

Canadian Triage and Acuity Scale: testing the mental health categories

Anne-Marie Brown¹
Diana E Clarke¹
Julia Spence²

¹College of Nursing, Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada; ²St Michael's Hospital, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Purpose: The study tested the inter-rater reliability and accuracy of triage nurses' assignment of urgency ratings for mental health patient scenarios based on the 2008 Canadian Triage and Acuity Scale (CTAS) guidelines, using a standardized triage tool. The influence of triage experience, educational preparation, and comfort level with mental health presentations on the accuracy of urgency ratings was also explored.

Methods: Study participants assigned urgency ratings to 20 mental health patient scenarios in randomized order using the CTAS. The scenarios were developed using actual triage notes and were reviewed by an expert panel of emergency and mental health clinicians for face and content validity.

Results: The overall Fleiss' kappa, the measure of inter-rater reliability for this sample of triage nurses (n=18), was 0.312, representing only fair albeit statistically significant ($P < 0.0001$) agreement. Kendall's coefficient of concordance for the sample was calculated to be 0.680 ($P < 0.0001$), which signifies moderate agreement. Although the sample reported high levels of education, comfort with mental health presentations, and experience, accuracy in urgency ratings measured by the percentage of correct responses ranged from 0.05% to 94% (mean: 54%). Greater accuracy in urgency ratings was recorded for triage nurses who used second-order modifiers and avoided the use of override.

Conclusion: Specific focus on the use of second-order modifiers in orientation and ongoing education of triage nurses may improve the reliability and validity of the CTAS when used to assign urgency ratings to mental health presentations.

Keywords: CTAS, mental health triage, inter-rater reliability

Introduction

Emergency department (ED) crowding has been identified as a significant patient safety issue as it has been associated with increased patient mortality, delayed resuscitation efforts, and increased length of stay in hospital.¹ Since ED triage processes influence the time to assessment, extent of patient workup, and length of stay, they play an important role in ED crowding and patient safety.¹ Despite growing numbers of mental health patients presenting to general hospital EDs for care,² triage nurses' accuracy for rating urgency levels of mental health patients is low.³ Moreover, studies testing Canadian Triage and Acuity Scale (CTAS) have included few mental health presentations in their designs.³⁻⁵

The CTAS is used by the majority of EDs in Canada.^{6,7} Agreement and accuracy study designs testing the CTAS have varied. While studies of the CTAS have found that inter-observer reliability is 'high' or 'significant',^{7,8} moderate or good levels of agreement are not uncommon.^{9,10}

Correspondence: Anne-Marie Brown
College of Nursing, Faculty of Health Sciences, University of Manitoba,
89 Curry Place, Winnipeg, MB R3T 2N2,
Canada
Tel +1 204 474 9080
Fax +1 204 474 7683
Email anne-marie.brown@umanitoba.ca

In 2004, CTAS was revised to include first- and second-order modifiers.⁶ These modifiers are objective observations that can alter the CTAS acuity level.⁶ As an example, first-order modifiers can include vital signs or pain scales, whereas second-order modifiers are more specific to the presenting complaint and help to determine the patient's risk.⁶ In 2008, significant revisions to the CTAS mental health category and the corresponding second-order modifiers were released.¹¹ This study contributes to the CTAS literature by testing these 2008 revisions. In particular, the study suggests that the consistent use of second-order modifiers may increase the accuracy of urgency ratings in mental health presentations.¹²

Purpose

The purpose of this study was to test the inter-rater reliability among triage nurses and their accuracy of assigning urgency ratings for standardized mental health patient scenarios, based on the 2008 CTAS guidelines.¹¹

Research questions

The following research questions were developed for this study:

1. What is the inter-rater/inter-observer reliability among triage nurses assigning levels of urgency to mental health patient scenarios, based on the 2008 CTAS guidelines, using a standardized triage tool (Emergency Department Information System [EDIS])?
2. How accurate are triage nurses in assigning levels of urgency to mental health patient scenarios, based on the 2008 CTAS guidelines, using a computerized tool (EDIS)?

