

A multicenter study on reliability and validity of a new triage system: the Triage Emergency Method version 2

Nicola Parenti,¹ Diego Sangiorgi,² Gianfranco Cervellini,³ Roberta Petrino,⁴ Gianni Rastelli,⁵ Gianni Ghetti,³ Mario Cavazza⁶

¹School of Healthcare, University of Parma; ²Biostatistic Research Department, University of Bologna; ³Department of Emergency Medicine, Parma University Hospital; ⁴Department of Emergency Medicine, Sant'Andrea Hospital, Vercelli; ⁵Department of Emergency Medicine, Fidenza Hospital; ⁶Emergency Department, S. Orsola-Malpighi University Hospital, Bologna, Italy

Abstract

In Italy there are many triage guidelines and methods based on consensus. But, to our knowledge, there are few data on the reliability and predictive validity of triage systems adopted by Italian emergency departments. The Triage Emergency Method version 2 (TEM v2) is a new four-level in-hospital triage system. This paper presents a before-and-after observational study performed using triage scenarios from June 2008 to September 2009 in 6 Italian emergency departments. Twelve nurses who received a 5-h training on TEM and a panel of experts on TEM assigned priority code to 66 scenarios. To test the inter-rater reliability among participants and the panel of experts (before and after the course), we used the weighted K statistic. We assessed the validity of TEM by calculating sensitivity, specificity and accuracy for predicting the reference standard's triage score. The TEM v2 showed good and very-good agreement among all 6 groups of nurses with a K range=0.61-1. Also, sensitivity, specificity and accuracy of nurses' triage rating for predicting the reference standard's triage code was good (accuracy range=78-90%). In this multicenter study, TEM v2 has a good inter-rater reliability for rating triage acuity among all groups of participating nurses, with a K value similar to the reference standard reliability (K=0.75). Thus, the Triage Emergency Method version 2 seems to be valid and accurate in predicting a reference standard rating.

Introduction

The Triage Emergency Method (TEM), is a new in-hospital triage method based on Italian guidelines.¹ It was developed in 2006. In a previous study, it showed a good inter- and intra-rater reliability for rating triage acuity and good accuracy in patient-admission prediction.²

TEM v2 contains one flowchart (Figure 1) and a table (Table 1). As shown in Figure 1, this new system is based on an acuity scale, resources used, the time the patient waits to be evaluated, and nursing procedures. It has 4 urgency categories (UC): 1 (Red), immediate assessment; 2 (Yellow), assessment within 20 min; 3 (Green), assessment within 60 min; and 4 (White), assessment within 120 min.

In 2008 TEM version 1² was reviewed by 25 Italian experts of triage: 19 nurses and 6 emergency physicians.³ After this study, following the experts' suggestions, TEM version 2 was developed.³ The new TEM v2 showed a good inter-rater reliability for rating triage acuity and good accuracy in predicting the rating of the reference standard also in a group of nursing students.^{4,5}

The aim of this study was to test whether TEM v2 shows a good inter-rater reliability for rating triage acuity and good accuracy in predicting the triage code rating of the reference standard in six groups of nurses from several Italian emergency departments.

Materials and Methods

Study design and setting

This is an observational study performed using triage scenarios from June 2008 to September 2009. Fundamental concepts of triage and the new Triage Emergency Method v2 (Figure 1) were shown, during a 5-h course, to 6 groups of nurses enrolled from 6 Italian emergency departments: Sant'Orsola University Hospital, Bologna; Santa Maria della Scaletta Hospital, Imola; Sant'Andrea Hospital, Vercelli; SS. Pietro e Paolo Hospital, Borgosesia; Parma University Hospital, Parma; and S. Secondo Hospital, Fidenza.

Data collection

We used 66 triage scenarios from a database used in previous studies.^{2,4,5} We recorded the following data: demographic and clinical characteristics, original nurse's triage category, admission status and site, and the data on triage forms completed by the nurse, namely, presenting complaint, mode and time of arrival, past diseases, vital signs, and pain score. Each case given to the study participants included the patient's age and gender, presenting complaint, a brief case scenario

Correspondence: Nicola Parenti, School of Healthcare, University of Parma, via Gramsci 14, 43100 Parma, Italy.
Tel/Fax: +39.051.9912065.
E-mail: nipar71@yahoo.it

Key words: triage, emergency department, reliability, predictive validity.

