Mediterranean Journal of Emergency Medicine
https://mjemonline.com/

MJEM
Mediterranean Journal of Emergency Medicine

Prospective evaluation of two emergency triage scales: the French Emergency Nurses Classification in Hospitals (FRENCH) and the Emergency Severity Index (ESI)

Citation: Clanet R, Aubrion A, Guigne A, Lecadet N, Audigie S, Creveuil Ch, Roupie E, Macrez R. Prospective evaluation of two emergency triage scales: the French Emergency Nurses Classification in Hospitals (FRENCH) and the Emergency Severity Index (ESI). Med Emergency, MJEM 2020; 28:8-12. DOI: 10.26738/MJEM.2017/MJEM28.2020/AA.OAE.130319

Key words: Accident and Emergency department, Emergency, ESI, FRENCH Triage.

ABSTRACT

Objective(s): Emergency Departments (ED) require a systematic approach to prioritize patient care depending on acuity. The Emergency Severity Index (ESI) scale is the most used. In France, the French Emergency Nurses' Classification in Hospitals (FRENCH) is used. The aim of this study was to evaluate inter-rater reliability and validity of the ESI and FRENCH triage scales.

Methods: We performed a prospective monocentric study in a French University College Hospital over the 2016 summer. All patients admitted to ED were evaluated to the triage area by two pairs of emergency physicians, each unaware of the triage results of the other pair. Reliability was estimated by a quadratic weighted Kappa. Validity was evaluated by the association between the level of triage and the following indirect criteria: rate of admission, type of admission amount of resources and length of stay in the emergency ward.

Results: Both the ESI and FRENCH triage systems showed strong reliability (weighted Kappa respectively 0.85 and 0.87) without any significant difference. An association was established for both scales between the level of emergency and each of the indirect criteria of validity. The associations between triage level and the amount of resources and length of stay criteria were significantly stronger for ESI scale. The area under the Receiver Operating Characteristic curve for prediction of an admission was 0.75 for ESI and 0.71 for FRENCH scale without any significant difference.

Conclusion: ESI and FRENCH scales have a strong inter-rater reliability and appear to have a good validity. Other studies, particularly multicenter studies including more qualitative criteria, would probably make it possible to decide on the most effective triage method.

Authors' affiliation:

Corresponding author: Antoine Aubrion, MD
Emergency medical service (SAMU 14), Caen University Hospital,
14000 Caen, France
aubrion-a@chu-caen.fr

Clanet R, MD[1,2], Aubrion A , MD[1,3], Guigne A, MD[1,4,5], Lecadet N, MD[1,2], Audigie S, MD[1,4], Creveuil Ch, MD[6], Roupie E, MD[1,5], Macrez R, MD[1,4,5,7]
1. Emergency medical service (SAMU 14), Caen University Hospital, Caen, France
2. Emergency department, Bayeux Hospital, Bayeux, France
3. Emergency department, Lisieux Hospital, Lisieux, France
4. Emergency department, Caen University Hospital, Caen, France
5. Normandy University, medical school, UNICAEN, Caen, France
6. Department of Biostatistics and Clinical Research, Caen University Hospital, Caen, France
7. Normandy University, UNICAEN, INSERM, INSERM UMR-S U1237, Physiopathology and Imaging for Neurological Disorders (PhIND), Cyceron, Caen, France

Article history / info:

Category: Original article
Received: Jan. 09, 2019
Revised: Feb. 13, 2019
Accepted: Mar. 13, 2019

Conflict of interest statement:

There is no conflict of interest to declare

Acknowledgements:

We thank Dr Mathieu Kozlik for his usefull advices concerning the study, Marie Aubrion and Ross Gail for her usefull reading of the manuscript.

INTRODUCTION

Triage is internationally recognized as one of the founding principles of accident and emergency medicine [1-4]. One finds it mainly in the use of triage scales in Emergency Departments (ED) on admission.

