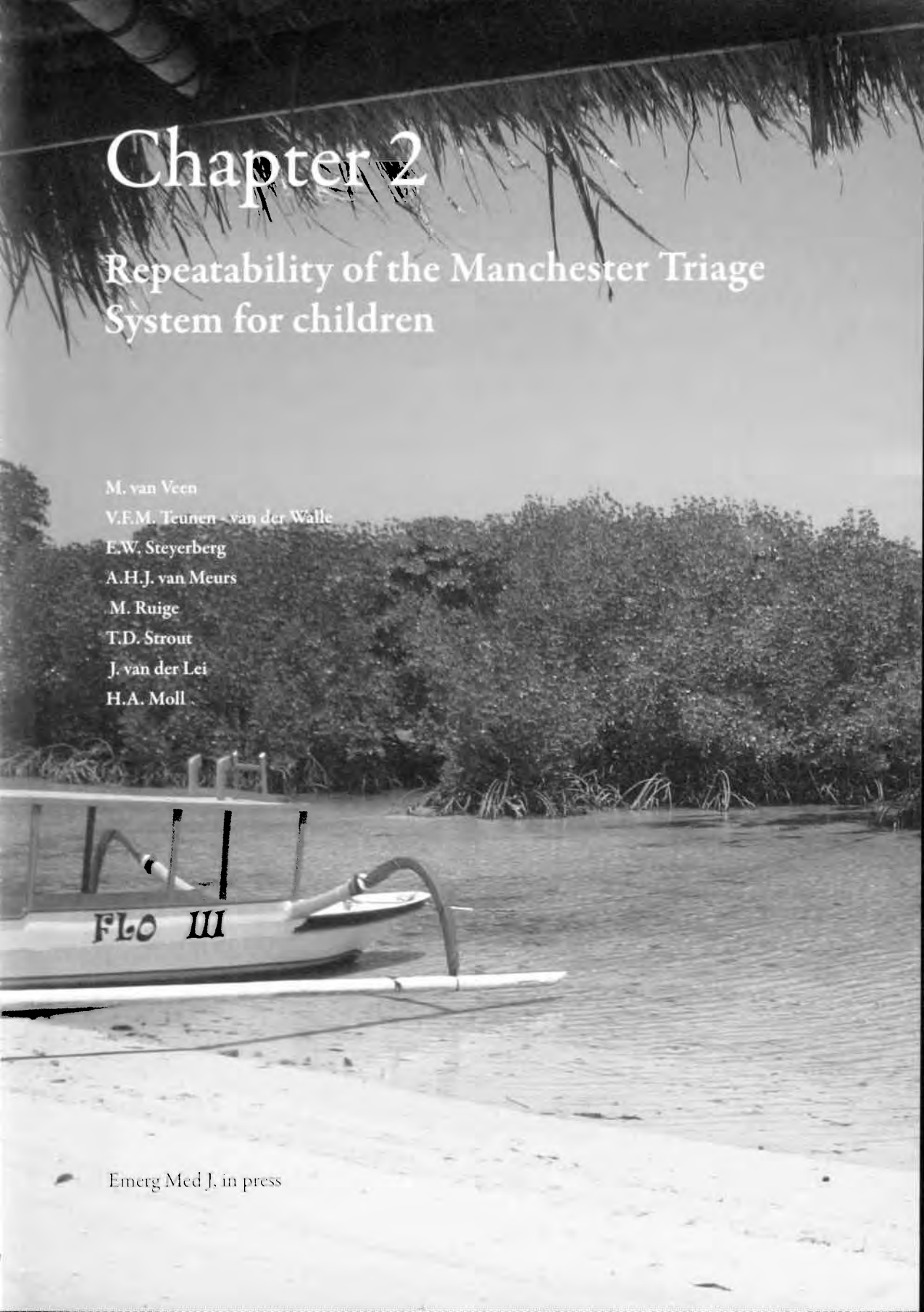
Chapter 2

Repeatability of the Manchester Triage System for children

M. van Veen
V.F.M. Teunen - van der Walle
E.W. Steyerberg
A.H.J. van Meurs
M. Ruige
T.D. Strout
J. van der Lei
H.A. Moll



Emerg Med J. in press

ABSTRACT

Objective We aimed to assess the repeatability of the Manchester Triage System (MTS) in children.

