

VALIDADE E CONFIABILIDADE DE UM NOVO SISTEMA DE CLASSIFICAÇÃO DE RISCO PARA EMERGÊNCIAS PEDIÁTRICAS: CLARIPED

Validity and reliability of a new triage system for pediatric emergency care: CLARIPED

Maria Clara de Magalhães-Barbosa^{a,*}, Arnaldo Prata-Barbosa^a,
Carlos Eduardo Raymundo^a, Antonio José Ledo Alves da Cunha^b,
Claudia de Souza Lopes^b

RESUMO

Objetivo: Avaliar a validade e a confiabilidade de um sistema de classificação de risco para emergências pediátricas (CLARIPED) desenvolvido no Brasil.

Métodos: Fase de validade: estudo observacional prospectivo em crianças de 0 a 15 anos atendidas consecutivamente no serviço de emergência pediátrica (SEP) de um hospital terciário entre 2 e 18 de julho de 2013. Avaliou-se a associação dos níveis de urgência com desfechos clínicos (utilização de recursos, taxa de admissão no SEP, taxa de hospitalização e tempo de permanência no SEP); e comparou-se o desempenho do CLARIPED com um padrão de referência. Fase de confiabilidade interobservadores: uma amostra de conveniência de pacientes atendidos entre abril e julho de 2013 foi submetida a duas classificações de risco consecutivas e independentes, e estimou-se o Kappa ponderado quadrático.

Resultados: Em 1.416 atendimentos realizados na fase de validade, a distribuição de níveis de urgência foi: 0,0% vermelho/emergência; 5,9% laranja/muito urgente; 40,5% amarelo/urgente; 50,6% verde/pouco urgente; e 3,0% azul/sem urgência. Os percentuais de pacientes que usaram dois ou mais recursos diminuíram do nível laranja para o amarelo, verde e azul (81, 49, 22 e 2%, respectivamente, $p < 0,0001$), assim como a taxa de admissão no SEP, o tempo de permanência no SEP e a taxa de admissão hospitalar. A sensibilidade para discriminar pacientes muito urgentes foi de 0,89 (intervalo de confiança de 95% — IC95% 0,78–0,95), e a taxa de subtriagem foi de 7,4%. A confiabilidade interobservadores em 191 pacientes classificados por duas enfermeiras foi substancial ($k_w^2 = 0,75$; IC95% 0,74–0,79).

ABSTRACT

Objective: To assess the validity and reliability of a triage system for pediatric emergency care (CLARIPED) developed in Brazil.

Methods: Validity phase: prospective observational study with children aged 0 to 15 years who consecutively visited the pediatric emergency department (ED) of a tertiary hospital from July 2 to 18, 2013. We evaluated the association of urgency levels with clinical outcomes (resource utilization, ED admission rate, hospitalization rate, and ED length of stay); and compared the CLARIPED performance to a reference standard. Inter-rater reliability phase: a convenience sample of patients who visited the pediatric ED between April and July 2013 was consecutively and independently double triaged by two nurses, and the quadratic weighted kappa was estimated.

Results: In the validity phase, the distribution of urgency levels in 1,416 visits was the following: 0.0% red (emergency); 5.9% orange (high urgency); 40.5% yellow (urgency); 50.6% green (low urgency); and 3.0% blue (no urgency). The percentage of patients who used two or more resources decreased from the orange level to the yellow, green, and blue levels (81%, 49%, 22%, and 2%, respectively, $p < 0.0001$), as did the ED admission rate, ED length of stay, and hospitalization rate. The sensitivity to identify patients with high urgency level was 0.89 (confidence interval of 95% [95%CI] 0.78–0.95), and the undertriage rate was 7.4%. The inter-rater reliability in 191 patients classified by two nurses was substantial ($k_w^2 = 0.75$; 95%CI 0.74–0.79).

*Autor correspondente. E-mail: mariaclaramb@globo.com (M.C. Magalhães-Barbosa).

^aInstituto D'Or de Pesquisa e Ensino, Rio de Janeiro, RJ, Brasil.

^bUniversidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.

Recebido em 15 de abril de 2017; aprovado em 24 de setembro de 2017; disponível on-line em 27 de novembro de 2018.

Conclusões: O sistema CLARIPED mostrou boa validade e substancial confiabilidade para classificação de risco em um serviço de emergência pediátrica.

Palavras-chave: Serviços médicos de emergência; Pediatria; Triage; Medição de risco; Reprodutibilidade dos resultados.

Conclusions: The CLARIPED system showed good validity and substantial reliability for triage in a pediatric emergency department.

Keywords: Emergency medical services; Pediatrics; Triage; Risk assessment; Reproducibility of results.

INTRODUÇÃO

A triagem ou classificação de risco (CR) no serviço de emergência pediátrica (SEP) é um desafio. Limitada capacidade de comunicação, apresentações subclínicas em crianças jovens, variabilidade na normalidade dos sinais vitais (SV), entre outros fatores, tornam a CR em pediatria uma tarefa complexa e difícil.¹

Os instrumentos mais utilizados no mundo para CR em crianças são o Sistema de Triage de Manchester (*Manchester Triage System* — MTS), a versão pediátrica da Escala Canadense de Triage e Acuidade (*Canadian Pediatric Triage and Acuity Scale* — PedCTAS), o Índice de Gravidade em Emergência (*Emergency Severity Index* — ESI) e a Escala Australiana de Triage (*Australasian Triage Scale* — ATS).^{2,3} Esses instrumentos foram inicialmente idealizados para adultos e, posteriormente, adaptados para crianças, que representam de 20 a 40% da população atendida em serviços de emergência.⁴

A validade e a confiabilidade desses sistemas de CR para crianças têm sido avaliadas predominantemente nos países nos quais foram idealizados ou em países desenvolvidos, com culturas semelhantes.⁵⁻²² Tais instrumentos são muito extensos ou complexos, e a maioria dos estudos em países com características sociodemográficas e/ou culturais diferentes mostra desempenho inferior.²³⁻²⁶ Diferenças de recursos humanos e tecnológicos, qualificação de profissionais e políticas de saúde podem interferir no desempenho. Algoritmos mais simples, desde que válidos e confiáveis, podem ser mais apropriados para países como o Brasil.

