

Original Article
Emergency &
Critical Care Medicine



Reliability of Korean Triage and Acuity Scale: Interrater Agreement between Two Experienced Nurses by Real-Time Triage and Analysis of Influencing Factors to Disagreement of Triage Levels

OPEN ACCESS

Received: May 5, 2019

Accepted: Jul 3, 2019

Address for Correspondence:

Tae Ho Lim, MD, PhD

Departments of Emergency Medicine,
College of Medicine, Hanyang University, 222
Wangsimni-ro, Seongdong-gu, Seoul 04763,
Korea.

E-mail: erthim@gmail.com

*Joon Bum Park and Juncheol Lee contributed
equally to this work.

© 2019 The Korean Academy of Medical
Sciences.

This is an Open Access article distributed
under the terms of the Creative Commons
Attribution Non-Commercial License ([https://
creativecommons.org/licenses/by-nc/4.0/](https://creativecommons.org/licenses/by-nc/4.0/))
which permits unrestricted non-commercial
use, distribution, and reproduction in any
medium, provided the original work is properly
cited.

ORCID iDs

Joon Bum Park

<https://orcid.org/0000-0003-3052-2339>

Juncheol Lee

<https://orcid.org/0000-0003-3058-4759>

Yu Jin Kim

<https://orcid.org/0000-0001-7449-9025>

Jin Hee Lee

<https://orcid.org/0000-0001-5479-1133>

Tae Ho Lim

<https://orcid.org/0000-0003-1045-413X>

Funding

This work was supported by the
Soonchunhyang University Research Fund.

Joon Bum Park ^{1*}, Juncheol Lee ^{2,3*}, Yu Jin Kim ⁴, Jin Hee Lee ⁵, and
Tae Ho Lim ⁶

¹Department of Emergency Medicine, Soonchunhyang University College of Medicine, Seoul, Korea

²Department of Emergency Medicine, Armed Forces Capital Hospital, Seongnam, Korea

³Graduate School, College of Medicine, Hanyang University, Seoul, Korea

⁴Department of Emergency Medicine, Seoul National University Bundang Hospital, Seoul, Korea

⁵National Emergency Medical Center, National Medical Center, Seoul, Korea

⁶Department of Emergency Medicine, Hanyang University College of Medicine, Seoul, Korea

ABSTRACT

Background: All emergency centers in Korea use the Korean Triage and Acuity Scale (KTAS) as their initial triage tool. However, KTAS has been used without verification of its reliability. In this study, we assess the interrater agreement of KTAS by two independent nurses in real-time and analyse the factors which have an effect on the disagreement of KTAS levels.

Methods: This study was a prospective observational study conducted with patients who visited an emergency department (ED). Two teams, each composed of two nurses, triaged patients and recorded KTAS level and the main complaint from the list of 167 KTAS complaints, as well as modifiers. Interrater reliability between the two nurses in each team was assessed by weighted-kappa. Pearson's χ^2 test was conducted to determine if there were differences between each nurse's KTAS levels, depending on whether they chose the same complaints and the same modifiers or not.

Results: The two teams triaged a total of 1,998 patients who visited the ED. Weighted-kappa value was 0.772 (95% confidence interval [CI], 0.750–0.794). Patients triaged by different chosen complaints showed (38.0%) higher inconsistency rate in KTAS levels than those triaged by the same complaint (10.9%, $P < 0.001$). When nurses chose the same complaint and different modifiers, the ratio of different levels (50.5%) was higher than that of the same complaint and same modifier (8.1%, $P < 0.001$).

Conclusion: This study showed that KTAS is a reliable tool. Selected complaints and modifiers are confirmed as important factors for reliability; therefore, selecting them properly should be emphasized during KTAS training courses.

Keywords: Reliability; Triage; Emergency Department

Disclosure

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: Lim TH. Data curation: Kim YJ. Investigation: Lee JH. Methodology: Kim YJ. Supervision: Lee JH. Writing - original draft: Park JB, Lee J. Writing - review & editing: Park JB, Lee J.

