Inter-observer agreement using the Canadian Emergency Department Triage and Acuity Scale

Daria Manos, BA;* David A. Petrie, MD;*† Robert C. Beveridge, MD;*‡ Stephen Walter, PhD;§
James Ducharme, MD*‡

ABSTRACT

Objective: To determine the inter-observer agreement on triage assignment by first-time users with diverse training and background using the Canadian Emergency Department Triage and Acuity Scale (CTAS).

Methods: Twenty emergency care providers (5 physicians, 5 nurses, 5 Basic Life Support paramedics and 5 Advanced Life Support paramedics) at a large urban teaching hospital participated in the study. Observers used the 5-level CTAS to independently assign triage levels for 42 case scenarios abstracted from actual emergency department patient presentations. Case scenarios consisted of vital signs, mode of arrival, presenting complaint and verbatim triage nursing notes. Participants were not given any specific training on the scale, although a detailed one-page summary was included with each questionnaire. Kappa values with quadratic weights were used to measure agreement for the study group as a whole and for each profession.

Results: For the 41 case scenarios analyzed, the overall agreement was significant (quadratic-weighted $\kappa = 0.77$, 95% confidence interval, 0.76–0.78). For all observers, modal agreement within one triage level was 94.9%. Exact modal agreement was 63.4%. Agreement varied by triage level and was highest for Level I (most urgent). A reasonably high level of intra- and inter-professional agreement was also seen.

Conclusions: Despite minimal experience with the CTAS, inter-observer agreement among emergency care providers with different backgrounds was significant.

RÉSUMÉ

Objectif : Déterminer le niveau de concordance entre observateurs lors de l'assignation du triage par des personnes de formations et d'expériences diverses qui utilisaient l'Échelle de triage et de gravité pour les départements d'urgence pour la première fois.

Méthodes : Vingt dispensateurs de soins d'urgence (5 médecins, 5 infirmières, 5 techniciens médicaux d'urgence [Soins immédiats en réanimation] et 5 techniciens médicaux d'urgence [Soins avancés en réanimation]) oeuvrant dans un important hôpital universitaire en milieu urbain participèrent à l'étude. Les observateurs utilisèrent l'ÉTG à 5 niveaux pour assigner indépendamment un niveau de triage à 42 scénarios cliniques tirés de cas réels reçus à l'urgence. Les scénarios comprenaient les signes vitaux des patients, le mode de transport vers l'urgence, la raison de consultation et les notes verbatim de l'infirmière de triage. Les participants ne reçurent aucune formation spécifique quant à l'utilisation de l'échelle, bien qu'un résumé détaillé d'une page leur ait été

remis avec chaque questionnaire. Des coefficients kappa avec pondération quadratique furent utilisés pour évaluer la concordance au sein du groupe d'étude dans son ensemble et au sein de chaque profession.

Résultats : Pour les 41 scénarios cliniques analysés, la concordance globale était significative (κ avec pondération quadratique = 0,77, intervalle de confiance à 95 %, 0,76–0,78). Pour tous les observateurs, la concordance modale à l'intérieur d'un même niveau de triage était de 94 %. La concordance modale exacte était de 63,4 %. La concordance variait selon le niveau de triage et était la plus élevée pour le Niveau I (Réanimation). On a également constaté un niveau relativement élevé de concordance intra et inter professionnelle.

Conclusions : Malgré une expérience minimale avec l'ÉTG, la concordance entre observateurs parmi les dispensateurs de soins d'urgence de formations diverses était significative.

Key words: triage, paramedics, emergency medical services

Introduction

Although the goals of triage have changed somewhat since its wartime inception, this process remains important in the safe and efficient functioning of emergency departments (EDs). The main role of triage is still to assign priority to patients who need more urgent care and to predict the nature and scope of care likely to be required. In spite of the recognized importance and worldwide ubiquity of triage, EDs have had difficulty obtaining an ideal triage tool. Triage scales and algorithms have been criticized for being insufficiently reliable or valid and for being unable to predict patient acuity.[1,2] In order to be of practical value in busy EDs, triage assessment scales must be easy to understand and amenable to being used quickly.

