Head-to-Head Comparison of Disaster Triage Methods in Pediatric, Adult, and Geriatric Patients

Keith P. Cross, MD, MS; Mark X. Cicero, MD

Study objective: A variety of methods have been proposed and used in disaster triage situations, but there is little more than expert opinion to support most of them. Anecdotal disaster experiences often report mediocre real-world triage accuracy. The study objective was to determine the accuracy of several disaster triage methods when predicting clinically important outcomes in a large cohort of trauma victims.

Methods: Pediatric, adult, and geriatric trauma victims from the National Trauma Data Bank were assigned triage levels, using each of 6 disaster triage methods: simple triage and rapid treatment (START), Fire Department of New York (FDNY), CareFlight, Glasgow Coma Scale (GCS), Sacco Score, and Unadjusted Sacco Score. Methods for approximating triage systems were vetted by subject matter experts. Triage assignments were compared against patient mortality at hospital discharge with area under the receiver operator curve. Secondary outcomes included death in the emergency department, use of a ventilator, and lengths of stay. Subgroup analysis assessed triage accuracy in patients by age, trauma type, and sex.

Results: In this study, 530,695 records were included. The Sacco Score predicted mortality most accurately, with area under the receiver operator curve of 0.883 (95% confidence interval 0.880 to 0.885), and performed well in most subgroups. FDNY was more accurate than START for adults but less accurate for children. CareFlight was best for burn victims, with area under the receiver operator curve of 0.87 (95% confidence interval 0.85 to 0.89) but mistriaged more salvageable trauma patients to "dead/black" (41% survived) than did other disaster triage methods (≈10% survived).

Conclusion: Among 6 disaster triage methods compared against actual outcomes in trauma registry patients, the Sacco Score predicted mortality most accurately. This analysis highlighted comparative strengths and weakness of START, FDNY, CareFlight, and Sacco, suggesting areas in which each might be improved. The GCS predicted outcomes similarly to dedicated disaster triage strategies. [Ann Emerg Med. 2013;61:668-676.]

Please see page 669 for the Editor's Capsule Summary of this article.

A **feedback** survey is available with each research article published on the Web at www.annemergmed.com.
A **podcast** for this article is available at www.annemergmed.com.

0196-0644/$-see front matter

http://dx.doi.org/10.1016/j.annemergmed.2012.12.023

INTRODUCTION

Since the first formalized battlefield triage by Napoleonic-era surgeons, health care providers have debated efficient, accurate systems to classify patients in mass casualty situations.[1] To date, most methods have been based on expert opinion with limited evidence.[1,2] The difficulties of prospectively studying mass casualty triage include the practical issues of rarity, nonrepeatability, randomizing subjects to different study groups, ethical considerations, logistics, and cost.[3] Consequently, simulation studies and postevent evaluations, rather than prospective trials, constitute much of the existing evidence base.[4-10]

Data from trauma registries offer a way to assess the performance of mass casualty triage systems on patient records without the practical limitations of a prospective study. Key to such registry-based analysis is tying initial, out-of-hospital patient information such as mental status and vital signs, used for triage assignment, to clinically meaningful outcomes in a large number of trauma patients. This approach has been used in some studies to date, albeit with scope restricted by either a limited number of patients or by few triage methods.[7,11-14]

For this study, we sought to compare the performance of 6 disaster triage methods. We applied these methods to a large registry of trauma patients and compared the calculated patient triage priorities with patient record outcomes. Our hypothesis was that some disaster triage methods predict important clinical outcomes more accurately than others.

MATERIALS AND METHODS

Study Design

This study used an existing trauma registry database of patient characteristics and outcomes to test the comparative performance of 6 candidate disaster triage methods. Each

Editor's Capsule Summary

What is already known on this topic

Objective evidence to support the accuracy of different methods of disaster triage is limited.

What question this study addressed

This study used the National Trauma Data Bank to assess the accuracy of 6 triage disaster methods, including simple triage and rapid treatment, Glasgow Coma Scale, and Sacco Score. Using explicit approximations, triage assignments were compared with hospital mortality.

What this study adds to our knowledge

Of the 530,695 records used, few represented the types or simultaneity of patients expected in disasters. The Sacco Score predicted mortality most accurately; however, the superiority may not be clinically significant.

How this is relevant to clinical practice

Although this study is unlikely to change practice, it highlights the need to objectively test triage methods.

method was retrospectively applied to each patient in the registry to obtain an assigned triage level, which was then compared against the patient's actual recorded outcome. For each triage method, across many patients, predictive accuracy was calculated. The performance of each method could then be compared head-to-head and further analyzed in specific subsets of patients.

Selection of Participants

The study patient population came from the 2007 to 2009 reporting years of the deidentified National Trauma Data Base (NTDB) (version 7.2; Chicago, IL), which collects out-of-hospital, emergency department (ED), inpatient, and discharge information about patients at participating trauma centers across the United States. It is managed by the Committee on Trauma of the American College of Surgeons.

The NTDB adds approximately 600,000 records per year. NTDB data are subject to the a priori inclusion, exclusion, and validation rules stated by the registry, as described in their documentation (available at http://www.facs.org/trauma/ntdb/ntdbapp.html). Participation in the NTDB is voluntary and has increased over the years. For example, in 2008, 435 US hospitals submitted data to the NTDB, including 75.6% of Level I and 59.4% of Level II trauma centers.[15] A current map reflecting US trauma center participation rates is available at http://www.ntdsdictionary.org/ntdbParticipants/stateInformation/index.html. Data in the NTDB may be used for a variety of purposes, including benchmarking processes (eg, ambulance response times, interfacility transfer rate), studying injury epidemiology (eg, the interaction of drug/alcohol use with motor vehicle–related trauma), and evaluating outcomes (eg, discharge mortality, hospital length of stay).

The institutional review boards for the authors' institutions found this study to be exempt from formal review because the NTDB data are preexisting and deidentified.

Data Collection and Processing and Primary Data Analysis

NTDB data records were downloaded into SPSS version 20 (IBM, Armonk, NY). The data were processed for inclusion or exclusion, triage level assignment, and data analysis, as outlined below. The data records are all from facilities with American College of Surgeons trauma accreditation.

For this study, registry patient records were included if they had a valid age and reported patient outcomes at both the ED disposition and final hospital discharge. A record was excluded if it represented an interfacility transfer or if it had any of several data irregularities, such as a total hospital length of stay that was shorter than the ED length of stay or a final hospital discharge outcome that was inconsistent with the ED disposition. Records were also excluded if they did not report the initial/scene vital signs of pulse, respirations, and Glasgow Coma Scale (GCS) score. Records that reported only vital signs obtained after arrival at a hospital ED were not acceptable for this study. Finally, records without an Injury Severity Score were excluded.

The primary statistical analysis was the development of receiver operator curves with associated area under the receiver operator curve (AUC) statistics. Additionally, descriptive statistics—mean, median, interquartile range, etc—were determined and reported for nonbinary outcomes. All these analyses were performed with SPSS software version 20 (SPSS, Inc.).

Outcome Measures

The primary outcome was mortality at hospital disposition as recorded in the NTDB. Mortality was compared against initial triage assignments with the receiver operator curve and AUC.

Secondary outcomes were death in the ED, defined as any ED disposition of "death," and use of a ventilator, defined as any time reported on a ventilator at any point in the hospital stay. These binary outcomes were compared against initial triage assignments with receiver operator curve and AUC.

This study also examined patients' hospital length of stay in days and ED length of stay in minutes compared with their initial triage assignments. For this data analysis, we reported all length of stay data as median and 25th and 75th percentiles (included in Appendix E1, available online at http://www.annemergmed.com).