INTRODUCTION

In Korea, since the early 1990s, studies on the classification of severity and acuity for emergent patients have been performed, and there has been a continuous need for improvement on this topic. In 2012, the Korean Society of Emergency Medicine (KSEM) first developed the hospital-level Korean Triage and Acuity Scale (KTAS), which was based on the Canadian Triage and Acuity Scale (CTAS).¹ CTAS is an excellent triage scale that was developed with an interest in national emergency medical systems not limited to the hospital level but also including the prehospital level.² Several countries have applied CTAS for their own national emergency medical systems.^{3,4}

Although KTAS is a triage instrument based on CTAS, which has been verified as a valid and reliable scale,⁵⁻⁷ it should be applied and interpreted differently according to its ethnic, cultural, language, and medical environment features. In Korea, KTAS researchers have tried to reduce mistakes in the simple translation of CTAS by distributing questionnaires and conducting Delphi surveys among emergency specialists during the development of KTAS.⁸ Hence, although CTAS's reliability had already been verified, the reliability of KTAS needs to be reverified. To our best knowledge, there are no studies that have analyzed the reliability of the triage tool by independent nurses in real-time or investigated the factors which have an effect on the disagreement between triage nurses.

Reliability between nurses during the triage process is an essential condition of a triage tool; however, KTAS has been used without verification of its reliability.⁹ In this study, we assess the interrater agreement of KTAS by two independent nurses and analyze the factors which have an effect on the disagreement of KTAS levels.

METHODS

This study was a prospective observational study conducted from June 15, 2018 to August 24, 2018, with patients who visited an emergency department (ED). Four nurses were enrolled. We created two teams which were composed of two nurses; each team was assigned to the ED at a different time.

Study design

All research nurses had more than three years of experience in the ED and had received KTAS training. When the hospital-affiliated triage nurse performed KTAS triage on actual patients who visited the ED, both research nurses observed the triage process and independently performed a triage using KTAS in real time. Research nurses did not ask patients any question directly to not affect the real ED process. Two research nurses who were grouped as a team independently triaged patients who visited the ED from 08:00 a.m. to 10:00 p.m., including weekends.

Data collection and definition of modifiers

Each team was composed of two nurses, and in total, two teams triaged patients in the ED at different times. Each research nurse recorded KTAS level and the main complaint of the patient, which nurses thought most fit the patient from the list of 167 KTAS complaints, as well as the 1st- or 2nd-order modifiers, which determine the final KTAS level. When research nurses chose the 1st-order modifiers to decide KTAS level, they recorded which modifier they selected from the list of pain, hemorrhagic disease, hemodynamic state, body

temperature, respiration, conscious state, and accident mechanism. Second-order modifiers are complaint-specific modifiers regardless of vital signs; in the case of “Burn”, percentage of burned area in body surface area (BSA) should be considered to evaluate the patient's severity or acuity. Therefore, as a 2nd-order modifier, “Burn 25% > BSA” is classified as KTAS level 2, “5%–25% of BSA” level 3 and “< 5% of BSA” level 4.

Statistical analysis

To analyze whether the different choice of patient's main complaints and modifiers between research nurses makes a difference in the decision of KTAS level, we divided cases with different levels among the two research nurses into cases with different chosen complaints and cases with the same chosen complaints. When the chosen patient's complaints were the same, cases were further divided according to whether the selected modifiers were the same or different among two research nurses.

Interrater reliability between the two research nurses was assessed by weighted-kappa. Weighted values were 1 in the same level, 0.75 in one level difference and 0.5 in two level difference. Fisher's exact test was conducted to determine if there were differences between each research nurse's KTAS levels, depending on whether they chose the same complaints and the same modifiers or not. Results were considered statistically significant when the significance level was less than 0.05. Statistical analysis was performed using SPSS statistics software, version 18.0 (SPSS Inc., Chicago, IL, USA) and R software, version 3.4.2 (R Project for Statistical Computing, Vienna, Austria).

Ethics statement

This study protocol was reviewed and approved by the Institutional Review Board (IRB) of Hanyang University (IRB No. 2017-11-032). Informed consent was obtained from four research nurses before they were enrolled.