Methods All emergency department (ED) nurses (n=43) from a general teaching hospital and a university children's hospital in the Netherlands triaged 20 written case scenarios using the Manchester Triage system. Secondly, at two EDs real-life simultaneous triage of patients (<16 years) was performed by ED nurses and two research nurses. The written case scenarios and the patients included in the real-life simultaneous triage study were representative of children attending the ED, in age, problem and urgency level. We assessed inter-rater agreement using quadratic weighted kappa values.

Results The weighted kappa between the nurses, triaging the case scenarios was 0.83 (95% C.I.:0.74–0.91). In total, 88% (N=198) of the eligible ED patients were triaged simultaneously, with a weighted kappa of 0.65 (95% C.I.: 0.56–0.72).

Conclusions The MTS showed good to very good repeatability in paediatric emergency care.

INTRODUCTION

As triage aims to see patients first who benefit most from immediate care, it is essential that triage is both objective and reproducible. Different triage systems are extensively used in emergency departments across the world. The Manchester Triage System (MTS) was described and published in 1997 and is nowadays adopted around the world.[1,2] Little research on repeatability and validity of triage systems in paediatric emergency care, has been conducted to date.[3-9] As triage systems are widely used and it is not yet clear if one system is preferred over the others, research on their repeatability and validity is important and must be performed.

The MTS was developed by expert opinion.[1] The Dutch Institute of Healthcare recommended using the MTS in the Netherlands.[10] It consists of 52 flowcharts all representing a presenting problem, of which 49 are suitable for children. Following flowchart selection, general (life threat, haemorrhage, pain, conscious level, temperature and acuteness) and specific discriminators are considered. For example, a patient with an affirmative response to the discriminator "Increased work of breathing?" is triaged into urgency level two. Patients are allocated into one of five urgency levels. The MTS prescribes maximum waiting time for each urgency category (0, 10, 60, 120 and 240 minutes).

In adults, the MTS was shown to be sensitive for those with chest pain (sensitivity 87%, 95% CI: 78–92 and specificity 72%, 95% CI: 61–82 to identify high risk cardiac chest pain)[11] and for those with a critical illness.[12] The Manchester pain scale, a part of the MTS, showed a strong concurrent validity when compared to the Oucher pain scale.[13] The inter-rater agreement of the MTS in all ages, demonstrated a quadratic weighted kappa of 0.62 (95% CI 0.60 to 0.65) when studied using written case scenarios.[14] In a large prospective observational study the MTS demonstrated moderate validity when used in paediatric emergency care. It errs on the safe side, with much more over-triage than under-triage compared with an independent reference standard for urgency.[8,9] The inter-rater agreement of the MTS for children in particular has not yet been evaluated.

The aim of this study was to evaluate repeatability of the MTS in paediatric emergency care, using both written case scenarios and simultaneous triages by ED nurses.

METHODS

Study Design

To study repeatability we performed two studies on inter-rater agreement. First, 20 written case scenarios were triaged by 43 ED nurses, from two different hospitals, using the MTS. (Part 1) Second, 198 patients presenting to the two study EDs were each triaged simultaneously using the MTS, by one out of 25 ED nurses and one out of two research nurses. (Part 2) Table 1 reviews our study design. The requirement for informed consent was waived by the institutional review board.

Table 1 | Study design

Part	Patients / scenarios	Nurses	Setting	Outcome
1	20 written case scenarios	43 nurses	ED general teaching hospital* ED university hospital**	Repeatability
2	198 real life simultaneous triage assessments	First triage: triage nurse† Second triage: research nurse†	ED general teaching hospital* ED university hospital**	Repeatability

* Erasmus University Medical Center- Sophia Children's hospital, Rotterdam, The Netherlands
** Haga Hospital- Juliana Children's hospital, The Hague, The Netherlands
† During the selected shifts, one out of 25 nurses performed triage and one out of two research nurses performed the second triage assessment. ED = Emergency Department.

Patients

The ED of the Erasmus University Medical Center-Sophia Children's Hospital, Rotterdam is a paediatric-specific ED and is visited by nearly 9,000 patients per year. The MTS was implemented in 2005. The ED of the Haga Hospital-Juliana Children's Hospital, The Hague is a general paediatric-adult ED in a large teaching hospital with approximately 30,000 patients visits yearly, including 15,000 paediatric visits. For this site, the MTS was implemented in 2003. Participating ED nurses were experienced in both paediatric nursing and ED nursing, with a median of 10 years of ED nursing experience (IQR:7–14 years) and a minimum of two years. Both studies were performed between November 2006 and February 2007.