Subsets of subjects were investigated to determine whether particular triage methods predict the primary outcome (mortality) better in some situations than in others. We performed subgroup analysis with 3 trauma types (blunt, penetrating, and burn), both sexes, and 4 age groups (0 to 8, 9 to 15, 16 to 64, and ≥65 years). The selection of these age subgroups was made according to guidelines from the American College of Surgeons to represent groups that may reasonably be expected to have the distinct physiology, the trauma epidemiology, and the preexisting comorbidities of young children, older children, adults, and geriatric patients, respectively.[16,17]

This study compared 6 disaster triage methods. Each patient in the study data set was retroactively assigned a triage level with each of these methods. The approach to approximating each triage method for assignment to NTDB records is discussed in detail in Appendix E1 (available online at http://www.annemergmed.com). Some triage algorithms use data that are not available in the NTDB. In these cases, we used approximations based on available surrogate information, as noted. We invited several disinterested experts in emergency medical services (EMS) and disaster management to vet the methods used to approximate the disaster triage strategies with the NTDB data. The experts were members of the American Academy of Pediatrics Disaster Preparedness Advisory Committee or the National Disaster Medical System or were medical directors of state or national EMS agencies. The experts were asked to review the methods as presented in this article in a post hoc manner.

Three of the experts reviewed the methods for approximating the triage methods. All 3 endorsed the methods used to approximate the disaster triage strategies as described in the methods and Appendix E1 (available online at http://www.annemergmed.com). Expert 1 suggested consideration of using the Revised Trauma Score rather than the Injury Severity Score in approximating triage levels. This method would have been reasonable but was not practical because NTDB data are often missing scene systolic blood pressure, a factor needed to calculate Revised Trauma Score. For generalizability of our findings, expert 2 suggested using the trauma age categories used by the American College of Surgeons. Expert 3 noted the limitations of using databank data but, given the limitations, endorsed our methods.

RESULTS

For this study, 530,695 NTDB patient records met inclusion and exclusion criteria. A breakdown of how these records were obtained from the NTDB appears in Figure 1.

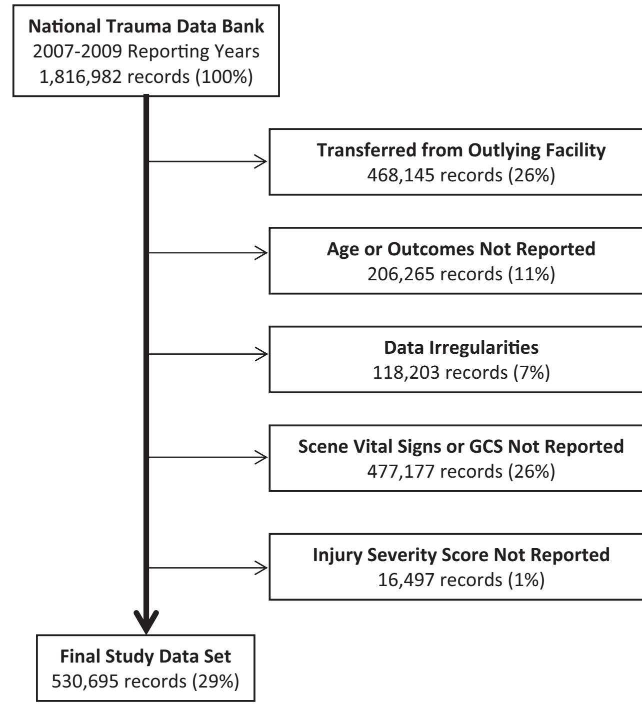

Figure 1. Application of study inclusion and exclusion criteria to source records from the NTDB to obtain the study data set.

Records that reported consistent age and outcome information but did not report initial scene vital signs, GCS information, or Injury Severity Score numbered 493,674 and were excluded from the analysis. A comparison of excluded records and records included in the study analysis is shown in Table 1.

Among the included records, there were 15,114 patients aged 0 to 8 years, 21,781 patients aged 9 to 15 years, 379,144 patients aged 16 to 64 years, and 114,656 patients aged 65 years or older. Women composed 35.1% of the records. Racially, 64.2% were white, 15.5% were black, 12.2% were Hispanic/Latino, and 7.5% had no race recorded. Victims with Injury Severity Score greater than 15 (major injury) were 24.4% of all patients. Blunt trauma was reported in 84.8% of records, penetrating trauma in 11.0%, and burn trauma in 0.9%; the remaining 3.3% of records did not report a trauma type.

There were 21,887 deaths, which was an overall mortality of 4.1%, and 2,198 patients died in the ED. Mortality in men was 4.5%; in women, 3.4%. There were 68,925 patients who received ventilator support at some point in their hospital stay, which was 13.0% of all included study patients. The median for hospital length of stay was 2.99 days (interquartile range 1.19 to 6.01 days) and for ED length of stay was 223 minutes (interquartile range 131 to 345 minutes). Table E4*A* and *B* (available online at http://www.annemergmed.com) shows the detailed results for hospital length of stay in days and ED length of stay in minutes as a function of triage method and resulting triage level assignment. The median and interquartile ranges are also reported.

The results for comparison of initial triage assignments to mortality outcomes are shown in Table 2.

The receiver operator curves for the 6 triage methods predicting mortality are shown in Figure 2, whereas the summary of area under the curves is shown in Table 3 for overall mortality at hospital discharge, death in the ED, and any ventilator use.

Sensitivity Analyses

For the simple triage and rapid treatment (START), Fire Department of New York (FDNY), and CareFlight triage methods, determination of "minor/green" versus

Table 1. Overview of the 1,024,396 scene transport records with consistent data and reported patient ages and outcomes.*

Demographic of Interest	Included (Complete) Records	Excluded (Incomplete) Records
Number of records	530,695	493,674
Median age (IQR), y	42 (24–61)	39 (21–59)
Age range, %, y		
0–8	2.8	7.9
9–15	4.1	5.7
16–64	71.4	65.8
≥65	21.6	20.6
White, %	69.4	68.1
Male, %	64.8	64.3
Deaths in ED, No. (%)	2,198 (0.4)	2,037 (0.4)
Deaths at hospital disposition, No. (%)	21,887 (4.1)	15,162 (3.1)
Median ED LOS (IQR), min	223 (131–345)	247 (152–377)
Median hospital LOS (IQR), days	2.99 (1.19–6.01)	2.15 (0.86–4.86)
Median Injury Severity Score (IQR)	9 (4–14)	9 (4–11)

IQR, Interquartile range; *LOS*, length of stay.
*Records are separated into those included in the study analysis versus those excluded for incomplete reporting of initial vital signs, Glasgow Coma Scale score, or Injury Severity Score. Note that 16,497 records in the "Excluded" column did not report an Injury Severity Score and were not included in the analysis for the final row of the table.

"delayed/yellow" assignment is based on whether a patient can ambulate, useful in disaster settings but neither typically recorded by EMS personnel nor available in the NTDB. As discussed at length in Appendix E1 (available online at http://www.annemergmed.com), this determination based on the ability to ambulate was approximated with a surrogate. Patients with Injury Severity Score less than or equal to 10 were assigned to minor/green and those with Injury Severity Score greater than 10 were assigned to delayed/yellow, with the results reported in Table 3. During sensitivity analysis, this Injury Severity Score threshold was increased to 12 (more patients assigned to minor/green) and decreased to 8 (more patients assigned to delayed/yellow), and the resulting AUC for predicting mortality was recalculated. For START, Injury Severity Score cutoffs at 8 and 12 resulted in AUCs of 0.840 (95% confidence interval [CI] 0.837 to 0.842) and 0.848 (95% CI 0.845 to 0.850), respectively. For FDNY, Injury Severity Score cutoffs at 8 and 12 resulted in AUCs of 0.861 (95% CI 0.858 to 0.863) and 0.851 (95% CI 0.848 to 0.853), respectively. For CareFlight, the resulting AUC was 0.844 (95% CI 0.841 to 0.847) and 0.854 (95% CI 0.851 to 0.857), respectively.

We assessed the performance of the 6 triage methods when evaluating specific subsets of patients. The results of this subgroup analysis are shown in Table E5*A*, *B*, and *C* (available online at http://www.annemergmed.com) for differing age groups, trauma types, and sexes.