A escala sul-africana de triagem (*South African Triage Scale* — SATS) é uma ferramenta simples e objetiva,²⁷ entretanto, só tem quatro níveis de urgência (NU) e três faixas etárias para avaliação de SV. A ferramenta recomendada pela Organização Mundial da Saúde (OMS) para países menos desenvolvidos, Avaliação de Triage e Tratamento na Emergência (*Emergency Triage Assessment and Treatment* — ETAT), prioriza a identificação de pacientes muito graves.²⁸ Essa característica não contempla 70 a 90% da população que afliu aos SEP públicos e privados brasileiros, que se encontram em níveis intermediários de urgência.

Com a finalidade de atender a essas demandas, uma equipe de especialistas em SEP no Brasil desenvolveu o sistema de classificação de risco para emergências pediátricas (CLARIPED),

uma ferramenta simples e objetiva, de fácil treinamento e utilização, com cinco NU e estratificação em cinco faixas etárias.²⁹ O propósito deste estudo foi avaliar a validade e a confiabilidade desse instrumento.

MÉTODO

Estudo observacional prospectivo realizado no SEP de um hospital terciário privado da cidade do Rio de Janeiro (Rio de Janeiro, Brasil), com 3 mil atendimentos por mês, equipe diária de 3 ou 4 médicos, 2 enfermeiras e 2 técnicas, além de residentes de pediatria em sistema de rodízio.

Para avaliar a validade do CLARIPED, todos os pacientes atendidos consecutivamente de 2 a 18 de julho de 2013 e que foram submetidos à classificação de risco foram incluídos. Os pacientes que não receberam classificação de risco foram excluídos. Os dados foram coletados a cada 24 horas dos boletins do dia anterior. Variáveis demográficas e clínicas (CR e evolução), como recursos diagnósticos/terapêuticos, tempo de permanência no SEP e destino, foram acrescentadas em formulários informatizados, e sua consistência foi avaliada.

Para avaliar a confiabilidade interobservadores, uma amostra de conveniência foi prospectivamente selecionada durante plantões diurnos (8 às 17h) entre abril e julho de 2013. Pacientes consecutivos e seus responsáveis foram convidados a participar de maneira voluntária, imediatamente após a CR realizada por uma das enfermeiras de triagem. Após concordância e assinatura do Termo de Consentimento Livre e Esclarecido e, quando pertinente, do Termo de Assentimento, tanto o paciente quanto o responsável eram encaminhados a outra sala, onde uma enfermeira da pesquisa realizava um novo procedimento completo de CR, com total desconhecimento sobre a primeira classificação. As enfermeiras de pesquisa faziam parte da equipe de triagem, tinham o mesmo nível de treinamento e participaram voluntariamente do estudo, nas horas extras de trabalho. Nessa fase, por razões éticas, foram excluídos pacientes que necessitavam de tratamento imediato, pela impossibilidade de submetê-los a dois processos de CR consecutivos.

O CLARIPED foi aplicado conforme anteriormente descrito.²⁹ Na primeira etapa, aferiram-se quatro sinais vitais, para cálculo do escore de sinais vitais em pediatria (escore VIPE),

de 0 a 12, ao qual se atribui classificação em cinco NU: 0=azul (sem urgência); 1 ou 2=verde (pouco urgente); 3 a 5=amarelo (urgente); 6 a 9=laranja (muito urgente); e 10 a 12=vermelho (emergência). Na segunda etapa, consultou-se uma tabela de discriminadores (sinais, sintomas ou queixas), também distribuídos em 5 NU. Caso se identificasse um discriminador correspondente a um NU mais elevado que o atribuído pelo escore VIPE, a classificação final corresponderia ao maior NU.²⁹

Na ausência de um padrão ouro para CR em SEP, utilizamos dois métodos para avaliação de validade:

1. a associação entre os NU atribuídos pelo CLARIPED e quatro desfechos clínicos (utilização de recursos diagnósticos/terapêuticos; taxa de admissão na sala de observação — SO/SEP); tempo de permanência no SEP e hospitalização), entendidos como equivalentes de urgência, de maneira semelhante a outros estudos;^{9-12,14,15} e
2. comparação da classificação dada pelo CLARIPED com a classificação atribuída por um padrão de referência.

O primeiro passo baseou-se na seguinte hipótese: se o CLARIPED discriminar bem os cinco NU, um gradiente decrescente na frequência dos desfechos deve ocorrer, do maior para o menor NU. O desfecho “recursos utilizados” incluía testes diagnósticos, procedimentos terapêuticos e pareceres de especialistas, segundo tabela previamente padronizada e adaptada de Gilboy et al.³⁰ Essa variável foi dicotomizada (<2 recursos e ≥2 recursos), à semelhança de outros estudos.^{12,14,15,19} A taxa de admissão na SO/SEP incluiu somente crianças que, após ocupação de um leito no SEP, receberam alta para o domicílio. O tempo de permanência no SEP foi computado a partir do início da consulta médica até a saída do SEP. Pacientes que evoluíram para internação hospitalar, mesmo que em outras instituições (transferências), foram computados na taxa de hospitalização.

