Received: 13 January 2020 | Revised: 27 October 2021 | Accepted: 10 November 2021

DOI: 10.1002/emp2.12613

JACEP OPEN WILEY

ORIGINAL RESEARCH

Disaster

Comparison of prehospital professional accuracy, speed, and interrater reliability of six pediatric triage algorithms

Tabitha Cheng MD[1] | Katherine Staats MD[1] | Amy H. Kaji MD PhD[2] | Nicole D'Arcy MD[3] | Kian Niknam BS[4] | J. Joelle Donofrio-Odmann DO[5]

[1] Department of Emergency Medicine, UCSD, San Diego, USA

[2] Department of Emergency Medicine, Harbor-UCLA Medical Center, Torrance, USA

[3] Department of Emergency Medicine, Santa Clara Valley Medical Center, San Jose, USA

[4] Department of Emergency Medicine, Stanford University Medical Center, Palo Alto, USA

[5] Departments of Emergency Medicine and Pediatrics, UCSD; Rady Children's Hospital of San Diego, San Diego, USA

Correspondence
Tabitha Cheng, MD, Department of Emergency Medicine, UCSD, San Diego, USA.
Email: tabitha.cheng@gmail.com

Funding and support: By *JACEP Open* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see www.icmje.org). The authors have stated that no such relationships exist.

Abstract

Objectives: We evaluated prehospital professionals' accuracy, speed, interrater reliability, and impression in a pediatric disaster scenario both without a tool ("No Algorithm"–NA) and with 1 of 5 algorithms: CareFlight (CF), Simple Triage and Rapid Treatment (START) and JumpSTART (J-START), Pediatric Triage Tape (PTT), Sort, Assess, Life-saving interventions, Treatment/Transport (SALT), and Sacco Triage Method (STM).

Methods: Prehospital professionals received disaster lectures, focusing on 1 triage algorithm. Then they completed a timed tabletop disaster exercise with 25 pediatric victims to measure speed. A predetermined criterion standard was used to assess accuracy of answers. Answers were compared to one another to determine the interrater reliability.

Results: One hundred and seven prehospital professionals participated, with 15–28 prehospital professionals in each group. The accuracy was highest for STM (89.3%; 95% confidence interval [CI] 85.7% to 92.2%) and lowest for PTT (67.8%; 95% CI 63.4% to 72.1%). Accuracy of NA and SALT tended toward undertriage (15.8% and 16.3%, respectively). The remaining algorithms tended to overtriage, with PTT having the highest overtriage percentage (25.8%). The 3 fastest algorithms were: CF, SALT, and NA, all taking 5 minutes or less. STM was the slowest. STM demonstrated the highest interrater reliability, whereas CF and SALT demonstrated the lowest interrater reliability.

Conclusions: This study demonstrates the most common challenges inherent to mass casualty incident (MCI) triage systems: as accuracy and prehospital professional interrater reliability improve, speed slows. No triage algorithm in our study excelled in all these measures. Additional investigation of these algorithms in larger MCI drills requiring collection of vital signs in real time or during a real MCI event is needed.

KEYWORDS
CareFlight, children, disaster, emergency medical services, JumpSTART, mass casualty triage, pediatric triage tape, prehospital, sacco triage method, SALT

Supervising Editor: Matthew Hansen, MD, MCR

1 | INTRODUCTION

1.1 | Background

Children comprise about 23% of the US population, and they are affected directly or indirectly by disasters; however, many disaster plans lack attention to pediatric specific issues.[1–4] Compared to adults, children are more vulnerable during mass casualty incidents (MCI), likely because of anatomic, physiologic, and developmental differences.[2]

Triage algorithms have been developed to address the differences in adult and pediatric physiology. Currently, there are 5 common triage tools that apply to pediatric patients in disaster settings: CareFlight (CF), Simple Triage and Rapid Treatment and JumpSTART (J-START), Pediatric Triage Tape (PTT), Sort-Assess-Lifesaving interventions-Treatment/transport (SALT), and Sacco Triage Method (STM).[5–7] A number of differences exist between these algorithms (Table 1). Some are adaptations of adult triage tools with pediatric-specific components. Some require more subjectivity and knowledge of the impact of the disaster on the local system, and others use more objective criteria, requiring more time to categorize patients but increasing overall interrater reliability.[6,8–13]

1.2 | Importance

General principles for the best triage algorithms balance accuracy, interrater reliability, speed, and ease of use. Accuracy ensures different professionals using the algorithm on the same patient calculate the same designation. Speed allows the highest number of patients to be sorted in the shortest amount of time. Ease of use affects the application of other principles and can account for how higher-level thinking skills degrade in high-stress situations like an MCI. Over- and undertriage of patients are other important algorithm considerations, the former being costly for resources and the latter being dangerous for patients. Although many studies have examined the accuracy of pediatric MCI triage tools in assigning appropriate triage categories, there are few studies that have assessed the accuracy of a prehospital professional to correctly use the triage algorithm on pediatric patients, none of which have compared more than 2 algorithms.[14–19]

1.3 | Goals of this investigation

This study was intended to evaluate prehospital professionals' accuracy, interrater reliability, speed, and impression in a pediatric disaster tabletop exercise both without a tool ("No Algorithm"–NA) and with 1 of 5 triage algorithms: CareFlight (CF), Simple Triage and Rapid Treatment (START) and JumpSTART (J-START), Pediatric Triage Tape (PTT), Sort, Assess, Life-saving interventions, Treatment/Transport (SALT), and Sacco Triage Method (STM).