Manchester Triage System

Children under 16 years of age visiting the ED were triaged using a computerised version of the MTS. Registered nurses selected an MTS flowchart that suits the problem the patient

presents with. Selection of the appropriate discriminator leads to allocation of an urgency level. The chosen flowchart and discriminator were documented by the software application during triage. We used the official, translated version of the MTS advocated by the Dutch Association of ED Nurses.[1,15] Triage difficulties identified by the nurse participants could be reported and were discussed at ED meetings.

Part 1: Written Case Scenarios

Twenty written case scenarios were obtained and translated from Baumann et al.[3] Case scenarios are based on children presenting to the emergency department. Age, gender and presenting symptoms of the case scenarios were comparable to the total population presenting at the two EDs (table 2).

Table 2 | Patient characteristics of the total population presenting to the emergency departments in 2006 and the patients selected for the real life simultaneous triage (Part 2) and the written case scenarios (Part 1)

Variable	Total population[9] n=13,554	Real life simultaneous triage (Part 2) n=198	Written case scenarios (Part 1) n=20
ED			
General hospital	6,923 (51)*	139 (70)	N.A.
University hospital	6,631 (49)**	59 (30)	
Age†	3.4 (1.2-8.0)	2.5 (0.8-6.1)	6.0 (1.3, 7.5)
Sex, male %	7,813 (58)	104 (52)	12 (86)†
MTS urgency level			
Immediate	205 (1.5)	0	2 (10)
Very urgent	2,872 (21)	58 (29)	9 (45)
Urgent	4,462 (33)	58 (29)	2 (10)
Standard	5,895 (43)	81 (41)	5 (25)
Non urgent	120 (1)	1 (1)	2 (10)
Patient problems			
Trauma	3,591 (26)	49 (25)	6 (30)
Fever of unknown origin	1,306 (10)	35 (18)	3 (15)
Gastro-intestinal	2,166 (16)	22 (11)	2 (10)
Respiratory tract	2,356 (17)	35 (18)	3 (10)
Other	4,135 (30)	57 (29)	6 (20)

Numbers represent median with interquartile range or N (%); * Inclusion period: 7 months, ** Inclusion period: 13 months, † Sex is unknown in four cases.

The high urgency patients were overrepresented; the cases contained more boys and were somewhat older.

44 nurses received a written description of the cases and triaged the cases using the digital MTS application. Each case provided the patient's age, gender, problem of encounter and a short description of the history and vital signs (table 3).

Table 3 | Example written case scenario (English translation)

'An 8-year-old female presents to triage with her mom. The child has a sore throat, vomiting, and a fever all day. Mom states her child has been having difficulty swallowing all day. The child is making grunting noises and her skin is warm and flushed. T 38.7 °C, HR 122/min, Resp Rate 22/min, BP 110/53, SpO2 99% on room air.

Part 2: Real-time Simultaneous Triage

Patients attending the ED were triaged by one of 25 ED nurses. One of the two research nurses was present during the triage assessment, but did not interfere. After the assessment, both nurses triaged the patient. Patients were included during 12 work shifts ranging in duration from seven to ten hours. The research nurses selected the shift on basis of their own availability and were not aware of the working schedule of the triage nurses. They triaged all consecutive patients presenting at the ED.

Data on patient characteristics were gathered prospectively by the ED nurse in the triage application.

Primary Data Analysis

The characteristics of included patients were compared to characteristics of the total group of patients presenting at the same two ED's during respectively 7 and 13 months in 2006/2007[9] (table 3). The agreement between the nurses in MTS urgency level, flowchart and discriminator was determined for all twenty cases. First, we considered the urgency, flowchart or discriminator with the highest percentage agreement between nurses per case and secondly, we calculated the median and interquartile range of the percentage agreement of all cases. We determined the quadratic weighted kappa (K_w) by calculating the intraclass correlation coefficient (ICC) for agreement in urgency level. The ICC is equivalent to the quadratic weighted kappa.[16] The quadratic weighted kappa uses increasing weights for more severe disagreement.[17] We used the two way mixed model, type consistency function to calculate the ICC, for two as well as for multiple raters. (SPSS 14.0.1, Chicago, IL) The

simple kappa was calculated for agreement in the chosen MTS flowchart and discriminator using Stata v 8.2 (College Station, TX).
Kappa values can be interpreted as poor if K< 0.20, fair if > 0.21 < K and < 0.40, moderate if >0.41 < K and < 0.60, good if > 0.61 < K and <0.80 and very good if K>0.80.[17]