Contributions: NP conceived the study, designed it, collected and managed the data. Other authors contributed equally.

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

Received for publication: 27 June 2013.
Revision received: 26 August 2013.
Accepted for publication: 26 August 2013.

This work is licensed under a Creative Commons Attribution 3.0 License (by-nc 3.0).

©Copyright N. Parenti et al., 2013
Licensee PAGEPress, Italy
Emergency Care Journal 2013; 9:e20
doi:10.4081/ecj.2013.e20

with mode and time of arrival, past diseases, vital signs, and pain score.

Study participants

Twelve nurses (2 for each hospital) who worked in the emergency departments were assigned to undergo a five-hour training in triage and TEM. Four senior triage nurses and one doctor, who had emergency nursing and teaching triage certification, more than 15 years of emergency nursing experience, and prolonged training in the new triage method (TEM), comprised the panel of triage experts. The panel independently assigned, using TEM, triage scores to the 66 scenarios. Their triage codes were the reference standard (RS) for the triage level in this study. They were blinded to the triage category assigned both by the original triage nurse and by the nurses involved in this study. The nurses enrolled in the study completed a questionnaire related to their demographics, education, and work experience.

Study protocol

Between June 2008 and September 2009 a 5-h course on triage (1 h) and TEM (4 h) was performed in the hospitals enrolled. Each participant independently assigned triage scores to the 66 scenarios before the course (T0) and three months after the course (T1). To prevent communication between participants, the students assigned triage codes in different rooms and in the presence of two investigators. The triage scenarios were given randomly to the participants. During the second test, three

months after the course, participants could consult the TEM flowchart (Figure 1 and Table 1). The data were collected and entered on a spreadsheet by an investigator who was blind to the aim of the study. The nurses' group remained concealed during data entry and analysis. Being a quality assurance investigation, without access to patient's data, the institutions exempted the study from formal review. The nurses involved in the study gave informed consent and permission to access their data.

Data analysis

To test the quality of rating triage code by the group of participants, we choose the inter-rater reliability before and after the course and we assessed the validity of TEM.

Reliability was measured with weighted K by comparing the triage nurses' rating (inter-rater) at T0 and T1. We also measured the inter-rater reliability between the group and its reference standard by measuring the weighted K against the urgency category assigned by the panel of triage experts. In accordance with the literature,⁶ we consider a K value between 0.00

to 0.20 to be poor agreement, a K value between 0.20 and 0.60 to be fair to moderate agreement, a K value between 0.60 and 0.80 to be good agreement, and a K value between 0.80 and 1 to be very good agreement.

We evaluated the validity of TEM by calculating sensitivity, specificity and accuracy for prediction of the reference standard's triage score. To analyze the predictive validity for the reference standard's triage score, for each scenario we considered the mode of the urgency category assigned by the nurses, and we used this code in all validity calculations. We evaluated the validity by calculating sensitivity and specificity for prediction of reference standard's triage score using the following cut-offs: true codes 1 and 2=patient sick and likely to be admitted; true code 3 and 4=less urgency.

We calculated participant and scenarios sample size according to Rotondi *et al.*,⁷ anticipating a standard error of 0.05. Statistical significance was tested at an alpha level=0.05. We used the STATA v9.2 software (Statacorp, College Station Texas, USA) for statistical analysis.

Results

Of the 66 patients included in triage scenarios, 44% were women, and the mean age was 39 years ($SD \pm 27.5$). The most frequent symptoms were minor trauma (19%) and pain (14%). Seven hospital admissions were recorded: six in non-intensive wards and one in intensive care units. The group of twelve participants had good experience in nursing (median: 6 years; range 3-15) with a median of 4 years in emergency triage.

The rate of urgency categories at T0 (without TEM v2) and T1 (with TEM v2) assigned to each scenario was similar between the triage panel experts and the nurses enrolled (Figure 2). The green code (code 3) was most prevalent.