LIMITATIONS

The main limitation of our study is the use of retrospective registry data. The data include mostly singleton trauma cases, rather than victims of mass casualty events in which resources are more constrained and patient outcomes would likely be worse. Moreover, individual data fields may be missing or inaccurately recorded. As shown in Table 1, there are differences between included (complete) records and excluded (incomplete) records, namely, that included records tended to have older age (median 42 versus 39 years), higher mortality risk (4.1% versus 3.1%), and a broader distribution of Injury Severity Scores (interquartile range 4 to 14 versus 4 to 11). Options to impute missing data for this analysis were considered but were not used. Choosing not to impute missing data avoids the layering of data approximations on top of triage method approximations. In our opinion, the bias toward including higher-risk, more significantly injured subjects in this study is possibly helpful because it enriches the study data set with the types of patients of greatest interest to medical providers in a mass casualty situation. Low-priority "green" patients—the numerous "walking wounded" disaster victims who can inundate and distract a trauma center if improperly handled—still represented between 49% and 73% of subjects in the study, depending on triage method, as shown in Table 2.

Records in the NTDB were submitted only from participating US trauma centers and are therefore not a nationally representative sample. Trauma centers may not submit data for all trauma patients; for example, an accident victim dead at the scene may not be recorded in the registry.

In addition to the data set limitations, our methodology of assigning triage levels to patients is imperfect. As noted in the "Materials and Methods" section (and described in detail in Appendix E1, available online at http://www.annemergmed.com), several triage methods use information that is not recorded in the NTDB registry (such as "Can this patient walk away from the trauma scene on command?"). Although the approximations we used seem reasonable, given the practical limitations of the data, and were vetted by outside experts, they are not precisely what is specified in the triage methods. Therefore, where we used approximations, the performance of a particular triage algorithm may be overestimated or underestimated. Sensitivity analysis of our most significant approximation, the use of Injury Severity Score less than or equal to 10 as a proxy for the ability to ambulate, showed AUC results either overlapping or significantly worse when alternate cutoffs of 8 or 12 were used. This outcome implies that an alternate threshold for Injury Severity Score would have little if any effect on our main findings and conclusions.

Mass casualty incidents may involve injury mechanisms not well represented in the NTDB, such as burn/blast injuries or illness from environmental, radioactive, chemical, or biologic exposures. Some triage methods may perform differently in such

Table 2. Resulting mortality outcomes at hospital discharge for all 530,695 study patients.

Triage Method	Assigned Triage Level	Lived	Died	Mortality, %
START/JumpSTART, FDNY, CareFlight, and GCS disaster triage methods				
START	Dead/black	218	2,089	90.6
	Immediate/red	66,431	13,422	16.8
	Delayed/yellow	122,043	4,330	3.4
	Minor/green	320,116	2,046	0.6
FDNY	Dead/black	218	2,089	90.6
	Immediate/red	64,928	13,183	16.9
	Urgent/orange	87,939	4,248	4.6
	Delayed/yellow	41,303	355	0.9
	Minor/green	314,420	2,012	0.6
CareFlight	Dead/black	2,775	3,923	58.6
	Immediate/red	53,800	11,033	17.0
	Delayed/yellow	126,698	4,807	3.7
	Minor/green	325,535	2,124	0.6
Glasgow Coma Scale	3	11,971	9,859	45.2
	4	1,706	668	28.1
	5	1,934	610	24.0
	6	3,626	825	18.5
	7	3,526	552	13.5
	8	3,880	476	10.9
	9	4,019	448	10.0
	10	5,025	434	8.0
	11	5,785	404	6.5
	12	9,042	495	5.2
	13	15,796	650	4.0
	14	56,778	1,426	2.5
	15	385,720	5,040	1.3
Resulting mortality outcomes at hospital discharge for Sacco and Unadjusted Sacco triage methods				
Sacco	−2	11	143	92.9
	−1	34	307	90.0
	0	117	1,514	92.8
	1	71	377	84.2
	2	237	642	73.0
	3	827	1,012	55.0
	4	1,935	1,416	42.3
	5	1,952	1,542	44.1
	6	2,412	1,517	38.6
	7	4,708	1,858	28.3
	8	15,130	2,724	15.3
	9	17,015	1,759	9.4
	10	88,519	3,878	4.2
	11	113,831	2,091	1.8
	12	241,173	1,089	0.4
	13	16,285	16	0.1
	14	4,551	2	0
Unadjusted Sacco	0	155	2,073	93.0
	1	44	270	86.0
	2	148	397	72.8
	3	676	830	55.1
	4	2,050	1,422	41.0
	5	1,846	1,591	46.3
	6	1,925	1,108	36.5
	7	3,389	1,394	29.1
	8	10,365	2,962	22.2
	9	10,260	1,168	10.2
	10	35,415	1,728	4.7
	11	58,191	1,953	3.2
	12	384,344	4,991	1.3

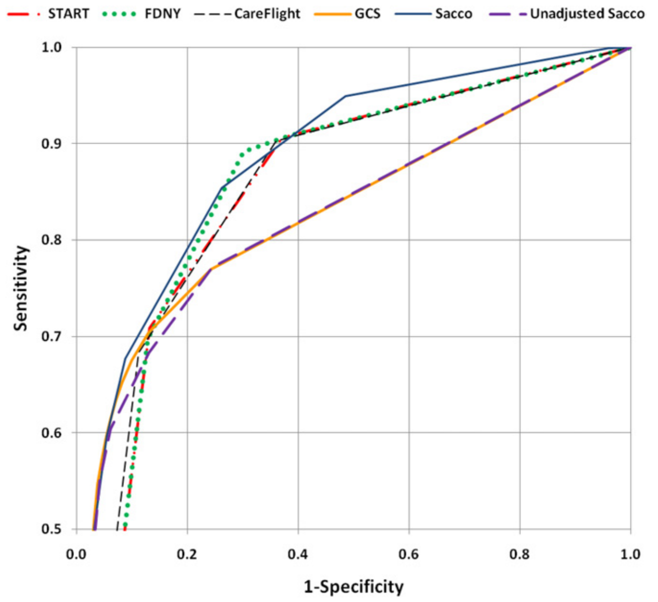

Figure 2. Receiver operator curves for the prediction of mortality at hospital discharge, using data from 530,695 NTDB patient records with complete outcome and scene data for all 6 triage methods. The graph is truncated at a *y* axis value of 0.5 to focus on the area of greatest interest at the inflection points of the 6 curves. Below the 0.5 level, the curves continue directly to the origin without interesting changes.

Table 3. AUC for primary outcome (mortality at hospital discharge) and secondary outcomes when predicted by each disaster triage method.*

			95% CI	
Outcome	**Triage Method**	**AUC**	**Lower Bound**	**Upper Bound**
Mortality at hospital discharge, N=21,887	START	0.846	0.843	0.849
	FDNY	0.851	0.848	0.853
	CareFlight	0.852	0.850	0.855
	GCS	0.825	0.822	0.829
	Sacco	0.883	0.880	0.885
	Unadjusted Sacco	0.824	0.821	0.828
Death in the ED (at disposition from the ED), N=2,198	START	0.950	0.946	0.954
	FDNY	0.951	0.947	0.955
	CareFlight	0.955	0.951	0.959
	GCS	0.952	0.947	0.957
	Sacco	0.967	0.963	0.971
	Unadjusted Sacco	0.970	0.965	0.974
Any ventilator use, N=68,925	START	0.799	0.797	0.801
	FDNY	0.805	0.803	0.807
	CareFlight	0.801	0.799	0.803
	GCS	0.744	0.742	0.746
	Sacco	0.714	0.711	0.716
	Unadjusted Sacco	0.735	0.733	0.738

*The 95% CI upper and lower bounds for AUC are shown. All analyses use 530,695 NTDB records; the number of patients with each outcome of interest (N) is shown.

specialized situations than we report here using general NTDB data.

This study does not reflect the effect of complexity on accuracy. That is, more complicated triage methods (eg, Sacco) may perform less well in low-frequency, high-stakes disasters when out-of-hospital care personnel must calculate the triage level, rather than a computer analyzing a research database.

Finally, although the comparison of triage methods using AUC gives clear statistical information, it is vague about the magnitude of clinically relevant differences. Any clinical differences between triage methods may be further blurred when one considers how inconsistently triage methods are applied by stressed first responders with limited disaster experience. Readers would be wise to consider the AUC differences less a reliable prediction of performance in the field and more a method for gaining insight into the relative design strengths and weaknesses of each triage method.

