Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Human Triage versus Computer-Based Triage Using a Predetermined Flowchart

Shruti silwal

shrutisilwal@pahs.edu.np

Patan Academy of Health Sciences

Kripa Maharjan

Patan Academy of Health Sciences

Abhishek Kafle

Patan Academy of Health Sciences

Anupama Gnawali

Patan Academy of Health Sciences

Amar Das

Patan Academy of Health Sciences

Sinchan Pokhrel

Patan Academy of Health Sciences

Zenish Niraula

Patan Academy of Health Sciences

Ritika Shrivastab

Patan Academy of Health Sciences

Research Article

Keywords: emergency triage, clinical decision support, WHO Interagency Integrated Triage Tool, inter-rater reliability, Cohen's kappa, Nepal

Posted Date: August 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10521238/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background: Emergency department (ED) triage is central to patient safety, yet human triage is prone to variability. We compared a rule-based computer application implementing the WHO Interagency Integrated Triage Tool (WHO-IITT) against routine human triage, using independent expert assessment as the reference standard.

Methods: In this prospective observational study at Patan Hospital ED, Nepal, patients aged ≥12 years were triaged by either routine human triage or a web-based WHO-IITT flowchart application, assigned by ED shift. All cases were independently reviewed by a blinded expert panel. Agreement was quantified with Cohen's kappa and 95% confidence intervals (CI); accuracy, under-triage, and over-triage rates were compared between methods using two-proportion z-tests.

Results: Of 737 patients, 359 underwent human triage and 378 underwent computer-based triage. Computer-based triage was more accurate than human triage (86.5%, 95% CI 83.1–90.0% vs. 67.4%, 95% CI 62.6–72.3%; p<0.001) and showed substantially stronger agreement with expert assessment (κ = 0.743, 95% CI 0.677–0.808) than human triage (κ = 0.386, 95% CI 0.294–0.477; difference p<0.001). Over-triage was lower with computer-based triage (7.7% vs. 23.7%; p<0.001); under-triage was numerically lower (5.8% vs. 8.9%) but did not reach statistical significance (p=0.11). For RED (emergency) patients, sensitivity was higher with computer-based triage (83.5%, 95% CI 76.9–90.1%) than human triage (71.7%, 95% CI 62.5–80.9%; p=0.039).

Conclusion: A standardized, flowchart-based computer triage application achieved substantially higher accuracy and inter-rater reliability than routine human triage, with fewer over-triage events and better identification of high-acuity patients. As a decision-support adjunct to clinical judgment, such tools may improve triage consistency and safety in high-volume, resource-limited EDs.

Background

Emergency departments (EDs) are essential components of health systems, providing immediate care across a broad spectrum of clinical severity. Triage is central to ED operations: it prioritizes patients by clinical urgency, supports timely intervention for critically ill patients, and directs limited resources efficiently. Inaccurate or inconsistent triage can delay care, worsen outcomes, and contribute to overcrowding and prolonged waiting times.[1]

Conventional triage relies on clinical judgment supported by structured tools, but human-led triage remains vulnerable to variability driven by clinical experience, workload, cognitive bias, and time pressure. Prior work has documented substantial inter-rater variation among emergency providers, raising concern about under-triage of patients with time-sensitive conditions.[8]

Computerized and electronic triage systems have been developed to standardize decision-making through structured, rule-based algorithms. Evidence to date suggests such systems can improve

agreement with expert assessment and identification of high-acuity patients relative to traditional triage, and may reduce mis-triage and improve documentation quality, although further validation across diverse settings is needed.[6,7,9]

In resource-limited settings, the World Health Organization's Interagency Integrated Triage Tool (WHO-IITT) offers a standardized, flowchart-based approach that classifies patients as RED (emergency), YELLOW (priority), or GREEN (non-urgent) based on predefined danger signs and presenting complaints. It is designed to be simple, rapid, and reproducible across cadres of health workers, and validation studies in low-resource EDs have reported acceptable performance and high inter-rater reliability.[10,11]

Patan Hospital ED is among the largest tertiary emergency care centers in Nepal, managing a high and diverse patient volume under conditions of crowding and limited resources conditions in which variability in human triage may contribute to mis-triage and delayed care. We evaluated a rule-based computer application implementing the WHO-IITT adult flowchart against routine human triage, using independent expert review as the reference standard.