RESULTS

Analysis of interrater agreement for KTAS levels

During the study period, two teams respectively triaged 1,000 and 998 patients who visited the ED. All research nurses performed triages independently in the ED at the same time for the same patient. **Table 1** compares KTAS levels of the total 1,998 cases triaged by the two teams. Weighted-kappa value was 0.772 (95% confidence interval [CI], 0.750–0.794) and showed substantial agreement.¹⁰

The number of cases in which two research nurses did not agree in KTAS level was 259 (13.0%) out of 1,998. In cases covering levels 3 and 4, which can include severe or mild patients, 137

Table 1. Weighted-kappa value between classifiers

Variables	Level 1	Level 2	Level 3	Level 4	Level 5	Total
Level 1	32	1	0	0	0	33
Level 2	2	293	28	6	0	329
Level 3	0	24	466	66	3	559
Level 4	0	4	71	664	28	767
Level 5	0	3	3	20	284	310
Total	34	325	568	756	315	1,998

Weighted-kappa = 0.772; 95% confidence interval, 0.750–0.794.

Table 2. Analysis of KTAS levels consistency ratio between consistency of complaints

Variables	Same complaints	Different complaints	Total	P value
Same level	1,361 (89.1)	101 (62.0)	1,462 (86.5)	< 0.001
Different level	166 (10.9)	62 (38.0)	228 (13.5)	
Total	1,527 (100.0)	163 (100.0)	1,690 (100.0)	

Data are presented as number (%).

KTAS = Korean Triage and Acuity Scale.

(10.8% of total levels 3 and 4) patient's KTAS level was inconsistent between the two research nurses (**Table 1**). There were cases that KTAS level disagreement between two research nurses was level 1 or 2 versus level 4 or 5. These different triage results may affect the course of a patient's treatment priority in an emergency situation. Thirteen cases showed these discrepancies. In seven cases, research nurses selected different complaints from each other; for example, 2 cases of 'severe trauma-penetration vs. laceration', 3 cases of 'cardiogenic chest pain vs. non-cardiogenic chest pain', and each case of 'post-operative complication vs. tooth gum problem' and 'general weakness vs. vertigo'. Although nurses chose the same complaint, six cases showed discrepancies by difference of modifiers. In two cases, each nurse triaged differently by the degree of respiratory distress in 'shortness of breath' and by the estimation of hemodynamic state in 'direct referral for consultation'. In another four cases, both nurses chose 2nd-order modifiers however, they triaged differently by the amount of 'epistaxis', the severity of 'localized swelling/redness' and the acuity of 'visual disturbance' in 2 cases.

Analysis of KTAS levels inconsistency according to complaints selected

We analyzed 1,690 cases excluding 308 cases where either of the two research nurses missed any modifier data. Among these 1,690 cases, 1,462 (86.5%) were triaged at the same level but 228 (13.5%) were not (**Table 2**). Among cases with the same chosen complaints, 1,361 (89.1%) were triaged at the same level, and 166 (10.9%) were not. Among cases with different chosen complaints, 101 (62.0%) were triaged at the same level, and 62 (38.0%) were not. Patients triaged by different chosen complaints showed (38.0%) higher inconsistency rate in KTAS levels than those triaged by the same complaint (10.9%, $P < 0.001$).

Analysis of inconsistency of KTAS levels according to selected modifiers

Even if two research nurses chose the same complaint, final levels may be different between two nurses when they chose different modifiers. When research nurses chose the same complaint and modifiers, the ratio of different KTAS levels was 8.1% (**Table 3**). However, when they chose the same complaint and different modifiers, the ratio of different KTAS levels was 50.5% ($P < 0.001$). This means that even if nurses selected the same complaint, the level decision can be affected depending on whether they choose the same modifiers or not. In the case of different selected modifiers, the most frequent pair was composed of a 2nd-order modifier and one of 1st-order modifiers (**Table 4**). In fifty-one pairs of total cases, thirty-nine cases (76.5%) included 2nd-order modifiers and the second most frequent modifier was pain, one of 1st-order modifiers, in 25 cases (49.0%).

Table 3. Analysis of KTAS levels according to consistency of modifiers by the same complaint

Variables	Same complaint		Total	P value
	Same modifier	Different modifier		
Same level	1,311 (91.9)	50 (49.5)	1,361 (89.1)	< 0.001
Different level	115 (8.1)	51 (50.5)	166 (10.9)	
Total	1,426 (100.0)	101 (100.0)	1,527 (100.0)	

Data are presented as number (%).