Para a segunda etapa do estudo, o padrão de referência baseou-se em uma matriz desenvolvida por especialistas para estudar o MTS na população pediátrica¹⁹ e adaptada para o presente estudo. Essa matriz utiliza dados extraídos dos prontuários (sinais vitais muito alterados, condições clínicas ameaçadoras da vida, exames laboratoriais e de imagem, abordagem terapêutica e destino do paciente), isoladamente ou em várias combinações possíveis, para caracterizar de modo retrospectivo a urgência adequada do paciente, a ser comparada com a urgência atribuída anteriormente pelo sistema de triagem.

Para a validade, estimou-se uma amostra de 1.385 atendimentos, com base em dados da literatura, referentes ao tempo de permanência no SEP, que era o desfecho que demandava a maior amostra. Assumindo erro alfa=0,05 e erro beta=0,80, utilizou-se diferença de 71 minutos entre o nível 2/muito urgente e

o nível 3/urgente, relatado em estudo sobre o ESI-4 (309 minutos; intervalo de confiança de 95% — IC95% 257–361, desvio padrão — DP) 225,5 versus 238 minutos; IC95% 223–251, DP=112,8, respectivamente).¹⁵

Para a confiabilidade interobservadores, o cálculo amostral fundamentou-se em estudo piloto com 61 atendimentos, que gerou o Kappa ponderado quadrático (k_w^2) de 0,57 (IC95% 0,51–0,68). Para reduzir a amplitude do intervalo de confiança para 0,10, foram feitas simulações com tamanhos crescentes de amostras, com a mesma distribuição de concordâncias e discordâncias entre os NU observada no estudo piloto, resultando numa estimativa de amostra de 183 atendimentos.

Associações entre NU e desfechos foram avaliadas com o teste do qui-quadrado/Fisher para variáveis categóricas e o teste de Mann-Whitney/Kruskal-Wallis para variáveis contínuas. Regressões logísticas foram usadas para estimar razões de chance (*odds ratio* — OR) resultantes da associação de NU (variáveis independentes) com hospitalização e utilização de recursos (variáveis dependentes), após ajustes para potenciais fatores de confusão (idade, hora e dia da semana do atendimento). Taxas de supertriagem, subtriagem, sensibilidade e especificidade para diagnosticar os casos muito urgentes foram calculadas mediante a comparação entre o CLARIPED e o padrão de referência. Estratificação por faixa etária e categorias diagnósticas foi feita em caráter exploratório.

Para confiabilidade interobservadores, optou-se pelo k_w^2 , porque essa estimativa leva em conta o grau de discordância entre categorias, além de ser a mais utilizada em outros estudos.

Nível de significância de 0,05 e IC95% foram considerados para as análises. Utilizaram-se os softwares estatísticos Stata 12.0 (StataCorp, College Station, Texas, Estados Unidos) e R 2.15.3 (R Foundation, Viena, Áustria). O projeto foi aprovado pelo Comitê de Ética em Pesquisa (CEP) da instituição.

RESULTADOS

Na fase de validade, foram incluídos 1.416 atendimentos consecutivos (80,2% dos elegíveis) e excluídos 28 boletins extravaiados (1,6%), 12 casos de abandono do SEP antes da CR (0,7%) e 310 casos de ausência de CR (17,6%). Na fase de confiabilidade, participaram 179 pacientes (93,7% dos convidados); 9 recusaram-se a integrar o estudo; e 3 foram excluídos por ausência de um responsável legal (Figura 1).

Na amostra de validade, observou-se a seguinte distribuição: vermelho 0,0%; laranja 5,9%; amarelo 40,5%; verde 50,6%; e azul 3,0% (Tabela 1). A utilização de recursos foi avaliada em 1.415 atendimentos; a admissão na SO/EP e a hospitalização em 1.413; e o tempo de permanência no SEP em 1.090 (Tabela 2). Os 326 atendimentos com dados faltantes

para esse último desfecho apontaram distribuição de NU semelhante aos 1.416 incluídos ($p=0,2181$), sugerindo perda aleatória. As frequências de todos os desfechos diminuíram com a redução dos NU: ≥ 2 recursos (laranja 80,9%; amarelo 48,5%; verde 21,8%; e azul 2,4%; $p<0,0001$), admissão na SO/SEP (20,2; 4,7; 0,4; e 0,0%; $p=0,0005$), hospitalização (19,0; 2,6; 0,0 e 0,0%; $p=0,0005$) e tempo de permanência no SEP (209; 106; 47 e 27 minutos; $p<0,0001$) (Tabela 2).

Os pacientes de nível laranja exibiram maior chance de utilizar ≥ 2 recursos (OR=4,67; IC95% 2,61–8,34, ajustado para idade e hora de atendimento) e maior chance de hospitalização (OR=10,77; IC95% 4,85–23,910, ajustado para idade), em relação aos pacientes de nível amarelo.

Na comparação entre o CLARIPED e o padrão de referência, houve concordância absoluta em 33,5% dos casos, supertriagem em 59,1% e subtriagem em 7,4%. A maioria das discordâncias representou diferenças de nível de urgência acima (49,4%), principalmente nos níveis azul e verde, ou de nível de urgência abaixo (7,3%), sobretudo no nível amarelo (Tabela 3). Não houve diferenças entre as faixas etárias nas taxas de supertriagem ($p=0,20$ a 0,98) e subtriagem ($p=0,13$ a 0,52), quando comparadas às taxas gerais. As taxas de supertriagem foram menores nas doenças respiratórias baixas (29,6%; $p<0,01$) e maiores nas doenças respiratórias altas (67,1%; $p=0,002$) e doenças do ouvido (76,1%; $p=0,0002$). Nenhuma categoria diagnóstica mostrou taxa de subtriagem diferente da taxa geral de 7,4%.