This study aimed to compare human triage and computer-based flowchart triage in terms of accuracy and inter-rater reliability against expert assessment, with particular attention to identification of RED (emergency) patients.

Methods

Study Design and Setting

This was a prospective observational study conducted in the Emergency Department of Patan Hospital, a tertiary referral center in Nepal. The study compared human triage and computer-based flowchart triage, with triage accuracy and inter-rater reliability against expert assessment as the primary outcomes. Reporting follows the STROBE guideline for observational studies.

Guideline Framework

The computer-based system operationalized the WHO-IITT, adapted for the Nepalese context by the Ministry of Health and Population.[11] The adult triage flowchart, applicable to patients aged ≥ 12 years, was used throughout.

Study Population

All patients aged ≥ 12 years presenting to the ED during the study period were eligible. Excluded were patients aged < 12 years, pregnant women, and polytrauma patients.

Triage Allocation

Patients were assigned to human triage or computer-based triage according to ED shift allocation rather than individual-level randomization. This pragmatic, non-randomized design reflects real-world implementation but means shift-level differences in patient case-mix, staffing, or time of presentation cannot be fully excluded as sources of confounding; we address this in the Limitations.

Human triage was performed by trained triage nurses or designated triage officers following routine hospital protocols. Computer-based triage was delivered through a web-based clinical decision-support application operationalizing the WHO-IITT adult flowchart via a deployed Google Apps Script backend. The flowchart proceeded stepwise danger signs, then moderate symptoms and risk factors, then low-risk features with predefined severity thresholds evaluated at each decision node until a final RED, YELLOW, or GREEN classification was generated. The system was used strictly as a decision-support tool and did not replace clinical judgment; clinicians retained authority to override its output.

Expert Review (Reference Standard)

All triage decisions ,human and computer-based — were independently reviewed by a panel of emergency physicians trained in WHO-IITT principles, blinded to which method had generated each classification. The panel's assessment served as the reference standard for accuracy and agreement analyses.

Outcome Measures

Primary outcomes were: (1) triage accuracy of each method against expert review; (2) inter-rater reliability (Cohen's kappa), with particular attention to RED-category agreement; and (3) rates of under-triage and over-triage relative to expert classification.

Data Collection and Statistical Analysis

Data were collected prospectively on a structured pro forma capturing demographics, presenting complaint, vital signs, assigned triage category, and expert review outcome, and entered into Microsoft Excel. Categorical variables are summarized as counts and percentages. Inter-rater reliability was quantified using unweighted Cohen's kappa with 95% confidence intervals (large-sample asymptotic standard error). Accuracy, under-triage, and over-triage proportions were compared between human and computer-based triage using two-proportion z-tests, with 95% CIs calculated by normal approximation. A two-sided $p < 0.05$ was considered statistically significant. Analyses were performed in Epi Info and cross-checked in Python 3.

Results

Study Participants

A total of 737 patients were included: 359 (48.7%) underwent human triage and 378 (51.3%) underwent computer-based triage. Baseline demographic characteristics were comparable between groups (Table 1). Females comprised 407 (55.2%) of the cohort, and most patients were aged 20–59 years (425, 57.7%).

Table 1
Baseline demographic characteristics (n = 737)

Characteristic	Human triage (n = 359)	Computer triage (n = 378)	Total (n = 737)
Male	156 (43.5)	174 (46.0)	330 (44.8)
Female	203 (56.5)	204 (54.0)	407 (55.2)
Age 13–19 y	20 (5.6)	38 (10.1)	58 (7.9)
Age 20–59 y	210 (58.5)	215 (56.9)	425 (57.7)
Age ≥ 60 y	129 (35.9)	124 (32.8)	253 (34.3)

Values are n (%). Patients aged < 12 years were excluded.