RESULTS

Part 1: Inter-rater Agreement: Written Case scenarios

All ED nurses (N=44) working at the two EDs, triaged each scenario. The results from one nurse were excluded due to a procedural error of the computer application. As a result, data from 43 nurses were included, 24 from the university hospital and 19 from the general hospital. The median agreement in urgency level was 81% (IQR: 60%, 90%) with a K_w of 0.83 (95% C.I.:0.74–0.91). For traumatic cases the K_w was 0.91, 95% C.I.:0.80–0.98 was and for non-traumatic cases 0.77, 95% C.I.:0.63–0.90

Part 2: Real-time simultaneous triage

During six shifts in December 2006 and six shifts in February 2007 (between 10 am and 6 pm or between 1 pm and 11 pm), 198 patients were triaged simultaneously (88% of eligible patients). 139 were included at the general hospital and 59 at the university hospital. No patients refused to participate. One research nurse was available per shift, and consequently some patients were missed because they entered the ED at the same time as other patients. The characteristics of the selected patients were comparable to the characteristics of the total ED population, except that the selected patient cohort contained slightly more patients with fever without a focus then the general patient population (table 3).
The agreement in MTS urgency level between the triage nurse and the research nurse was 66% with a K_w of 0.65 (95% CI: 0.56–0.72). In most cases of disagreement in urgency level, the disagreement was one level (28%, N=56) (table 4).

Table 4 | Real life simultaneous triage (Part 2): agreement in MTS urgency level between the triage nurse and the research nurse.

		Research nurse					Total
		Emergent	Very urgent	Urgent	Standard	Non urgent	
Triage nurse	Emergent	0	0	0	0	0	0
	Very urgent	1	48	7	2	0	58
	Urgent	0	12	28	18	0	58
	Standard	1	8	16	55	1	81
	Non urgent	0	0	0	1	0	1
Total		2	68	51	76	1	198

The used MTS flowchart and discriminator to triage patients and decide on urgency were available in 190 versus 181 patients, respectively.

The agreement in MTS flowchart and discriminator was 64% and 28% respectively, with simple kappa scores of 0.60 (95% C.I.:0.55–0.64) and 0.26 (95% C.I.:0.23–0.29). Pain score was documented at triage in 60% of the cases, with nurses agreeing on pain score in 24% of cases, (K_w=0.44,95% C.I.:0.28–0.58). Disagreement in urgency level between triage nurse and research nurse was strongly related to disagreement in discriminator and not related to disagreement in flowchart. (agreement/disagreement in urgency versus agreement/disagreement in discriminator OR=(52/3)/(69/57),=14, 95% C.I.4.2–48, agreement/disagreement in urgency versus agreement/disagreement in flowchart, OR= (83/39)/(43/25)=1.2, 95% C.I.:0.7–2.3).

Nurses agreed in 66% in both patients with a traumatic problem (N=56, K_w 0.45, 95% C.I.:0.22–0.64) and patients with a non-traumatic presenting problem (N=142, K_w=0.60,95% C.I.:0.48–0.69). Disagreement did not depend on the patient's age. (Median age 2.47 and 2.69 years, Mann Whitney test, p=0.55)

DISCUSSION

This study showed adequate repeatability of the MTS when applied to paediatric emergency care. The MTS demonstrated good to very good inter-rater agreement when studied using written case scenarios and real-time simultaneous triage.

Compared to the inter-rater agreement of other triage systems studied in children using written case scenarios, the inter-rater agreement found for the MTS in our study is high (table 5).

Table 5 | Inter-rater agreement of triage systems for children.

Triage system	Method	N	Measure	Value (95% C.I.)
ESI version 3[3]	Written case scenarios	20 scenarios	Quadratic weighted kappa	0.84–1.00
ESI version 3[3]	Simultaneous triage	272 patients	Quadratic weighted kappa	0.59 (0.55–0.63)
3-level triage system[18]	Written case scenarios	12 scenarios	Kappa	0.29
Paediatric CTAS[4]	Written case scenarios	55 scenarios	Weighted kappa*	0.51 (0.50–0.52)
Paediatric CTAS[7]	Simultaneous triage	499 patients	Linear weighted kappa	0.55 (0.48–0.61)
			Quadratic weighted kappa	0.61 (0.42–0.80)
4-level triage scale[5]	Written case scenarios	55 scenarios	Weighted kappa*	0.45 (0.45–0.46)
Soterion Rapid Triage System[19]	Simultaneous triage	117 patients	Weighted kappa*	0.90 (0.83–0.96)

* Unknown whether linear or quadratic weighted kappa was used

The agreement found at simultaneous triage of the MTS is somewhat higher compared to simultaneous triage using the ESI[3] and the paedCTAS, in children[7] and lower compared to the Soterion Rapid Triage System.[19] However, the studies on the ESI and paedCTAS studies performed the triage assessment twice, which may explain a lower agreement.