Worldwide, the Emergency Severity Index (ESI) scale is one of the most used [4-6]. This 3-step scale results in classifying the patient into one of five levels of decreasing severity. This tool has been approved for children [2], elderly patients [1] and has been much studied [1;7-12]. It particularly relies upon the nurses' intuitive abilities [11]. In France, the FRench Emergency Nurses' Classification in Hospitals (FRENCH) is the most customary scale [4]. Briefly, nurses first record the vital signs and perform a short medical history taking to know the reason of admission. Then, they rank patients into five decreasing levels of severity in accordance to a correspondence table (**Table 1**). This scale has been approved for adults only [12;13]. Its use is limited in France. We therefore aimed at comparing FRENCH scale with one of the most used tool in the world for the same population and under the same conditions, as some authors have previously advised [3]. Such a comparison has never been done. The main objective of this study was to evaluate the inter-rater reliability of the ESI scale and the FRENCH scale. The secondary objectives were to evaluate the validity of the ESI and the FRENCH scale by the use of indirect criteria.

METHODS

Study design:

We proceeded in a monocentric, prospective observational study. This study took place in the triage area of the ED of a French Hospital. This ED admitted almost 53,000 adults' patients annually. Our observations were led from June to July 2016, over daytime sessions for five to ten hours. The study was approved by our institutional ethics committee and all patients received information about the study.

Data collection:

The study has been conducted as follows: four emergency physicians independently sorted every patient admitted into the ED after receiving training concerning these two scales. Two physicians used the FRENCH-version 2. Two others the ESI-version 4. Each physician was unaware of either the rankings chosen by his partner, or of those chosen by the physicians working with the second scale. All adult patients arriving during a data-collection session were included in the study whatever their means of admission was. Paediatric emergencies were not included because the FRENCH system has not been approved for this population.

Data analysis:

To assess a triage scale, we commonly accepted to consider two parameters: reproducibility and validity. The inter-rater reproducibility represents the reliability of the tool to give a uniform measurement for a given patient with different triage scales. Validity represents the ability of the tool to assess the seriousness of the patient's state. To evaluate reproducibility and validity, we used indirect criteria: admission rate, place of admission (conventional service, continuing care department or intensive care unit), time spent in ED (minutes between assessment by the triage nurse and the conclusion of diagnosis) and resources (Labs, X-rays, etc) according to the ESI scale.

We assessed inter-rater reproducibility through a quadratically weighted Kappa with its 95% confidence interval, as described in literature [1;7;9;12;14-16]. The correlation was considered to be: low if the Kappa was less than 0.2; correct, between 0.2 and 0.4; good, from 0.4 to 0.6; very good, from 0.6 to 0.75 and excellent if over 0.75 [9;10;12;17]. Sample size was determined in order to statistically detect an excellent agreement (Kappa 0.75) between the two physicians of each pair [17]. We used sample size formulas designed for intra-class correlation coefficients [17–19].

Table 1: FRENCH scale: general description of triage and the actions considered.

Triage	Description	Action
1	Immediately life-threatening	Actions focused on support of one or more vital functions Immediate medical and paramedical intervention
2	Marked impairment of a vital organ or imminently life-threatening or functionally disabling traumatic lesion	Actions focused on treatment of the vital function or traumatic lesion Immediate paramedical and medical intervention within 20 min
3	Functional impairment or organic lesions likely to deteriorate within 24 h or complex medical situation justifying the use of several hospital resources	Multiple actions focused on diagnostic evaluation and prognostic evaluation in addition to treatment Medical intervention within 60 min ± followed by paramedical intervention
4	Stable, noncomplex functional impairment or organic lesions, but justifying the urgent use of at least one hospital resource	Consultation with limited diagnostic and/or therapeutic procedures Medical intervention within 120 min ± followed by paramedical intervention
5	No functional impairment or organic lesion justifying the use of hospital resources	Consultation with no diagnostic or therapeutic procedure Medical intervention within 240 min
*	Intense symptom or abnormal vital parameter justifying rapid corrective action	Specific action within 20 min The star can complete a triage 3 or 4

With a two-tailed alpha error of 0.05, we estimated that a total sample size of 300 patients was needed.
The confidence interval of the difference between the two Kappa coefficients obtained for each scale was also calculated using non-parametric bootstrap, to compare the two scales.

Statistical analysis:
For quantitative criteria, number of resources used and length of stay of patients, medians and quartiles were calculated for each triage level and compared by a Kruskal Wallis test. Spearman correlation coefficients with their 95% confidence interval were also calculated between the triage levels and each variable; a non-parametric bootstrap was used to estimate the confidence interval of the difference of the Spearman correlation coefficients obtained for each scale.