Distribution of Triage Categories

By expert assessment, 213 patients (28.9%) were RED, 424 (57.5%) YELLOW, and 100 (13.6%) GREEN (Table 2). Both triage methods under-called the GREEN category relative to experts, and human triage assigned proportionally more patients to YELLOW.

Table 2
Triage category assigned by human triage, computer-based triage, and expert assessment

Category	Human triage (n = 359)	Computer triage (n = 378)	Expert (n = 737)
RED	97 (27.0)	111 (29.4)	213 (28.9)
YELLOW	247 (68.8)	245 (64.8)	424 (57.5)
GREEN	15 (4.2)	22 (5.8)	100 (13.6)

Agreement with Expert Assessment

Agreement between human triage and expert assessment was fair (unweighted $\kappa = 0.386$, 95% CI 0.294–0.477; Table 3, Table 5). Agreement between computer-based triage and expert assessment was substantial ($\kappa = 0.743$, 95% CI 0.677–0.809; Table 4, Table 5). The difference in kappa between methods was statistically significant ($z = 6.22$, $p < 0.001$).

Table 3
Human triage vs. expert assessment

Human \ Expert	RED	YELLOW	GREEN	Total
RED	66	30	1	97
YELLOW	24	169	54	247
GREEN	2	6	7	15
Total	92	205	62	359

Table 4
Computer-based triage vs. expert assessment

Computer \ Expert	RED	YELLOW	GREEN	Total
RED	101	10	0	111
YELLOW	19	207	19	245
GREEN	1	2	19	22
Total	121	219	38	378

Triage Accuracy

Computer-based triage correctly classified 327 of 378 cases (86.5%, 95% CI 83.1–90.0%), significantly more than human triage, which correctly classified 242 of 359 cases (67.4%, 95% CI 62.6–72.3%; $z = 6.18$, $p < 0.001$) (Table 5).

Table 5
Accuracy, agreement, and misclassification by triage method (95% CI)

Metric	Human triage	Computer triage	p-value*
Correct classification, n (%)	242 (67.4)	327 (86.5)	< 0.001
95% CI	62.6–72.3%	83.1–90.0%	
Under-triage, n (%)	32 (8.9)	22 (5.8)	0.107
95% CI	6.0–11.9%	3.5–8.2%	
Over-triage, n (%)	85 (23.7)	29 (7.7)	< 0.001
95% CI	19.3–28.1%	5.0–10.4%	
Cohen's κ (vs. expert)	0.386	0.743	< 0.001
95% CI	0.294–0.477	0.677–0.809	

**Two-proportion z-test for accuracy, under-triage, and over-triage; z-test on the kappa difference for agreement. CI = confidence interval.*

RED (Emergency) Category Performance

Among expert-classified RED cases, computer-based triage correctly identified 101 of 121 (83.5%, 95% CI 76.9–90.1%), compared with 66 of 92 (71.7%, 95% CI 62.5–80.9%) for human triage — a statistically significant difference (z = 2.06, p = 0.039) (Table 6). Misclassification of RED patients into a lower-acuity category was less frequent with computer-based triage (16.5% vs. 28.3%).

Table 6
Sensitivity for RED-category patients

Triage method	Correct RED, n (%)	95% CI	Misclassified, n (%)	Total RED (expert)
Human triage	66 (71.7)	62.5–80.9%	26 (28.3)	92
Computer triage	101 (83.5)	76.9–90.1%	20 (16.5)	121

Discussion

In this prospective comparison at a high-volume tertiary ED in Nepal, a standardized computer-based triage application achieved substantially higher accuracy and inter-rater reliability against expert assessment than routine human triage, with fewer over-triage events and better sensitivity for RED (emergency) patients. Under-triage was numerically lower with computer-based triage but the difference did not reach statistical significance, which we discuss below.

