Validity of triage systems for paediatric emergency care: a systematic review

Maria Clara de Magalhães-Barbosa,[1] Jaqueline Rodrigues Robaina,[1] Arnaldo Prata-Barbosa,[1,2] Claudia de Souza Lopes[3]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/emermed-2016-206058).

[1]Department of Paediatrics, Instituto D'Or de Pesquisa e Ensino (IDOR), Rio de Janeiro, RJ, Brazil
[2]Department of Paediatrics, School of Medicine, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brazil
[3]Department of Epidemiology, Instituto de Medicina Social (IMS), Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, RJ, Brazil

Correspondence to
Dr Maria Clara de Magalhães-Barbosa, Department of Paediatrics, Instituto D'Or de Pesquisa e Ensino (IDOR), Rio de Janeiro, RJ 22281-080, Brazil; mariaclaramb@globo.com

Received 15 May 2016
Revised 31 July 2017
Accepted 1 August 2017
Published Online First
4 October 2017

ABSTRACT

Aim To present a systematic review on the validity of triage systems for paediatric emergency care.

Methods Search in MEDLINE, Cochrane Library, Latin American and Caribbean Health Sciences Literature (LILACS), Scientific Electronic Library Online (SciELO), Nursing Database Index (BDENF) and Spanish Health Sciences Bibliographic Index (IBECS) for articles in English, French, Portuguese or Spanish with no time limit. Validity studies of five-level triage systems for patients 0–18 years old were included. Two reviewers performed data extraction and quality assessment as recommended by PRISMA statement.

Results We found 25 studies on seven triage systems: Manchester Triage System (MTS); paediatric version of Canadian Triage and Acuity Scale (PedCTAS) and its adaptation for Taiwan (paediatric version of the Taiwan Triage and Acuity System); Emergency Severity Index version 4 (ESI v.4); Soterion Rapid Triage System and South African Triage Scale and its adaptation for Bostwana (Princess Marina Triage Scale). Only studies on the MTS used a reference standard for urgency, while all systems were evaluated using a proxy outcome for urgency such as admission. Over half of all studies were low quality. The MTS, PedCTAS and ESI v.4 presented the largest number of moderate and high quality studies. The three tools performed better in their countries or near them, showing a consistent association with hospitalisation and resource utilisation. Studies of all three tools found that patients at the lowest urgency levels were hospitalised, reflecting undertriage.

Conclusions There is some evidence to corroborate the validity of the MTS, PedCTAS and ESI v.4 for paediatric emergency care in their own countries or near them. Efforts to improve the sensitivity and to minimise the undertriage rates should continue. Cross-cultural adaptation is necessary when adopting these triage systems in other countries.

► http://dx.doi.org/10.1136/emermed-2016-206295
► http://dx.doi.org/10.1136/emermed-2017-206973

To cite: de Magalhães-Barbosa MC, Robaina JR, Prata-Barbosa A, *et al*. *Emerg Med J* 2017;**34**:711–719.

INTRODUCTION

The purpose of an emergency triage system is to establish a safe and effective hierarchy of care, based on clinical risk, by prioritising the more urgent cases.[1] Paediatric triage is a complex task, which presents many challenges to the triage team due to communication difficulties with young children and their parents and high variability over a wide range of factors within each age group, such as physiological parameters, epidemiology and clinical presentation of various diseases.[2]

The most widely used triage systems are the Australasian Triage Scale (ATS),[3 4] the Canadian Triage and Acuity Scale (CTAS),[5 6] the Manchester Triage System (MTS)[7] and the Emergency Severity Index (ESI)[8 9] developed in USA. They are complex five-level triage systems, which have demonstrated better validity when compared with three-level systems.[10] Studies on these triage systems have been performed predominantly in their respective home countries and in the adult population. The South African Triage Scale (SATS) is a more recent and less complex scale, developed in an emergent country, but there are few studies on its effectiveness, particularly in the paediatric population.[11 12]

According to the American College of Emergency Physicians (ACEP) and the Emergency Nurses Association (ENA), the ideal triage scale must demonstrate the characteristics of reliability, validity, utility and relevance.[13] The triage process must be easily understood, rapidly applied, have high rates of inter-observer agreement, facilitate appropriate placement, correlate with ED resource use requirements and predict clinical outcomes, including severity of illness and mortality rate.[13]

The validity of triage systems depends on their ability to discriminate different levels of urgency. Criterion validity, that is, comparison with a reference standard, which is the preferred method for validating diagnostic tests, is a challenge as there is no reference standard for 'urgency'.[14] The literature evaluating the validity of triage systems has relied on one of the two methods: (1) the comparison of the performance of the triage system with a reference standard developed by experts (an approximation of the criterion validity) (see online supplementary files 1A, 1B); (2) the association of levels of urgency with outcomes proxy variables of urgency, mainly hospitalisation, resource utilisation and length of stay in the ED.[15] The expert-developed reference standard uses data such as clinical picture and vital signs at presentation and outcomes related to diagnostic tests performed, treatment received and patient's destination, to determine retrospectively the 'true' urgency level of the patient, to be compared with the level assigned by the triage system. Several combinations of outcomes may be associated to some levels of urgency.[16]

Recent reviews suggest that there are many gaps regarding the validity of triage systems, particularly in the paediatric population.[15 17–20] The aim of this study was to perform a systematic review on the validity of triage systems for paediatric emergency

care, assessed by either an expert developed reference standard or the association with proxy outcomes.

METHODS

Search of literature

From July 2014 to September 2015, we searched for original articles, systematic reviews, government and medical society documents in several databases (MEDLINE, Cochrane Library, Latin American and Caribbean Health Sciences Literature (LILACS), Scientific Electronic Library Online (SciELO), Nursing Database Index (BDENF), Spanish Health Sciences Bibliographic Index (IBECS)), in the reference lists of selected articles, in Google and Google Scholar. The search included articles published in English, French, Spanish or Portuguese with no time restriction. The concepts used were *emergency department, child, triage and validity or reliability,* according to the PICO strategy (PRISMA guidelines).[21] We added the name of each triage system found in the first step to broaden the search (online supplementary file 2A).