In adults, the inter-rater agreement (weighted kappa) of 5-level triage systems studied by simultaneous triage ranged from 0.66–0.87.[19-22] Two studies used written case scenarios and demonstrated a weighted kappa of 0.80 and 0.71.[22,23]

In several studies weighted kappa values were calculated to determine inter-rater agreement. However, from these papers it is often not clear if linear or quadratic weighted kappa values were calculated (table 5). A quadratic weighted kappa gives a somewhat higher weight if raters disagree with only one level compared to the lineair weighted kappa. In our study we determined quadratic weighted kappa values.[16]

We argue that the inter-rater agreement of triage systems depends roughly on three criteria. First, nurses must be experienced with the signs and symptoms of patients presenting at the ED. Second, the nurses must be well trained in the particular triage system in order to use the correct definitions belonging to the discriminators. The nurses working in the studied hospitals all met these criteria.

Third, the triage system must be unambiguous and should contain discriminators numerous enough to match the diversity of patients visiting the ED. For example, one written case scenario had a very low agreement in urgency level, since nurses had chosen fourteen different discriminators to triage the case. It presented a 2-month-old boy with a short period of apnoea. The fact that this presentation (incident or Apparent Life-Threatening Event, ALTE) is not exactly covered in the MTS, probably explains the low agreement for this case. Agreement could potentially be improved with ongoing training for ED nurses. After finishing the study, investigators discussed the disconcordant cases with the ED nurses in order to improve the triage process.

Our results showed that the agreement on flowchart level (representing the patient's presenting problem) is moderate (K 0.60, 95% C.I.:0.55–0.64) and on discriminator level fair. (K 0.26, 95% C.I.:0.23–0.29) The low agreement at discriminator level did not result in a low agreement in urgency level. Since more MTS discriminators can lead to one urgency level, the low agreement in MTS discriminator has little influence on the urgency level. This provides evidence supporting a high internal consistency for the MTS.

In the case scenarios study (part 1), nurses performed a somewhat higher agreement in traumatic cases compared to non-traumatic cases. However, 0.91 and 0.77 represent a very good and good inter-rater agreement. The difference between good and very good agreement is not considered as clinical important.

To appreciate the results, some limitations should be considered.

The set of written case scenarios was obtained from another study group so we had no influence on the selection of cases. That's why selection bias does not seem likely. To check if the cases were representative of our population we compared patient characteristics of the cases with our population (table 2). The cases were comparable with our population.

The triage of written case scenarios is not an exact substitute for evaluation of the actual triage process. The nuance of the nurse's interpretation of each patient's signs and symptoms is an important part of the triage process, and this essence is not captured using the written case scenarios method. We attempted to address this shortcoming of the paper scenarios

with the addition of real-time simultaneous triages. This method still demonstrated a good inter-rater agreement.

The written case scenario method is often used to assess the inter-rater agreement of triage systems. A recent study showed moderated to high agreement between simultaneous triage and paper case scenarios.[24]

During the real-time simultaneous triage (part 2) we did not perform the triage assessment twice. The research nurse was present during the assessment of the triage nurse. Subsequently, both nurses triaged the patient blinded using the MTS in a separate room and did not discuss the patients' signs and symptoms with each other.

A double independent triage assessment might better evaluate the actual triage process. However, such a method was not possible because of the possible impact on patient management and waiting times. Using a double assessment method, the nurse's translation from the patient's signs and symptoms to a triage decision, would be incorporated.

CONCLUSION

The MTS has a good to very good inter-rater agreement when applied to paediatric emergency patients. Good repeatability is an essential requirement for valid triage.

Acknowledgements

We thank the ED nurses of the Erasmus MC- Sophia children's hospital and the Haga Hospital-Juliana Children's hospital for their cooperation in this study, Marcel de Wilde, BSc, Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands for technical support, and M.J.C. Eijkemans, PhD, Department of Public Health, Erasmus MC, University Medical Center, Rotterdam, The Netherlands for statistical support.