Methodology

Following a review of previous work on inter-rater reliability involving triage nurses, a consistent approach to research design was not found.^{9,13} Although 'real-time' design was considered, the limits on time and variation among patient presentations across the participating sites were considered significant challenges. While paper-based scenarios have raised methodological concerns,¹⁴ they offer the ability to obtain responses from all participants based on the same information, thus reducing the variability of the patient's presentation that occurs in real-time design.^{13,15} Accordingly, paper-based scenarios were used as a 'suitable estimate' of inter-rater reliability with live triage cases.¹⁶

Sample

The sample of 18 triage nurses was drawn from three participating sites, including one community and two tertiary

acute care hospitals. Each participant completed a demographic questionnaire that asked about years of triage experience, specific education or training, and overall comfort level and confidence with triaging specific mental health presentations.

Scenario development

The patient scenarios used in this study were developed by referencing triage notes from mental health patients' presentations from a teaching hospital ED. The abstracted data for the patient scenarios included: a) Canadian Emergency Department Information System (CEDIS) category, b) CTAS chief complaint, c) age, d) mode of arrival, e) vital signs, and f) details of the triage note. The abstracted data also indicated whether or not the patient attended the department alone as well as the patient's involuntary status under a provincial 'Mental Health Act', prior history of mental health diagnoses, level of risk, and displays of agitation and/or aggressive behavior. The abstracted data were used to form a template for the mental health patient scenarios and served as the first stage of scenario development. The data's primary use was to determine the components of mental health triage assessments, for example, 'Mental Health Act' status.

A total of 38 scenarios were distributed to an expert panel comprised of emergency physicians and academic researchers with expertise on the topic who reviewed the scenarios for both face and content validity. The expert panel directed revisions and established the accurate triage ratings (that is, 'correct' urgency level). Following the process of expert review, 20 scenarios were chosen to represent the revised guidelines and a typical distribution of presentations. Included were: two CTAS 1 (immediate), five CTAS 2 (emergent), six CTAS 3 (urgent), five CTAS 4 (less urgent), and two CTAS 5 (not urgent) scenarios.

Methods

Ethical approval for this study was obtained from the University of Manitoba, Education/Nursing Research Ethics Board. All study participants utilized the EDIS, a computerized tool that facilitates application of CTAS guidelines during the triage process. Although the EDIS prompts the nurse through a series of decisions, the triage nurse maintains the ability to override the final CTAS generated score and choose a CTAS rating that best matches their clinical assessment. The CTAS guidelines include the use of CEDIS entrance complaints and second-order modifiers. Each participant rated the 20 scenarios in randomized order. As the participants rated each of the scenarios, they had the opportunity

to select a second-order modifier to assist in their decision making. In circumstances where the participant chose a second-order modifier, the modifier chosen was recorded in the EDIS and data sheets were printed for each scenario triaged. At the beginning of each session, participants were provided scripted information, including the assumption that the presenting patient was medically stable. Participants were asked to enter their triage assessments as true to a real-life scenario as possible.

Statistical analysis

The inter-rater reliability of the CTAS has been measured in previous studies using written patient scenarios and analyzed using an unweighted and weighted kappa statistic.^{7,8} Unlike previous studies,^{7,8} which compared raters between groups the current study tested the agreement within one group. This distinction necessitated the use of a statistical analysis method capable of measuring agreement among multiple raters.

The present study used Fleiss' kappa statistic and Kendall's coefficient of concordance. Fleiss' kappa is bounded between 1 and -1 where 0 represents no agreement and perfect agreement is considered to be at values over 0.80.¹⁸⁻²⁰ Kendall's coefficient calculation ranges between 0-1, where 0 represents no agreement and 1 represents complete agreement.²¹ Since Kendall's coefficient partially accounts for the degree of closeness of responses, overall agreement scores calculated using Kendall's statistic may be higher than those using the Fleiss calculation.