For qualitative criteria, means and place of admission of the patient, the distributions of each variable were estimated by triage level and compared by Mantel-Haenszel chi-square tests. For patient admission, a Receiver Operating Characteristic (ROC) curve analysis allowed to estimate the area under the curve, a measure of the predictive value of the triage scale. The areas under the curves obtained for each of the two scales were compared statistically. The analysis was made by the Unit of Biostatistics and Clinical Research department of our hospital with the IBM SPSS and R software programs.

RESULTS

Three hundred patients were assessed simultaneously with the two scales, which was the number of subjects we required. One hundred percent of the included subjects' data were collected. Our population included 52.3% of women and 47.7% of patients aged over 50 years old. The reason for consultation was medical in almost 75% of the cases. The median for time spent in ED was 205 min and an average of 2.5 resource units per case was necessary.
We first observed the triage's distribution for each scale. We found that the distribution is similar from one scale to another (**Table 2**). The inter-rater reliability for FRENCH was excellent with a weighted Kappa at 0.87, 95% CI [0.82;0.91]. The result was similar for ESI with a Kappa at 0.85, 95% CI [0.81;0.89]. The calculation of the difference between the two Kappa returned a value of 0.016, 95% CI [-0.071;0.038] which did not enable us to find a significant difference between the Kappa of the two scales (**Table 2**).
About the validity, for the ESI and the FRENCH scales, we showed an association between the level of triage and the number of resources as well as the length of stay ($p < 0.001$).

Table 2: Correspondence table for FRENCH and ESI scales for inter-rater reliability. Triage had been performed by two senior physicians. A/ ESI scale; B/ FRENCH scale.There is no significant difference between the two triage scales.

A

Physician 1 \ Physician 2	Level 1	Level 2	Level 3	Level 4	Level 5	Total
Level 1	4	0	0	0	0	**4**
Level 2	0	41	9	1	0	**51**
Level 3	0	8	96	23	1	**128**
Level 4	0	2	5	67	5	**79**
Level 5	0	0	0	11	27	**38**
Total	**4**	**51**	**110**	**102**	**33**	**300**

B

Physician 1 \ Physician 2	Level 1	Level 2	Level 3	Level 4	Level 5	Total
Level 1	4	0	1	0	0	**5**
Level 2	0	27	8	1	0	**36**
Level 3	0	4	93	12	2	**111**
Level 4	0	2	10	80	8	**98**
Level 5	0	0	1	8	41	**50**
Total	**4**	**31**	**113**	**101**	**51**	**300**

Table 3: Criteria of validity for FRENCH and ESI scales. To evaluate validity, we studied the amount of resources, lenght of stay, hospitalisation and department of admission. A) Criteria of validity for ESI scale ; B) Criteria of validity for FRENCH scale.

A

ESI	Amount of resources			Length of stay			Hospitalisation					Hospitalisation						
	Median	IQR	p	Median	IQR	p	No		Yes		p	Conventional		Continuing		Intensive		p
							n	%	n	%		n	%	n	%	n	%	
1	4.5	1.5;7.5		386	114;601		0	0	4	100		2	50.0	1	25.0	1	25.0	
2	4.0	4.0;6.0		362	253;449		24	47.1	27	52.9		22	81.5	4	14.8	1	3.7	
3	4.0	2.0;5.0	< 0.001	296	179;419	< 0.001	69	53.9	59	46.1	< 0.001	56	94.9	3	5.1	0	0	0.002
4	1.0	1.0;2.0		103	56;189		71	89.9	8	10.1		8	100	0	0	0	0	
5	0.0	0.0;1.0		66	42;110		37	97.4	1	2.6		1	100	0	0	0	0	

B

FRENCH	Amount of resources			Length of stay			Hospitalisation					Hospitalisation						
	Median	IQR	p	Median	IQR	p	No		Yes		p	Conventional		Continuing		Intensive		p
							n	%	n	%		n	%	n	%	n	%	
1	4.0	3.0;4.0		225	186;348		1	20.0	4	80.0		2	50.0	1	25.0	1	25.0	
2	5.0	4.0;6.0	< 0.001	372	295;486	< 0.001	16	44.4	20	55.6	< 0.001	19	95.0	1	5.0	0	0	0.030
3	4.0	2.0;5.0		273	149;407		63	56.8	48	43.2		41	85.4	6	12.5	1	2.1	
4	1.0	1.0;3.0		142	78;280		73	74.5	25	25.5		25	100	0	0	0	0	
5	0.0	0.0;1.0		72	44;133		48	96.0	2	4.0		2	100	0	0	0	0	