The Bottom Line

In this article Dr. Cheng et al performed a comparative evaluation of various disaster triage tools, including best clinical judgment among paramedics. They found that there was no "ideal" tool and there were inherent tradeoffs between speed and accuracy among the tools tested. As accuracy of the tools improved, the speed decreased. This study was conducted using tabletop simulations where some information, such as vital signs, was already provided to the participants and future studies should utilize more realistic simulations where all the information needs to be collected by participants.

2 | METHODS

2.1 | Study design

This was an interventional study involving prehospital professionals available during normally scheduled 3-hour continuing education (CE) sessions. Groups of prehospital professionals were scheduled in the morning or afternoon sessions over a 3-day training period at 2 locations. The prehospital professionals were divided into 6 groups based on convenience sampling by date and time of their normally scheduled CE training. Before the study, each of the 6 algorithms being studied was randomly assigned to 1 of the 6 sessions. A sample size of 15 prehospital professionals per group, total of 90 prehospital professionals, was calculated for a power of > 90% to detect a difference between the 6 groups. After consent, all eligible prehospital professionals were enrolled into the study. CE credits and educational experience were provided regardless of whether they participated in the tabletop portion of the study. University of California, San Diego human research protections program institutional review board approved this study before implementation.

2.2 | Study population

Prehospital professionals were recruited from emergency medical services (EMS) agencies within California Service Area 17, which includes the cities of Del Mar, Encinitas, Solana Beach, and Rancho Santa Fe in the county of San Diego. At the time of the study, all agencies used J-START as their MCI triage algorithm. Trainees were not excluded from participation in the study.

2.3 | Study flow

Figure 1 includes a flow diagram describing the 6 sessions of the study. All 6 groups of prehospital professionals were provided an educational

TABLE 1 Comparison of pediatric mass casualty incident triage tools used by the prehospital professionals during the tabletop scenario[5–13,18,19,24,31,32]

Triage tool	Triage designations	Category descriptions	Vital signs counted	Interventions	Geographic use	Time to triage (prior study estimates)	Benefits/challenges
CareFlight (CF)	Delayed: Green Urgent: Yellow Immediate: Red Unsalvageable: Black	Green–Walks Yellow–Obeys commands + palpable radial pulse Red–Obeys commands + non-palpable radial pulse OR does not obey commands + breathes with open airway Black–Does not obey commands or breathe after opening airway	None	Open airway	Australia Indonesia	15 seconds (No studies, but algorithm states)	Benefits: – Relies on qualitative data, no measurements Challenges: – No evidence showing superiority over other triage systems
JumpSTART (J-START)	Minor: Green Delayed: Yellow Immediate: Red Deceased: Black	Green–Walks or non-ambulatory at baseline and no yellow or red criteria Yellow–Respiratory rate 15-45 bpm + palpable pulse + age-appropriate mental status Red - Respiratory rate < 15 or > 45 bpm OR no peripheral pulse OR age-inappropriate mental status OR Breathing only after airway manuever or assisted respirations Black- Does not obey commands or breathe with open airway	Respirations	Open airway No breathing but radial pulse, give 5 rescue breaths	North America	26–70 s	Benefits: – Most commonly used algorithm today for pediatrics in the United States – Emphasizes an apneic child is more likely to have a primary respiratory problem Challenges: – Not based on evidence – A lot of time goes into counting respirations – No evidence showing superiority over other triage systems
Pediatric Triage Tape (PTT)	Delayed: Green Urgent: Yellow Immediate: Red Dead: Black	Green–Alert and moving all limbs OR walking Yellow–Breathing + Respiratory rate within range for height + Capillary refill < 2 s OR Pulse rate within range for height Red–Breathing only after airway manuever OR Respiratory rate out of range for height OR Pulse rate out of range for height Black–Not breathing after opening airway	Respirations Pulse rate	Open airway with jaw thrust	United Kingdom Australia India South Africa	No studies on PTT alone. One study including PTT and adult triage sieve estimated 30 s per patient.	Benefits: – Tape with triage sieve on it for reference – Used widely in the United Kingdom Challenges: – Relies on counting vital signs – Physical tape needed for triage – Height does not always predict physiological parameters – No evidence showing superiority over other triage systems

(Continues)

TABLE 1 (Continued)

Triage tool	Triage designations	Category descriptions	Vital signs counted	Interventions	Geographic use	Time to triage (prior study estimates)	Benefits/challenges
Sort Assess Life-Saving Treatment (SALT)	Minimal: Green Delayed: Yellow Immediate: Red Expectant/Dead: Black	Green–Breathing, obeys commands or has purposeful movement, has peripheral pulses, no respiratory distress, and minor injuries only Yellow–Breathing, obeys commands or has purposeful movement, has peripheral pulses, no respiratory distress, major hemorrhage controlled, and has major injuries Red–Breathing spontaneously or after opening airway or rescue breaths; does not obey commands or have purposeful movement, OR does not have peripheral pulses, OR has respiratory distress, OR hase major hemorrhage uncontrolled, AND is likely to survive given current resources Black–Not breathing after opening airway or 2 rescue breaths (Dead) OR immediate category but unlikely to survive given current resources (Expectant)	None	Control major hemorrhage Open airway (in children consider 2 rescue breaths) Chest decompression Autoinjector antidotes	United States	28–34 seconds	Benefits: – Few steps – No counting of vital signs – Developed using beneficial aspects of existing algorithms to standardize MCI triage in the United States Challenges: – Multiple aspects to evaluate in some steps – Subjective to experience – Tends to undertriage patients – Time to triage longer in 1 study – No evidence showing superiority over other triage systems
Sacco Triage Method (STM)	11–14: Green 9–10: Yellow 2–8: Red 0–1: Black Scores based on: – Respiratory rate (R) – Pulse (P) – Motor response (M)	Black 0–1: Likely Expectant Red 2–4: Critical, very low survival probability Red 5–8: Compromised/salvageable, deterioration without definitive care Yellow 9–10: Delayed/slow, high survival probability, little deterioration expected in the first 60 min Green 11–14: Likely minor, high survival probability	Respirations Pulse Motor score	Open airway Decompress pneumothorax Stop bleeding	North America	45–71 s	Benefits: – Evidence-based triage method – Only tool based on outcome data Challenges: – Proprietary tool – Designed to rely on computer system – Complicated calculations
"No Algorithm" (NA)	**Green** **Yellow** **Red** **Black**	Green–May need medical attention in a few days. No poor outcome without treatment. Yellow–Needs medical attention, but treatment can be delayed a few hours Red–Needs medical attention within 1–2 hours or will perish Black–Unlikely to survive with resources available	None	Based on "gestalt"	None	No studies	Benefits: – "Gestalt" based – No counting of vital signs Challenges: – No objective criteria to follow – Few studies have evaluated this method