Results

The participants were experienced triage nurses; over 40% of the sample reported more than 10 years of triage experience, with high levels of education in emergency nursing (see Table 1) and confidence in mental health triage (see Table 2). In contrast to earlier findings, participants did not report lower rates of confidence with aggressive behaviors.^{22,23} No

relationship was found between level of confidence and years of experience.

Data analysis

The overall Fleiss' kappa, the measure of inter-rater reliability for this sample of triage nurses (n=18), was 0.312. Based on guidelines developed by Landis and Koch,²⁰ this kappa represents only fair agreement, but is nonetheless statistically significant ($P < 0.0001$). Additional analysis showed the agreement among triage nurses based on CTAS level. Moderate agreement was shown for CTAS level 1 (kappa = 0.459) and 4 (kappa = 0.500). The kappa statistic for CTAS level 2 was 0.107, CTAS level 3 was 0.218, and CTAS level 5 was 0.022. Kendall's coefficient for this sample was 0.680 ($P < 0.0001$), indicating moderate agreement.

In order to calculate the accuracy of the urgency ratings, a custom code in SAS was used to compute the 'Light' statistic²⁴ to determine whether the raters agree with correct responses significantly more than by chance alone. A calculated P -value < 0.001 demonstrated that triage nurses' agreement with the correct level was not purely random (Dufault B, University of Manitoba, personal communication, February, 2011). Accuracy has also been presented as a correct response by urgency level in earlier studies.¹⁷ A similar calculation was conducted in this study to determine which CTAS levels produced the highest reported accuracy (see Table 4).

Discussion

Since the overall agreement in this study was calculated statistically to be 'fair', significant variability in the urgency ratings existed. In this sample, the highest rate of agreement occurred at the highest (CTAS level 1) and lower ends of the scale (CTAS level 4) (see Table 3). Despite the significant variability in urgency ratings, the influence of the second-order modifiers was an important finding. Triage scales like CTAS are designed to support decision making, guiding the triage nurse to a correct decision.²⁵ The addition of second-order modifiers to the CEDIS entrance complaints in the mental health category are intended to further aid in the triage nurses' decision making. As an example, for the CEDIS entrance complaint of 'depression, suicidal, or deliberate self harm', the second-order modifiers are: 1) attempted suicide or clear suicide plan, 2) active suicidal intent, 3) uncertain flight or safety risk, 4) suicidal ideation, no plan, and 5) depressed, no suicidal ideation.¹¹ The first three second-order modifiers listed above would result in a CTAS score of 2, whereas the fourth modifier would result in a CTAS score of 3, and the fifth modifier would result in a CTAS score

Table 1 Specialized mental health training or experience

Educational courses completed	Number of triage nurses	Percentage of nurses who completed the specified educational course
None	2	11.1
CTAS training	1	5.6
Regional triage orientation	1	5.6
Advanced emergency course	1	5.6
Other educational course	1	5.6
More than one educational course	12	66.7
Total	18	100.0

Abbreviation: CTAS, Canadian Triage and Acuity Scale.

Table 2 Triage nurses' comfort level with mental health patient presentations

Presentation	Not at all comfortable	Mildly confident	Moderately confident	Very confident	Total
Triage of mental health patients	0	4	7	7	18
Psychotic symptoms	0	6	8	4	18
Manic symptoms	0	4	9	5	18
Anxiety	0	0	9	9	18
Depression	0	1	9	8	18
Suicidal ideation	0	1	10	7	18
Aggressive behaviors	0	4	9	5	18
Behavior/personality disorders	2	7	5	4	18

of 4.¹¹ The additional cues provided by these modifiers may help to increase the structure of the tasks at triage, thereby allowing for greater analysis by the nurse while making their decision. Arguably, not utilizing second-order modifiers places greater emphasis on intuition or gut feelings, which may impact triage accuracy. If more participants in this study had relied on the second-order modifiers, the overall rate of agreement and accuracy may have been higher.