DISCUSSION

This study compared the performance of 6 mass casualty triage methods, using more than 500,000 records from a national trauma registry. Each method's initial triage assignments were compared with patient record outcomes. There were significant differences in the predictive performance of these triage methods, as assessed with receiver operator curves, for both primary and secondary outcomes. Subgroup analysis in specific patient subsets revealed further differences among methods.

The START triage method in this study performed particularly well among burn and penetrating trauma patients but much less well among patients with blunt trauma, who composed the majority of the study population. START's main limitation appears to be the total of 4 triage levels. Additional triage strata allow greater discrimination of patient urgency and clinical need. This increased discrimination comes at the cost of increased complexity. Other methods with more risk strata (FDNY, GCS, and Sacco) frequently outperformed START.

In this study, FDNY triage method performance was statistically similar to or marginally better than that of START, on which it is based. However, its AUC was significantly worse than the AUC for all other methods in younger children (aged 0 to 8 years), as shown in Table E5*A* (available online at http://www.annemergmed.com). The simplification of vital sign criteria for pediatric patients may have cost FDNY some accuracy. Additionally, in 1,522 infants in this data set, only 14.8% had abnormal respirations by START criteria and only 2.4% were initially pulseless, whereas 66.1% had Injury Severity Scores of 10 or less (indicating minor injury only). Therefore, the automatic assignment of "immediate/red" to all infants may be overtriage that further worsens FDNY AUC predictive value in young children.

In contrast, the use of "urgent/orange" to further segment adult patients appeared to be an effective strategy; the FDNY AUC was significantly better in adults (aged 16 to 64 years) than the START AUC. The mortality outcomes in Table 2 show that the FDNY strategy successfully identified riskier patients as urgent/orange (4,248 deaths; 4.6% mortality). Compared with START delayed/yellow (4,330 deaths; 3.4% mortality), the FDNY delayed/yellow group had significantly fewer deaths (355 deaths; 0.9% mortality).

The CareFlight method's simplicity and its relatively strong performance in a smaller head-to-head study (3,461 pediatric patients) make it an attractive candidate for analysis and use in disasters.[13] CareFlight generally predicted outcomes well in our study. However, it overtriages many viable patients to "dead/black," resulting in a low 58.6% mortality for this triage level, as shown in Table 2. Only the GCS triage method, with a mortality of 45.2% in its highest-risk group (GCS score=3), had less accuracy than CareFlight when identifying unsalvageable patients in this trauma registry.

The difference between our current findings and the more decisively positive results for CareFlight in the previous study by Wallis and Carley[13] may be due to different choices of endpoints. Our study used clinical endpoints: mortality at discharge, death in the ED, receiving ventilator support. Wallis and Carley[13] and others[14] instead used proxy measures (Injury Severity Score, New Injury Severity Score) and medical interventions (Garner criteria) as the outcomes against which they compared CareFlight. Wallis and Carley[13] also used point estimates of sensitivity and specificity for immediate/red assignments only, rather than the performance of a receiver operator curve across all triage levels. Finally, for Wallis and Carley[13] the study population was children only.

Though the GCS was not originally intended for mass casualty triage, this study showed that it performs well in many situations. It appeared to be particularly accurate in young children (aged 0 to 8 years), where its AUC exceeded that of other triage methods, although all triage methods had excellent AUC greater than 0.9 in this patient subset. This finding suggests that mental status assessment in young children may have greater triage predictive value than determining other vital signs and comparing them to normal values by age.

The weakest aspect of GCS was its performance at the extremes, as shown in Table 1*A*. A GCS score of 3 (presumably dead) had a mortality of only 46.4% compared with START and FDNY dead/black mortality and Sacco Score "0 or less" mortality, which are all greater than 90%. At the other extreme, a GCS score of 15 (normal) had a mortality of 1.3% compared with START, FDNY, and CareFlight minor/green mortality of 0.6% and Sacco Score "12 or more" mortality of 0.4%.

A clear conclusion from these findings is that although mental status is a useful predictor of both mortality and the secondary outcomes investigated here, particularly for children, other factors appear to add discriminating power to triage algorithms, particularly for adult and geriatric patients.

Much like the findings of previous studies of Sacco Score in trauma patients,[7,12,18] in our study the Sacco Score correlated to a steeply graded mortality curve: from greater than 90% at one end to less than 0.5% at the other end, with nearly uninterrupted decreasing steps in between. In our data, an exception was at a score of 5, with a slightly higher mortality than the preceding score of 4, as shown in Table 2.

Consequently, when judged by AUC for predicting mortality, the Sacco Score performed as well as and often significantly better than other disaster triage methods both broadly across all trauma cases (Tables 2 and 3) and narrowly in most subanalyses (Table E5*A*, *B*, and *C*, available online at http://www.annemergmed.com). Sacco Scores were significantly more accurate than other methods in patients with blunt trauma, as shown in Table E5*B* (available online at http://www.annemergmed.com). However, the situation reversed in penetrating trauma cases, in which START and FDNY performed better.

The age adjustment term appeared to add little when children were triaged because the Unadjusted Sacco Score did just as well in this group. However, the adjustment appeared to improve accuracy significantly in geriatric patients, as shown in Table E5*A* (available online at http://www.annemergmed.com). In light of these results, we endorse simplifying the Sacco Score by dropping the pediatric portion of the age adjustment.

In regard to the elderly, all the triage methods here studied were noticeably less accurate than in other age groups, and Sacco accuracy in the elderly lagged significantly behind that of START and FDNY. Several recent studies of other triage systems (eg, field triage, Emergency Severity Index) found problems with accuracy in elderly patients as well.[19,20] We speculate that chronic illness and comorbidity complicate outcome prediction in geriatric trauma patients. An area of future research is how Sacco and other methods can better triage elderly patients.

A notable and difficult-to-explain finding was the difference in predictive accuracy observed in men versus women. As shown in Table E5*C* (available online at http://www.annemergmed.com), all 6 triage methods have statistically better accuracy, judged by AUC, in men than in women. It is not clear whether this statistical difference crosses the threshold of clinical relevancy. There are clear underlying differences between men and women in overall trauma epidemiology. In the data set used for this study, men were more often trauma victims (64.7% of patients in the data set versus 35.1%). Furthermore, women were more likely to survive than men (3.4% mortality versus 4.5% mortality), had a lower mean Injury Severity Score than men (10 versus 12 respectively), but had a similar median Injury Severity Score (9 for both), suggesting that men represent a broader distribution with more severe injuries. How these differences affect triage accuracy is unclear and is a suitable subject for further research.

Viewed in total, these results highlight strengths and weaknesses of each triage method but stop short of demonstrating a clear winner. Some triage methods work better in some circumstances but not as well in others, which suggests that to move the science of triage forward, the disaster medicine community should blend the best elements of each system, and perhaps some novel elements as well, to create a practical, efficient, and accurate method whose empiric performance is robust across a wide range of mass casualty situations. Differences by subgroup (eg, age, sex, trauma type) may need to be part of the optimized algorithm, such as recognizing the increased mortality risk associated with elderly victims or penetrating trauma mechanisms and upgrading the triage assignment in such cases. Much work remains to blend, optimize, and implement evidence-based disaster triage.

Among 6 disaster triage methods compared against outcomes in trauma registry patients, the Sacco Score tended to predict mortality most accurately. This analysis highlights comparative strengths and weakness of START/JumpSTART, FDNY, CareFlight, the GCS, and the Sacco Score, suggesting areas in which each triage method might be further improved.

Future research on disaster triage should focus on the characteristics of mistriaged patients—why false-positive and false-negative cases were missed—and how the best elements of various triage methods may be blended and optimized to create a robust, evidence-based, and practical disaster triage system. Further, though there are challenging logistics to testing disaster triage methods during simulated or actual disasters, such testing will more firmly establish the relative merits of the methods evaluated in this study.