Abbreviation: MCI, mass casualty incident.

JACEP OPEN WILEY

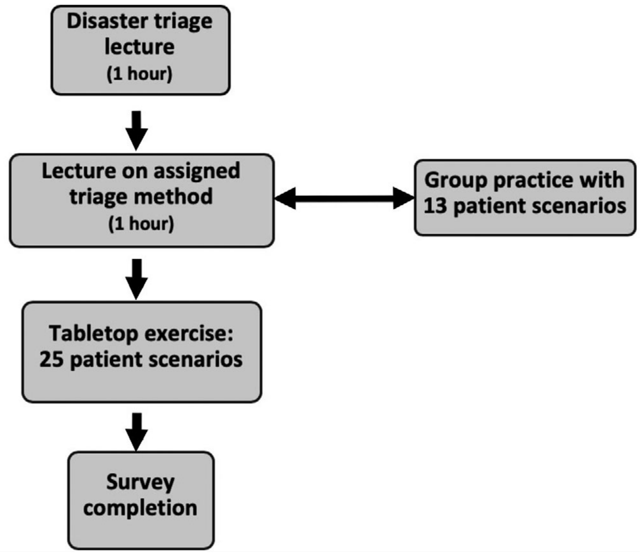

FIGURE 1 Flow diagram

1-hour lecture on disaster triage, followed by a 1-hour lecture focusing on the triage method assigned based on the date and time of their normally scheduled CE training. They were provided 13 practice patients as a group before testing, with feedback after each practice patient. Each prehospital professional was provided a reference sheet containing the designated triage algorithm for use in the tabletop exercise. For the PTT algorithm the physical tape was not provided, but a reference sheet with each of the height-based sieves (or tiers) was provided. For NA, a gestalt-based method was taught during the 1-hour educational session, teaching the prehospital professionals to triage pediatric patients based on their status and anticipated time-sensitive interventions needed for survival. For this training, the criteria outcomes tool from a paper by Donofrio et al. was primarily used to help the prehospital professionals understand in which triage categories (Black, Red, Yellow, Green) to designate the patients.[20] During the tabletop exercise for NA, the professionals had ready access to reference the definitions for the 4 triage categories. NA, as a gestalt and experience-centric algorithm, was an intentional addition to the study based on research showing "gut feeling" is often a driving force during triage designations in a disaster, although few studies have evaluated its efficacy.[21]

A precollected de-identified data set, from a prior study of pediatric disaster triage was used to develop the 25 patient case descriptors. The data set contained singleton patients under 14 years of age brought in by ambulance as a trauma activation to a level I trauma center. Field data (including prehospital narratives and treatments), and precompleted outcomes for each triage algorithm (performed by an EMS and pediatric emergency medicine fellowship-trained physician), as defined by the Criteria Outcomes Tool (COT) was included in the data set.[20] Each case was coded and the field description of the patient along with the original field vital signs were used for the tabletop scenario of a bus crash.

For the tabletop exercise, each of the 6 groups of prehospital professionals were placed in a large classroom setting with individual desks. On each desk was a stopwatch and a large sheet of paper with 25 patient scenarios represented with boxes. Following a shared start time, the participants each began their own stopwatches and read the form. Each box contained the patient scenario with a description of the patient (Example: 13-year-old male, facial trauma, unresponsive, lying in street. 50 kg. No respirations, faint palpable pulse at 40 bpm. No change with rescue breaths. No movement with stimulation). Age, chief complaint, pertinent physical exam findings, limited vital signs, and interventions with the patient's response were included. The prehospital professionals wrote their triage level designation to the right of the patient description. Although the prehospital professionals were asked not to change responses once the patient was triaged, this was not actively monitored during the scenario. Following the tabletop scenario, the prehospital professionals stopped their stopwatch and completed a survey on personal demographics, previous disaster training or experience, and algorithm feedback. Prehospital professionals were instructed not to discuss the list of patients or the disaster scenario to minimize communication between prehospital professionals with those who had not completed the drill.

The demographic data collected included gender, age, years of EMS experience, level of certification, and previous disaster medicine training. The prehospital professionals were provided a survey with Likert scales from 1 (low) to 5 (high) to rate their impressions of ease of understanding, ease of performing, and ease of remembering, as well as usefulness in MCIs with 10, 50, 100, and 1000 patients for their assigned triage algorithm. Each triage algorithm, including NA, was presented and tested in the same manner. Each group was given the same 25 patients to triage with their designated triage method to allow direct comparison between the algorithms.