Brillman and colleagues[3] found only moderate levels of nurse and physician agreement (κ = 0.45 and 0.21 respectively) in the initial categorization of patients according to a 4-level scale of urgency. Similar results for triage on presentation have been obtained for urgent vs. non-urgent assignment[4] and for a British 4-level system.[5] When retrospective assignment of triage level is included in the assessment of inter-observer agreement, results generally show even less agreement.[4–6]

These data are discouraging given the increasing pressure to expand the role of triage. It has been proposed that standardized triage assignments be used as an ED case-mix measurement and a workload and performance indicator.[7–10] Administrators argue that documentation of admission rates for each level of acuity would help determine whether true case-mix differences exist between EDs. The combination of triage level and disposition data may provide a "fingerprint" or standardized descriptor of the ED and the hospital and allow for national and international comparisons between similar facilities. In addition, triage is increasingly being used (and misused) as a means to refer patients out of the ED.[1,11–13]

It has also been proposed that paramedics apply triage in the out-of-hospital setting to improve understanding of the acuity of patients being transferred to receiving health care professionals and possibly to affect trip destination and "hospital bypass" decisions. For this to occur, it is imperative that paramedics be able to apply a triage scale with a reasonably high level of inter-observer agreement.

Before any consideration can be given to non-traditional uses of triage, we must first be confident that we have a triage tool that is easily understood by users and that has a high level of inter-observer agreement. In 1997, the Canadian Association of Emergency Physicians and the National Emergency Nurses' Affiliation jointly endorsed a new Canadian Triage and Acuity Scale (CTAS) and recommended its use in all Canadian EDs.[7] The implementation guidelines for CTAS were subsequently published as a supplement to *CJEM*, in October 1999.[14] The CTAS is modified from the Australasian 5-level triage scale.[9] The CTAS outlines 5 triage levels, each consisting of degree of illness and acuity, time objectives to medical assessment and intervention, and examples of usual clinical presentations and sentinel diagnoses. Level I is considered most urgent and Level V least urgent.

Although preliminary studies have shown that the Australasian 5-level triage scale has high inter-observer agreement rates,[8,15] only one published study has examined the Canadian version of this scale.[16] The authors of that study (including R.C.B.) used 10 ED case presentations (written summaries of actual cases) from each of the 5 acuity levels and found high levels of agreement between physicians and nurses: for all 20 participants, chance-corrected agreement was 0.80, based on quadratically weighted kappa. The raters were previously unfamiliar with the CTAS, so these results suggest that the scale is relatively easy to use.[16]

The objectives of the present study were (1) to determine whether high intra- and inter-professional agreement among nurses and physicians could be reproduced in an-

Table 1. Summary of the Canadian ED Triage and Acuity Scale (CTAS)