The authors acknowledge Kendra Sikes, BS, for her indispensable help with administrative and research compliance matters; the American College of Surgeons and the staff at the National Trauma Data Bank for creating and maintaining the data set used for this study; Charles Woods, MD, for his review and insightful comments; Ronald Paul, MD, and Gerard Rabalais, MD, for their ongoing support of our research efforts; Lou Romig, MD, for her kind provision of the START/JumpSTART figure (Figure E1); and the outside experts who reviewed, advised, and vetted our approach to approximating triage methods in the face of specific data limitations (reviewer 1: Kathleen Adelgais, MD, MPH, associate professor of pediatrics at the University of Colorado School of Medicine; reviewer 2: James Parker, MD, assistant professor of pediatrics at the University of Connecticut School of Medicine; and reviewer 3: Richard Aghababian, MD, professor of emergency medicine at the University of Massachusetts Medical School).

Supervising editor: Robert A. De Lorenzo, MD, MSM

Author affiliations: From the Department of Pediatrics, Kosair Children's Hospital, University of Louisville, Louisville, KY (Cross); and the Department of Pediatrics, Yale–New Haven Hospital, Yale School of Medicine, New Haven, CT (Cicero).

Author contributions: KPC and MXC conceived and designed the study, obtained institutional review/approval, and prepared the article. KPC performed the data analysis. KPC takes responsibility for the paper as a whole.

Funding and support: By *Annals* 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. This study was internally funded by the Department of Pediatrics at the University of Louisville School of Medicine.

Publication dates: Received for publication July 8, 2012. Revisions received October 2, 2012; November 19, 2012; and December 11, 2012. Accepted for publication December 19, 2012. Available online March 7, 2013.

Address for correspondence: Keith P. Cross, MD, MS, E-mail keith.cross@louisville.edu.

REFERENCES

1. Jenkins JL, McCarthy ML, Sauer LM, et al. Mass-casualty triage: time for an evidence-based approach. *Prehosp Disaster Med*. 2008;23:3-8.
2. SALT mass casualty triage: concept endorsed by the American College of Emergency Physicians, American College of Surgeons Committee on Trauma, American Trauma Society, National Association of EMS Physicians, National Disaster Life Support Education Consortium, and State and Territorial Injury Prevention Directors Association. *Disaster Med Public Health Prep*. 2008;2:245-246.
3. Antommaria AH, Powell T, Miller JE, et al. Ethical issues in pediatric emergency mass critical care. *Pediatr Crit Care Med*. 2011;12 (6 suppl):S163–168.
4. Cao H, Huang S. Principles of scarce medical resource allocation in natural disaster relief: a simulation approach. *Med Decis Making*. 2012;32:470–476.
5. Kahn CA, Schultz CH, Miller KT, et al. Does START triage work? an outcomes assessment after a disaster. *Ann Emerg Med*. 2009;54:424-430, 430.e421.
6. de Ceballos JP, Turegano-Fuentes F, Perez-Diaz D, et al. 11 March 2004: the terrorist bomb explosions in Madrid, Spain—an analysis of the logistics, injuries sustained and clinical management of casualties treated at the closest hospital. *Crit Care*. 2005;9:104-111.
7. Sacco WJ, Navin DM, Waddell RK 2nd, et al. A new resource-constrained triage method applied to victims of penetrating injury. *J Trauma*. 2007;63:316-325.
8. Montan KL, Khorram-Manesh A, Ortenwall P, et al. Comparative study of physiological and anatomical triage in major incidents using a new simulation model. *Am J Disaster Med*. 2011;6:289–298.
9. Ingrassia PL, Carenzo L, Barra FL, et al. Data collection in a live mass casualty incident simulation: automated RFID technology versus manually recorded system. *Eur J Emerg Med*. 2012;19:35–39.
10. Cone DC, Serra J, Kurland L. Comparison of the SALT and Smart triage systems using a virtual reality simulator with paramedic students. *Eur J Emerg Med*. 2011;18:314–321.
11. Thompson DO, Hurtado TR, Liao MM, et al. Validation of the Simplified Motor Score in the out-of-hospital setting for the

prediction of outcomes after traumatic brain injury. *Ann Emerg Med.* 2011;58:417–425.
12. Sacco WJ, Navin DM, Fiedler KE, et al. Precise formulation and evidence-based application of resource-constrained triage. *Acad Emerg Med.* 2005;12:759-770.
13. Wallis LA, Carley S. Comparison of paediatric major incident primary triage tools. *Emerg Med J.* 2006;23:475-478.
14. Garner A, Lee A, Harrison K, et al. Comparative analysis of multiple-casualty incident triage algorithms. *Ann Emerg Med.* 2001;38:541-548.
15. Nathens A, Fantus R. National Trauma Data Bank 2008 annual report version 8.0. American College of Surgeons [annual report]. Available at: http://www.ntdsdictionary.org/referenceMaterials/documents/ntdbannualreport2008.pdf. Accessed September 12, 2012.
16. Resources of optimal care of the injured patient—hospital prereview questionnaire. Available at: http://www.facs.org/trauma/prq.doc. Accessed September 12, 2012.
17. Lehmann R, Beekley A, Casey L, et al. The impact of advanced age on trauma triage decisions and outcomes: a statewide analysis. *Am J Surg.* 2009;197:571-574; discussion 574–575.
18. Cross K, Cicero M. Independent application of the Sacco disaster triage method to pediatric trauma patients. *Prehosp Disaster Med.* 2012;27:306-311.
19. Grossmann FF, Zumbrunn T, Frauchiger A, et al. At risk of undertriage? testing the performance and accuracy of the Emergency Severity Index in older emergency department patients. *Ann Emerg Med.* 60:317–325.e313.
20. Nakamura Y, Daya M, Bulger EM, et al. Evaluating age in the field triage of injured persons. *Ann Emerg Med.* 60:335–345.

APPENDIX E1.

Additional study methods information.

START and JumpSTART

First responders have used variations on the START (for adults) and JumpSTART (aged <9 years) triage methods for mass casualty triage for years (see Figure E1).[1-3] Despite its widespread use, START has not been validated, and postevent reviews of actual disasters have suggested that it may not be specific or discriminating.[4-6] START has been criticized for classifying too many patients as immediate/red and for having minimal evidence basis.[7]

When the START/JumpSTART algorithm is approximated retrospectively on the NTDB, there are several limitations. When START is used, the minor/green triage level in the algorithm was assigned to any patient who can walk away from the incident site; however, data about who can ambulate and who cannot were not recorded in the registry. Therefore, to approximate START in this study, minor/green was assigned to registry patients who did not meet criteria for dead/black or immediate/red and had Injury Severity Scores (ISSs) recorded as 10 or less. The ISS has been used elsewhere in the literature on minor trauma[8] and has been shown to correlate with START triage levels in a study of elderly trauma victims.[9] ISS is defined as the sum of the squares of the 3 highest Abbreviated Injury Scores and ranges from 1 to 75.[10] For this study, we used the ISS calculated by each trauma center locally for each patient and reported to the NTDB. The use of ISS is not clinically practical (it would be a rare field medic who calculates ISS as part of an initial assessment), but we believe it was a reasonable substitute in this analysis for the specific ability of a patient to walk at the incident scene. This approximation, in particular, was vetted with outside experts for this study. Additionally, we performed a sensitivity analysis for the chosen ISS threshold of 10, determining results when the threshold was decreased to 8 or increased to 12.

More specifically, to approximate START/JumpSTART performance in the field, we used the following approach to assign triage levels to registry data:

1. Any patients with initial pulse and initial respirations=0 → dead/black.
2. For all remaining patients:
   a. Any patient aged greater than or equal to 9 years and with respirations greater than 30 breaths/min or=0 → immediate/red
   b. Any patient aged 8 years or younger and with respirations greater than 45 or less than 15 breaths/min → immediate/red
   c. Any patient with initial pulse=0 → immediate/red
   d. Any patient with GCS-Verbal score less than 4 and GCS-Motor score less than 6 → immediate/red
3. For all remaining patients:
   a. If ISS greater than 10 → delayed/yellow
   b. All other patients (ISS ≤10) → minor/green

Fire Department of New York Method (FDNY)

The FDNY in collaboration with academic groups recently modified their disaster triage from a traditional START/JumpSTART approach to add an urgent/orange classification and to simplify pediatric rules. These modifications were intended to promote more useful risk stratification and to minimize over- and undertriage of moderately injured patients.[11] An outline of the FDNY triage levels appears in Table E1.