2.4 | Data analysis

Descriptive statistics summarized demographic variables, accuracy (%), over- and undertriage (%), interrater reliability, and time to triage 25 patients for the 6 algorithms. The triage designations for CF, J-START, PTT, SALT, and STM triage were compared with the precompleted provided triage designation to assess prehospital professional accuracy in the use of the triage algorithm. Accuracy was defined as having precise matching triage designations (ie, prehospital professional triage of immediate matches the precompleted triage designation of immediate). The COT was used to assess accuracy for the NA group.[20] Given that the NA group, taught as a gestalt-based triage method, is not previously studied in comparison to other pediatric triage algorithms, we analyzed this group as a triage algorithm group and not as a control, or gold standard. Accuracy was defined by the number of correct triage designations by the prehospital professional out of the 25 patient scenarios. The accuracy (in percentile) of each triage algorithm was calculated with its associated 95% confidence interval (CI). The overtriage and undertriage were calculated based on the comparison between the prehospital professional's triage designation and the correct triage designation for each of the 25 patients (ie, the prehospital professional undertriaged the patient using SALT

TABLE 2 Characteristics of prehospital professional participants for all 6 algorithm groups before participation in the tabletop scenarios

	CF n = 28	J-START n = 15	PTT n = 18	SALT n = 15	STM n = 15	NA n = 16
Males	28	15	16*	15	15	16
Age: Median (IQR)	38 (30.5–46.5)	37 (35.0–41.0)	39.5 (34.0–46.5)	37 (33.0–48.0)	40 (37.0–52.0)	37.5 (34.0–47.0)
Years in EMS: Median (IQR)	14.5 (10.0–28.0)	15 (10.0–15.0)	17 (11.0–22.0)	11 (10.0–17.0)	15 (8.0–22.0)	14.5 (12.5–21.0)
Level of certification	20 (71)	15 (100)	15 (83)	12 (80)	11 (73)	14 (88)
Paramedic (% of group)	20 (71)	15 (100)	15 (83)	12 (80)	11 (73)	14 (88)
AEMT	1	0	1	1	1	1
EMT-B	7	0	2	2	3	1
Disaster medicine training (%)	5 (17.9)	2 (13.3)	8 (44.4)	0 (0.0)	0 (0.0)	3 (18.8)

* 2 declined to answer.
Abbreviations: AEMT, advanced emergency medical technician; CF, CareFlight; EMS, emergency medical services; EMT-B, emergency medical technician-basic; IQR, interquartile range; J-START, JumpSTART (Simple Triage and Rapid Treatment); NA, no algorithm; PTT, Pediatric Triage Tape; SALT, Sort, Assess, Life-saving interventions, Treatment/Transport; STM, Sacco Triage Method.

TABLE 3 Prehospital professional accuracy and over- and undertriage percentages by algorithm with 95% confidence intervals

	% Undertriaged				% Overtriaged		
Algorithm	−2	−1	Total	% Accuracy (95% CI)	+1	+2	Total
CF	0.3	10.0	10.3	74.0 (70.5–77.2)	14.6	1.1	15.7
J-START	0.0	8.6	8.6	80.3 (75.8–84.2)	10.8	0.3	11.1
PTT	0.4	6.0	6.4	67.8 (63.4–72.1)	23.3	2.4	25.8
SALT	0.8	15.5	16.3	70.4 (65.5–75.0)	13.1	0.3	13.3
STM	0.0	2.1	2.1	89.3 (85.7–92.2)	7.5	1.1	8.6
NA	0.8	15.0	15.8	77.3 (72.8–81.3)	6.8	0.3	7.0

Abbreviations: CF, CareFlight; CI, confidence interval; J-START, JumpSTART (Simple Triage and Rapid Treatment); NA, no algorithm; PTT, Pediatric Triage Tape; SALT, Sort, Assess, Life-saving interventions, Treatment/Transport; STM, Sacco Triage Method.

because delayed was selected when the correct designation was immediate). Medians and interquartile ranges were calculated to describe reliability among prehospital professionals. Time to triage 25 patients was reported as medians and range, with a calculated median time per patient to allow for comparison to previous studies. When comparing Likert score ratings for prehospital professional impressions of the algorithms, medians, and interquartile ranges were described, and the Kruskal-Wallis test was used to test for a statistical difference among the 6 methods.

Data were analyzed using Microsoft Excel 2016 v.16.16.4 (Redmond, WA), and SAS v.9.3 (SAS Institute, Cary, NC) using DBMS/Copy (Dataflux Corporation, Cary, NC) for analysis. GraphPad Prism v8.0.1 (GraphPad software Inc, La Jolla, CA) was used to construct the boxplot for interrater reliability.

3 | RESULTS

One hundred seven prehospital professionals were present for the CE sessions, and 100% participated in the study, ranging from 15 to 28 prehospital professionals in each algorithm group. All 107 prehospital professionals were present for their respective triage method lecture and the tabletop simulation scenario. Eight prehospital professionals participated twice because of the study's schedule and their 2 algorithm participations were included as separate events in 2 separate triage algorithm groups. Separate analysis was done excluding these repeaters and found minimal changes in the data, none of which affected the order of results. Table 2 shows the participant characteristics, including sex, age, years of EMS experience, level of certification, and previous disaster training. Sex, age, years of EMS experience, and level of certification were similar across all 6 algorithm groups.