Selection of studies

Two reviewers (MCMB and APB) performed the selection of articles, based on the inclusion and exclusion criteria. Both researchers are paediatricians with extensive experience in paediatric emergency care. Instances of disagreement were discussed to meet consensus.

Inclusion and exclusion criteria

We initially selected original articles on the validity and reliability of five-level triage instruments applied to general paediatric population or paediatric subgroups, aged 0–18 years old, who were triaged in hospital ED. Because of size limitations, we split the resultant material into two groups: studies with validity assessment to be included in the present review and studies with reliability assessment to be included in another review. For this review on validity, we included prospective or retrospective studies with two different designs: (1) those comparing levels of urgency assigned by triage systems to a reference standard developed by experts; (2) those assessing the association of urgency levels with outcomes proxies of urgency, such as resource utilisation, hospitalisation, admission to the paediatric intensive care unit, ED length of stay (LOS) or severe bacterial infection.

We excluded studies of prehospital care, mass casualty events or telephone triage.

Data extraction and quality assessment

Two reviewers (MCMB and JRR), both with a PhD in epidemiology, independently performed data extraction and elaborated a list of items by consensus to evaluate the methodological quality of the articles (see online supplementary file 2B). The items were based on three tools: Quality Assessment of Diagnostic Accuracy Studies (QUADAS),[22] Statement for Reporting of Diagnostic Accuracy (STARD)[23] and an instrument developed by Hayden *et al* for prognostic studies.[24] The two reviewers independently classified the risk of bias related to participants, attrition, measurement, outcome and statistical analysis, into high, uncertain or low categories, as recommended by QUADAS. They also rated the quality of the discussion section as good, moderate or poor. Agreement between the two reviewers for each type of bias and the discussion assessment was estimated by the quadratic-weighted kappa (kw^2). Instances of disagreement were solved by consensus. We classified the methodological quality in each study as low, moderate or high according to the amount of risk of bias.

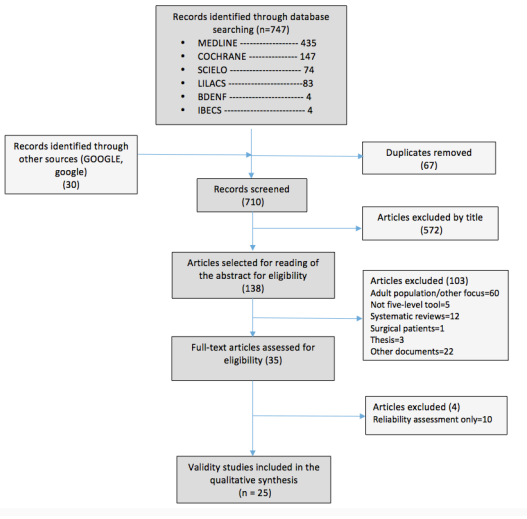

Figure 1 Algorithm for the selection of references according to PRISMA guidelines. BDENF, Base de Dados de Enfermagem (Nursing Database Index); IBECS, Índice Bibliográfico Español de Ciencias de la Salud (Spanish Health Sciences Bibliographical Index); LILACS, Literatura Latino-Americana em Ciências da Saúde (Latin American and Caribbean Health Sciences Literature); SciELO, Scientific Electronic Library Online.

In this review, we did not try to pool data from studies of the same triage systems, because there was great heterogeneity in the sampling, methods of validation and definitions of outcome variables.

We used the statistical software Stata V.12.0 (Stata, Texas, USA).

RESULTS

The search strategy located 25 articles on five original and two adapted triage systems for paediatric emergency care: MTS (n=9); the paediatric version of CTAS (PedCTAS) (n=8); the pediatric version of Taiwan Triage System (PedTTAS), an adaptation of the PedCTAS (n=1); ESI (n=4); Soterion Rapid Triage System (SRTS) (n=1); SATS (n=1) and Princess Marina Triage Scale (PATS), an adaptation of the SATS (n=1) (see online supplementary file 2A, figure 1, table 1).

Agreement between the two reviewers for the risk of the various types of bias varied from a kw^2 of 0.848 (95% CI 0.722 to 0.965) to 0.573 (95% CI 0.242 to 0.667) (see online supplementary file 2C). The validity assessments were rated high quality in two articles, moderate quality in 11 and low quality in 12 (table 2).

There were 14 retrospective and 11 prospective observational studies assessing validity (table 1). Most of them used proportional sampling, including the five levels of urgency with the same frequency as they occurred in the source population. Two studies used disproportionate sampling to ensure a minimum number of patients in the most urgent levels or the same number of patients in all levels of urgency (table 3). The detailed characteristics of these studies are summarised in the online supplementary file 2D.

Table 3 includes only studies that used the reference standard method and reported estimates of sensitivity, specificity, overtriage and undertriage rates. Table 4 presents results for those studies that used proxy outcomes giving estimates of

Table 1 Characteristics of validity studies of triage systems for paediatric emergency care