Over-triage and under-triage are important considerations in triage literature. Over-triage occurs when a patient is seen faster than is required, while under-triaging results in a patient waiting longer than is considered appropriate.¹³ The influence of the nurse's experience level on over- or under-triaging has been reported with mixed results.^{13,26} Nurses in the present study were more likely to over-triage. The slight tendency toward over-triaging may have been influenced by the atypical conditions, including: more time to gather data, absence of department pressures, and the knowledge that they were participating in a study. Moreover, specific scenarios influenced the rate of over- and under-triaging in this study. In particular, a high percentage of participants over-triaged scenarios 16 and 20 (see Table 4). Furthermore, nearly 80% of the participants assigned CTAS level 3 to scenario 20, which may suggest the content of the scenario more accurately represented that level of urgency. Half of the participants assigned CTAS level 3 rather than the 'correct' score of 2

for both scenarios 13 and 18. This finding may suggest distinguishing between CTAS level 2 and 3 can pose challenges to nurses when triaging mental health presentations.

Based on this sample, the influence of education, confidence, and comfort level on the accuracy or inter-rater reliability of urgency ratings among triage nurses is not yet clear. Overall, the inter-rater reliability was fair to moderate despite high levels of reported confidence and educational preparation with triaging mental health presentations. The influence of second-order modifiers on the accuracy or urgency ratings may be significant. While it is possible that the scenarios in this study lacked the specific triage cues that triage nurses rely on when triaging mental health scenarios, participants who consistently used second-order modifiers had higher accuracy ratings. Nurses who assigned urgency ratings that matched the 'correct' response more than 60% of the time used second-order modifiers for the majority, and in some instances, all of the 20 scenarios (see Table 5). Nurses in this group also refrained from or avoided entirely the use of override to change the computer-generated CTAS score. In contrast, nurses that assigned the correct score less often (less than 40% of the time) were less likely to use second-order modifiers or avoided their use altogether. Some individuals in this second group of nurses also utilized override more often than triage nurses who had higher frequencies of 'correct' responses. This provides some support for the notion that the use of heuristics and intuitive decision making in the emergency setting²⁷ may introduce a cognitive bias that could compromise patient outcomes.²⁷

Table 3 Percentage correct by CTAS urgency level

CTAS urgency level	Number of scenarios	Number of correct responses	Number of responses	Percentage correct
1	2	27	36	75%
2	5	29	90	32.2%
3	6	71	108	65.7%
4	5	63	90	70%
5	2	4	36	11.1%

Note: The figures in bold represent the highest percentage correct by CTAS urgency level.

Abbreviation: CTAS, Canadian Triage and Acuity Scale.

Limitations

Several limitations of the present study exist. Participants used paper-based scenarios to assign their levels of urgency and conducted their ratings in an environment that was in stark contrast to their typical work settings. These paper-based scenarios were reviewed by an expert committee but were not piloted with a sample of ED triage nurses. As

Table 4 Accuracy, over-triage, and under-triage by one level

Scenario number	Correct answer	Percentage of correct responses (accuracy) n=18	Number of nurses who over-triaged by one level	Number of nurses who under-triaged by one level	Total n=18
1	3	72.2%	1	3	17
2	4	94%	1	0	18
3	2	22.2%	8	6	18
4	4	77.7%	4	0	18
5	2	22.2%	6	6	14
6	3	72.2%	3	2	18
7	4	94%	0	1	18
8	1	77.7%	0	4	18
9	2	38.8%	6	4	17
10	3	72.2%	3	1	17
11	3	61.1%	2	5	18
12	1	72.2%	0	2	15
13	2	44.4%	1	9	18
14	3	66.6%	0	6	18
15	4	77.7%	3	1	18
16	5	11.1%	11	0	13
17	5	11.1%	2	0	4
18	2	33.3%	0	9	15
19	3	50%	8	1	18
20	4	0.05%	14	0	15

an example, the fact that nearly 80% of the participants assigned a rating of 3 to scenario 20 suggests that if the scenario was piloted by practicing triage nurses, it may have undergone revision. While a pilot of the scenarios may have led to additional revisions, a pilot may have reduced the sample of participants available for the study. The sample

and research design were also limitations of the study and must be considered when reviewing these findings. The participants in this study identified themselves as confident and comfortable with mental health presentations. Future studies should consider mixing mental health scenarios in amongst medical and trauma presentations to reduce the focus on mental health, potentially reducing the effect of self-selection.