A sensibilidade e a especificidade do CLARIPED para discriminar níveis muito urgentes foram 0,89 (IC95% 0,78–0,95) e 0,98 (IC95% 0,97–0,99), respectivamente. A estratificação dessas estimativas por categorias diagnósticas e faixa etária

ficou prejudicada pelo pequeno número de casos nos subgrupos (Tabela 4).

Na fase de confiabilidade, 15 enfermeiras com o mesmo nível de treinamento no CLARIPED participaram em pares do processo de dupla CR: 13 na primeira classificação e 2 na segunda. A mediana de idade das enfermeiras foi de 28 anos (intervalo interquartil — IQ: 26,0–29,5). Quatro enfermeiras tinham mais de 5 anos de experiência em SEP, incluindo as 2 enfermeiras da pesquisa, enquanto 11 tinham menos de 5 anos. A distribuição na primeira CR foi: laranja 7,3%; amarelo 39,1%; verde 41,9%; e azul 11,7%, e na segunda, laranja 7,3%; amarelo 41,9%; verde 34,6%; e azul 16,2% (Tabela 1). A concordância absoluta foi 68,7% e o κ^2 de 0,75 (IC95% 0,73–0,79) (Tabela 5).

DISCUSSÃO

Além de boa validade e confiabilidade, um sistema de CR ideal deve ser viável e efetivo. É fundamental assegurar adesão da equipe e um procedimento de curta duração. O CLARIPED apresentou boa validade, demonstrada por forte associação entre NU e desfechos evolutivos, além de substancial concordância interobservadores. A associação com desfechos foi comparável à observada em estudos semelhantes com outros sistemas de triagem. A chance de hospitalização no nível laranja foi quase 11 vezes maior que no amarelo; e maior que a estimativa obtida em estudo multicêntrico sobre o PedCTAS (OR=4,93; IC95% 2,95–8,25),¹⁰ entretanto, a taxa de hospitalização (2,2%) foi inferior às levantadas em outros estudos (5–10%),^{9,10,12,15,26,27} podendo refletir em diferenças nas populações ou nas políticas institucionais.

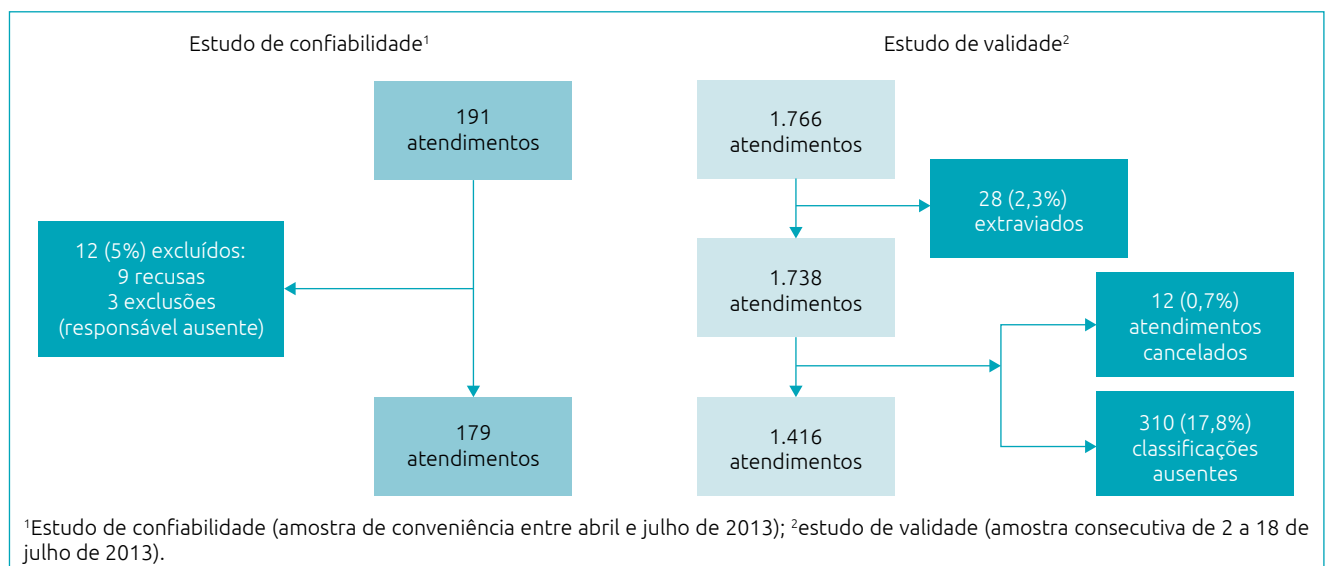


Figura 1 Algoritmo de seleção de pacientes para o estudo de confiabilidade e validade.

Tabela 1 Características basais da população geral e das amostras de validade e confiabilidade e classificação de risco das amostras de validade e confiabilidade.