Table 3 compares prehospital professional accuracy of triage algorithm application. Prehospital professionals were able to apply the STM algorithm group most accurately (89.3%; 95% CI 85.7% to 92.2%), whereas the PTT algorithm group had the lowest accuracy in using the algorithm (67.8%; 95% CI 63.4% to 72.1%). NA and SALT tended toward under- rather than overtriage (NA 15.8% vs 7.0%, SALT 16.3% vs 13.3%). The remaining algorithms tended towards over- rather than undertriage (CF 15.7% vs 10.3%, J-START 11.1% vs 8.6%, STM 8.6% vs 2.1%), with PTT having the highest overtriage at 25.8%.

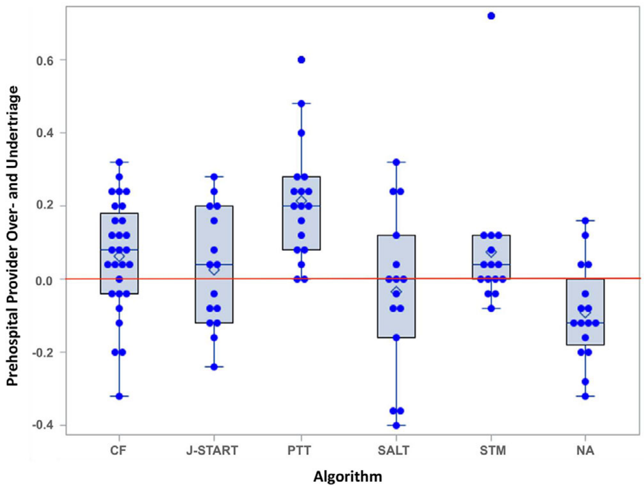

FIGURE 2 Prehospital professional interrater reliability by algorithm with each data point representing a prehospital professional's tendency to over- or undertriage. The boxplots represent the minimum and maximum range excluding outliers (whiskers), the interquartile range (box), the sample median (line), and the sample mean (diamond)
Abbreviations: CF, CareFlight; J-START, JumpSTART (Simple Triage and Rapid Treatment); NA, no algorithm; PTT, Pediatric Triage Tape; SALT, Sort, Assess, Life-saving interventions, Treatment/Transport; STM, Sacco Triage Method

TABLE 4 Prehospital professional median time to triage of 25 patients with range in minutes and seconds with calculated median time to triage per patient in seconds by algorithm

Algorithm	Median time to triage of 25 patients (min:sec, range)		Calculated median time to triage per patient (seconds)
CF	04:50	(2:23–6:40)	11.6
J-START	05:39	(3:06–8:24)	13.56
PTT	07:43	(4:53–20:11)	18.52
SALT	04:46	(3:35–6:41)	11.44
STM	10:53	(6:01–14:48)	26.12
NA	05:00	(3:58–6:22)	12

Abbreviations: CF, CareFlight; J-START, JumpSTART (Simple Triage and Rapid Treatment); NA, no algorithm; PTT, Pediatric Triage Tape; SALT, Sort, Assess, Life-saving interventions, Treatment/Transport; STM, Sacco Triage Method.

Prehospital professional over- or undertriage by more than 1 triage designation was low, with PTT having the highest percentage of overtriage by 2 triage designations (2.4%). Prehospital professional interrater reliability by algorithm is shown in Figure 2. STM demonstrated the highest interrater reliability, and CF and SALT demonstrated the greatest variability regarding professional scoring consistency.

The speed for the different algorithms varied tremendously (Table 4). The 3 fastest algorithms to perform were CF, SALT and NA, all with times of 5 minutes or less to triage 25 patients, an average of 12 seconds or less per patient. STM's median time to triage 25 patients was more than double these times and greater than 3 minutes longer than the nearest algorithm median time. PTT had 3 outliers with increased time to triage 25 patients; no other algorithm had this variation. NA had a narrow range of time to triage 25 patients (3:58–6:22), whereas J-START, PTT, and STM had the widest ranges (3:06–8:24; 4:53–20:11; 6:01–14:48; respectively).

Table 5 compares the median Likert score ratings for the prehospital professional impression of ease of understanding, performing, and remembering, as well as usefulness in MCIs with 10, 50, 100, and 1000 patients for each of the 6 triage algorithms. CF and NA, the algorithms that did not require vital signs calculations and had the least number of steps, had the highest scores in all domains. STM was rated lowest for ease of remembering and usefulness in MCIs with 10–1000 patients.

4 | LIMITATIONS

This study has several limitations. Overall, the study numbers of prehospital professionals and patient scenarios were small. It is possible that a larger number of prehospital professionals or number of patients to triage would have demonstrated differences in accuracy or impression for each algorithm. The prehospital professionals were selected for each group based on a convenience sample, given their continuing education schedule; therefore, there was a varying number of prehospital professionals in each group (from 15 prehospital professionals in J-START, SALT, and STM to 28 prehospital professionals in CF). Further, although all groups had similar demographics, considerably more subjects in the PTT group (8/18, 44%) had prior disaster training, whereas the other groups had little to no prehospital professionals with prior disaster training. It is worth considering whether a PTT group with less disaster training would have still performed more accurate triage or had different Likert score impression ratings than our PTT group; lack of this serves as an additional limitation of the study.

In addition, unlike a high-fidelity simulation, the paper-based nature of a tabletop simulation limits the external validity of the results. Notably for the algorithms requiring quantitative physiologic parameters, these values were provided, and the time, difficulty, and interrater variability of calculation under the stress of an MCI was not a factor for the participants of our study. J-START requires calculation of a respiratory rate, and Sacco Triage Method requires an interactive neurological exam and counting of respirations and pulse. Compared with a tool like SALT, which requires none, these items potentially facilitated triage when provided but could prove cumbersome if the onus were on professionals to count. Therefore, the time to triage for these algorithms was likely shorter in our study for these algorithms. Likewise, the accurate application of the algorithm and prehospital professional impression Likert scores may have been artificially elevated for the algorithms requiring such data.