Table 5 Use of second-order modifiers

Rater	Frequency of use of second-order modifiers (N=20)	Use of override when rating scenarios (N=20)	Percentage of correct responses
1	20 (100%)	0	70% (n=14)
2	20 (100%)	0	55% (n=11)
3	1 (5%)	8 (40%)	40% (n=8)
4	16 (80%)	0	60% (n=12)
5	19 (95%)	2 (10%)	50% (n=10)
6	14 (70%)	0	55% (n=11)
7	20 (100%)	0	55% (n=11)
8	15 (75%)	2 (10%)	65% (n=13)
9	18 (90%)	0	60% (n=12)
10	0	0	35% (n=7)
11	19 (95%)	0	55% (n=11)
12	17 (85%)	0	65% (n=13)
13	20 (100%)	0	55% (n=11)
14	12 (60%)	3 (15%)	50% (n=10)
15	19 (95%)	7 (35%)	35% (n=7)
16	20 (100%)	0	65% (n=13)
17	20 (100%)	6 (30%)	55% (n=11)
18	20 (100%)	0	45% (n=9)

Conclusion

Overcrowding, rising patient acuity, and longer lengths of stay have increased the pressure on health care systems to devise and implement triage systems that are both fast and accurate.²⁸ Interestingly, although mental health patients are presenting to general hospital EDs for care in greater numbers, inter-rater reliability and the accuracy of triage ratings may be lower than for medical presentations.^{2,3,28} Although lack of confidence may influence ratings, the participants in this study were highly experienced and rated themselves as comfortable with mental health presentations. Despite this, inter-rater reliability was low, particularly in the mid ranges of the CTAS score (level 2 and 3) when second-order modifiers were not used as a triage aid. Specific focus on the use of second-order modifiers in the orientation and ongoing education of triage nurses may improve the reliability and validity of the CTAS when used to assign urgency ratings to mental health presentations.

Disclosure

The authors report no conflicts of interest in this work.