Triage Level	Acuity level	Time to physician	Usual presentation	Sentinel diagnoses
Level I	Resuscitation	Immediate	Code arrest Major shock Shock states Near-fatal asthma Severe respiratory distress Altered mental state (unconscious or delirious)	Traumatic shock Pneumothorax (traumatic or tension) Facial burns with airway compromise Severe burns > 30% body surface area Overdose with hypotension or unconsciousness AMI with complications (CHF or hypotension) Status asthmaticus Head injury (major or unconscious) Status epilepticus
Level II	Emergent	≤ 15 min	Head injury (risk features with or without altered mental state) Severe trauma Altered mental state (lethargic, drowsy, agitated) Signs of serious infection (purpuric rash, toxic) Allergic reaction (severe) Chemical exposure (eyes) Nontraumatic, visceral chest pain (with or without associated symptoms) Vomiting or diarrhea, suspicion of dehydration Overdose (but conscious) or drug withdrawal Abdominal pain (age > 50 yr) with visceral symptoms Sexual assault GI bleeding with abnormal vital signs CVA with major deficit Severe asthma (peak expiratory flow rate <40%) Moderate or severe dyspnea Acute vaginal bleeding (pain scale > 5 with or without abnormal vital signs) Neonate (age ≤ 7 days) Fever (age ≤ 3 mo), with rectal temp > 38.0°C Acute psychotic episode or extreme agitation Diabetic hypoglycemia or hyperglycemia Headache, with pain scale 8–10/10 Chemotherapy or immunocompromise Pain scale 8–10/10 (abdominal, costovertebral angle, back, eye)	Head injury Trauma involving multiple sites Multiple rib fractures Neck or spinal cord injury Anaphylaxis Alkaline or caustic ocular burns AMI, unstable angina or CHF Chest pain NOS Gastroesophageal reflux Unspecified drug or medicinal overdose Abdominal aortic aneurysm Appendicitis, cholecystitis GI bleeding with hypotension CVA Severe asthma or COPD Croup Spontaneous abortion Ectopic pregnancy or rupture Epiglottitis, meningitis, sepsis Acute psychotic episode, agitation or DTs Diabetic ketoacidosis Hypoglycemia, hyperglycemia or migraine Renal colic Keratitis
Level III	Urgent	≤ 30 min	Head injury: alert with vomiting Moderate trauma Abuse, neglect or assault Signs of infection Mild or moderate asthma (peak expiratory flow rate ≥ 40%) Mild or moderate dyspnea Cheat pain with no visceral symptoms (sharp or MSK, no previous heart disease) GI bleeding with normal vital signs Acute vaginal bleeding with normal vital signs Seizure (alert on arrival) Acute psychosis with or without suicidal ideation Pain scale 8–10/10 with minor injuries Pain scale 4–7/10 with headache, costovertebral angle or back pain Vomiting and diarrhea (age ≤ 2 yr) without dehydration Dialysis problems	Head injury Anterior shoulder dislocation Tibia or fibula fracture Bimalleolar or trimalleolar ankle fracture Pyelonephritis or sepsis Asthma without status or COPD Bronchiolitis or croup Pneumonia Unspecified chest pain NOS (MSK, GI, respiratory) Uncomplicated GI bleeding Spontaneous abortion Seizure Acute psychosis with or without suicidal ideation Low back pain, strain (disk) Migraine

other busy Canadian ED, and (2) to expand the analysis to investigate inter-observer agreement among and between paramedics, nurses and physicians. Although in-hospital paramedics are responsible for the triage of ED patients in some Canadian centres, their reliability in using CTAS has never been assessed.

Methods

Setting and subjects

This study was conducted in the summer of 1997 at the Queen Elizabeth II (QE2) Health Sciences Centre in Halifax, Nova Scotia, a large urban teaching hospital with approximately 70 000 adult ED visits per year. Five emergency physicians (CCFP-EM or FRCPC), 5 emergency nurses with more than 5 years of ED experience, 5 Basic Life Support (BLS) paramedics and 5 Advanced Life Support (ALS) paramedics were chosen from a convenience sample of ED employees. All BLS paramedics were employed performing triage at the hospital, and all ALS paramedics had at least 5 years of experience in the field. Informed consent was obtained from each participant.

The study was conducted before the CTAS was put into general use at the QE2 Health Sciences Centre, and all 20 participants were first-time users of the scale. They received no formal training on the scale, although a detailed 1-page summary was included with each questionnaire (see Table 1). Because there were no specific interventions and no patient identifiers and the participants were volunteers, the study was exempt from ethical review at our institution.

Triage cases

The 20 participants used the CTAS to independently rate the same 42 case scenarios abstracted from actual ED case presentations. These cases were randomly chosen from a database of scenarios used in a previously published inter-observer agreement study.[16] The patient scenarios included vital signs, mode of arrival to the ED, presenting complaint and verbatim triage nursing notes.