The local prehospital professionals' previous knowledge and use of J-START may have impacted the study results, as the cohort of prehospital professionals were employed in an EMS system that uses J-START. In one study on the J-START and Smart triage tools,

TABLE 5 Median Likert scores ratings for each triage algorithm with interquartile ranges and Kruskal-Wallis *P* value for difference

	CF n = 28	J-START n = 15	PTT n = 18	SALT n = 15	STM n = 15	NA n = 16	Kruskal-Wallis *P* value
Easy to understand	5 (4–5)	4 (4–4)	4 (4–4)	4 (4–4)	4 (2–4)	5 (4.5–5)	<0.0001
Easy to perform	5 (4–5)	4 (4–4)	4 (3–4)	4 (3–4)	4 (3–4)	5 (4–5)	<0.0001
Easy to remember	4 (4–5)	4 (3–4)	4 (2–4)	4 (4–4)	2 (2–3)	5 (4–5)	<0.0001
Useful for 10	4 (4–5)	4 (4–5)	4 (4–4)	4 (4–4)	3 (2–4)	5 (4–5)	<0.0001
Useful for 50	5 (4–5)	4 (3–4)	4 (2–4)	4 (4–4)	2 (2–3)	5 (4–5)	<0.0001
Useful for 100	5 (4–5)	4 (3–4)	4 (1–4)	4 (3–4)	2 (1–3)	5 (4–5)	<0.0001
Useful for 1000	5 (4–5)	4 (3–4)	4 (1–4)	4 (3–4)	2 (1–2)	5 (4–5)	<0.0001

Abbreviations: CF, CareFlight; J-START, JumpSTART (Simple Triage and Rapid Treatment); NA, no algorithm; PTT, Pediatric Triage Tape; SALT, Sort, Assess, Life-saving interventions, Treatment/Transport; STM, Sacco Triage Method.

education did improve triage accuracy by ∼10% (from 80% to 90%).[22] The triage algorithms were presented to the prehospital professionals just before the tabletop exercise, potentially limiting the ability to understand the long-term retention of the previously unknown or unfamiliar algorithms.

Lastly, 8 participants also contributed in 2 cohorts because of scheduling redundancies. Separate analysis was undertaken to determine the impact of these 8 repeaters on the overall data. Although there were differences found in the data analysis when excluding the second participation event of the 8 repeaters, the changes were minimal with changes of accuracy of up to 0.4% and 2 changes in medians of 0.5 and 1 point. These minimal changes did not affect the overall results. There were also no fundamental differences in the demographics of the repeaters verses the other study participants. Finally, 100% of the reporting participants were male, which may limit the generalizability to other agencies (2 participants declined disclosure of sex classification).

5 | DISCUSSION

This is the first study comparing prehospital professional accuracy, interrater reliability, and speed both without a tool ("No Algorithm") and with 1 of 5 common pediatric triage algorithms during a tabletop exercise. None of these algorithms excelled in all 3 areas of prehospital professional accuracy, interrater reliability, and speed during this exercise. Our study found J-START to have high accuracy and Likert impression ratings, though it performed with moderate speed and interrater reliability. SALT and CF were found to be fast with high Likert impression ratings but with lower accuracy, higher over-/undertriage rates, and the lowest prehospital professional interrater reliability. This is the first known pediatric MCI triage study comparing a gestalt-based method, NA, to the other pediatric MCI triage algorithms. NA had the third highest accuracy and speed; however, NA had a high level of under-triage (15.8%), second only to SALT (16.3%).

Currently, no standards define the acceptable accuracy for MCI triage algorithms used on pediatric patients. STM was applied with the highest accuracy in our study (89.3%, 95% CI 85.7% to 92.2%) and is the second highest accuracy seen in any previous studies for other pediatric MCI triage algorithms (highest 92.4%).[18,19,22–27] Our study accuracy percentages for J-START (80.3%, 95% CI 75.8% to 84.2%) and SALT (70.4%, 95% CI 65.5% to 75.0%) were similar to results from prior MCI studies for these algorithms (64.3% to 92.4% and 66% to 85.9%, respectively), showing that prehospital professionals can accurately apply J-START and SALT in the majority of patients.[18,19,22–27] No previous studies were found examining prehospital professional accuracy for the remaining algorithms. Our study data can act as a starting point for accuracy and overtriage/undertriage comparisons for future studies on these pediatric triage algorithms. Likewise, as there are no known previous studies comparing prehospital professional interrater reliability by algorithm, our study data can be used as a foundation for future studies with STM demonstrating the highest interrater reliability, whereas CF and SALT demonstrated the lowest interrater reliability.