Triage Accuracy and Standardization

The higher accuracy observed with computer-based triage is consistent with prior reports that standardized, rule-based decision logic reduces subjective variation relative to manual triage.[1,2,4] Alumran et al. found that electronic triage systems improved documentation quality and reduced inter-observer variability compared with manual processes.[1] The comparatively lower accuracy of human triage in our cohort likely reflects the well-documented influence of clinician experience, cognitive load, and ED crowding on triage decisions.[3]

Agreement with Expert Assessment

Kappa agreement with expert assessment was substantial for computer-based triage but only fair for human triage, and this difference was statistically significant. This pattern echoes prior work reporting higher inter-rater reliability for algorithm-driven triage than for clinician-led triage.[6] Pourasghar et al. similarly found improved concordance between electronic triage and physician assessment, attributing this to the reproducibility of standardized algorithms.[6]

Under-Triage, Over-Triage, and Patient Safety

Over-triage was significantly lower with computer-based triage, driven mainly by fewer GREEN patients being escalated to YELLOW; excessive over-triage can strain ED throughput and contribute to crowding.[3]

Under-triage — the safety-critical error — was numerically lower with computer-based triage (5.8% vs. 8.9%) but this difference did not reach statistical significance (p = 0.107), likely reflecting more limited statistical power for this less frequent outcome. This finding should therefore be interpreted cautiously: while the point estimates favor the computer-based system, our data do not permit a firm claim that it reduces under-triage, and this is an important target for confirmation in larger, adequately powered studies. Decision-support systems have elsewhere been associated with improved recognition of high-acuity presentations by enforcing consistent application of predefined criteria.[2,5]

Performance for High-Acuity (RED) Patients

Correct identification of RED patients is arguably the most safety-critical function of any triage system. Computer-based triage showed significantly higher sensitivity for RED patients than human triage (83.5% vs. 71.7%, p = 0.039), with fewer RED patients down-classified to YELLOW or GREEN. This is consistent with prior reports that computerized and AI-assisted triage tools improve sensitivity for high-acuity cases, particularly in overcrowded settings.[2,4]

Implications for Practice

These findings suggest that a standardized, flowchart-based computer application can improve triage accuracy, consistency, and RED-patient detection in a high-volume, resource-limited ED. Such tools appear best positioned as decision-support adjuncts rather than replacements for clinical judgment, with clinicians retaining oversight for atypical or complex presentations not fully captured by rule-based logic.[5]

Strengths and Limitations

Strengths include the prospective design, use of a blinded expert panel as reference standard, adequate sample size across acuity strata, and quantification of agreement using Cohen's kappa with confidence intervals rather than point estimates alone.

Several limitations warrant emphasis. First, and most importantly, allocation to triage method followed ED shift rather than individual-level randomization; unmeasured differences in patient case-mix, staffing ratios, or time-of-day acuity between shifts could confound the comparison, and this design cannot fully separate the effect of the triage method itself from shift-level factors. Second, this was a single-center study, which may limit generalizability to EDs with different patient populations, staffing models, or resource levels. Third, although expert panel assessment is an accepted reference standard, it retains an element of subjective judgment and was not itself validated against a fully independent gold standard. Fourth, we assessed triage accuracy and agreement but did not evaluate downstream clinical outcomes such as time to intervention, length of stay, morbidity, or mortality, so the clinical impact of improved triage accuracy remains inferential. Fifth, the computer-based system is rule-based and may not fully capture atypical or complex presentations, underscoring the continued need for clinical oversight. Finally, the non-significant difference in under-triage should not be over-interpreted as equivalence; the study may have been underpowered for this specific comparison.

Conclusions

In this prospective, single-center comparison, computer-based triage using a predetermined WHO-IITT flowchart achieved significantly higher accuracy, stronger agreement with expert assessment, less over-triage, and better sensitivity for RED-category patients than routine human triage. Under-triage was numerically but not statistically significantly lower. Used as a decision-support adjunct alongside clinical judgment not a replacement for it standardized computer-based triage may improve consistency and safety in high-volume, resource-limited emergency departments, though confirmation in multicenter, randomized, and outcome-linked studies is needed before broader adoption.

Recommendations

For Clinical Practice

• Introduce computer-based triage as a decision-support adjunct to improve accuracy and consistency, while preserving clinician authority to override its output when clinically indicated.

• Provide structured training for triage staff on appropriate use and limitations of the tool, particularly for atypical presentations.