There was an overtriage and undertriage (with respect to the reference standard assignment) of 8 and 11% among nurses who used TEM v2. In particular, 2 of the 7 undertriage cases were for code 2, and 1 of the 7 were for code 1. All overtriage cases were for code 4.

Table 1. Triage emergency method patient's main complaints.

Code	Situation guide	Symptom
Red	Very high risk situations	Severe acute pain (VAS=9-10); shock; AMI arrhythmia: actual chest pain with syncope and/or arrhythmia and/or dyspnea; aphasia and/or numbness <3 h; dyspnea with wheeze or laryngospasm; major trauma: penetrating injury, severe facial trauma or cranial trauma with GCS≤14, thoracic trauma with volet or dyspnea, abdominal trauma with SBP ≤90 mmHg, rachis trauma with sensitive-motor deficit, amputation of long bones, open fractures of long bones, 2 nd or 3 rd degree skin burns (>30% adult or >20% babies), eyes or airway burns, RTS≤10; multiple trauma with major mechanisms (fall from 5 m; ejection outside a vehicle, or pedestrian run down, age <5 years); severe intoxication (quantity or kind of substance or substance unknown) with dysphonia, dysphagia, chest or abdominal pain; status epilepticus; headache with altered level of consciousness or meningism, seizures or syncope; hematemesis or other severe haemorrhage in action; severe allergic reactions with dyspnea, dysphonia or severe hypotension; emergency delivery; eclampsia; severe vaginal bleeding Children: severe dehydration, headache with GCS≤14 or lethargy or hypotonia, abuse
Yellow	High risk situations	Cardiac: chest pain≤6 h, syncope, arrhythmia, limbs ischemia, hypertension crisis (SBP>200 mmHg), unstable hypotension (SBP<100mmHg) Pneumology: mild dyspnea (SaO2 88-92%) Abdomen: acute abdominal pain, vomit and diarrhea with dehydration, hypovolemia (hypotension-tachycardia), hematemesis, melena, severe rectal bleeding in action Neurology: headache or with SBP>200 mmHg, headache after cranial trauma if anticoagulants; consciousness alterations (agitation, drowsiness, acute confusion), GCS 8-14, aphasia or sensitive-motor deficit <3 h, convulsions; alcohol or drug abuse; severe dizziness/vertigo or with headache or motor deficit; moderate pain (VAS=7-8); minor intoxication, severe allergic reactions (extensive nettle rash, dysphonia, angioedema, multiple hymenoptera stings in history of anaphylaxis) Infections: fever with lethargy, severe infection (rash or purpura), temperature >39°C, fever in immunodeficiency; trauma: concussive cranial trauma or anticoagulants, long bones, dislocation, bone deformation, open fractures, severe lacerations, crush syndrome, limb trauma without pulse, multiple trauma, major dynamic; acute lumbar pain (if age >40yrs or in case of hypertension); severe glycemic failure in diabetes (40 mg/dL<glycemia>300 mg/dL); sexual assault, severe or painful haemorrhage or mild but persistent in anticoagulants or hypertension crisis; renal-genitourinary: scrotum pain, anuria or oliguria Gynaecology and obstetrics: vaginal bleeding in the elderly or pregnancy with pain, delivery with active contractions <5 min, pre-eclampsia, pelvic pain Eye: eye injury with alkali or acid, visual deficit; psychiatric symptoms in patient suffering from mental illness; severe or mild haemorrhage (any cause) Child: newborn <3 months; moderate dehydration, severe vomiting or diarrhea; recent trauma <12 h; neonatal crying, recent convulsion

VAS, visual analogic scale; AMI, acute myocardial ischemia; GCS, Glasgow coma score; SBP, systolic blood pressure; RTS, revised trauma score; SaO2 (%), percentage of oxygen saturation.

Complete disagreement (when nurses of the same group assigned to the same scenario triage codes that differed in more than two priority levels, e.g. one nurse assigned level 4, and another, level 1) and complete agreement (when all nurses enrolled assigned the same triage code) occurred in 8 and 17%, respectively, of scenarios evaluated with TEM v2.