						Study design
Triage system	**Author, year**	**Country**	**Number of centres**	**Sample size**	**Study population**	**Validity assessment**
1—PedCTAS	Gravel J *et al*[25]	Canada	Multicentre (12)	n=550 940	<18 years	Retrospective
2—PedCTAS	Gravel J *et al*[26]	Canada	Multicentre (9)	n=1464	<18 years	Prospective
3—PedCTAS	Gravel J *et al*[27]	Canada	One centre	n=58 529	<18 years	Retrospective
4—PedCTAS	Gravel J *et al*[47]	Canada	One centre	n=28 349	<18 years	Retrospective
5—PedCTAS	Ma W *et al*[28]	Canada	One centre	n=1618	0–19 years	Retrospective
6—PedCTAS	Gouin S *et al*[29]	Canada	One centre	n=1281	0–19 years	Prospective
7—PedCTAS	Fernandez *et al*[30]	Spain	One centre	n=57 617	0–14 years	Retrospective
8—PedCTAS	Al-Hindi AA *et al*[31]	Saudi Arabia	One centre	n=3337	0–12 years	Prospective
9—TTAS	Chang YC *et al*[32]	Thailand	One centre	n=84 874	<18 years	Retrospective
10—ESI v.4	Green NA *et al*[33]	USA	One centre	n=780	<18 years	Retrospective
11—ESI v.4	Travers DA *et al*[34]	USA	One centre	n=1000	<18 years	Prospective
12—ESI v.4	Baumann MR *et al*[35]	USA	One centre	n=510	<14 years	Retrospective
13—ESI v.4	Jafari-Rouhi AH *et al*[36]	Iran	One centre	n=1104	<18 years	Prospective
14—MTS	Roukema J *et al*[16]	Netherlands	One centre	n=1065	<16 years	Retrospective
15—MTS	Van Veen M *et al*[37]	Netherlands	Multicentre (2)	n=13 554	<16 years	Prospective
16—MTS	Seiger N *et al*[38]	Netherlands	Multicentre (2)	n=13 408	<16 years	Retrospective
17—MTS	Van Veen M *et al*[39]	Netherlands	Multicentre (2)	n=11 260	<16 years	Prospective
18—MTS	Nijman RG *et al*[40]	Netherlands	One centre	n=1255	1 month to 16 years	Prospective
19—MTS	Van Veen M *et al*[41]	Netherlands	Multicentre (2)	n=3975	<16 years	Prospective
20—MTS	Seiger N *et al*[42]	Netherlands	One centre	n=8592	<16 years	Prospective
21—MTS	Seiger N *et al*[43]	Netherlands, Portugal and England	Multicentre (3)	n=60 375	<16 years	Retrospective
22—MTS	Madramany AA *et al*[44]	Spain	One centre	n=23 173	<14 years	Retrospective
23—SRTS	Maningas PA *et al*[45]	USA	One centre	n=7077	<13 years	Retrospective
24—SATS	Twomey M *et al*[12]	South Africa	Multicentre (6)	n=2014	<13 years	Prospective
25—PATS	Mullan PC *et al*[46]	Botswana	One centre	n=35 948	<13 years	Retrospective

ESI v.4, Emergency Severity Index version 4; MTS, Manchester Triage System; PATS, Princess Marina Hospital Triage Scale; PedCTAS, paediatric version of Canadian Triage and Acuity Scale; SATS, South African Triage Scale; SRTS, Soterion Rapid Triage System; TTAS, Taiwan Triage and Acuity Scale.

sensitivity, specificity, overtriage and undertriage (for the proxy outcome), while table 5 looks at studies using proxy outcomes but presents the frequency of clinical outcomes in each level of urgency. Therefore, two studies were included twice: Roukema *et al*[16] used the two methods of validation (tables 3 and 4) and Travers *et al* (2009) reported two types of estimates (tables 3 and 5).

Five studies performed in Netherlands compared the MTS to an expert developed reference standard to assess the ability of the MTS to detect high urgency cases (levels 1 and 2) (table 3). The original MTS presented a moderate sensitivity of 63% and high overtriage rates (40%–54%).[16 37] A modified version of the MTS increased the specificity from 78%–79% to 87%, but not the sensitivity (64%), resulting in the reduction of overtriage (47%) without a parallel increase in undertriage (15%).[39] The other two studies assessed subgroups of patients and showed an undertriage rate of 2% in patients levels 1 and 2[38] and poorer sensitivity (58% vs 74%) and higher undertriage rate (17% vs 11%) in patients with symptoms of infection and chronic disease compared with those without chronic disease.[42]

Twenty-one studies assessed the association of levels of urgency with one or more proxy outcomes of urgency (tables 4 and 5). Seven of these studies assessed one outcome, such as hospitalisation or severe bacterial infection (one study) and reported sensitivity, specificity, over/undertriage rates or other estimates to predict the outcome (table 4). These results could not be compared with the results of the MTS studies in table 3, because the definitions they are based on were completely different.

Fifteen of those 21 studies assessed the frequency of at least two of the following three outcomes (hospitalisation rates, resource utilisation and LOS) across the five-triage levels (table 5). In the first nine validity studies in table 5, triage systems were assessed in the countries where they were developed; in the last six validity studies, triage systems were assessed outside their own countries. Level 1 (immediate urgency) represented less than 1% of the visits, while levels 3 (urgent) and 4 (low urgency) together contributed around 70%–90%, in most of the studies which used proportionate sampling. The distribution of urgency levels was more similar in the Canadian and US studies compared with studies in other countries. In most of these studies, the frequency of hospitalisation decreased from the higher to the lower level of urgency. The decreasing gradient was more evident in studies performed in the countries where the triage systems were developed, such as Canada (PedCTAS) and US (ESI-4). The combined frequency of hospital admission in levels 4 and 5 with the PedCTAS varied from 2.6% to 4% in Canadian studies and from 1.5% to 25% in other countries. In the ESI v.4 studies, it varied from 1.8% to 6% in US studies and was 3.3% in the only study performed in Iran. The two MTS studies showed a combined hospital admission rate in levels 4 and 5 of 0.9% in Netherlands and 5% in Spain.