Características	População total atendida entre abril e julho de 2013	Participantes da avaliação de validade*	Não participantes da avaliação de validade	Participantes da avaliação de confiabilidade**
Total	13.453	1.416	322	179
Masculino	6.917 (51,4)	719 (50,8)	199 (61,8)	94 (52,5)
Faixa etária (anos)				
<1	†6.336 (47,1)	145 (10,2)	41 (12,7)	17 (9,5)
1-4		723 (51,1)	148 (46,0)	84 (46,9)
5-11	5.304 (39,4)	459 (32,4)	95 (29,5)	59 (33,0)
12-15	1.674 (12,4)	89 (6,3)	38 (11,8)	19 (10,6)
NA	139 (1,0)	0	0	0
Horário de atendimento				
7-13h	4.408 (32,8)	491 (34,7)	57 (17,7)	159 (88,8)
13-19h	4.670 (34,7)	532 (37,6)	79 (25,5)	20 (11,2)
19-1h	3.391 (25,2)	314 (22,2)	145 (45,0)	0
1-7h	746 (5,5)	79 (5,6)	41 (12,7)	0
NA	238 (1,8)	0	0	0
Dias da semana				
Segunda-feira	2.207 (16,4)	203 (14,3)	35 (10,9)	21 (11,7)
Terça-feira	1.919 (14,3)	239 (16,9)	54 (16,8)	10 (5,6)
Quarta-feira	1.870 (13,9)	242 (17,1)	51 (15,8)	82 (45,8)
Quinta-feira	1.717 (12,8)	232 (16,4)	475 (14,6)	18 (10,1)
Sexta-feira	1.689 (12,6)	146 (10,3)	22 (7,1)	48 (26,8)
Sábado	1.993 (14,8)	170 (12,0)	63 (19,6)	0
Domingo	2.058 (15,3)	184 (13,0)	49 (15,2)	0
Principais diagnósticos finais				
Doença respiratória alta		466 (32,9)	82 (25,5)	61 (34,1)
Doença respiratória baixa		179 (12,6)	41 (12,7)	38 (21,2)
Gastrointestinal		179 (12,6)	23 (7,1)	16 (8,9)
Doenças do ouvido		134 (9,5)	37 (11,5)	19 (10,6)
Causas externas		148 (10,5)	44 (13,7)	2 (1,1)
Gerais e inespecíficos		113 (8,0)	26 (8,1)	8 (4,5)
Destino				
Alta após prescrição		776 (54,8)	164 (50,9)	115 (64,2)
Alta após medicação		559 (39,5)	99 (30,7)	58 (32,4)
Alta após sala de observação		47 (3,3)	17 (5,3)	6 (3,4)
Internação em quarto	‡279 (2,1)	£ 17 (1,2)	23 (7,1)	0
Internação em UTIP		£ 4 (0,3)	2 (0,6)	0
Transferência	39 (0,3)	£ 10 (0,7)	3 (0,9)	0
Evasão antes da consulta	129 (1,0)		-	
Alta à revelia pós-consulta	189 (1,2)		-	
NA	89 (0,7)	3 (0,2)	14 (4,3)	0
Classificação de risco				1ª / 2ª ^α
Vermelho		0		0/0
Laranja		84 (5,9)		13 (7,3)/13 (7,3)
Amarelo		573 (40,5)		70 (39,1)/75 (41,9)
Verde		717 (50,6)		75 (41,9)/62 (34,6)
Azul		42 (3,0)		21 (11,7)/29 (16,2)

*Amostra consecutiva; **amostra de conveniência; NA: nenhuma das alternativas; UTIP: unidade de terapia intensiva pediátrica; †agrupando os atendimentos de todas as crianças <5 anos; ‡agrupando todas as internações de quarto e UTIP=2,1%; £somando as internações em quarto (1,2%), UTIP (0,3%) e transferências (0,7%), a taxa de internação total =2,2%; 1ª e 2ª classificação de risco da avaliação de confiabilidade.

A utilização de recursos mostrou-se um desfecho mais apropriado para a população estudada. A frequência de pacientes que usaram ≥ 2 recursos diminuiu do maior para o menor NU (81,0; 48,5; 21,8; e 2,4%; $p < 0,0001$). Considerando que não houve pacientes classificados como vermelho, esses resultados assemelharam-se aos relatados com o ESI-4 (100, 70, 45, 17 e 4%)¹⁴ e são mais discriminativos do que os observados com o MTS (41,7; 25,4; 30,2; 16,6; e 3,7%).¹⁹ O nível laranja conferiu chance quase 5 vezes maior de utilizar ≥ 2 recursos, enquanto o verde apresentou chance 5 vezes menor, quando comparados ao nível amarelo. Essa associação também foi muito semelhante à relatada com o PedCTAS — OR (nível muito urgente)=4,67; IC95% 2,61–8,34 e OR (nível pouco urgente)=0,21; IC95% 0,17–0,28; sendo a referência nível urgente.¹⁰

O tempo de permanência no SEP foi computado por meio do atendimento médico, e não pela chegada no SEP, como na maioria dos estudos. A intenção foi evitar distorção na associação entre esse desfecho e os NU, uma vez que a CR determina que, quanto menos urgente for o paciente, maior o tempo de espera tolerado para o atendimento médico. A distribuição do tempo de permanência mostrou um gradiente decrescente do maior para o menor NU, corroborando a boa validade do instrumento (209; 106; 47; 27 minutos; $p < 0,0001$). Desconsiderando o nível 1 (vermelho), esse gradiente também foi mais discriminativo do que os gradientes evidenciados em dois estudos sobre o PedCTAS (191; 250; 191; 96; 66 minutos; $p < 0,0001$ e 309; 238; 186; 160 minutos)^{9,10} e em dois estudos sobre o ESI-4 (334; 221; 207; 151; 132 minutos; $p < 0,001$ e 156; 236; 259; 117; 99 minutos; $p < 0,0001$),^{12,14} entretanto, a diferença de definição desse desfecho nesses estudos pode ter contribuído para resultados menos consistentes.