We also found a significant link between seriousness and hospitalization ($p < 0.001$). At least, we observed a significant association between the level of triage and the place of admission ($p = 0.002$ for the ESI and $p = 0.03$ for the FRENCH; **Table 3**). Using the Spearman correlation coefficient, the analysis of the severity level for both scales, on the one hand, and of the quantitative variables, on the other hand, allowed us to establish a moderate but significant link ($p < 0.001$) regarding the number of resources, with $r = -0.65$ for the ESI and $r = -0.58$ for the FRENCH, and the length of stay as well, with respectively $r = -0.59$ and $r = -0.51$ for the ESI and the FRENCH. The calculation of the correlation coefficients' difference for the amount of resources (0.074, 95% CI [0.016;0.14]) and the length of stay (0.08, 95% CI [0.014;0.15] establishes a significant difference in favor of the ESI scale.

Finally, for prediction of patient's admission, the area under the ROC curve was 0.75 ($p < 0.001$) for ESI and 0.71 ($p < 0.001$) for FRENCH. Comparing these results did not reveal any significant difference ($p = 0.06$).

DISCUSSION

The aim of the study was to compare the FRENCH scale with one of the most used triage scale in the world, the ESI scale. We showed a strong inter-rater reliability for both scales and good validity without significant difference.

While literature found Kappa values about 0.64 [13] and 0.77 [14] between nurses, here we obtained with the second FRENCH triage system a weighted Kappa of 0.87. This difference can be partly explained by the fact that we assessed the reliability between trained physicians. If we may assume that a medical staff is more capable of assessing the seriousness of the state of a patient, we can imagine that their evaluations converge more. We found no study concerning the reliability of the FRENCH system between medical staff in the existing literature. As for the ESI system, our study reached a weighted Kappa at 0.85 for values ranging from 0.78 to 0.98 in version 4, according to the previous studies [9;11;12].

The secondary objective was to evaluate the validity of these tools through internationally recognized indirect criteria. The ESI offered better discrimination to predict the amount of resources, length of stay, hospitalization rate, and place of admission depending on the severity level. This data is comparable to existing literature [1;7;11;15]. The calculation of the area under the curve regarding predictability of admission in our study was also comparable to the available data [1;11]. The FRENCH was also associated with all indirect criteria used in our study. Even though the literature was less abundant than for the ESI, results were similar in available data [14;16] apart for place of admission for which no data was found. The area under the curve for predictability of admission was 0.71 in our study, lower than the data found in the existing literature 0.858 [14;16].

A comparison between these two scales was made. We did not find any significant difference in reliability. Nevertheless, a difference in favor of the ESI system could be hypothesized on the secondary criteria of validity: resources and length of stay. The comparison of the areas under the curve did not reveal any significant difference ($p = 0.058$). However, this lack of difference could be due to a lack of power, the sample size having been calculated for the main objective.

This study has a main limit. The main weakness of our work is that the patients' evaluation was carried out by physicians whereas these existing tools were developed for nurses. This bias probably minimized the risks of under and over triage and self-corrected the defects of each scale. Nevertheless, previous studies showed agreements between senior physicians' and nurses' triage choices for both the FRENCH [13] and the ESI [8], allowing us to think that this bias is not major.

However this work has lot of strength. First, a simultaneous evaluation, within the same time frame, of both the ESI and the FRENCH triage systems on the same given population has never been done before. This study's prospective continuity limits the selection bias and number of patients lost to follow-up. This work also has the advantage to be performed in a real-life situation, which had not been the case much among the studies available when writing this article - reinforcing the knowledge as regards the ESI and the FRENCH. Finally, studies dealing with the FRENCH triage system are not common and the assessment of the ESI in France has rarely been described. Thus this work also improves general knowledge regarding these two scales especially for FRENCH comparatively to the ESI, one of the internationally recognized gold standards. This work pointed out the strong results of the inter-rater reliability for the FRENCH and the ESI triage scales on adult population. No significant difference was found between them. The analysis of the secondary criteria also reveals a validity of both these scales. This work allows us to validate the safety of the FRENCH system's use.