Inter-rater reliability among nurses who assigned priority code to 66 scenarios using the triage systems of their emergency departments and TEM v2 is shown in Table 2. TEM v2 showed good and very good agreement among all emergency departments with a K range=0.61-1 (Table 2). Also sensitivity, specificity and accuracy of nurses' triage rating for predicting the reference standard's triage code was good (Table 3).

Discussion

In all six groups of nurses, TEM v2 showed good to very good inter-rater reliability and good accuracy in predicting the reference standard's rating. Its performance was good despite the nurses' lack of experience using TEM. In addition, the Triage Emergency Method has the advantage of predicting resource utilization in the emergency department.

After a brief course, TEM v2 reached reliability performances similar to those of traditional triage systems (Table 2).

The six groups of nurses who used TEM proved accurate in predicting the reference standard's triage code: accuracy range=78-90% (Table 3). Few previous studies used a reference standard to test the validity of a triage system.^{8,9} However, it's very difficult to establish validity criteria for triage acuity classification in the absence of a clear reference standard. For this reason we tried to develop a surrogate *gold standard* based on a panel consensus, and we tested the predictive validity of our triage system against this gold standard.

It is difficult to compare our results on validity with previous studies because of the differences in the setting and the type of triage system (five levels compared to four levels). Nevertheless, our results on validity and reliability of TEM are similar to previous studies on emergency severity index v4¹⁰⁻¹² and the Canadian triage acuity scale.¹³

In accordance with a previous Italian study, the most frequent priority code was green. TEM is a four-level acuity triage system, based on Italian guidelines.¹ It was developed by 25 Italian experts of triage (emergency doctors and nurses). It is original because its sample structure (one flow chart and one table) is clear and sample to teach, memorize and consult.²⁻⁵ Moreover, it is the first Italian triage model that includes a resources evaluation; it is simple and