References

- O'Connor E, Gatien M, Weir C, Calder L. Evaluating the effect of emergency department crowding on triage destination. *Int J Emerg Med.* 2014;7:16.
- Broadbent M, Moxham L, Dwyer T. Issues associated with the triage of clients with a mental illness in Australian emergency departments. *Australas Emerg Nurs J.* 2010;13(4):117–123.
- Happell B, Summers M, Pinikahana J. The triage or psychiatric patients in the hospital emergency department: a comparison between emergency department nurses and psychiatric nurse consultants. *Accid Emerg Nurs.* 2002;10:65–71.
- Fernandes CM, McLeod S, Krause J, 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):227–232.
- Andersson AK, Omberg M, Svedlund M. Triage in the emergency department – a qualitative study of the factors which nurses consider when making decisions. *Nurs Crit Care.* 2006;11(3):136–145.
- Murray M, Bullard M, Grafstein I; CTAS National Working Group; CEDIS National Working Group. Revisions to the Canadian Emergency Triage Acuity Scale (CTAS) implementation guidelines. *CJEM.* 2004;6(6):421–427.
- Beveridge R, Ducharme J, Janes L, Beaulieu S, Walter S. Reliability of the Canadian emergency department triage and acuity scale: interrater agreement. *Ann Emerg Med.* 1999;34(2):155–159.
- Manos D, Petrie DA, Beveridge RC, Walter S, Ducharme J. Inter-observer agreement using the Canadian Emergency Department Triage and Acuity Scale. *CJEM.* 2002;4(1):16–22.
- Dong SL, Bullard MJ, Meurer DP, Blitz S, Holroyd BR, Rowe BH. The effect of training on nurse agreement using an electronic triage system. *CJEM.* 2007;9(4):260–266.
- Dong SL, Bullard MJ, Meurer DP, et al. Reliability of computerized emergency triage. *Acad Emerg Med.* 2006;13(3):269–275.
- Bullard MJ, Unger B, Spence J, Grafstein E; CTAS National Working Group. Revisions to the Canadian Emergency Department Triage and Acuity Scale (CTAS) adult guidelines. *CJEM.* 2008;10(2):136–151.
- Brown AM. Testing the Reliability and Accuracy of Urgency Ratings Determined by Triage Nurses for Mental Health Scenarios, Using the Canadian Triage and Acuity Scale [master's thesis]. Winnipeg: University of Manitoba; 2011. Available from: <http://hdl.handle.net/1993/4493>. Accessed September 19, 2014.
- Considine J, Ung L, Thomas S. Clinical decisions using the National Triage Scale: how important is postgraduate education? *Accid Emerg Nurs.* 2001;9(2):101–108.
- Dallaire C, Poitras J, Aubin K, Lavoie A, Moore L. Emergency department triage: do experienced nurses agree on triage scores? *J Emerg Med.* 2012;42(6):736–740.
- Dallaire C, Poitras J, Aubin K, Lavoie A, Moore L, Audet G. Interrater agreement of Canadian Triage Acuity Scale scores assigned by base hospital and emergency department nurses. *CJEM.* 2010;12(1):45–49.
- Worster A, Sardo A, Eva K, Fernandes CM, Upadhye S. Triage tool inter-rater reliability: a comparison of live versus paper case scenarios. *J Emerg Nurs.* 2007;33(4):319–323.
- Olofsson P, Gellerstedt M, Carlström ED. Manchester Triage in Sweden – interrater reliability and accuracy. *Int Emerg Nurs.* 2009; 17(3):143–148.
- Braga-Baiak A, Shah A, Pietrobon R, Braga L, Neto AC, Cook C. Intra- and inter-observer reliability of MRI examination of intervertebral disc abnormalities in patients with cervical myelopathy. *Eur J Radiol.* 2008;65(1):91–98.
- Fleck BW, Tandon A, Jones PA, Mulvihill AO, Minns RA. An interrater reliability study of new 'zonal' classification for reporting the location of retinal haemorrhages in childhood for clinical, legal and research purposes. *Br J Ophthalmol.* 2010;94(7):886–890.
- Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics.* 1977;33(2):363–374.
- Sigler EA, Tallent-Runnels MK. Examining the validity of scores from an instrument designed to measure metacognition of problem solving. *J Gen Psychol.* 2006;133(3):252–276.
- Farrell GA, Shafiei T, Salmon P. Facing up to 'challenging behaviour': a model for training in staff-client interaction. *J Adv Nurs.* 2010; 66(7):1644–1655.
- Clarke DE, Brown AM, Hughes L, Motluk L. Education to improve the triage of mental health patients in general hospital emergency departments. *Accid Emerg Nurs.* 2006;14(4):210–218.
- Light RJ. Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychol Bull.* 1971;76(5):365–377.
- Farrokhnia N, Castrén M, Ehrenberg A, 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.
- Chung JY. An exploration of accident and emergency nurse experiences of triage decision making in Hong Kong. *Accid Emerg Nurs.* 2005;13(4):206–213.
- Croskerry P. From mindless to mindful practice – cognitive bias and clinical decision making. *N Engl J Med.* 2013;368(26):2445–2448.
- Ruger JP, Lewis LM, Richter CJ. Identifying high-risk patients for triage and resource allocation in the ED. *Am J Emerg Med.* 2007; 25(7):794–798.

Open Access Emergency Medicine

Publish your work in this journal

Open Access Emergency Medicine is an international, peer-reviewed, open access journal publishing original research, reports, editorials, reviews and commentaries on all aspects of emergency medicine. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <http://www.dovepress.com/open-access-emergency-medicine-journal>

Dovepress

Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.