As in the START method (on which the FDNY method was based), certain elements of the FDNY scheme were difficult to apply to retrospective registry data. Ambulatory patients with minor injuries were not specifically noted in the data set, and we therefore assigned minor/green to patients who did not meet other triage level criteria and had an ISS of 10 or less. This cutoff at ISS less than or equal to 10 was further evaluated in a sensitivity analysis. Assignment to urgent/orange in the FDNY system depends on delayed assessment information not specifically stated in the data set. It was approximated by assigning any patient not meeting criteria for black or red to orange if he or she had significantly altered mental status (GCS score <14) or evidence of major injury (ISS >15).

Thus, to approximate the FDNY triage assignments in the field, we used the following specific approach in the registry data set:

1. Any patients with initial pulse and initial respirations of 0 are assigned dead/black.
2. For all remaining patients:
   a. If age is aged 0 years (infants younger than 1 year)→ immediate/red
   b. If respirations greater than 30 or less than 10 breaths/min → immediate/red
   c. If initial GCS-Motor score less than 6 (unable to follow commands) → immediate/red
3. For all remaining patients:
   a. If GCS-Total score less than 14 → urgent/orange
   b. If ISS greater than 15 → urgent/orange
4. For all remaining patients:
   a. If ISS greater than 10 → delayed/yellow
   b. Otherwise (ISS ≤10) → minor/green

CareFlight Triage Method

The CareFlight Triage Method resembles the START method and has 4 triage assignment levels. However, it is simpler than START and does not have pediatric rules. Like some other methods, its initial step is to separate ambulating patients. It then considers mental status and radial pulse to separate immediate/red from delayed/yellow patients, as shown in Figure E2.[12-14]

To approximate CareFlight triage assignments in the field, we used the following specific approach in the registry data set:

1. All patients with initial ISS less than or equal to 10 are assigned to minor/green and those with ISS greater than 10, to delayed/yellow.
2. If GCS-Motor score less than 6 (unable to follow commands) or pulse=0, patient is reassigned to immediate/red.
3. If GCS-Motor score less than 6 (unable to follow commands) and respirations=0, patient is reassigned to dead/black.

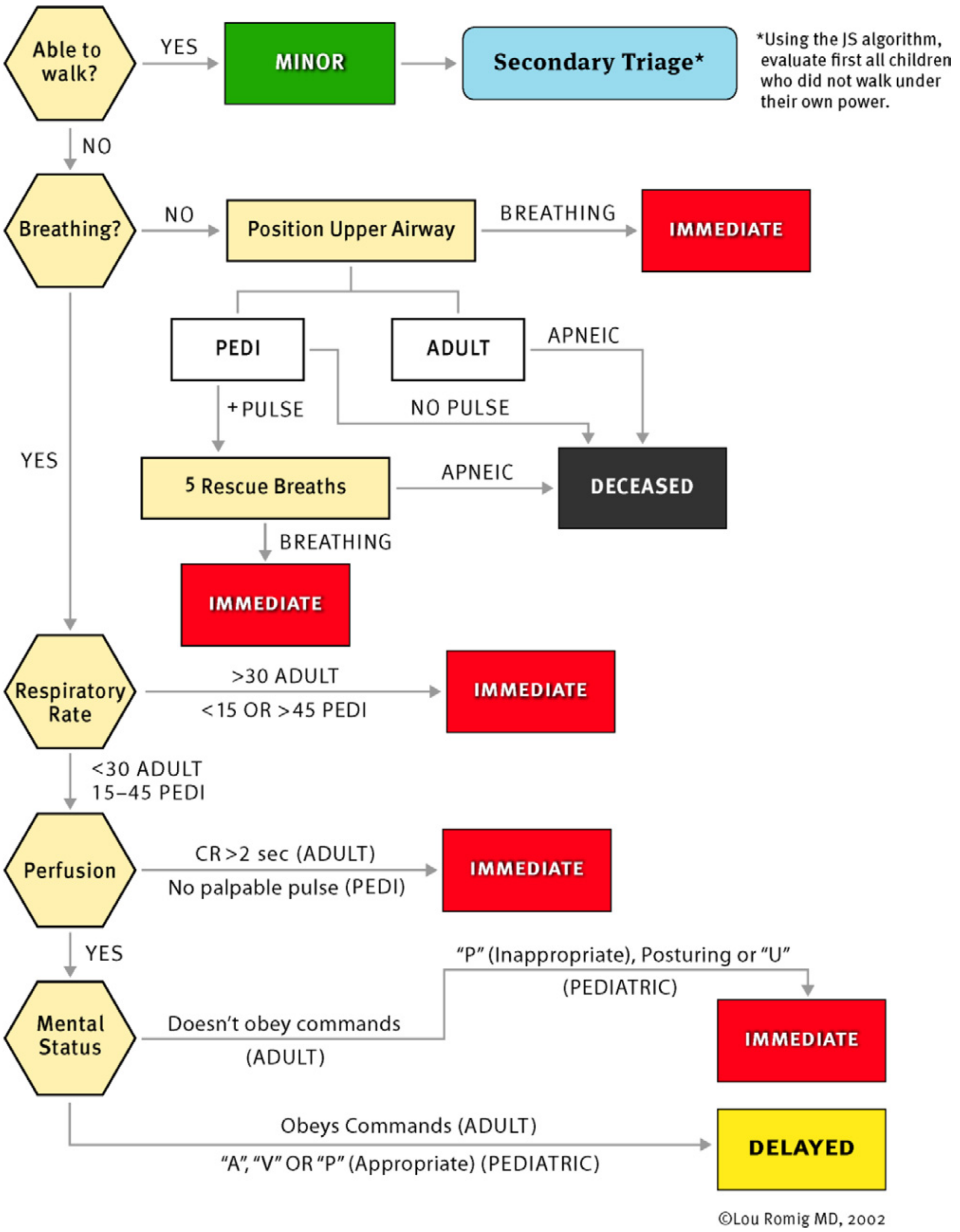

Figure E1. START/JumpSTART triage method. The figure is reproduced with the permission of Lou E. Romig, MD, from a source file available at http://www.jumpstarttriage.com/uploads/Combined_Algorithm.pdf, accessed July 5, 2012.

GCS Method

The GCS was not initially created for triage, but it is familiar to medical providers. Both the total GCS and its components have been found to be good predictors of outcomes.[15,16] An outline of GCS appears in Table E2. The initial scene GCS was recorded on most NTDB records and may be used directly to assign triage priority. This method is not specifically age dependent; however, the exact evaluation of GCS subcomponents is different in young children than in adults.[17,18] For this analysis, we used the scene-reported GCS, regardless of how it was calculated by the first responder.

Sacco Triage Method

This triage method was designed specifically for mass casualty and disaster situations. It was originally proposed for adults but was subsequently modified for assessment of children. The method has been validated against adult trauma registry data in patients with both blunt and penetrating trauma.[19,20] A simulation study using military medical registry data from adult patients in combat trauma situations found the Sacco triage method to have survival performance at least 22% better than that of START.[21] Recent work using the NTDB has also demonstrated the utility of the Sacco triage method for

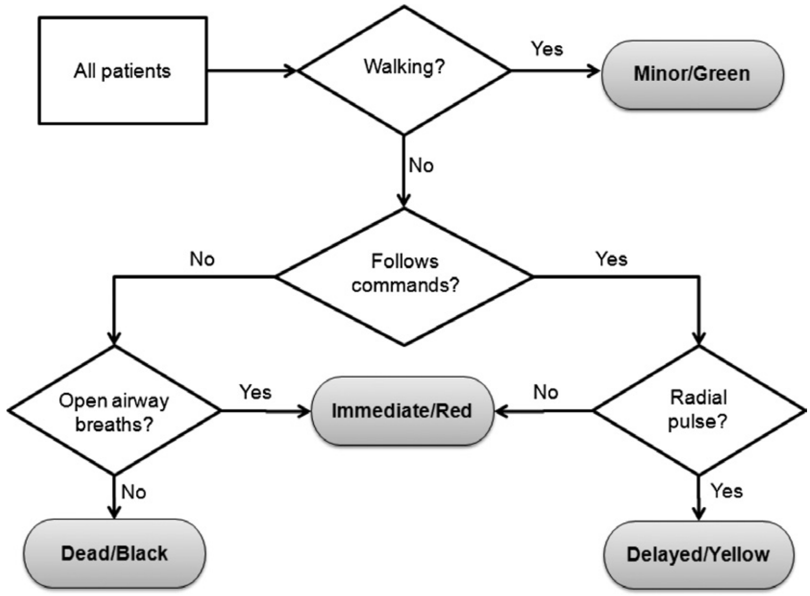

Figure E2. The CareFlight triage algorithm.[14]

