 2(4), 158–15861. https://doi.org/10.22037/emergency.v2i4.6449

FitzGerald, G., Jelinek, G. A., Scott, D., & Gerdtz, M. F. (2010). Emergency department triage revisited. Emergency Medicine Journal, 27(2), 86–92. https://doi.org/10.1136/emj.2009.077081

Fry, M., & Burr, G. (2002). Review of the triage literature: Past, present, future? Australian Emergency Nursing Journal, 5(2), 33–38. doi:10.1016/s1328-2743(02)80018-9

Gräff, I., Goldschmidt, B., Glien, P., Bogdanow, M., Fimmers, R., & Hoeft, A., et al. (2014). The German version of the Manchester triage system and its quality criteria - First assessment of validity and reliability. PLoS ONE, 9(2), 1–12. https://doi.org/10.1371/journal.pone.0088995

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. Tutorial Quantitative Methods Psychology, 8(1), 23–34. https://doi.org/10.1093/carcin/9.8.1355

Health Policy Priorities Principal Committee. (2011). Australian triage process review. Adelaide: Health Policy Priorities Principal Committee.

Hinson, J. S., Martinez, D. A., Schmitz, P. S. K., Toerper, M., & Radu, D., et al. (2018). Accuracy of emergency department triage using the Emergency Severity Index and independent predictors of under-triage and over-triage in Brazil: a retrospective cohort analysis. International Journal of Emergency Medicine, 11(1), 3. https://doi.org/10.1186/s12245-017-0161-8

Hodge, A., Hugman, A., Varndell, W., & Howes, K. (2013). A review of the quality assurance processes for the Australasian Triage Scale (ATS) and implications for future practice. Australasian Emergency Nursing Journal, 16, 21–29. https://doi.org/10.1016/j.aenj.2012.12.003

Iserson, K. V., & Moskop, J. C. (2007). Triage in Medicine, Part I: Concept, History, and Types. Annals of Emergency Medicine, 49(3), 275–281. https://doi.org/10.1016/j.annemergmed.2006.05.019

Jordi, K., Grossmann, F., Gaddis, G. M., Cignacco, E., & Denhaerynck, R., et al. (2015). Nurses' accuracy and self-perceived ability using the Emergency Severity Index triage tool: A cross-sectional study in four Swiss hospitals. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, 23(1). https://doi.org/10.1186/s13049-015-0142-y

Lampi, M. (2018). Triage Management of the Trauma Patient. Linköping University Medical Dissertations (Vol. 32). https://doi.org/10.1186/s13049-014-0050-6

Manchester Triage Group. (2014). Emergency Triage. Third edition (3rd ed.; K. Mackway-Jones, J. Marsden, & J. Windle, eds.). West Sussex: John Wiley & Sons, Ltd.

Mirhaghi, A., Heydari, A., Mazlom, R., & Hasanzadeh, F. (2015). Reliability of the emergency severity index: Meta-analysis. Sultan Qaboos University Medical Journal, 15(1), e71–e77.

Mirhaghi, A., Kooshiar, H., Esmaeili, H., & Ebrahimi, M. (2015). Outcomes for emergency severity index triage implementation in the emergency department. Journal of Clinical and Diagnostic Research, 9(4), OC04-OC07. https://doi.org/10.7860/JCDR/2015/11791.5737

Monash Institute of Health Services Research. (2001). Consistency of Triage in Victoria's Emergency Departments: Guidelines for Triage Education and Practice. Victoria: Monash Institute of Health Services Research

Olofsson, P., Gellerstedt, M., & Carlström, E. D. (2009). Manchester Triage in Sweden - Interrater reliability and accuracy. International Emergency Nursing, 17(3), 143–148. https://doi.org/10.1016/j.ienj.2008.11.008

Parenti, N., Reggiani, M. L. B., Iannone, P., Percudani, D., & Dowding, D. (2014). A systematic review on the validity and reliability of an emergency department triage scale, the Manchester Triage System. International Journal of Nursing Studies, 51(2014), 1062–1069. https://doi.org/10.1016/j.ijnurstu.2014.01.013

Qureshi, N. A. (2010). Triage systems: a review of the literature with reference

to Saudi Arabia. Eastern Mediterranean Health Journal, 16(6), 690–698. https://doi.org/10.26719/2010.16.6.690

Qureshi, N. A. (2010). Triage systems: a review of the literature with reference to Saudi Arabia. Eastern Mediterranean Health Journal, 16(6), 690–698. https://doi.org/10.26719/2010.16.6.690

Saeed, A. Bin, Al-Fayyadh, F. M., Alshomar, K. M., Zekry, Z. W., & Alamiri, N. N., et al. (2017). Validating the Implementation of the Triage System in an Emergency Department in a University Hospital. Journal of Health Specialties, 5, 73–79. https://doi.org/10.4103/jhs.JHS

Souza, C. C. de, Chianca, T. C. M., Júnior, W. C., Rausch, M. do C. P., & Nascimento, G. F. L. (2018). Reliability analysis of the manchester triage system: Inter-observer and intra-observer agreement. Revista Latino-Americana de Enfermagem, 26. https://doi.org/10.1590/1518-8345.2205.3005

Storm-Versloot, M. N., Ubbink, D. T., Kappelhof, J., & Luitse, J. S. K. (2011). Comparison of an Informally Structured Triage System, the Emergency Severity Index, and the Manchester Triage System to Distinguish Patient Priority in the Emergency Department. Academic Emergency Medicine, 18(8), 822–829. https://doi.org/10.1111/j.1553-2712.2011.01122.x

Twomey, M. (2011). Performance characteristics of the South African Triage Scale (Adult Version) (Doctoral dissertation). School of Public Health and Family Medicine University of Cape Town.

Varndell, W., Hodge, A., Ryan, E., & Fry, M. (2019). The use and feasibility of an online software system to support assessment of accuracy and consistency in applying the Australasian Triage Scale. Australasian Emergency Care. 10.1016/j.auec.2019.06.003.

Yuksen, C., Sawatmongkornkul, S., Suttabuth, S., Sawanyawisuth, K., & Sittichanbuncha, Y. (2016). Emergency severity index compared with 4-level triage at the emergency department of Ramathibodi University Hospital. Asian Biomedicine, 10(2), 155–161. https://doi.org/10.5372/1905-7415.1002.477