Despite different definitions and cut-off points used for the outcome 'resource utilisation' (diagnostic and therapeutic resources or only laboratory or radiological tests or hospital costs), most studies showed decreasing frequency from the

Table 2 Risk of bias in validity studies of paediatric triage systems

		Bias type/kappa						
		Participants	Attrition/losses	Measurement	Outcome measures	Statistical analysis		
Triage system	Author/year/country	kw^2 0.70 (0.58–0.78)	kw^2 0.66 (0.54–0.80)	kw^2 0.77 (0.56–0.89)	kw^2 0.57 (0.24–0.67)	kw^2 0.85 (0.72–0.97)	Quality of discussion	Methodological quality
PedCTAS	Gravel *et al*[25], 2013, Canada[25]	Low risk	Low risk	Low risk	Uncertain	Uncertain	Good	Moderate
PedCTAS	Gravel *et al*, 2012, Canada[26]	Uncertain	Low risk	Low risk	Uncertain	Low risk	Good	Moderate
PedCTAS	Gravel *et al*, 2009, Canada[27]	Low risk	Uncertain	Low risk	Uncertain	Uncertain	Good	Moderate
PedCTAS	Gravel *et al*, 2008, Canada[47]	Low risk	Low risk	Low risk	Low risk	Low risk	Good	High
PedCTAS	Ma *et al*, 2009, Canada[28]	Uncertain	Low risk	Low risk	Uncertain	High risk	Good	Low
PedCTAS	Gouin *et al*, 2005, Canada[29]	Uncertain	High risk	Low risk	Uncertain	Uncertain	Good	Low
PedCTAS	Fernandez *et al*, 2010, Spain[30]	Low risk	Low risk	Uncertain	Uncertain	High risk	Fair	Low
PedCTAS	Al-Hindi *et al*, 2014, Saudi Arabia[31]	High risk	Uncertain	Uncertain	Uncertain	Uncertain	Fair	Low
TTAS	Yu-Che Chang *et al*, 2013, Thailand[32]	Low risk	Low risk	High risk	Uncertain	Uncertain	Fair	Low
ESI v.4	Green *et al*, 2012, USA[33]	Low risk	Low risk	Low risk	High risk	Uncertain	Good	Moderate
ESI v.4	Travers *et al*, 2009, USA[34]	Low risk	Low risk	Low risk	Uncertain	Uncertain	Good	Moderate
ESI v.4	Baumann *et al*, 2005, USA[35]	High risk	Low risk	Low risk	High risk	Uncertain	Good	Low
ESI v.4	Jafari-Rouhi *et al*, 2013, Iran[36]	Uncertain	Uncertain	Low risk	High risk	High risk	Poor	Low
MTS	Roukema *et al*, 2006 Netherlands[16]	Low risk	Low risk	Low risk	Uncertain	Uncertain	Good	Moderate
MTS	Van Veen *et al*, 2008, Netherlands[37]	Low risk	Uncertain	Low risk	Uncertain	Low risk	Good	Moderate
MTS	Seiger *et al*, 2011, Netherlands[38]	Uncertain	Low risk	Low risk	Uncertain	Low risk	Good	Moderate
MTS	Van Veen *et al*, 2012, Netherlands[39]	Low risk	Low risk	Low risk	Uncertain	Low risk	Good	High
MTS	Nijman *et al*, 2011, Netherlands[40]	Uncertain	Low risk	Low risk	High risk	Low risk	Good	Moderate
MTS	Van Veen *et al*, 2011, Netherlands[41]	Low risk	Uncertain	Uncertain	Uncertain	Uncertain	Good	Low
MTS	Seiger *et al*, 2013, Netherlands[42]	Uncertain	Low risk	Uncertain	Uncertain	Low risk	Good	Moderate
MTS	Seiger *et al*, 2014, Netherlands/Portugal/England[43]	Low risk	Uncertain	Uncertain	Uncertain	Low risk	Good	Moderate
MTS	Madramany *et al*, 2013, Spain[44]	Uncertain	High risk	High risk	Uncertain	Uncertain	Fair	Low
SATS	Twomey *et al*, 2013[12]	Uncertain	Low risk	Low risk	Uncertain	High risk	Fair	Low
PATS	Mullan *et al*, 2014[46]	Uncertain	Uncertain	Low risk	Uncertain	Uncertain	Good	Low
SRTS	Maningas *et al*, 2006, USA[45]	Uncertain	Uncertain	Low risk	Uncertain	Uncertain	Good	Low

ESI v.4, Emergency Severity Index version 4; kw^2, quadratic weighted kappa for the interrater agreement between two reviewers for each type of risk of bias, with 95% CI between parenthesis; MTS, Manchester Triage System; PATS, Princess Marina Hospital Triage Scale; PedCTAS, paediatric version of Canadian Triage and Acuity Scale; SATS, South African Triage Scale; SRTS, Soterion Rapid Triage System; TTAS, Taiwan Triage and Acuity Scale.

highest to the lowest level of urgency. Again, this gradient was clearer in the Canadian and US studies (table 5).

The outcome LOS did not show a consistently decreasing gradient across the five levels of urgency. The LOS for level 1 was shorter than for level 2 in two PedCTAS studies in Canada and two ESI v.4 studies in USA, one of which also had a shorter LOS for level 2 than for level 3 (table 5).

DISCUSSION

The present review found 25 validity studies on five original and two adapted triage systems in paediatric emergency care. Five studies (all MTS) used an expert-developed reference standard, while 21 studies involving all triage systems (MTS, PedCTAS, TTAS, ESI v.4, SATS, PATS and SRTS) used proxy outcomes of urgency. The MTS, the PedCTAS and the ESI v.4 were the triage systems with the largest number of moderate or high quality studies. There were no studies on the validity of the ATS in the paediatric population and very few and low quality studies of the SATS and the STRS.