REFERENCES

1. Mackway-Jones K. Emergency Triage, Manchester Triage Group: London: BMJ Publishing Group 1997.
2. Mackway-Jones K, Marsden J, Windle J. Emergency Triage, Manchester Triage Group. Second edition ed: Oxford: Blackwell Publishing Ltd 2006.
3. Baumann MR, Strout TD. Evaluation of the Emergency Severity Index (version 3) triage algorithm in pediatric patients. Acad Emerg Med. 2005 Mar;12(3):219-24.
4. Bergeron S, Gouin S, Bailey B, Amre DK, Patel H. Agreement among pediatric health care professionals with the pediatric Canadian triage and acuity scale guidelines. Pediatr Emerg Care. 2004 Aug;20(8):514-8.
5. Bergeron S, Gouin S, Bailey B, Patel H. Comparison of triage assessments among pediatric registered nurses and pediatric emergency physicians. Acad Emerg Med. 2002 Dec;9(12):1397-401.
6. Gouin S, Gravel J, Amre DK, Bergeron S. Evaluation of the Paediatric Canadian Triage and Acuity Scale in a pediatric ED. Am J Emerg Med. 2005 May;23(3):243-7.
7. Gravel J, Gouin S, Manzano S, Arsenault M, Amre D. Interrater Agreement between Nurses for the Pediatric Canadian Triage and Acuity Scale in a Tertiary Care Center. Acad Emerg Med. 2008 Oct 17.
8. Roukema J, Steyerberg EW, van Meurs A, Ruige M, van der Lei J, Moll HA. Validity of the Manchester Triage System in paediatric emergency care. Emerg Med J. 2006 Dec;23(12):906-10.
9. van Veen M, Steyerberg EW, Ruige M, van Meurs AH, Roukema J, van der Lei J, et al. Manchester triage system in paediatric emergency care: prospective observational study. BMJ. 2008;337:a1501.
10. Richtlijn Triage op de spoedeisende hulp: http://www.cbo.nl/Downloads/406/rl_triage_08.pdf
11. Speake D, Teece S, Mackway-Jones K. Detecting high-risk patients with chest pain. Emerg Nurse. 2003 Sep;11(5):19-21.
12. Cooke MW, Jinks S. Does the Manchester triage system detect the critically ill? J Accid Emerg Med. 1999 May;16(3):179-81.
13. Lyon F, Boyd R, Mackway-Jones K. The convergent validity of the Manchester Pain Scale. Emerg Nurse. 2005 Apr;13(1):34-8.
14. van der Wulp I, van Baar ME, Schrijvers AJ. Reliability and validity of the Manchester Triage System in a general emergency department patient population in the Netherlands: results of a simulation study. Emerg Med J. 2008 Jul;25(7):431-4.
15. Mackway-Jones K. Triage voor de spoedeisende hulp, Manchester Triage Group: Maarsen: Elsevier gezondheidszorg 2002.
16. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coeficient as measures of reliability. Educational and Psychological Measurement. 1973;33:613-9.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar;33(1):159-74.
18. Maldonado T, Avner JR. Triage of the Pediatric Patient in the Emergency Department: Are We All in Agreement? Pediatrics. 2004 August 1, 2004;114(2):356-60.
19. Maningas PA, Hime DA, Parker DE. The use of the Soterion Rapid Triage System in children presenting to the Emergency Department. J Emerg Med. 2006 Nov;31(4):353-9.

20. Dong SL, Bullard MJ, Meurer DP, Blitz S, Ohinmaa A, Holroyd BR, et al. Reliability of computerized emergency triage. Acad Emerg Med. 2006 Mar;13(3):269-75.

21. Grafstein E, Innes G, Westman J, Christenson J, Thorne A. Inter-rater reliability of a computerized presenting-complaint-linked triage system in an urban emergency department. Cjem. 2003 Sep;5(5):323-9.

22. Wuerz RC, Travers D, Gilboy N, Eitel DR, Rosenau A, Yazhari R. Implementation and Refinement of the Emergency Severity Index. Acad Emerg Med. 2001 February 1;8(2):170-6.

23. Goransson K, Ehrenberg A, Marklund B, Ehnfors M. Accuracy and concordance of nurses in emergency department triage. Scand J Caring Sci. 2005 Dec;19(4):432-8.

24. 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 Aug;33(4):319-23.