Em que pese a diferença de tamanho amostral entre este estudo ($n=1.416$) e dois estudos com o MTS ($n=13.554$ e 11.260),^{20,22} a utilização de um padrão de referência semelhante nas três pesquisas permite alguma comparação no desempenho das ferramentas. O CLARIPED mostrou concordância absoluta similar ao MTS

Tabela 3 Classificações de risco atribuídas pelo padrão de referência *versus* classificação de risco para emergências pediátricas (CLARIPED): estudo de validade ($n=1.416$).

CLARIPED	Padrão de referência					
	Vermelho	Laranja	Amarelo	Verde	Azul	Total
Vermelho	0	0	0	0	0	0
Laranja	0	55	13	12	4	84
Amarelo	0	6	175	271	121	573
Verde	0	1	86	214	416	717
Azul	0	0	0	11	31	42
Total	0	62	274	508	572	1.416

■ Concordância absoluta
■ Subtriagem de 1 categoria
■ Supertriagem de 1 categoria
■ Subtriagem >1 categoria
■ Supertriagem >1 categoria

Concordância absoluta = 33,5% (475/1.416)	
Supertriagem total = 59,1% (837/1.416)	Subtriagem total = 7,3% (104/1.416)
Supertriagem (laranja) = 0,0%	Subtriagem (laranja) = 11,3% (7/62)
Supertriagem (amarelo) = 4,7% (13/274)	Subtriagem (amarelo) = 31,4% (86/274)
Supertriagem (verde) = 55,7% (283/508)	Subtriagem (verde) = 2,2% (11/508)
Supertriagem (azul) = 94,5% (541/1.416)	Subtriagem (azul) = 0

Tabela 2 Distribuição dos desfechos evolutivos para diferentes categorias de urgência: estudo de validade ($n=1.416$).

Classificação de risco	Abordagem terapêutica e destino				
	Total n (%)	≥ 2 recursos ¹ n (%)	Admissão na sala de observação ² n (%)	Internação hospitalar ³ n (%)	Média de tempo de permanência ⁴ min (IC95%)
Vermelho	0 (0,0)	–	–	–	–
Laranja	84 (5,9)	68 (81,0)	17 (20,2)	16 (19,0)	209 (160–258)
Amarelo	573 (40,5)	278 (48,5)	27 (4,7)	15 (2,6)	106 (93–119)
Verde	717 (50,6)	125 (21,8)	3 (0,4)	0	47 (43–51)
Azul	42 (3,0)	1 (2,4)	0 (0,0)	0	27 (19–35)
Total	1.416 (100,0)	472 (33,4)	47 (3,3)	31 (2,2)	
p-valor		* $< 0,0001$	**0,0005	**0,0005	*** $< 0,0001$

*Teste do qui-quadrado; **teste exato de Fisher; ***teste de Kruskal-Wallis; ¹total de 1.415 atendimentos avaliados; ²total de 1.413 atendimentos avaliados; ³total de 1.413 atendimentos avaliados; ⁴total de 1.090 atendimentos avaliados; IC95%: índice de confiança de 95%.

(33,5 *versus* 34,0%), maior taxa de supertriagem (59,1% *versus* 54,0%), maior sensibilidade (89,0 *versus* 63,0%) e especificidade (98,0 *versus* 79,0%) e menor taxa de subtriagem (7,4 *versus* 12,0%)²⁰. Após modificações em alguns discriminadores pediátricos do MTS, houve aumento na especificidade (87%) e diminuição na supertriagem (47%), sem modificação da sensibilidade (64%) e com discreto aumento na subtriagem (15%).²² Todavia, é preciso questionar se o padrão de referência confere o nível de urgência adequado em todos os casos. Por exemplo, um lactente que chega ao SEP chorando, com irritabilidade e dor intensa seria adequadamente classificado como laranja ou amarelo pelo MTS e pelo CLARIPED. Se o diagnóstico final for otite média aguda, tão comum em pediatria, o paciente receberá um analgésico no setor e alta com prescrição, sendo considerado verde pelo padrão de referência. Deveras, a taxa de supertriagem do CLARIPED foi particularmente elevada em doenças do ouvido. Da mesma forma, um paciente com uma ferida corto-contusa extensa, necessitando sutura, seria classificado como amarelo pelo CLARIPED e pelo MTS, mas seria verde pelo padrão de referência. Esses e outros exemplos semelhantes podem justificar a baixa concordância absoluta, assim como a elevada taxa de supertriagem das duas ferramentas comparadas ao padrão de referência. De fato, o padrão de referência não foi validado.

Para estimar a confiabilidade, este estudo incluiu somente pacientes tratados em tempo real, em vez de cenários hipotéticos,

comumente usados,^{5,12,16,17,21,27} que não reproduzem as dificuldades do processo de CR, sendo sujeitos a vieses. A confiabilidade interobservadores alcançada foi substancial ($kw^2=0,75$; IC95% 0,74–0,79). Esse resultado é melhor do que os obtidos nos primeiros estudos com pacientes reais com outros instrumentos: MTS ($kw^2=0,65$; IC95% 0,56–0,72)²⁰, PedCTAS ($kw^2=0,61$; IC95% 0,42–0,80)⁸ e ESI-4 ($kw^2=0,57$; IC95% 0,52–0,62).¹⁴ Estudos mais recentes mostraram melhor confiabilidade com o PedCTAS¹⁰ ($kw^2=0,74$;

Tabela 5 Classificações de risco atribuídas pela enfermeira da pesquisa *versus* enfermeira da triagem — estudo de confiabilidade (n=179).

Enfermeira da triagem	Enfermeira da pesquisa					
	Vermelho	Laranja	Amarelo	Verde	Azul	Total
Vermelho	0	0	0	0	0	0
Laranja	0	10	3	0	0	13
Amarelo	0	3	56	11	0	70
Verde	0	0	15	44	16	75
Azul	0	0	1	7	13	21
Total	0	13	75	62	29	179

Concordância absoluta =68,7%; $kw^2=0,75$ (IC95% 0,73–0,79)

kw^2 : Kappa ponderado quadrático; IC95%: intervalo de confiança de 95%.