The use of a reference standard seems to be advantageous, because as it establishes objective criteria to define each level of urgency, it ensures a more robust assessment of the 'true' urgency of patients and more consistent comparison between studies. However, there is no scientific evidence of the validity of this reference standard.[14] Besides, the criteria used to define the levels of urgency include clinical outcomes similar to the proxy outcomes of urgency used in the other type of study. The difference is that, with a reference standard, some levels of urgency are defined by many possibilities of combinations of these outcomes. These outcomes are satisfactory markers of complexity and severity, but they do not always account for all sets of urgency. Moreover, they can be influenced by variables related to the quality and efficacy of the treatment given.[14 15] Twomey *et al* suggested a Delphi process to achieve consensus among experts to serve as a reference standard in validity studies. This could eliminate the limitations associated with proxies of urgency and the biases inherent to individual groups of specialists.[14]

The evidence on the validity of triage systems in paediatric patients is better for the PedCTAS, the ESI v.4 and the MTS, but remains insufficient. The three triage systems showed unacceptably high rates of hospital admission in the less urgent levels in several studies, suggesting undertriage. The MTS is the most intensively studied triage system in the paediatric population, but the sensitivity to detect high urgency was modest, despite an elevated overtriage and a moderate undertriage rate. Although there are no recommendations about the safe limits of sensitivity, undertriage and overtriage rates for emergency triage systems, an effective screening tool is expected to prioritise sensitivity and a low undertriage rate. On the other hand, a high overtriage rate might affect the flow of patients,

Table 3 Outcomes resulting from the comparison of the MTS with an expert developed reference standard: sensitivity, specificity, positive and negative likelihood ratios, absolute agreement, over/undertriage rates and diagnostic OR of high urgency

Author	Quality	Systems	Method	Observation	Sens % (95% CI)	Sp % (95% CI)	LR (+) % (95% CI)	LR (–) % (95% CI)	Absolute agreement %	Overtriage %	Undertriage %	DOR (95% CI)
Roukema *et al*, 2006[16]	++/+++	MTS	MTS versus Reference Standard	Ability to detect high urgency	63	78			45	40	15	
Van Veen, 2008[37]	++/+++	MTS	MTS versus Reference Standard	Ability to detect high urgency	63 (59 to 66)	79 (79 to 80)	3.0 (2.8 to 3.2)	0.47 (0.43 to 0.52)	34	54	12	
Van Veen, 2012[39]	+++/+++	ModifiedMTS	MTS versus Reference Standard	Ability to detect high urgency	64 (60 to 68)	87 (86 to 87)			37	47	15	DOR 11.5 (9.6 to 14)
Seiger, 2011[38]	++/+++	MTS	MTS versus Reference Standard	Percentage of undertriage in levels 1 and 2							2 0.9* 0.65**	
Seiger, 2013[42]	++/+++	MTS	MTS versus Reference Standard	Ability to detect high urgency in febrile patients with versus without chronic disease	58 (53 to 62) x 74 (70 to 78)	78 (76 to 79) x 75 (74 to 77)	2.6 (2.4 to 2.9) x 1.1 (1.1 to 1.2)	0.71 (0.68 to 0.74) x 0.80 (0.73 to 0.87)	35×30	48×59	17×11	DOR 4.8 (3.9 to 5.9) x 8.7 (7.1 to 11)

*Percentage of undertriage >1 category; **Percentage ofserious undertriage according to experts.
DOR, diagnostic OR of high urgency = (sensitivity/1-sensitivity)/(1-specificity/specificity); LR(+), positive likelihood ratio; LR(-), negative likelihood ratio; MTS, Manchester Triage System; NPP, negative predictive value; PPV, positive predictive value; Sens, sensitivity; Sp, specificity.

Table 4 Outcomes resulting from the association of levels of urgency of triage systems with hospitalisation or severe bacterial infection: sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, over/undertriage rates, area under the curve and OR of hospital admission

Author	Quality	Systems	Method	Observation	Sens % (95% CI)	Sp % (95% CI)	PPV % (95% CI)	NPV% (95% CI)	LR (+) % (95% CI)	LR (–) % (95% CI)	Overtriage %	Undertriage %	AUC (95% CI)	OR (95% CI)
Nijman, 2011[40]	++/+++	MTS	Association with serious bacterial infection	Ability to detect serious bacterial infection (SBI)	42 (35 to 41)	69 (66 to 72)	14 (10 to 17)	91 (89 to 93)	1.35 (1.08 to 1.69)	0.84 (0.73 to 0.98)			0.57 (0.52 to 0.62)	
Seiger, 2014[43]	++/+++	Modified MTS 1 MTS 2	MTS1 & MTS2 versus original MTS Association with hospitalisation	Ability to predict hospital admission					1.1 (1.1 to 1.2) x 2.6 (2.4 to 2.9)	0.80 (0.73 to 0.87) x 0.71 (0.68 to 0.74)			0. 56 (0.55 to 0.58) x 0.66 (0.64 to 0.67)	1.4 (1.2 to 1.6) x 3.7 (3.3 to 4.7)
Twomey *et al*, 2013[12]	+/+++	SATS	Association with hospitalisation	Ability to predict hospital admission	91	54.5	37.5	95.3			45.5	9		
Mullan, 2014[46]	+/+++	PATS	Association with hospitalisation	Overtriage and undertriage rates (see definitions below)*							28.8*	21.9*		
Travers, 2009[34]	++/+++	ESI v.4	Association with hospitalisation	Overtriage and undertriage rates (see definitions below)**							16†	11†		
													Hospital admission rate % (95% CI for the difference) RD % (SD)	
Gravel, 2008[47]	+++/+++	PedCTAS	Association with hospital admission in febrile infants 3–36 months old	Hospital admission rate in downtriaged level 3 versus remaining level 3 and all level 4									2.4×2.8 (0.3 to 1.1) (downtriaged level 3 x all level 4) RD 13.1 to 2.4=10.7 (±1.9, p<0.001) (remaining level 3 - downtriaged level 3)	
													OR (95% CI)	
Van Veen, 2011[41]	+/+++	MTS	Association of levels 4 and 5 with hospitalisation	Complaints associated with greater chance of hospital admission									Hospital admission 3.5% OR 3.0 (2.2 to 4.1)<1 year OR 2.5 (1.5 to 4.1) dyspnoea OR 3.5 (2.5 to 4.9) GI problem OR 2.8 (1.1 to 7.2) fever w/o source	