Table E1. FDNY triage levels and criteria.[11]

Triage Level	Criteria
Dead/black	No pulse, no respirations, not ambulating, no response to 5 breaths with bag-valve-mask
Immediate/red	Not ambulating, and pulse present±respirations with: Respirations >30 or <10 breaths/min, and if 0 does respond to 5 breaths with bag-valve-mask, or: Unable to follow commands, or: Infant (age appears to be <12 mo)
Delayed/yellow	Not ambulating, pulse present, 30>respirations>10 breaths/min, and follows simple commands
Minor/green	Ambulatory, and breathing and mentating well
Urgent/orange (secondary assignment)	Patient initially assigned delayed/yellow or minor/green who subsequently is found to have signs of: Respiratory distress or failure, or: Altered mental status, or: Major injury to the head or torso

pediatric victims of blunt and penetrating trauma.[22] Although the Sacco triage method has more evidence base than other methods, its complexity and unfamiliarity have limited its adoption to date.[23]

The Sacco triage method uses a grid to score patients. Respirations, pulse, and motor response (similar to the GCS motor subscore) are used for a primary score totaling 0 to 12 points, which is then adjusted for patient age, yielding a final Sacco Score from –2 (dead) to 14 (healthy). The grid appears in Table E3.

To approximate the Sacco triage method's performance in the field, we used the following approach when assigning a Sacco Score to NTDB registry data:

1. All patients were assigned a respirations subscore 0 to 4 per Table E3, using their initial respiratory rate.
2. All patients were assigned a pulse subscore 0 to 4 per Table E3, using their initial pulse rate.
3. All patients were assigned a motor subscore 0 to 4, using their initial GCS-Motor score:
   a. GCS-Motor=6 → 4
   b. GCS-Motor=5 → 3
   c. GCS-Motor=4 → 2
   d. GCS-Motor=2 or 3 → 1
   e. GCS-Motor=1 → 0
4. All patients were assigned an age adjustment subscore per Table E3.
5. The respiration, pulse, motor, and age adjustment subscores were summed to give a final Sacco triage score between –2 and 14 for each patient.

Unadjusted Sacco Score

The Sacco Triage Score can be calculated without the age adjustment term. Indeed, this approach was used in its initial validation in adult patients.[19,20] Age adjustment was thought to add accuracy in predicting outcomes in pediatric and geriatric casualties at the expense of more complexity for medical personnel performing mass casualty triage. It is unclear whether the added accuracy of the age adjustment justifies its added complexity.

For this study, the Unadjusted Sacco Score was calculated in the same manner as the regular Sacco Score, simply omitting the age adjustment term. For victims of all ages, this approach results in scores ranging from 0 (dead) to 12 (healthy).

Additional Results Information

The detailed results for length-of-stay analysis (Table E4*A* and *B*) and for subgroup analysis (Table E5*A*, *B*, and *C*) are presented here.

Table E2. GCS scoring.*

GCS Subscore →	1	2	3	4	5	6
Eyes response	No eye opening	Opens to pain	Opens to voice	Opens spontaneously	NA	NA
Verbal response	Makes no sounds	Incomprehensible	Inappropriate	Confused	Oriented, normal	NA
Motor response	No movement	Extension	Flexion	Withdraws	Localizes	Obeys commands

NA, Not applicable.
*Final score is the sum of 3 subscores and ranges from 3 (coma/dead) to 15 (normal mental status).[24]

Table E3. Sacco triage method scoring grid.*

Sacco Subscore →	−2	−1	0	1	2	3	4
Respirations/min	NA	NA	0	1–9	36+	25–35	10–24
Heartbeat/min	NA	NA	0	1–40	41–60	121+	61–120
Motor response	NA	NA	No response	Extension or flexion	Withdraws	Localizes	Obeys commands
Age, y	75+	55–74	15–54	8–14	0–8	NA	NA

NA, Not applicable.
*For each of 4 rows (factors), a subscore is assigned. The sum of the 4 subscores gives a total Sacco Score of −2 to 14. Excluding the age adjustment gives an Unadjusted Sacco Score of 0 to 12. A higher score indicates a healthier patient.[20]

Table E4. Hospital and ED length of stay median and interquartile ranges as a function of triage assignments.

Triage Method	Assigned Triage Level	Patients	Hospital Length of Stay, Days		ED Length of Stay, Minutes	
			Median	Interquartile Range	Median	Interquartile Range
START, FDNY, CareFlight, and GCS methods						
START	Dead/black	2,307	0.05	0–1.08	26	7–100
	Immediate/red	79,853	4.00	1.29–11.43	152	75–262
	Delayed/yellow	126,373	4.80	2.40–9.00	225	130–353
	Minor/green	322,162	2.35	1.00–4.60	239	151–360
FDNY	Dead/black	2,307	0.05	0–1.08	26	7–100
	Immediate/red	78,111	3.92	1.22–11.34	151	75–260
	Urgent/orange	92,187	5.12	2.51–10.03	214	120–340
	Delayed/yellow	41,658	3.78	1.97–6.63	246	152–375
	Minor/green	316,432	2.38	1.01–4.62	240	151–360
CareFlight	Dead/black	6,698	1.86	0.10–11.17	75	27–161
	Immediate/red	64,833	4.03	1.31–11.72	150	75–260
	Delayed/yellow	131,505	4.84	2.42–9.12	223	127–350
	Minor/green	327,659	2.35	1.00–4.60	238	150–359
GCS	3	21,830	2.95	0.47–13.59	101	50–191
	4	2,374	5.41	1.24–17.03	113	62–202
	5	2,544	6.91	1.77–19.15	118	63–208
	6	4,451	6.06	1.78–16.60	125	65–215
	7	4,078	5.49	1.71–14.99	133	67–230
	8	4,356	5.02	1.68–13.34	144	74–246
	9	4,467	4.48	1.59–11.30	161	80–275
	10	5,459	4.30	1.61–10.96	166	89–286
	11	6,189	3.80	1.52–9.01	176	93–300
	12	9,537	3.48	1.44–8.03	188	103–309
	13	16,446	3.26	1.30–7.29	201	115–320
	14	58,204	2.85	1.14–5.97	225	136–347
	15	390,760	2.94	1.20–5.61	237	146–359
Sacco and Unadjusted Sacco triage methods						
Sacco Score	−2	154	0.38	0.01–1.85	92	17–195
	−1	341	0.21	0.01–2.83	62	10–155
	0	1,631	0.03	0–0.70	20	6–76
	1	448	0.32	0.02–3.41	52	16–130
	2	879	0.99	0.10–7.60	74	31–160
	3	1,839	2.88	0.56–15.88	99	48–183
	4	3,351	4.21	0.69–15.97	104	50–192
	5	3,494	3.58	0.62–15.64	105	55–191
	6	3,929	4.84	1.05–15.82	120	60–220
	7	6,566	4.99	1.24–15.12	135	67–240
	8	17,854	4.71	1.80–11.09	181	90–300
	9	18,774	4.65	1.91–10.18	197	105–320
	10	92,397	4.18	2.25–7.05	252	162–370
	11	115,992	3.38	1.51–6.64	236	143–361
	12	242,262	2.39	1.02–4.85	222	132–348
	13	16,301	1.55	0.83–2.84	197	127–291
	14	4,553	1.13	0.79–2.12	195	129–280
Unadjusted Sacco Score	0	2,228	0.04	0–0.95	24	7–93
	1	314	0.16	0.01–2.57	40	15–102
	2	545	0.94	0.08–7.19	67	23–140
	3	1,506	2.53	0.48–15.69	90	45–170
	4	3,472	4.67	0.78–16.71	102	49–191
	5	3,437	3.39	0.48–16.10	104	53–189
	6	3,033	5.31	0.98–17.37	103	55–186
	7	4,783	5.01	1.07–16.69	112	60–205
	8	13,327	4.62	1.26–14.54	134	68–234
	9	11,428	4.33	1.48–11.58	157	81–267
	10	37,143	3.70	1.49–7.81	200	112–320
	11	60,144	3.04	1.17–6.94	196	110–315
	12	389,335	2.92	1.20–5.52	240	149–362