Inter-observer agreement

The kappa statistic is widely used to assess inter-observer agreement beyond that expected by chance.[17] In this study, inter-observer agreement was analyzed by means of the quadratically weighted kappa statistic rather than a simple (unweighted) kappa. The calculations for quadratically weighted kappa were based on the work of Fleiss,[18] as described in a previously published triage study.[16] Statistical analysis was performed using PC Agree, a program based on a paper by Holman.[19] This program calculates the qua-

Table 1. Summary of the Canadian ED Triage and Acuity Scale (CTAS) (cont'd)				
Triage level	Acuity level	Time to physician	Usual presentation	Sentinel diagnoses
Level IV	Less urgent	≤ 1 h	Head injury: alert with no vomiting Minor trauma Acute abdominal pain Vomiting and diarrhea (age > 2 yr) without dehydration Headache: not migraine, not sudden Earache Chest pain, minor trauma or MSK injury: no distress Suicidal ideation or depression Corneal foreign body Minor allergic reaction Chronic back pain URI symptoms Pain scale 4–7/10	Head injury: alert with no vomiting Colles' fracture Ankle sprain Appendicitis Cholecystitis URI Otitis media or otitis externa Chest pain NOS (MSK, GI, respiratory) Gastroesophageal reflux Suicidal ideation or depression Urticaria Corneal foreign body Low back pain or strain
Level V	Non-urgent	≤ 2 h	Minor trauma: not necessarily acute Sore throat without respiratory symptoms Diarrhea alone, without dehydration Vomiting alone, with normal mental status and no dehydration Menses Minor symptoms Chronic abdominal pain Psychiatric complaints Pain scale < 4/10	Low back pain or strain URI Gastroenteritis Vomiting Disorders of menstruation Dressing changes or cast changes Constipation Neurotic, personality and nonpsychotic mental disorders Unspecified superficial laceration(s)
AMI = acute myocardial infarction, CHF= congestive heart failure, GI = gastrointestinal, CVA = cerebrovascular accident, NOS = not otherwise specified, COPD = chronic obstructive pulmonary disease, DTs = delirium tremens, MSK = musculoskeletal, URI = upper respiratory infection.				

dratically weighted kappa statistic and its standard error, and also computes the probability of agreement between pairs of participants for each response category.

Results

All 20 participants returned the questionnaire. One physician rated only 41 of the 42 case scenarios because of an administrative oversight. Agreement probabilities and kappa statistics were therefore calculated for 41 scenarios, in order to use a balanced data set.

Overall, 63.4% of the assigned triage levels exactly matched the modal (most common) response, and 94.9% were within one level of this. The agreement varied slightly by triage level, with the highest agreement for the most urgent level. Quadratically weighted kappa statistics for agreement within each of the 4 professional groups were 0.80 for nurses (95% confidence interval [CI], 0.76–0.84), 0.82 for physicians (95% CI, 0.78–0.86), 0.76 for BLS paramedics (95% CI, 0.72–0.80) and 0.73 for ALS paramedics (95% CI, 0.68–0.77). The overall level of agreement among the 20 participants was 0.77 (95% CI, 0.76–0.78).

Table 2 shows the probability of 2 randomly chosen observers having exact or near agreement for a given case. Values marked with a dagger (†) correspond to exact agreement as to triage level, whereas those marked with a double dagger (‡) show near agreement (within one triage level). These data also show that levels of agreement were highest for triage Level I.

Discussion

Despite minimal experience with the CTAS, inter-observer agreement among emergency care providers with different

backgrounds was good. Agreement levels were also good among paramedics, who have not previously been studied.