Tabela 4 Sensibilidade e especificidade do CLARIPED para diagnosticar urgência elevada na população geral e nos subgrupos por faixa etária e principais categorias diagnósticas.

Subgrupo	Nº de pacientes	% de elevada urgência*		Sensibilidade†	Especificidade†
		CLARIPED	Referência		
Total	1.416	84 (5,9%)	62 (4,4%)	0,89 (0,78–0,95)	0,98 (0,97–0,99)
Faixa etária (anos)					
<1	145	9 (6,2%)	9 (6,2%)	0,89 (0,52–1,00)	0,99 (0,96–1,00)
1–4	723	31 (4,3%)	25 (3,5%)	0,84 (0,64–0,95)	0,99 (0,97–0,99)
5–11	459	34 (7,4%)	22 (4,8%)	0,91 (0,71–0,99)	0,97 (0,95–0,98)
12–15	89	10 (11,2%)	6 (6,7%)	1,00 (0,42–1,00)	0,95 (0,88–0,99)
Diagnóstico final					
Respiratório alto	466	5 (1,1%)	4 (1,0%)	0,75 (0,19–0,99)	1,00 (0,98–1,00)
Respiratório baixo	179	37 (20,7%)	31 (17,3%)	1,00 (0,84–1,00)	0,96 (0,92–0,99)
Gastrointestinal	179	17 (9,5%)	14 (7,8%)	0,79 (0,49–0,95)	0,96 (0,92–0,99)
Ouvido	134				
Causas externas	148	11 (7,4%)	2 (1,4%)	1,00 (0,09–1,00)	0,94 (0,89–0,97)
Generais e inespecíficos	113	4 (3,5%)	4 (3,5%)	0,50 (0,07–0,93)	0,98 (0,94–1,00)

*elevada urgência: categoria vermelha e laranja; sensibilidade†: casos de elevada urgência (vermelho e laranja) atribuída pelo CLARIPED/casos de elevada urgência atribuída pelo padrão de referência; especificidade†: casos de baixa urgência (amarelo, verde e azul) atribuída pelo CLARIPED/casos de baixa urgência atribuída pelo padrão de referência.

IC95% 0,71–0,76) e com o ESI-4^{15,26} (k [linear]=0,92; p<0,001 e k [não especificado]=0,82; IC95% 0,67–0,84). A melhoria na confiabilidade ao longo do tempo provavelmente reflete o aprimoramento dessas ferramentas e a qualificação das equipes. Nesse sentido, a confiabilidade alcançada nessa primeira avaliação do CLARIPED pode ser considerada promissora.

Este estudo apresenta algumas limitações. Ele foi realizado em um único centro, e a equipe de saúde pode ter sido altamente motivada pelos investigadores, resultando na superestimação da validade e confiabilidade do CLARIPED, no entanto, a fácil assimilação e execução da ferramenta sugerem que ela pode ser adequada em muitos ambientes semelhantes, incluindo serviços de urgência não hospitalares.

Outra limitação foi que os participantes da fase de validade representaram 80,2% dos pacientes elegíveis e a taxa de internação foi maior entre os não participantes (8,6 contra 2,2%) (Tabela 1). A razão mais plausível para essa diferença é que esse SEP recebe pacientes encaminhados para internação por seus pediatras assistentes. Essas crianças são diretamente encaminhadas à SO/SEP para início de tratamento, sem serem submetidas ao processo de CR, entretanto, os participantes do estudo não apresentaram diferenças em relação à população total do SEP (Tabela 1).

Mais uma limitação refere-se ao fato de não ter havido pacientes classificados como vermelho no período estudado, mas essa limitação não invalida os resultados alcançados nos outros quatro níveis de urgência, que constituem cerca de 99% dos atendimentos pediátricos. Dois estudos de validade sobre o PedCTAS também não incluíram pacientes necessitando cuidados imediatos.^{6,10} Esses pacientes são muito pouco frequentes na maioria dos SEP em todo o mundo^{7,9,11,23,25,26} e, na prática diária, não passam pela sala de triagem, sendo diretamente encaminhados à sala de reanimação e classificados retrospectivamente. Por outro lado, um dos principais desafios de uma ferramenta de CR é discriminar os pacientes de NU intermediários, como nível 3 (urgente) e nível 4 (pouco urgente), que compreendem a grande maioria dos que fluem aos SEP.

Pacientes de nível 3 são aqueles com potencial para agravar a sua condição, se esperarem muito tempo para o atendimento médico, mas que podem não ser facilmente identificados sem uma avaliação objetiva. São exemplos casos de dor abdominal cirúrgica (apendicite ou invaginação intestinal), casos com risco de desidratação severa (diarreia profusa ou vômitos incoercíveis) ou infecção bacteriana grave (febre alta em crianças pequenas).

Finalmente, o presente estudo utilizou desfechos evolutivos como equivalentes de urgência para determinar a validade de constructo convergente, no entanto o objetivo dos sistemas de CR não é prever desfechos clínicos, que são bons marcadores de complexidade e gravidade, mas nem sempre refletem todas as situações de urgência, além de serem influenciados pela qualidade do tratamento e por políticas institucionais. Por exemplo, um paciente com convulsão (vermelho) ou um paciente com crise de asma (laranja) pode ter alta da SO/SEP algumas horas após o tratamento, sem necessidade de internação. Por outro lado, um paciente com uma doença crônica grave pode se apresentar ao SEP com uma intercorrência de baixa gravidade (verde) e ser internado em razão da doença de base.