CONCLUSION

In the absence of a gold standard scale, further studies are needed to compare ESI and FRENCH. Using referent clinical cases composed by the authors of each scale allowed to test direct-validity of each scale, and inter-observer sorting with many evaluators. Next version of FRENCH scale will include pediatric patients. It will allow to compare all ages patient triage. Since a few years it has appeared e-triage scales, composed of a random model applied to triage data (vital signs, chief complaint, clinical context, medical history) to determine a triage score. These new tools need to be compared with other triage scales based on physician or nurse experiences.

REFERENCES

1. Baumann MR, Strout TD. Triage of geriatric patients in the emergency department: validity and survival with the emergency severity index. Ann Emerg Med 2007; 49:234-40.
2. Widgren BR, Jourak M. Medical emergency triage and treatment system (METTS): a new protocol in primary triage and secondary priority decision in emergency medicine. J Emerg Med 2011; 40:623-8.
3. Farrohknia N, Castrén M, Ehrenberg A, Lind L, Oredsson S, Jonsson H, et al. Emergency department triage scales and their components: a systematic review of the scientific evidence. Scand J Trauma Resusc Emerg Med 2011; 19:42.
4. Chauvin A, Truchot J, Ouled N, Plaisance P, Segal N. Evaluation of a new triage scale on patients' classification in a teaching emergency department: a retrospective study. Med Emergency, MJEM 2016; 24:5-10.

5. Taboulet P, Fontaine JP, Afdjei A, Tran C, Le Gall JR. Triage aux urgences par une infirmière d'accueil et d'orientation. Influence sur la durée d'attente et la satisfaction des consultants aux urgences. Réa Urgences 1997; 4:433-42.
6. Fernandes CMB, Tanabe P, Gilboy N, Johnson LA, McNair RS, Rosenau AM, et al. Five-level triage: a report from the ACEP/ENA Five-level triage task force. J Emerg Nurs 2005; 31:39-50.
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 Me 2004; 11:59-65.
8. Esmailian M, Zamani M, Azadi F, Ghasemi F. Inter-Rater agreement of emergency nurses and physicians in Emergency Severity Index (ESI) Triage. Emerg (Tehran). 2014; 2:158-61.
9. Mirhaghi A, Heydari A, Mazlom R, Hasanzadeh F. Reliability of the emergency severity index: Meta-analysis. Sultan Qaboos Univ Med J. 2015; 15:e71-7.
10. Christ M, Grossmann F, Winter D, Bingisser R, Platz E. Modern triage in the emergency department. Dtsch Arzteblatt Int 2010; 107:892-8.
11. Grossmann FF, Nickel CH, Christ M, Schneider K, Spirig R, Bingisser R. Transporting clinical tools to new settings: cultural adaptation and validation of the emergency severity index in German. Ann Emerg Med 2011; 57:257-64.
12. Bergs J, Verelst S, Gillet J-B, Vandijck D. Evaluating implementation of the emergency severity index in a Belgian hospital. J Emerg Nurs 2014; 40:592-7.
13. Fehre A, Chauvin A, Durand-Stocco C, Ouled N, Plaisance P, Segal N. Évaluation de l'application d'un triage par la classification infirmière des malades aux urgences par des infirmiers organisateurs de l'accueil en comparaison avec un triage réalisé par un médecin. Ann Fr Med Urg 2015; 5:150-8.
14. Taboulet P, Moreira V, Haas L, Porcher R, Braganca A, Fontaine J-P, et al. Triage with the french emergency nurses classification in hospital scale: reliability and validity. Eur J Emerg Med 2009; 16:61-7.
15. Gilboy N, Tanabe T, Travers D, Rosenau AM. Emergency severity index (ESI): A triage tool for emergency department care, version 4. Implementation handbook 2012 edition. Rockville, MD: Agency for healthcare research and quality; 2011.
16. Peschanski N, Barré M, Taboulet P, Joly LM. Validation monocentrique de la classification infirmière des malades aux urgences (CIMU V2). SFMU; 2009. http://nereja.free.fr/files/programme_2010_congres_urgences_paris_2_4_06_2010.pdf
17. Cicchetti D, Bronen R, Spencer S, Haut S, Berg A, Oliver P, et al. Rating scales, scales of measurement, issues of reliability: resolving some critical issues for clinicians and researchers. J Nerv Ment Dis 2006; 194:557-64.
18. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Stat Med 2012; 31:3972-81.
19. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973; 33:613-61.