KTAS = Korean Triage and Acuity Scale.

Table 4. Details of the cases with same complaint but different modifiers in different levels

Pair of modifiers	No. of pair(s)	Total pair(s) containing each modifier (%)
Accident mechanism		1/51 (2.0)
Temperature	1	
Consciousness		1/51 (2.0)
2nd modifier	1	
Hemodynamics		16/51 (31.4)
Pain	7	
Temperature	2	
2nd modifier	7	
Hemorrhagic disease		1/51 (2.0)
2nd modifier	1	
Pain		25/51 (49.0)
Hemodynamics	7	
Temperature	1	
2nd modifier	17	
Respiration		9/51 (17.6)
Temperature	1	
2nd modifier	8	
Temperature		10/51 (19.6)
Accident mechanism	1	
Hemodynamics	2	
Pain	1	
Respiration	1	
2nd modifier	5	
2nd-order modifier		39/51 (76.5)
Consciousness	1	
Hemodynamics	7	
Hemorrhagic disease	1	
Pain	17	
Respiration	8	
Temperature	5	
Total	51^a	

Each shaded box is a duplicate of an unshaded box only in the opposite order of modifiers.

^aSum of either shaded or unshaded boxes.

DISCUSSION

Several studies have verified the reliability of triage tools. In one study, researchers used paper-based case scenarios or medical records of triage nurses.^{7,11} In another study, authors compared triage levels defined by two independent nurses and the weighted-kappa value was 0.87.¹² However, in this study, they compared triage level of senior nurses with one of junior nurses. Therefore, disagreement between two research nurses cannot but be affected by their experience. To reduce this bias, the present study included two independent nurses with similar ED experience and triaged the same patient in the same place and at the same time with the same information.

Another study related to KTAS reliability compared KTAS levels of triage nurses with that of poly-clinic students.¹³ Although poly-clinic students received one-hour training by an emergency physician, their clinical experience was not enough to guarantee the quality of the triage. Therefore, in our study we evaluated KTAS levels defined by two independent experienced nurses which will guarantee a more valid assessment of KTAS reliability.

The weighted-kappa value of KTAS levels in this study showed substantial agreement. Amir et al.¹⁴ performed a meta-analysis study about interrater agreement of CTAS and the range of

weighted-kappa values ranged from 0.40 to 0.84. Unlike our study, this meta-analysis research indicated that most reliability studies had used patient-scenarios, not real patients; and the weighted-kappa value of our study is relatively high compared with those of other studies.

This study was designed to analyze not only the overall agreement between the two nurses but also to assess the effect of factors that caused nurses to triage the same patient differently. According to the algorithm of CTAS-based KTAS, complaints and the modifiers are key factors to decide KTAS levels. Our results confirmed that there were statistically significant differences in KTAS levels created by the inconsistency between complaints or modifiers.

In KTAS, modifiers are composed of 1st- and 2nd-order modifiers. Second-order modifiers are much more complaint specific and can change the initial KTAS level, which is based on the initial 1st-order modifiers, to the higher level. These 2nd-order modifiers are likely to have been determined by expert opinions when triage levels were assigned to each item. Further research and revisions are required to ensure that the same level of 1st- and 2nd-order modifiers have the same degree of severity and acuity.

This study focused on the fact that CTAS is a complaint-oriented triage tool and aimed to confirm that the degree of agreement between levels can be affected by the selected main complaints and modifiers. Based on the results of this study, it should be emphasized in KTAS education courses that triage nurse's choice of more proper complaint and modifier is important to decide the most proper KTAS level. In addition, some factors of the KTAS, like mismatch between 1st- and 2nd-order modifiers, should be corrected by further research.

This study has some limitations. First, we did not design a study where research nurses recorded the patient number and linked to the electrical medical record of the hospital. The research nurses were positioned behind the patient on each side where they could not see the monitor of triage nurse to assess the patient number. As a result, the emergent care process of patients was not available to be analyzed. Second, the research nurses did not have enough time to record all the items because they were not allowed to skip any patients to eliminate selection bias. Further research with a different study design should include the result of emergent care and the reason why nurses choose any particular complaint and modifier. Third, each research nurse triaged patients only with the information gathered by the triage nurse. It could not reflect difference in the examination method between two research nurses. The results of this study might have been different if each research nurse was able to obtain the information respectively from the patient. If research nurses asked their own questions, we may have been able to get more realistic results.