Triage algorithms commonly prioritize overtriage, rather than undertriage, potentially overusing resources but ensuring the largest group of patients are selected that could require urgent intervention.[30] Our over- and undertriage percentages for J-START (11.1% and 8.6%, respectively) are less than in previous studies (17.8% to 23% and 11% to 12.6%, respectively) with a tendency to overtriage.[18,23] The algorithms with the highest overtriage (PTT 25.8%, CF 15.7%) are worrisome as these algorithms may be more costly for resources and divert medical attention away from more critically injured children.[28,29] Our data show an opposite tendency for SALT, toward under- rather than overtriage (16.3% vs 13.3%, respectively).[18,24,26] In fact, SALT had the highest undertriage in our study, followed closely by NA (15.8%). Notably, these undertriage percentages are higher than in any previous study (highest previously reported 12.6%).[18,23,24,26]

Our time to triage was fastest for SALT, and these times were considerably faster than previous studies (11.44 seconds vs 28–34 seconds per patient).[18,24] Whereas our second fastest algorithm (CF) had times comparable to previous literature (11.6 seconds vs 10–15 seconds per patient).[12,13] STM, perhaps given the amount of data points needed to compute STM's scoring, was the slowest algorithm in our study, although a previous study found STM to be much slower (26.12 seconds vs 70.9 seconds per patient).[11] Providing respiratory rate and pulse rate to the prehospital professionals, instead of having them count these, likely shortened this time to triage for STM.

6 | CONCLUSION

This study demonstrates the most common challenges inherent to MCI triage systems: as prehospital professional accuracy and interrater reliability improve, speed slows. No triage algorithm in our study excelled in all 3 of these measures. This is the first study directly comparing the prehospital professional accuracy, interrater reliability, and speed both without a tool ("No Algorithm") and with 1 of 5 common pediatric triage algorithms. Our data can serve as comparison data for future investigation of prehospital professional use of these algorithms in larger MCI drills requiring vital signs to be collected in real time or during a real MCI event.

ACKNOWLEDGMENTS

Thank you to Dr. Ilene Claudius, Dr. Genevieve Santillanes, Dr. Marianne Gausche-Hill, and Dr. Todd Chang for their assistance in editing the manuscript. Thank you to Dr. Nancy Glober for assistance with constructing Figure 2. Thank you to Sabir Rashid for helping for his contributions to the on-site operation of the study.

MEETINGS PRESENTED

2019 National Association of EMS Physicians Annual Meeting, Austin, TX and 2019 Society for Academic Emergency Medicine Annual Meeting, Las Vegas, NV.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

J. Joelle Donofrio-Odmann, Katherine Staats, Nicole D'Arcy, and Tabitha Cheng created lectures and facilitated the tabletop scenarios for the study. J. Joelle Donofrio-Odmann, Katherine Staats, and Tabitha Cheng performed the literature review. Amy H. Kaji and Kian Niknam performed the statistical analysis. Tabitha Cheng wrote the manuscript. Nicole D'Arcy wrote the methods section with contributions from Amy H. Kaji on the statistical methods. Tabitha Cheng, Katherine Staats, Amy H. Kaji, Nicole D'Arcy, Kian Niknam, and J. Joelle Donofrio-Odmann critically revised the manuscript.

ORCID

Tabitha Cheng MD https://orcid.org/0000-0002-9976-2448
Katherine Staats MD https://orcid.org/0000-0001-8632-2606
Amy H. Kaji MD PhD https://orcid.org/0000-0003-4588-7939
J. Joelle Donofrio-Odmann DO https://orcid.org/0000-0003-3256-1331

REFERENCES

1. Hewes H, Krug S, Fisher M. American Academy of pediatrics textbook of pediatric care. *Am Acad Pediatr*. Second ed.
2. Gausche-Hill M. Pediatric disaster preparedness: are we really prepared?. *J Trauma*. 2009;67(2 Suppl):S73-6. https://doi.org/10.1097/TA.0b013e3181af2fff.
3. Cu J, Phan P, O'Leary FM. Knowledge and attitude towards paediatric cardiopulmonary resuscitation among the carers of patients attending the emergency department of the Children's Hospital at Westmead. *Emerg Med Australas*. 2009;21(5):401-406. https://doi.org/10.1111/j.1742-6723.2009.01217.x.
4. Kouliev T. Objective triage in the disaster setting: will children and expecting mothers be treated like others?. *Open Access Emerg Med*. 2016;8:77-86. https://doi.org/10.2147/OAEM.S96913.
5. Nadeau NL, Cicero MX. Pediatric disaster triage system utilization across the United States. *Pediatr Emerg Care*. 2017;33(3):152-155. https://doi.org/10.1097/PEC.0000000000000680.
6. Wallis LA, Carley S. Validation of the paediatric triage tape. *Emerg Med J*. 2006;23(1):47-50. https://doi.org/10.1136/emj.2005.024893.
7. Bazyar J, Farrokhi M, Khankeh H. Triage systems in mass casualty incidents and disasters: a review study with a worldwide approach. *Open Access Maced J Med Sci*. 2019;7(3):482-494. https://doi.org/10.3889/oamjms.2019.119.
8. Romig LE. Pediatric triage. A system to JumpSTART your triage of young patients at MCIs. *JEMS*. 2002;27(7):52-, 60–63.
9. Hodgetts T, Hall J, Maconochie I, Smart C. Paediatric Triage Tape. *Prehospital Immediate Care*. 1998;2:155-159.
10. Lerner EB, Schwartz RB, Coule PL, et al. Mass casualty triage: an evaluation of the data and development of a proposed national guideline. *Disaster Med Public Health Prep*. 2008;2(Suppl 1):S25-34. https://doi.org/10.1097/DMP.0b013e318182194e.
11. Jain TN, Ragazzoni L, Stryhn H, Stratton SJ. Della Corte F. Comparison of the sacco triage method versus START triage using a virtual reality scenario in advance care paramedic students. *CJEM*. 2016;18(4):288-292. https://doi.org/10.1017/cem.2015.102.
12. Garner A, Lee A, Harrison K, Schultz CH. Comparative analysis of multiple-casualty incident triage algorithms. *Ann Emerg Med*. 2001;38(5):541-548. https://doi.org/10.1067/mem.2001.119053.
13. Jenkins JL, McCarthy ML, Sauer LM, et al. Mass-casualty triage: time for an evidence-based approach. *Prehosp Disaster Med*. 2008;23(1):3-8.
14. Heffernan RW, Lerner EB, McKee CH, et al. Comparing the accuracy of mass casualty triage systems in a pediatric population. *Prehosp Emerg Care*. 2019;23(3):304-308. https://doi.org/10.1080/10903127.2018.1520946.
15. Wallis LA, Carley S. Comparison of paediatric major incident primary triage tools. *Emerg Med J*. 2006;23(6):475-478. https://doi.org/10.1136/emj.2005.032672.
16. Cross KP, Cicero MX. Head-to-head comparison of disaster triage methods in pediatric, adult, and geriatric patients. *Ann Emerg Med*. 2013;61(6):668-676.e7. https://doi.org/10.1016/j.annemergmed.2012.12.023.
17. Price CL, Brace-McDonnell SJ, Stallard N, Bleetman A, Maconochie I, Perkins GD. Performance characteristics of five triage tools for major incidents involving traumatic injuries to children. *Injury*. 2016;47(5):988-992. https://doi.org/10.1016/j.injury.2015.10.076.
18. Jones N, White ML, Tofil N, et al. Randomized trial comparing two mass casualty triage systems (JumpSTART versus SALT) in a pediatric simulated mass casualty event. *Prehosp Emerg Care*. 2014;18(3):417-423. https://doi.org/10.3109/10903127.2014.882997.
19. Claudius I, Kaji AH, Santillanes G, et al. Accuracy, efficiency, and inappropriate actions using JumpSTART triage in MCI simulations. *Prehosp Disaster Med*. 2015;30(5):457-460. https://doi.org/10.1017/S1049023X15005002.
20. Donofrio JJ, Kaji AH, Claudius IA, et al. Development of a pediatric mass casualty triage algorithm validation tool. *Prehosp Emerg Care*. 2016;20(3):343-353. https://doi.org/10.3109/10903127.2015.1111476.
21. Newgard CD, Kampp M, Nelson M, et al. Deciphering the use and predictive value of "emergency medical services provider judgment" in out-of-hospital trauma triage: a multisite, mixed methods assessment. *J Trauma Acute Care Surg*. 2012;72(5):1239-1248. https://doi.org/10.1097/TA.0b013e3182468b51.