*Overtriage=PATS level 1 or 2 which were not admitted to hospital; undertriage=PATS level 4, which were admitted to hospital.
†Overtriage=ESI level 1, 2 or 3 which used <2 resources, or ESI level 1 which were not admitted to hospital; undertriage=ESI level 4 or 5 which used ≥2 resources or which were admitted to hospital.
Quality— +/+++: low quality; ++/+++: moderate quality; +++/+++: high quality.
AUC, area under the curve; ESI v.4, Emergency Severity Index version 4; LR (+), positive likelihood ratio; LR (–), negative likelihood ratio; MTS, Manchester Triage System; MTS1, modification in discriminators; MTS2, modification in vital signs; NPP, negative predictive value; PATS; Princess Marina Hospital Triage Scale; OR, odds ratio of prediction of hospital admission; PPV, positive predictive value; RD, risk difference; SATS, South African Triage Scale; Sens, sensitivity; Sp, specificity;

Table 5 Distribution of the study population and the outcomes 'hospital admission', 'resource utilisation' and 'length of stay' across the five levels of urgency in the validity studies of paediatric triage systems

Author/year	Country	System	Quality*	Study population (%)					Hospital admission rate (%)					Resource utilisation					Length of stay (min)				
				Urgency levels					Urgency levels					Urgency levels					Urgency levels				
				L1	L2	L3	L4	L5	L1	L2	L3	L4	L5	L1	L2	L3	L4	L5	L1	L2	L3	L4	L5
Gravel, 2013[25]	Canada	PedCTAS	++/+++	0.6	11.0	37.0	44.0	7.0	61.0	30.0	10.0	2.0	0.9						280	282	232	182	76
														≥1 diagnostic/therapeutic resource (%)									
Gravel, 2012[26]	Canada	PedCTAS	++/+++	–	7	44	46	3	–	30.0	8.0	2.0	2.0	–	78.0	59.0	42.0	31.0	–	309	238	186	160
Gravel, 2009[27]	Canada	PedCTAS	++/+++	1.2	6.8	31.5	46.9	13.5	62.9	37.4	13.8	1.9	0.7						191	250	191	96	66
Gouin, 2005[29]	Canada	ESI v.4	+/+++	–	5	48	37	10	–	30.0	12.0	4.0	2.0										
														≥2 diagnostic/therapeutic resource (%)									
Green, 2012[33]	USA	ESI v.4	++/+++	0.3	9.3	37.1	32.2	21.2	100	42.5	14.9	1.2	0.6	100.0	57.5	39.4	14.7	7.3	102	247	262	152	128
Travers, 2009†[34]	USA	ESI v.4	++/+++	16.5	21.3	20.0	21.8	20.2	83.0	46.0	17.0	4.0	0.0	100.0	70.0	45.0	17.0	4.0	156	236	259	117	99
Baumann, 2005[35]	USA	ESI v.4	+/+++	2.9	18.6	34.1	36.7	7.7	80.0	34.0	13.0	3.0	0.0	93.0	51.0	22.0	6.0	0.0	334	221	207	151	132
														Hospital costs (Canadian‡ and American§ dollar, respectively)									
Ma, 2008‡[28]	USA	PedCTAS	+/+++	0.4	9.6	48	38.8	3.3						199 (L1 and L2)		151	119	106					
Maningas, 2006§[45]	USA	SRTS	+/+++											2673	1563	1112	477	258					
Jafari-Rouhi, 2013[36]	Iran	ESI v.4	+/+++	0.5	29.0	10.0	30.0	30.0	100	30.7	3.5	2.7	0.6						–	–	–	–	–
														≥2 diagnostic/therapeutic resource (%)									
Roukema, 2006†[16]	Netherlands	MTS	++/+++	0.8	19.7	23.9	40.8	0.8	53.5	28.6	16.2	6.0	0.9	41.7	25.4	30.2	16.6	3.7	–	–	–	–	–
														≥1 diagnostic test									
Fernandez, 2010[30]	Spain	PedCTAS	+/+++	0.1	1.7	43.1	50.0	5.1	84.0	?	?	1.6	1.6	63.6	41.4	16.8	9.4	8.0	–	–	–	–	–
														Laboratory/Radiological tests									
Al-Hindi, 2014[31]	Saudi Arabia	PedCTAS	+/+++	0.1	12.0	22.0	60.0	6.0	75.0	17.0	5.0	1.0	0.5	100/50	58/41	22/17	7/9	3/2	–	–	–	–	–
Madramany, 2012[44]	Spain	MTS	+/+++	0.5	0.9	3.0	94.0	1.7	82.0	20.0	21.0	4.0	5.0	91/43	57/31	55/28	31/13	32/14	–	–	–	–	–
														Hospital costs (NT dollar)									
Chan, 2013[32]	Taiwan	PedTTAS	+/+++	14.0	9.0	50.0	22.0	5.0	28.0	36.0	22.0	16.0	9.0	2295	2573	1510	1293	1055	–	–	–	–	–

*Quality – +/+++, low quality; ++/+++, moderate quality; +++/+++, high quality.
†Studies that used samples with a disproportionate distribution of patients among the five levels of urgency.
‡Study that assessed costs with Canadian dollar.
§Study that assessed costs with American dollar.
L1; L2; L3; L4; L5: urgency levels 1; 2; 3; 4 and 5.
ESI v.4, Emergency Severity Index version 4; MTS, Manchester Triage System; PedCTAS, paediatric version of Canadian Triage and Acuity Scale; PedTTAS, paediatric version of the Taiwan Triage and Acuity System.

ultimately compromising the care of the most urgent patients. An ideal triage system must balance between safety and accuracy.[9] These findings raise questions about the safety of triage systems, especially if used to divert the least urgent patients to outpatient care.