Table E5. *A*, Age-specific subgroup analysis is reported as the AUC for primary outcome (mortality at hospital discharge) when predicted by each disaster triage method.*

	AUC (95% CI)			
Triage Method	**0–8 years**	**9–15 years**	**16–64 years**	**≥65 years**
START	0.931 (0.924–0.938)	0.942 (0.936–0.948)	0.888 (0.885–0.890)	0.791 (0.785–0.797)
FDNY	0.891 (0.880–0.902)	0.946 (0.941–0.951)	0.894 (0.892–0.897)	0.797 (0.791–0.803)
CareFlight	0.952 (0.946–0.958)	0.955 (0.949–0.960)	0.897 (0.894–0.899)	0.791 (0.785–0.797)
GCS	0.964 (0.954–0.974)	0.949 (0.936–0.93)	0.880 (0.876–0.883)	0.733 (0.726–0.740)
Sacco	0.961 (0.951–0.971)	0.961 (0.951–0.972)	0.900 (0.896–0.903)	0.733 (0.727–0.740)
Unadjusted Sacco	0.963 (0.953–0.972)	0.958 (0.945–0.970)	0.892 (0.889–0.896)	0.713 (0.706–0.720)

*The 95% CI upper and lower bounds for AUC are shown.

B, Trauma type–specific subgroup analysis is reported as the AUC for primary outcome (mortality at hospital discharge) when predicted by each disaster triage method.*

	AUC (95% CI), Trauma Type		
Triage Method	**Blunt**	**Penetrating**	**Burn**
START	0.826 (0.822–0.829)	0.926 (0.922–0.930)	0.86 (0.84–0.88)
FDNY	0.831 (0.827–0.834)	0.930 (0.927–0.934)	0.83 (0.81–0.86)
CareFlight	0.831 (0.828–0.835)	0.934 (0.931–0.938)	0.87 (0.85–0.89)
GCS	0.807 (0.803–0.811)	0.900 (0.893–0.906)	0.71 (0.68–0.75)
Sacco	0.877 (0.874–0.880)	0.916 (0.910–0.921)	0.84 (0.82–0.86)
Unadjusted Sacco	0.801 (0.797–0.806)	0.909 (0.903–0.915)	0.69 (0.66–0.73)

*The 95% CI upper and lower bounds for AUC are shown. There were 450,062 patients with blunt trauma, patients with 58,328 penetrating trauma, and 4,676 burn patients. For 17,629 records in this study, no trauma type was reported, and these records are not included in the results in this table.

C, Sex-specific subgroup analysis is reported as the AUC for primary outcome (mortality at hospital discharge) when predicted by each disaster triage method.*

	AUC (95% CI)	
Triage Method	**Male**	**Female**
START	0.851 (0.848–0.854)	0.827 (0.822–0.833)
FDNY	0.857 (0.854–0.860)	0.830 (0.825–0.836)
CareFlight	0.859 (0.856–0.862)	0.832 (0.826–0.838)
GCS	0.834 (0.830–0.838)	0.800 (0.793–0.808)
Sacco	0.892 (0.889–0.895)	0.865 (0.860–0.869)
Unadjusted Sacco	0.838 (0.834–0.843)	0.787 (0.780–0.794)

*The 95% CI upper and lower bounds for AUC are shown. There were 186,499 female patients and 343,491 male patients. An additional 705 records did not report the patient's sex and were not included in the results in this table.

REFERENCES

1. Romig L. Pediatric triage. A system to JumpSTART your triage of young patients at MCIs. *JEMS*. 2002;27:52-58, 60–53.
2. Mace S, Bern A. Needs assessment: are disaster medical assistance teams up for the challenge of a pediatric disaster? *Am J Emerg Med*. 2007;25:762-769.
3. Benson M, Koenig KL, Schultz CH. Disaster triage: START, then SAVE—a new method of dynamic triage for victims of a catastrophic earthquake. *Prehosp Disaster Med*. 1996;11:117-124.
4. Kahn CA, Schultz CH, Miller KT, et al. Does START triage work? an outcomes assessment after a disaster. *Ann Emerg Med*. 2009;54:424-430, 430.e421.
5. de Ceballos JP, Turegano-Fuentes F, Perez-Diaz D, et al. 11 March 2004: the terrorist bomb explosions in Madrid, Spain—an analysis of the logistics, injuries sustained and clinical management of casualties treated at the closest hospital. *Crit Care*. 2005;9:104-111.
6. Zoraster R. Disaster triage: is it time to stop START? *Am J Disaster Med*. 2006;1:7-9.
7. Navin DM, Sacco WJ, McCord TB. Does START triage work? the answer is clear! *Ann Emerg Med*. 55:579–580; author reply 580–581.
8. Hedges JR, Osterud HR, Mullins RJ. Adult minor trauma patients: good outcome in small hospitals. *Ann Emerg Med*. 1992;21:402-406.
9. Hong R, Sierzenski PR, Bollinger M, et al. Does the simple triage and rapid treatment method appropriately triage patients based on trauma Injury Severity Score? *Am J Disaster Med*. 2008;3:265-271.
10. Baker SP, O'Neill B, Haddon W Jr, et al. The Injury Severity Score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma*. 1974;14:187-196.
11. The New York City Pediatric Disaster Coalition and the New York City Department of Health and Mental Hygiene. Recommendations for NYC Pediatric Disaster Triage and Transport. New York; 2009:1–40.
12. Wallis LA, Carley S. Comparison of paediatric major incident primary triage tools. *Emerg Med J*. 2006;23:475-478.
13. Garner A, Lee A, Harrison K, et al. Comparative analysis of multiple-casualty incident triage algorithms. *Ann Emerg Med*. 2001;38:541-548.

14. Nocera A, Garner A. An Australian mass casualty incident triage system for the future based upon triage mistakes of the past: the Homebush Triage Standard. *Aust N Z J Surg*. 1999;69:603–608.
15. Thompson DO, Hurtado TR, Liao MM, et al. Validation of the Simplified Motor Score in the out-of-hospital setting for the prediction of outcomes after traumatic brain injury. *Ann Emerg Med*. 2011;58:417–425.
16. Udekwu P, Kromhout-Schiro S, Vaslef S, et al. Glasgow Coma Scale score, mortality, and functional outcome in head-injured patients. *J Trauma*. 2004;56:1084-1089.
17. Newgard CD, Cudnik M, Warden CR, et al. The predictive value and appropriate ranges of prehospital physiological parameters for high-risk injured children. *Pediatr Emerg Care*. 2007;23:450–456.
18. Cohen J. Interrater reliability and predictive validity of the FOUR Score Coma Scale in a pediatric population. *J Neurosci Nurs*. 2009;41:261-267; quiz 268–269.
19. Sacco WJ, Navin DM, Waddell RK 2nd, et al. A new resourceconstrained triage method applied to victims of penetrating injury. *J Trauma*. 2007;63:316-325.
20. Sacco WJ, Navin DM, Fiedler KE, et al. Precise formulation and evidence-based application of resource-constrained triage. *Acad Emerg Med*. 2005;12:759-770.
21. Navin DM, Sacco WJ, McGill G. Application of a new resource constrained triage method to military-age victims. *Mil Med*. 2009; 174:1247-1255.
22. Cross K, Cicero M. Independent application of the Sacco disaster triage method to pediatric trauma patients. *Prehosp Disaster Med*. 2012;27:306-311.
23. Jenkins JL, McCarthy ML, Sauer LM, et al. Mass-casualty triage: time for an evidence-based approach. *Prehosp Disaster Med*. 2008;23:3-8.
24. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet*. 1974;2:81-84.