For Health Systems and Policy

• Consider phased scale-up of standardized flowchart-based triage tools in high-volume EDs, especially in resource-limited settings, alongside quality-improvement monitoring of under- and over-triage rates.

For Future Research

• Conduct multicenter, ideally randomized, studies to confirm generalizability and to adequately power comparisons of under-triage.

• Link triage classification to downstream outcomes (time to intervention, length of stay, morbidity, mortality).

• Evaluate cost-effectiveness of implementation in low- and middle-income country EDs.

Abbreviations

ED
Emergency Department
WHO
World Health Organization
WHO-IITT
World Health Organization Interagency Integrated Triage Tool
CI

confidence interval
κ
Cohen's kappa

Declarations

Ethics, Consent to Participate, and Consent to Publish

This study was approved by the Institutional Review Committee of Patan Academy of Health Sciences (Reference No. FLW2506062021). As this was an observational study using routinely collected triage data with no additional intervention, the requirement for individual written informed consent was waived by the committee. Not applicable for consent to publish, as no identifying individual patient data, images, or case details are included in this manuscript.

Consent for Publication

Not applicable.

Availability of Data and Materials

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions

SS conceived and designed the study. RS developed and implemented the computer-based triage application. AA, AK, AD, SP and ZN collected patient triage data and conducted the expert panel review used as the reference standard. SS and AA performed the statistical analysis. SS drafted the manuscript. SS and KM critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Alumran A, Alkhaldi O, Aldroorah Z, Alsayegh Z, Alsafwani F, Almaghraby N. Utilization of an electronic triage system by emergency department nurses. *Journal of Multidisciplinary Healthcare*. 2020;13:339–344. |DOI||Website||Google Scholer||Pubmed|
2. Tyler S, Olis M, Aust N, Patel L, Simon L, Triantafyllidis C, et al. Use of artificial intelligence in triage in hospital emergency departments: a scoping review. *Cureus*. 2024;16(5):e12345. |DOI||Website| |Google Scholer| |Pubmed |
3. Furmaga J, McDonald SA. Impact of rapid medical evaluation on patient flow in an urban emergency department. *Journal of Medical Systems*. 2021;45(6):63. |DOI| Website| Google Scholer | Pubmed|
4. Porto BM. Improving triage performance in emergency departments using machine learning and natural language processing: a systematic review. *BMC Emergency Medicine*. 2024;24:219. |DOI| Website | Google Scholer| Pubmed|
5. Chen Z, Liang N, Zhang H, Li H, Yang Y, Zong X, et al. Harnessing the power of clinical decision support systems: challenges and opportunities. *Open Heart*. 2023;10:e002432. |DOI |Website| Google Scholer |Pubmed |
6. Pourasghar F, Daemi A, Tabrizi JS, Ala A. Inter-rater reliability of triages performed by the electronic triage system. *Bulletin of Emergency & Trauma*. 2015;3(4):134–137. |Pubmed |Google Scholer|
7. Ebrahimi M, Heydari A, Mazlom R, Mirhaghi A. The reliability of the Emergency Severity Index: a meta-analysis. *Advanced Journal of Emergency Medicine*. 2019;3(1):e7.
8. |Google Scholer| Pubmed|
9. Christ M, Grossmann F, Winter D, Bingisser R, Platz E. Modern triage in the emergency department. *Deutsches Ärzteblatt International*. 2010;107(50):892–898 |DOI]|Google| Scholer| Pubmed| Website|
10. Göransson KE, Ehrenberg A, Marklund B, Ehnfors M. Accuracy and concordance of nurses in emergency department triage. *Scandinavian Journal of Caring Sciences*. 2005;19(4):432–438. |DOI | Website |Pubmed | Google Scholer|
11. Obermeyer Z, Emanuel EJ. Predicting the future — big data, machine learning, and clinical medicine. *New England Journal of Medicine*. 2016;375:1216–1219. |DOI| Website| Google Scholer| Pubmed|
12. World Health Organization. *Integrated Interagency Triage Tool (IITT): a tool for emergency departments in low-resource settings*. Geneva: WHO; 2019.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GRAPHICABSTRACT.png