The levels of agreement for the CTAS in our study are similar to those reported for the 5-level Australasian triage scale[15] and to the single published CTAS inter-observer study,[16] but higher than those reported for other triage systems.[3–6,20,21] Most studies either did not present kappa statistics[15] or yielded kappa values below 0.5, suggesting only moderate inter-observer agreement.[4,6,20,21] Although differences in the triage scales themselves, in terms of clarity, use of examples and level of detail, may contribute to the differences in agreement seen, study methodology factors, including statistical tools, study size and experimental design, can also influence results.

Measuring agreement

The higher level of agreement reported in this study is partly a function of the statistical tool used. Most of the previous studies used the simple (unweighted) kappa statistic, which is designed to measure agreement for nominal categories.[17] However, unweighted kappas acknowledge only exact agreement, giving no credit for near agreement. If the outcome has several ordinal categories, as do triage scales, the simple kappa fails to adequately represent all of the information contained in the data. In contrast, the "weights" incorporated in the quadratically weighted kappa statistic reflect the closeness or distance of disagreement. The quadratically weighted kappa therefore provides a better interpretation of the data, taking into account whether disagreements are "near misses" that deserve some credit (e.g., in adjacent triage categories) or marked disagreements that warrant little or no credit.[18]

The quadratically weighted kappa statistic is the most common statistical choice for this type of agreement analysis.[16,18] It is approximately equivalent to the intra-class correlation coefficient that would apply if the underlying construct for the data were a normally distributed quantity. However, no distributional assumption is required for this analysis.

A kappa value of 0 corresponds to agreement no more than that expected by chance, whereas a kappa of 1 shows perfect agreement. When Landis and Koch[17] originally described the use of simple (unweighted) kappa statistics for nominal categories, they arbitrarily interpreted values of 0 to 0.2 as slight agreement, 0.2 to 0.4 as fair, 0.4 to 0.6 as moderate, 0.6 to 0.8 as substantial and greater than 0.8 as near-perfect agreement, but we should use caution in applying the same descriptors to an analysis of quadratically weighted kappas.

Because major disagreement is not possible for a triage

Table 2. Distribution of triage level assignments (assuming 2 randomly selected observers)*

Level chosen by first observer	Level chosen by second observer; % probability				
	I	II	III	IV	V
I	78†	20‡	2	0	0
II	18‡	49†	22‡	8	3
III	2	26‡	37†	26‡	9
IV	0	9	24‡	41†	26‡
V	0	4	12	35‡	49†

*Data derived from all 20 subjects (physicians, nurses and paramedics).
†These 5 values represent the probabilities that a second randomly selected observer will agree exactly with the level assigned by the first observer.
‡These 8 values represent the probabilities that a second randomly selected observer will agree within one level of the level assigned by the first observer.

scale with 2 or 3 levels (since every miss would be a near miss), kappa values with quadratic weights cannot be used for these types of scales. While the different types of kappa values can be loosely compared, standard and weighted kappas should not be interpreted as identical statistics, because the incorporation of the weights changes the underlying distribution of the kappa estimate. Comparing the quadratically weighted kappa values of this study with the unweighted kappa values of other studies may, therefore, be misleading.

Although the kappa statistic provides an overall measure of the extent of agreement between health care workers (or observers), we can examine the patterns of agreement in more detail by examining the probability that 2 observers would select the same triage category for a given patient or the probability of disagreements of various kinds. Table 2 shows these probabilities, estimated from the observed data. In the first row of the table, we see that if one randomly chosen observer assigns a patient to triage Level I, there is a 78% chance that a second randomly chosen observer will agree. There is also a 20% probability that the second observer will assign the patient to Level II and a 2% chance of assignment to Level III. The sum of these probabilities is 100%.

Similarly, if the first observer assigns a patient to triage Level IV, the chance of agreement by a second observer is 41%. In this case, there are moderate probabilities of 24% and 26% that the second observer will disagree by one triage level, assigning the patient to Level III or Level V respectively. There is also a small chance (9%) of assigning the patient to Level II.