22. Cicero MX, Whitfill T, Overly F, et al. Pediatric disaster triage: multiple simulation curriculum improves prehospital care providers' assessment skills. *Prehosp Emerg Care*. 2017;21(2):201-208. https://doi.org/10.1080/10903127.2016.1235239.
23. Sapp RF, Brice JH, Myers JB, Hinchey P. Triage performance of first-year medical students using a multiple-casualty scenario, paper exercise. *Prehosp Disaster Med*. 2010;25(3):239-245.
24. Lerner EB, Schwartz RB, Coule PL, Pirrallo RG. Use of SALT triage in a simulated mass-casualty incident. *Prehosp Emerg Care*. 2010;14(1):21-25. https://doi.org/10.3109/10903120903349812.
25. Deluhery MR, Lerner EB, Pirrallo RG, Schwartz RB. Paramedic accuracy using SALT triage after a brief initial training. *Prehosp Emerg Care*. 2011;15(4):526-532. https://doi.org/10.3109/10903127.2011.569852.
26. Lee CW, McLeod SL, Peddle MB. First responder accuracy using salt after brief initial training. *Prehosp Disaster Med*. 2015;30(5):447-451. https://doi.org/10.1017/S1049023X15004975.
27. Claudius I, Kaji A, Santillanes G, et al. Comparison of computerized patients versus live moulaged actors for a mass-casualty drill. *Prehosp Disaster Med*. 2015;30(5):438-442. https://doi.org/10.1017/S1049023X15004963.
28. Cone DC, Serra J, Burns K, MacMillan DS, Kurland L, Van Gelder C. Pilot test of the SALT mass casualty triage system. *Prehosp Emerg Care*. 2009;13(4):536-540. https://doi.org/10.1080/10903120802706252.
29. Faul M, Wald MM, Sullivent EE, et al. Large cost savings realized from the 2006 field triage guideline: reduction in overtriage in U.S. trauma centers. *Prehosp Emerg Care*. 2012;16(2):222-229. https://doi.org/10.3109/10903127.2011.615013.
30. Trauma. *ACoSCo. Resources for Optimal Care of the Injured patient.*: American College of Surgeons, Committee on Trauma; 2014. p. 31.
31. Tran MD, Garner AA, Morrison I, Sharley PH, Griggs WM, Xavier C. The Bali bombing: civilian aeromedical evacuation. *Med J Aust*. 2003;179(7):353-356. https://doi.org/10.5694/j.1326-5377.2003.tb05592.x.
32. Rehn M, Andersen JE, Vigerust T, Krüger AJ, Lossius HM. A concept for major incident triage: full-scaled simulation feasibility study. *BMC Emerg Med*. 2010;10:17. https://doi.org/10.1186/1471-227X-10-17.

AUTHOR BIOGRAPHY

Tabitha Cheng, MD, is board certified in emergency medicine and the subspecialty of emergency medical services. She is pursuing a second fellowship in pediatric emergency medicine at Harbor-UCLA.

How to cite this article: Cheng T, Staats K, Kaji AH, et al. Comparison of prehospital professional accuracy, speed, and interrater reliability of six pediatric triage algorithms. *JACEP Open*. 2022;3:e12613. https://doi.org/10.1002/emp2.12613