Em conclusão, este é o primeiro estudo sobre validade e confiabilidade de um sistema de classificação de risco pediátrico no Brasil. O CLARIPED mostrou-se um instrumento válido e confiável no centro em que foi desenvolvido. Um estudo multicêntrico faz-se necessário para corroborar esses achados preliminares, apontar a necessidade de adaptações para diferentes contextos de saúde e garantir a validade externa do instrumento.

Financiamento

- Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brasil, processo nº 448855/2014-3
- Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), Brasil, processo nº E-26/010.000196/2015

Conflito de interesses

Os autores declaram não haver conflito de interesses.

REFERÊNCIAS

1. Hohenhaus SM, Travers D, Mecham N. Pediatric triage: a review of emergency education literature. *J Emerg Nurs*. 2008;34:308-13.
2. Christ M, Grossmann F, Winter D, Bingisser R, Platz E. Modern triage in the emergency department. *Dtsch Arztebl Int*. 2010;107:892-8.
3. Farrokhnia N, Castrén M, Ehrenberg A, Lind L, Oredsson S, Jonsson H, 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.
4. Tang N, Stein J, Hsia RY, Maselli JH, Gonzales R. Trends and characteristics of US emergency department visits, 1997-2007. *JAMA*. 2010;304:664-70.
5. 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;20:514-8.

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;23:243-7.
7. Ma W, Gafni A, Goldman RD. Correlation of the Canadian Paediatric Emergency Triage and Acuity Scale to ED resource utilization. *Am J Emerg Med.* 2008;26:893-7.
8. 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;15:1262-7.
9. 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.
10. Gravel J, Gouin S, Goldman RD, Osmond MH, Fitzpatrick E, Boutis K, et al. The Canadian Triage and Acuity Scale for children: a prospective multicenter evaluation. *Ann Emerg Med.* 2012;60:71-7.
11. Gravel J, Fitzpatrick E, Gouin S, Millar K, Curtis S, Joubert G, et al. Performance of the Canadian Triage and Acuity Scale for children: a multicenter database study. *Ann Emerg Med.* 2013;61:27-32.
12. Baumann MR, Strout TD. Evaluation of the Emergency Severity Index (version 3) triage algorithm in paediatric patients. *Acad Emerg Med.* 2005;12:219-24.
13. Durani Y, Brecher D, Walmsley D, Attia MW, Loiselle JM. The Emergency Severity Index Version 4: reliability in pediatric patients. *Pediatr Emerg Care.* 2009;25:751-3.
14. Travers DA, Waller AE, Katznelson J, Agans R. Reliability and validity of the emergency severity index for pediatric triage. *Acad Emerg Med.* 2009;16:843-9.
15. Green NA, Durani Y, Brecher D, DePiero A, Loiselle J, Attia M. Emergency Severity Index version 4: a valid and reliable tool in paediatric emergency department triage. *Pediatr Emerg Care.* 2012;28:753-7.
16. Considine J, LeVasseur SA, Villanueva E. The Australasian Triage Scale: examining emergency department nurses' performance using computer and paper scenarios. *Ann Emerg Med.* 2004;44:516-23.
17. Crellin DJ, Johnston L. Poor agreement in application of the Australasian Triage Scale to paediatric emergency department presentations. *Contemp Nurse.* 2003;15:48-60.
18. Durojaiye L, O'Meara M. A study of triage of paediatric patients in Australia. *Emerg Med (Fremantle).* 2002;14:67-76.
19. Roukema J, Steyerberg EW, Meurs A, Ruige M, Lei J, Moll HA. Validity of the Manchester Triage System in paediatric emergency care. *Emerg Med J.* 2006;23:906-10.
20. Veen M, Steyerberg EW, Ruige M, Meurs AH, Roukema J, Lei J, et al. Manchester triage system in paediatric emergency care: prospective observational study. *BMJ.* 2008;337:a1501.
21. Veen M, Walle VF, Steyerberg EW, Meurs AH, Ruige M, Strout TD, et al. Repeatability of the Manchester Triage System for children. *Emerg Med J.* 2010;27:512-6.
22. Veen M, Steyerberg EW, Klooster MV, Ruige M, Meurs AH, Lei JV, et al. The Manchester triage system: improvements for paediatric emergency care. *Emerg Med J.* 2012;29:654-9.
23. Landaluce AF, Pijoan JI, Mingegi S, Benito FJ. Evaluación de la escala canadiense de triaje pediátrico en un servicio de urgencias de pediatría europeo. *Emergencias.* 2010;22:355-60.
24. Chang YC, Ng CJ, Wu CT, Chen LC, Chen JC, Hsu KH. 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.
25. Al-Hindi AA, Al-Akhfash AA, Fareed AM, Alhusainan KS, Algasomy SF, Althowainy IR. Efficacy of implementation of a 5 scale pediatric triage and acuity scale in pediatric emergency, Saudi Arabia. *Saudi Med J.* 2014;35:999-1004.
26. Jafari-Rouhi AH, Sardashti S, Taghizadieh A, Soleimanpour, Barzegar M. 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.
27. Twomey M, Cheema B, Buys H, Cohen K, Sa A, Louw P, 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.
28. Robertson MA, Molyneux EM. Description of cause of serious illness and outcome in patients identified using ETAT guidelines in urban Malawi. *Arch Dis Child.* 2001;85:214-7.
29. Magalhães-Barbosa MC, Prata-Barbosa A, Cunha AJ, Lopes CS. CLARIPED: a new tool for risk classification in pediatric emergencies. *Rev Paul Pediatr.* 2016;34:254-62.