Patterns of agreement

In general terms, the chance of agreement is slightly better if the first observer classifies a patient at one extreme or the other of the triage scale, especially if Level I is chosen. In contrast, if the first observer assigns a patient to a central category, there is a somewhat lower chance of exact agreement by the second observer and moderate levels of assignment to adjacent levels on either side. This is a typical pattern for data of this kind, because when the first observer classifies a patient into an extreme category, disagreements by the second observer can occur in only one direction, toward the centre of the distribution. Also, there tends to be less clinical disagreement between observers in extreme cases, when the patients in question are critically ill or minimally ill. However, if the first observer picks a central category, then disagreement can occur in both directions, which increases the probability that such disagreement will occur.

Case scenarios vs. real patients

Although we felt that standardized, abstracted case scenarios were necessary to ensure that raters had access to identical information, not all studies in the triage literature have been conducted this way. Abstracted case scenarios are somewhat artificial in that they do not allow for the instinctive, qualitative nonverbal clues gained in the interview process. To try to minimize this phenomenon, we did not remove subjective information from the scenarios. For example, if the nurse who saw the actual patient said that the patient appeared "ill," this descriptor was included in the triage scenario.

Interestingly, one recent study found that visual triage assessment resulted in more pronounced inter-observer disagreement.[21] In addition, some triage studies have compared triage assessments at different times in the patient care continuum (for example, nurse and physician assessment on presentation vs. physician chart review after patient discharge).[4,6] For exampe, in the study by Gill and colleagues[4] retrospective agreement with the original triage decision was rated only as fair ($\kappa = 0.38$, 95% CI, 0.30–0.46). It is not surprising that lower levels of inter-observer agreement are found when observers have access to different amounts of information.

Compared with other triage systems, the CTAS gives a more detailed description of each triage level. For example, one triage scale[6] asks observers to rate patients as urgent or non-urgent on the basis of just 2 questions: "Do you feel there is a threat to the patient's life if he or she does not receive treatment within an hour?" and "Do you feel the patient needs care within a few hours to prevent the problem from becoming serious?" The study by Gill and colleagues[4] simply asked participants to rate cases as urgent or non-urgent. While these systems directly address the most important issue of triage, the questions are vague enough to leave room for significant disagreement. In comparison, CTAS users may assign triage levels on the basis of the patient's symptoms, and the scale then matches the symptoms to an urgency level. The 1-page summary (Table 1) of the CTAS also includes examples of typical patient presentations for each triage level. Other triage scales do not include this information and simply require that users predict urgency.

Limitations

Although this study included more participants than many similar studies,[3,4,6,21] the relatively small number of observers and case scenarios may be seen by some as a limitation. However, the small standard errors reported for the quadratically weighted kappas (see Results section) sug-

gest that the study provided precise estimates for inter-observer agreement.

Despite the limitations of this study and the difficulty of comparing its results with those of studies of other triage tools, this work contributes to the broadening appraisal of triage systems used in EDs. The consistency with which the CTAS was understood suggests that it is a reliable tool. In addition, because paramedics had rates of agreement that were similar to those of nurses and doctors, this study adds evidence-based justification for the role of specifically trained paramedics in hospital triage. In fact, the CTAS may be a promising tool in the development of a multi-option emergency medical services system.[22,23] Only a triage scale backed by extensive evidence, including appropriate application by paramedics, could guide decisions about safe bypass to tertiary care centres or less urgent care areas such as walk-in clinics. This should be an area of future research.

Finally, with evidence of the reliability of CTAS growing, issues of validity (for example, correlation with hospital admission, morbidity and mortality rates or comparison with retrospective patient acuity assessment) should be the next area of critical attention. Knowledge of the ability of this triage tool to accurately predict patient need and acuity is essential if its use is to be recommended.

Conclusions

Despite minimal experience with the CTAS, inter-observer agreement among emergency care providers with different backgrounds was high. This reliability study is an important step in an evidenced-based approach to triage systems. Further studies assessing validity and acceptability are warranted.