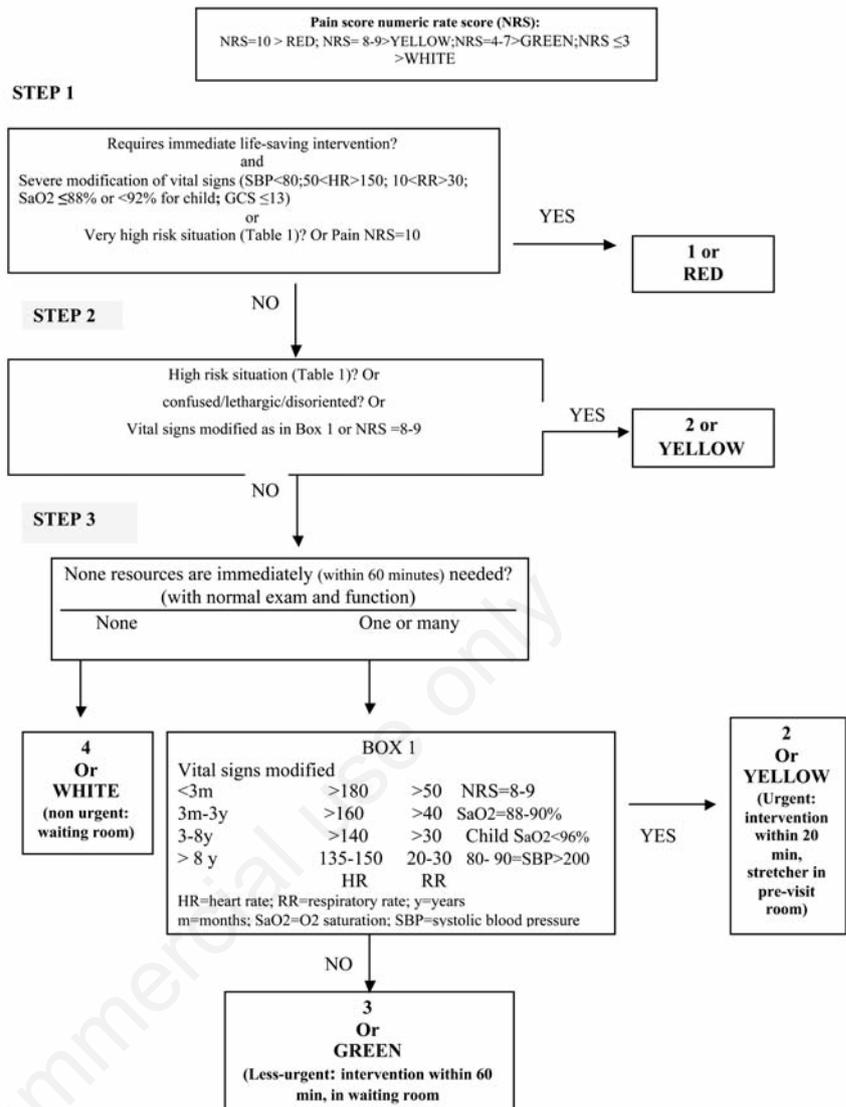


Figure 1. Flowchart of the Triage Emergency Method version 2.

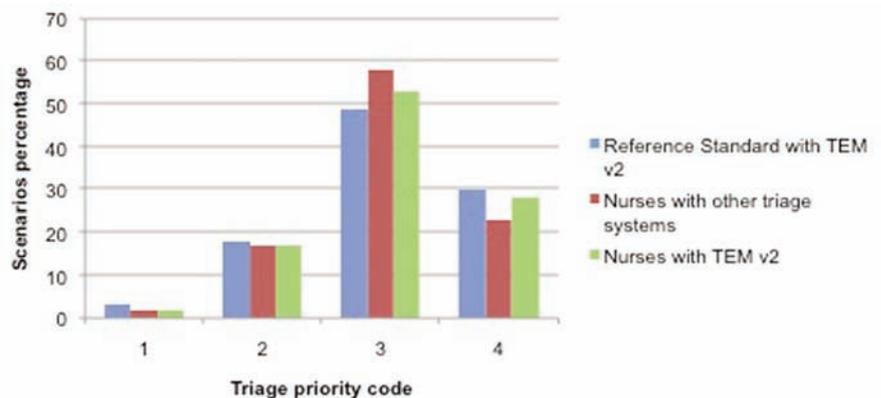


Figure 2. Rate of urgency categories assigned to each scenario.

Table 2. Inter-rater reliability of triage methods and Triage Emergency Method version 2.

	Group 1		Group 2		Group 3		Group 4		Group 5		Group 6	
	Met 1	TEM	Met 2	TEM	Met 3	TEM	Met 4	TEM	Met 5	TEM	Met 6	TEM
Weighted K inter-range*	0.71	0.75	0.80	1	0.69	0.61	0.76	1	0.98	0.84	0.25	0.79
Complete agreement (%)°	64	73	68	100	68	35	74	100	94	83	41	82
Complete disagreement (%)#	2	0	4	0	0	3	0	0	0	0	2	2

Met, triage method used in 12 emergency departments; TEM, triage emergency method version 2. *K inter-range 0-1 (0= scarce agreement, 1=maximum agreement); °when all 12 nurses enrolled assigned the same triage code; #when nurses of the same group assigned to the same scenario triage codes that differed in more than two priority levels, e.g. one nurse assigned level 4 and another assigned level 1.

Table 3. Validity of the Triage Emergency Method version 2 for prediction of the reference standard's rating.

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Accuracy (95% CI)	0.78 (0.51-1.05)	0.79 (0.57-1.00)	0.90 (0.71-1.09)	0.79 (0.57-1.00)	0.83 (0.62-1.04)	0.86 (0.67-1.04)
Sensitivity (95% CI)	1.00 (1.00-1.00)	1.00 (1.00-1.00)	1.00 (1.00-1.00)	0.98 (0.94-1.02)	1.00 (1.00-1.00)	1.00 (1.00-1.00)
Specificity (95% CI)	0.97 (0.92-1.01)	0.95 (0.90-1.00)	0.98 (0.94-1.02)	0.94 (0.88-1.00)	0.97 (0.92-1.01)	0.97 (0.92-1.01)
PPV (95% CI)	1.00 (1.00-1.00)	1.00 (1.00-1.00)	1.00 (1.00-1.00)	0.92 (0.76-1.07)	1.00 (1.00-1.00)	1.00 (1.00-1.00)
NPV (95% CI)	0.96 (0.91-1.01)	0.94 (0.88-1.01)	0.97 (0.92-1.02)	0.94 (0.88-1.01)	0.96 (0.91-1.01)	0.96 (0.91-1.01)

CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value. We tested the accuracy, sensitivity, specificity, PPV, and NPV among groups of nurses in each emergency department using TEM v2 to predict the reference standard's rating.

fast to assign the lower priority codes because the assignment is based on the estimation of resources necessary to the patient evaluated at triage. Finally, to our knowledge, the TEM v2 is the first Italian triage model who showed a good validity and reliability in more studies conducted with several kind of participants (student, nurses) in more settings. The main limitation of our study is that it was conducted with paper scenarios and not with patients; however, this procedure has been used and validated in other studies on inter-rater reliability of triage tools.¹⁰⁻¹³ The main limits of TEM v2 are that it is a theoretical model without feasibility studies and that the table for higher priority codes (red and yellow) could slow down the assignment.

Conclusions

In conclusion, to our knowledge, this is the first Italian multicenter study that tests the reliability and validity of a 4-level triage system. It seems that TEM v2 has good inter-rater reliability for rating triage acuity among all groups of nurses enrolled (K inter range=0.6-1). This K value is similar to the reference standard reliability (K=0.75). In this study, the TEM v2 seems to be valid and accurate in predicting a reference standard rating. Other studies with a wider sample of triage scenarios and more nurses (raters) participating would be necessary to confirm these results.

References

1. Italian Ministry of Health. [Triage intraospedaliero nel sistema dell'emergenza-urgenza sanitaria]. [Regulation in Italian]. Regulation n. 285/7712/2001. Rome: Italian Ministry of Health ed.; 2001.
2. Parenti N, Ferrara L, Bacchi Reggiani ML, et al. Reliability and validity of two four-level emergency triage systems. *Eur J Emerg Med* 2009;16:115-20.
3. Parenti N, Serventi V, Miglio R, et al. [Il Triage Emergency Method versione 2 (TEM v2). Un nuovo modello di triage intraospedaliero]. *Emerg Care J* 2011;3:27-34.
4. Parenti N, Bacchi Reggiani ML, Sangiorgi D, et al. [Affidabilità e validità di un nuovo metodo di triage intraospedaliero: il Triage Emergency Method in un gruppo di studenti universitari]. [Article in Italian]. *SIMEU Journal* 2011;3:20-1.
5. Parenti N, Bacchi Reggiani ML, Sangiorgi D, et al. Effect of a triage course on quality of rating triage codes in a group of university nursing students: a before-after observational study. *World J Emerg Med* 2013;4:20-5.
6. Fernandes C, Tanabe P, Gilboy N, et al. Five-level triage: a report from the ACEP/ENA five-level triage task force. *J Emerg Nurs* 2005;31:39-50.
7. Rotondi MA, Donner A. A confidence interval approach to sample size estimation for interobserver agreement studies with multiple raters and outcomes. *J Clin Epidemiol* 2012;65:778-84.
8. Storm-Versloot MN, Ubbink DT, Chin a Choi V, et al. Observer agreement of the Manchester triage system and the emergency severity index: a simulation study. *Emerg Med J* 2009;26:556-60.
9. van der Wulp I, Schrijvers A JP, van Stel HF. Predicting admission and mortality with the emergency severity index and the Manchester triage system: a retrospective observational study. *Emerg Med J* 2009;26:506-9.
10. Eitel DR, Travers DA, Rosenau AM, et al. The emergency severity index triage algorithm version 2 is reliable and valid. *Acad Emerg Med* 2003;10:1070-80.
11. Tanabe P, Gimbel R, Yarnold PR, et al. Reliability and validity of scores on the emergency severity index version 3. *Acad Emerg Med* 2004;11:59-65.
12. Wuerz RC, Milne LW, Eitel DR, et al. Reliability and validity of a new five-level triage instrument. *Acad Emerg Med* 2000;7:236-42.
13. Beveridge R. CAEP issues. The Canadian triage and acuity scale: a new and critical element in health care reform. *J Emerg Med* 1998;16:507-11.