Although the PedCTAS and ESI v.4 consistently predicted hospital admission and resource utilisation in the countries where they were developed, the performances in other countries such as Spain, Iran and Taiwan were lower. This can be detected in table 5, where the frequency of the five levels of urgency in the study populations of the Canadian and US studies were very similar, while in other countries these frequencies were more heterogeneous. We could not determine if this heterogeneity was due to actual differences in the characteristics of the study populations or differences in the knowledge and training of the healthcare professionals, which may have contributed to misclassification and lower performance of the instruments in those countries. Indeed, the low methodological quality of some studies may have accounted for the inconsistency of the outcomes observed. Therefore, caution is necessary when applying inferences from studies performed in the countries where the triage systems were developed to other countries with highly diverse healthcare contexts. A myriad of factors, including the morbidity and mortality features of the target population, the quality and amount of technical and human resources, the professional training and skills, sociocultural factors and health policies, among others may play a role.[14]

ED LOS did not consistently decrease across the five levels of urgency in the six studies that analysed this outcome. This should not be surprising as less urgent patients have longer waiting times,[13] while it is important to promptly admit the most urgent patients. Furthermore, mean LOS may be distorted by aberrant values. The use of this outcome should be avoided in future studies.

This review has some limitations. First, almost half of the studies were rated low quality, especially those performed in countries distant from the country where the triage systems were developed. However, as the main triage systems have been adopted in many countries around the world, these studies were purposely included to give an idea of the amount and quality of evidence of the validity of each triage system, and how safe it is to generalise the results to other countries. Second, there was great heterogeneity among the studies, even those that used the same method of validation. Different cut-off points and definitions were used to assess outcomes such as hospitalisation, resource utilisation, ED LOS, sensitivity, specificity, undertriage and overtriage rates. These differences precluded pooling of the data to give the reader some summarise statistics as well as comparing the performance among different triage systems. Third, we could not include the reliability assessment of triage systems because the review would be too extensive. Good reproducibility with high interobserver reliability reinforces the validity of the instruments.[13]

In conclusion, there is some evidence to corroborate the validity of the MTS, PedCTAS and ESI v.4 for paediatric emergency care, particularly in or near the countries where these instruments were developed. However, further efforts are needed to decrease the undertriage rates in the three tools to ensure safety. Diligent cross-cultural adaptation and rigorous training followed by local validity and reliability studies are necessary when adopting these triage systems for paediatric emergency care in countries with different socioeconomic and cultural context. Finally, consensus on the best methods and outcome definitions for validity studies of triage systems among experts from different countries would be very useful to enable comparison of results.

Contributors MCMB conceived and designed the study, contributed to obtain research funding, participated as the first reviewer of the search of the literature, the selection of the articles and the extraction and analysis of data, drafted the initial manuscript and approved the final manuscript as submitted. APB obtained research funding, participated as a second reviewer in the selection of the articles, reviewed and revised the manuscript and approved the final manuscript as submitted. JRR contributed to obtain funding, participated as a second reviewer in the extraction of data, in the design of the quality assessment instruments, in the assessment of the risk of bias of the selected articles, and approved the final manuscript as submitted. CdSL supervised the conduct of the study, reviewed and revised the manuscript and approved the final manuscript as submitted.

Funding Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – 448855/2014-3; Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) – E-26-200.991-2015.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data extraction files and spreadsheets with quality assessment of the articles analysed may be accessible by request to the corresponding author (mariaclaramb@globo.com).

REFERENCES

1 FitzGerald G, Jelinek GA, Scott D, *et al*. Emergency department triage revisited. *Emerg Med J* 2010;27:86–92.
2 Hohenhaus SM, Travers D, Mecham N. Pediatric triage: a review of emergency education literature. *J Emerg Nurs* 2008;34:308–13.
3 Gerdtz MF, Collins M, Chu M, *et al*. Optimizing triage consistency in Australian emergency departments: the emergency triage education kit. *Emerg Med Australas* 2008;20:250–9.
4 Considine J, LeVasseur SA, Charles A. Development of physiological discriminators for the Australasian Triage Scale. *Accid Emerg Nurs* 2002;10:221–34.
5 Warren DW, Jarvis A, LeBlanc L, *et al*. Revisions to the Canadian triage and acuity scale paediatric guidelines (PaedCTAS). *CJEM* 2008;10:224–32.
6 Bullard MJ, Unger B, Spence J, *et al*. Revisions to the Canadian Emergency Department Triage and Acuity Scale (CTAS) adult guidelines. *CJEM* 2008;10:136–42.
7 Mackway-Jones K, Marsden J, Windle J, eds. Emergency triage. In *Manchester triage group*. 2nd ed. Oxford: Blackwell Publishing Ltd, 2006.
8 McHugh M, Tanabe P, McClelland M, *et al*. More patients are triaged using the Emergency Severity Index than any other triage acuity system in the United States. *Acad Emerg Med* 2012;19:106–9.
9 Gilboy N, Tanabe P, Travers D, *et al*. eds. *Emergency severity index, version 4:implementation handbook*. Rockville, MD: Agency for Heathcare Research and Quality, 2005:72.
10 Travers DA, Waller AE, Bowling JM, *et al*. Five-level triage system more effective than three-level in tertiary emergency department. *J Emerg Nurs* 2002;28:395–400.
11 Bruijns SR, Wallis LA, Burch VC. A prospective evaluation of the Cape triage score in the emergency department of an urban public hospital in South Africa. *Emerg Med J* 2008;25:398–402.
12 Twomey M, Cheema B, Buys H, *et al*. Vital signs for children at triage: a multicentre validation of the revised South African Triage Scale (SATS) for children. *S Afr Med J* 2013;103:304–8.
13 Fernandes CM, Tanabe P, Gilboy N, *et al*. Five-level triage: a report from the ACEP/ENA Five-level triage task force. *J Emerg Nurs* 2005;31:39–50.
14 Twomey M, Wallis LA, Myers JE. Limitations in validating emergency department triage scales. *Emerg Med J* 2007;24:477–9.
15 van Veen M, Moll HA. Reliability and validity of triage systems in paediatric emergency care. *Scand J Trauma Resusc Emerg Med* 2009;17:38.
16 Roukema J, Steyerberg EW, van Meurs A, *et al*. Validity of the Manchester triage System in paediatric emergency care. *Emerg Med J* 2006;23:906–10.
17 Christ M, Grossmann F, Winter D, *et al*. Modern triage in the emergency department. *Dtsch Arztebl Int* 2010;107:892–8.
18 Farrohknia 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.
19 Parenti N, Reggiani ML, Iannone P, *et al*. A systematic review on the validity and reliability of an emergency department triage scale, the Manchester Triage System. *Int J Nurs Stud* 2014;51:1062–9.
20 Azeredo TR, Guedes HM, Rebelo de Almeida RA, *et al*. Efficacy of the Manchester triage system: a systematic review. *Int Emerg Nurs* 2015;23:47–52.