Competing interests: None declared.

References

1. Harden RD. Critical appraisal of papers describing triage systems. Acad Emerg Med 1999;6:1166-71.

2. Williams RM. Triage and emergency department services. Ann Emerg Med 1996;27:506-8.

3. Brillman JC, Doezema D, Tandberg D, Sklar DP, Davis KD, Simms S. Triage: limitations in predicting need for emergent care and hospital admission. Ann Emerg Med 1996;27:493-500.

4. Gill JM, Reese CL, Diamond JJ. Disagreement among health care professionals about the urgent care needs of emergency department patients. Ann Emerg Med 1996;28:474-8.

5. Read S, George S, Westlake L, Williams B, Glasgow J, Potter T. Piloting an evaluation of triage. Int J Nurs Stud 1992;29(3):275-88.

6. George S, Read S, Westlake L, Fraser-Moodie A, Pritty P, Williams B. Differences in priorities assigned to patients by triage nurses and by consultant physicians in accident and emergency departments. J Epidemiol Commun Health 1993;47:312-5.

7. Beveridge R. CAEP issues. The Canadian Triage and Acuity Scale: a new and critical element in health care reform. Canadian Association of Emergency Physicians. J Emerg Med 1998;16(3):507-11.

8. Cleary MI, Ashby RH, Jelinek GA, Lagaida R. The future of casemix in emergency medicine and ambulatory care. Med J Aust 1994;161(Suppl):S30-3.

9. Australasian College for Emergency Medicine. The Australasian Triage Scale [policy document]. Emerg Med 1994;6:145-6. Also available: www.acem.org.au/open/documents/triage.htm (accessed 2001 Nov 27).

10. Beveridge R. The president's letter. CAEP Communiqué 1995;Apr:1.

11. Derlet RW, Nishio DA. Refusing care to patients who present to an emergency department. Ann Emerg Med 1990;19:262-7.

12. Derlet RW, Kinser D, Ray L, Hamilton B, McKenzie J. Prospective identification and triage of nonemergency patients out of an emergency department: a 5-year study. Ann Emerg Med 1995;25:215-23.

13. Lowe PA, Bindman AB, Ulrich SK, Norman G, Scaletta TA, Keane D, el al. Failure to validate a predictive model for refusal of care to emergency department patients. Acad Emerg Med 1994;1:213-7.

14. Beveridge R, Clarke B, Janes L, Savage N, Thompson J, Dodd G, et al. Canadian Emergency Department Triage and Acuity Scale implementation guidelines. CJEM 1999;1(3 Suppl).

15. Jelinek G, Little M. Interrater reliability of the National Triage Scale. Emerg Med 1996;8:226-30.

16. 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;34:155-9.

17. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

18. Fleiss J. Statistical methods for rates and proportions. 2nd ed. New York (NY): Wiley; 1981.

19. Holman CD. Analysis of interobserver variation on a programmable calculator. Am J Epidemiol 1984;120:154-60.

20. Butler WR. ED patient classification matrix: development and testing of one tool. J Emerg Nursing 1986;12:279-85.

21. Brillman JC, Doezema D, Tandberg D, Sklar DP, Skipper BJ. Does a physician visual assessment change triage? Am J Emerg Med 1997;15:29-33.

22. Neely KW, Drake ME, Moorhead JC, Schmidt TA, Skeen DT, Wilson EA. Multiple options and unique pathways: A new direction for EMS? Ann Emerg Med 1997;30:797-9.

23. Neely KW. Ambulance transports: What are the alternatives? Acad Emerg Med 1997;4:1137-41.

Correspondence to: Dr. David A. Petrie, Department of Emergency Medicine, Queen Elizabeth II Health Sciences Centre, 1796 Summer St., Halifax NS B3H 3A7; fax 902 494-1625; dapetrie@is.dal.ca