This study showed that KTAS is a reliable tool. Selected complaints and modifiers were confirmed as important factors; therefore, selecting them properly should be emphasized during KTAS training course. Concerning the revision of KTAS, the list of complaints and guidelines for using modifiers to improve the reliability of KTAS should be verified. We expect that further studies of complaints and modifiers in KTAS will contribute to improve its reliability.

REFERENCES

1. Park J, Lim T. Korean Triage and Acuity Scale (KTAS). *J Korean Soc Emerg Med* 2017;28(6):547-51.

2. Leeies M, Ffrench C, Strome T, Weldon E, Bullard M, Grierson R. Prehospital application of the Canadian Triage and Acuity Scale by Emergency Medical Services. *CJEM* 2017;19(1):26-31.
[PUBMED](#) | [CROSSREF](#)
3. Kuriyama A, Ikegami T, Kaihara T, Fukuoka T, Nakayama T. Validity of the Japan Acuity and Triage Scale in adults: a cohort study. *Emerg Med J* 2018;35(6):384-8.
[PUBMED](#) | [CROSSREF](#)
4. Chang W, Liu HE, Goopy S, Chen LC, Chen HJ, Han CY. Using the five-level Taiwan Triage and Acuity Scale computerized system: factors in decision making by emergency department triage nurses. *Clin Nurs Res* 2017;26(5):651-66.
[PUBMED](#) | [CROSSREF](#)
5. Dong SL, Bullard MJ, Meurer DP, Blitz S, Akhmetshin E, Ohinmaa A, et al. Predictive validity of a computerized emergency triage tool. *Acad Emerg Med* 2007;14(1):16-21.
[PUBMED](#) | [CROSSREF](#)
6. Jiménez JG, Murray MJ, Beveridge R, Pons JP, Cortés EA, Garrigós JB, et al. Implementation of the Canadian Emergency Department Triage and Acuity Scale (CTAS) in the principality of andorra: can triage parameters serve as emergency department quality indicators? *CJEM* 2003;5(5):315-22.
[PUBMED](#) | [CROSSREF](#)
7. Fernandes CM, McLeod S, Krause J, Shah A, Jewell J, Smith B, et al. Reliability of the Canadian Triage and Acuity Scale: interrater and intrarater agreement from a community and an academic emergency department. *CJEM* 2013;15(4):227-32.
[PUBMED](#) | [CROSSREF](#)
8. Park J, Choi H, Kang B, Kim C, Kang H, Lim T. A nationwide survey of Korean emergency department triage systems and scales: a first step towards reform of the emergency medical service system. *J Korean Soc Emerg Med* 2014;25(5):499-508.
9. Hardern RD. Critical appraisal of papers describing triage systems. *Acad Emerg Med* 1999;6(11):1166-71.
[PUBMED](#) | [CROSSREF](#)
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-74.
[PUBMED](#) | [CROSSREF](#)
11. Alquraini M, Awad E, Hijazi R. Reliability of Canadian Emergency Department Triage and Acuity Scale (CTAS) in Saudi Arabia. *Int J Emerg Med* 2015;8(1):80.
[PUBMED](#) | [CROSSREF](#)
12. Ng CJ, Yen ZS, Tsai JC, Chen LC, Lin SJ, Sang YY, et al. Validation of the Taiwan triage and acuity scale: a new computerised five-level triage system. *Emerg Med J* 2011;28(12):1026-31.
[PUBMED](#) | [CROSSREF](#)
13. Kim JY, Hong DY, Kim SY, Kim JW, Park SO, Lee KR, et al. Reliability of Korean Triage and Acuity Scale-based Triage system as a severity index in emergency patients. *J Korean Soc Emerg Med* 2017;28(6):552-6.
14. Mirhaghi A, Heydari A, Mazlom R, Ebrahimi M. The reliability of the Canadian Triage and Acuity Scale: meta-analysis. *N Am J Med Sci* 2015;7(7):299-305.
[PUBMED](#) | [CROSSREF](#)