21 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
22 Mann R, Hewitt CE, Gilbody SM. Assessing the quality of diagnostic studies using psychometric instruments: applying QUADAS. *Soc Psychiatry Psychiatr Epidemiol* 2009;44:300–7.
23 Bossuyt PM, Reitsma JB, Bruns DE, *et al*. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003;138:W1–12.
24 Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006;144:427–37.
25 Gravel J, Fitzpatrick E, Gouin S, *et al*. Performance of the Canadian triage and acuity scale for children: a multicenter database study. *Ann Emerg Med* 2013;61:27–32.
26 Gravel J, Gouin S, Goldman RD, *et al*. The Canadian triage and acuity scale for children: a prospective multicenter evaluation. *Ann Emerg Med* 2012;60:71–7.
27 Gravel J, Manzano S, Arsenault M. Validity of the Canadian paediatric triage and acuity scale in a tertiary care hospital. *CJEM* 2009;11:23–8.
28 Ma W, Gafni A, Goldman RD. Correlation of the Canadian pediatric emergency triage and acuity scale to ED resource utilization. *Am J Emerg Med* 2008;26:893–7.
29 Gouin S, Gravel J, Amre DK, *et al*. Evaluation of the paediatric Canadian triage and acuity scale in a pediatric ED. *Am J Emerg Med* 2005;23:243–7.
30 Fernández A, Pijoan JI, Ares MI, *et al*. Evaluación de la escala canadiense de triaje pediátrico en un servicio de urgencias de pediatría europeo. *Emergencias* 2010;22:355–60.
31 Al-Hindi AA, Al-Akhfash AA, Fareed AM, *et al*. Efficacy of implementation of a 5 scale pediatric triage and acuity scale in pediatric emergency, Saudi Arabia. *Saudi Med J* 2014;35:999–1004.
32 Chang YC, Ng CJ, Wu CT, *et al*. Effectiveness of a five-level paediatric triage system: an analysis of resource utilisation in the emergency department in Taiwan. *Emerg Med J* 2013;30:735–9.
33 Green NA, Durani Y, Brecher D, *et al*. Emergency Severity Index version 4: a valid and reliable tool in pediatric emergency department triage. *Pediatr Emerg Care* 2012;28:753–7.
34 Travers DA, Waller AE, Katznelson J, *et al*. Reliability and validity of the emergency severity index for pediatric triage. *Acad Emerg Med* 2009;16:843–9.
35 Baumann MR, Strout TD. Evaluation of the emergency severity index (version 3) triage algorithm in pediatric patients. *Acad Emerg Med* 2005;12:219–24.
36 Jafari-Rouhi AH, Sardashti S, Taghizadieh A, *et al*. The emergency severity index, version 4, for pediatric triage: a reliability study in Tabriz children's hospital, Tabriz, Iran. *Int J Emerg Med* 2013;6:36.
37 van Veen M, Steyerberg EW, Ruige M, *et al*. Manchester triage system in paediatric emergency care: prospective observational study. *BMJ* 2008;337:337–a1501.
38 Seiger N, van Veen M, Steyerberg EW. Undertriage in the Manchester triage system: an assessment of severity and options for improvement. *Arch Dis Child* 2011 96:653–7.
39 van Veen M, Steyerberg EW, Van't Klooster M, *et al*. The Manchester triage system: improvements for paediatric emergency care. Emergency medicine journal. *EMJ* 2012;29:654–9.
40 Nijman RG, Zwinkels RL, van Veen M, *et al*. Can urgency classification of the Manchester triage system predict serious bacterial infections in febrile children? *Arch Dis Child* 2011;96:715–22.
41 van Veen M, Steyerberg EW, Lettinga L, *et al*. Safety of the Manchester triage system to identify less urgent patients in paediatric emergence care: a prospective observational study. *Arch Dis Child* 2011;96:513–8.
42 Seiger N, van Veen M, Steyerberg EW, *et al*. Accuracy of triage for children with chronic illness and infectious symptoms. *Pediatrics* 2013;132:e1602–e1608.
43 Seiger N, van Veen M, Almeida H, *et al*. Improving the Manchester triage system for pediatric emergency care: an international multicenter study. *PLoS One* 2014;9:e83267.
44 Madramany AA, Reina PM, Peredo DC. Sistema de triaje y evaluatión de Urgencias pediátricas. *Rev Esp Pediatr* 2012;68:29–34.
45 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;31:353–9.
46 Mullan PC, Torrey SB, Chandra A, *et al*. Reduced overtriage and undertriage with a new triage system in an urban accident and emergency department in Botswana: a cohort study. *Emerg Med J* 2014;31:356–60.
47 Gravel J, Manzano S, Arsenault M. Safety of a modification of the triage level for febrile children 6 to 36 months old using the paediatric Canadian triage and acuity scale. *CJEM* 2008;10:32–7.