Published in final edited form as:

Prehosp Emerg Care. 2015 ; 19(2): 267–271. doi:10.3109/10903127.2014.959222.

A Consensus-Based Gold Standard for the Evaluation of Mass Casualty Triage Systems

E. Brooke Lerner, PhD¹, Courtney H. McKee, BS¹, Charles E. Cady, MD¹, David C. Cone, MD², M. Riccardo Colella, DO, MPH¹, Arthur Cooper, MD, MS³, Phillip L. Coule, MD⁴, Julio R. Lairet, DO⁵, J. Marc Liu, MD, MPH¹, Ronald Pirralo, MD, MHSA¹, Scott M. Sasser, MD⁵, Richard Schwartz, MD⁴, Greene Shepherd, PharmD⁶, and Raymond E. Swienton, MD⁷

¹Medical College of Wisconsin, Milwaukee, WI

²Yale University School of Medicine, New Haven, CT

³Columbia University Medical Center Affiliation at Harlem Hospital, New York, NY

⁴Georgia Regents University, Augusta, GA

⁵Emory University School of Medicine, Atlanta, GA

⁶University of North Carolina, Eshelman School of Pharmacy, Asheville, NC

⁷University of Texas Southwestern Medical Center, Dallas, Texas

Abstract

Introduction—Accuracy and effectiveness analyses of mass casualty triage systems are limited because there are no gold standard definitions for each of the triage categories. Until there is agreement on which patients should be identified by each triage category, it will be impossible to calculate sensitivity and specificity or to compare accuracy between triage systems.

Objective—To develop a consensus-based, functional gold standard definition for each mass casualty triage category.

Methods—National experts were recruited through the lead investigators' contacts and their suggested contacts. Key informant interviews were conducted to develop a list of potential criteria for defining each triage category. Panelists were interviewed in order of their availability until redundancy of themes was achieved. Panelists were blinded to each other's responses during the interviews. A modified Delphi survey was developed with the potential criteria identified during the interview and delivered to all recruited experts. In the early rounds, panelists could add, remove, or modify criteria. In the final rounds edits were made to the criteria until at least 80% agreement was achieved.

Results—Thirteen national and local experts were recruited to participate in the project. Six interviews were conducted. Three rounds of voting were performed, with 12 panelists participating

Address for Correspondence and reprints: E. Brooke Lerner, Ph.D.; Department of Emergency Medicine; Medical College of Wisconsin; 9200 W. Wisconsin Ave.; Milwaukee, WI 53226, **Phone:** (414) 805-0113, **Fax:** (414) 805-6464, eblerner@mcw.edu.

Conflicts of Interest: The authors have no relevant financial conflicts of interest to report.

Presented: National Association of EMS Physicians Annual Meeting, Bonita Springs, FL; January 2013

in the first round, 12 in the second round, and 13 in the third round. After the first two rounds, the criteria were modified according to respondent suggestions. In the final round, over 90% agreement was achieved for all but one criterion. A single e-mail vote was conducted on edits to the final criterion and consensus was achieved.

Conclusion—A consensus-based, functional gold standard definition for each mass casualty triage category was been developed. These gold standard definitions can be used to evaluate the accuracy of mass casualty triage systems after an actual incident, during training, or for research.

Introduction

Mass casualty incidents (MCI) occur when the medical needs of multiple victims overwhelm the available medical resources. These events range from motor vehicle crashes to large-scale disasters. Due to resource constraints during an MCI, healthcare providers must alter their approach from providing maximal care for a single victim to providing the best care for the greatest number of victims. Mass casualty triage systems exist to assist medical providers in prioritizing MCI patients for treatment and transport so that finite resources can be appropriately allocated. Accurately triaging MCI victims may improve survival and outcomes.¹

Multiple mass casualty triage systems are currently used throughout the United States and around the world.² The Model Uniform Core Criteria (MUCC) for Mass Casualty Triage was developed in an effort to standardize mass casualty triage across the United States and improve interoperability.³ The development of MUCC was based on the available literature, but that process also revealed the paucity of literature in this area.⁴ The lack of a defined gold standard for what is the “correct” triage category given a patient’s ultimate resource use is one barrier to advancing the science of mass casualty triage, as it is impossible for investigators to effectively evaluate or compare existing mass casualty triage systems. Most studies of triage accuracy are based upon the expected answer as assigned by an experienced user of the triage system, rather than the actual medical resource use by the patient.⁵⁻¹¹ If a functional gold standard definition for each mass casualty triage category existed, it could be objectively determined if a patient should have received the category he or she was assigned.

There have been a few studies that have attempted to evaluate triage systems using a functional gold standard. In 2001, Garner and colleagues modified previously developed criteria for identifying critically injured trauma patients and used it to determine the sensitivity and specificity of a variety of mass casualty triage systems in identifying patients who should be triaged as “immediate”.¹² Wallis and Carley replicated this methodology in a study of pediatric triage and additionally used an Injury Severity Score of 15 or greater as a secondary outcome measure for identifying immediate patients.^{13,14} However, the outcome measure used in these studies was limited to the evaluation of one triage category (immediate) when the objective of mass casualty triage is to sort patients into five categories that prioritize them for treatment and transport. In 2007, Kahn et al. expanded this work by further modifying the trauma criteria used by Garner and colleagues to study the accuracy of mass casualty triage following a train crash disaster. Kahn and his colleagues also expanded on Garner’s work by defining the triage category “delayed” as being the correct assignment

if the patient was admitted to the hospital for at least 24 hours.¹⁵ To our knowledge, the outcome definitions used by Kahn and his colleagues were author-defined and were not vetted by other experts in the field or externally validated.

Currently, there is no uniformly accepted gold standard for defining mass casualty triage categories, and this hinders the scientific advancement of mass casualty triage. The objective of this study was to develop a consensus-based, functional gold standard definition for each mass casualty triage category defined in the MUCC (i.e., immediate, delayed, minimal, expectant, dead).

METHODS

The literature on field trauma triage and mass casualty triage was reviewed and a list of previously used triage category criteria was generated. Simultaneously, local and national experts were recruited to participate in the project through the lead author's contacts and their suggested contacts. To be considered for the panel participants needed to have an established career in disaster response at the national or local level. Each person was sent an e-mail inviting him or her to participate on the panel.

Semi-structured key informant interviews using the findings of the literature review as a basis for discussion were then conducted to build an initial list of potential criteria for consideration as the gold standard definitions for each of the mass casualty triage categories. Participants were interviewed by phone or in person in order of their availability by the two leaders of the project (EBL and CHM). The interviewees were blinded to the responses of the other panelists. The interviewers read each potential criteria for each category and asked if it should be used as a gold standard or modified. The interviewees were also asked to suggest any additional criteria. After each interview, the additional criteria were added to the list but none of the criteria were removed. The interviews continued until no new criteria were suggested.

After the interviews were completed, a modified Delphi technique was used to refine and reach consensus on the list of criteria obtained during the interviews for each triage category. The Delphi survey method is used to solicit the opinions of diverse groups of experts and to provide an objective measure of consensus.^{16,17} Our survey was delivered to the panelists using Survey Monkey®. In the initial voting rounds, participants could add, remove, or modify criteria. After each voting round, edits were made to the criteria and participants were again asked to add, remove, or modify criteria. The voting rounds continued until no new criteria were suggested and at least 80% agreement was achieved for each of the criteria.

RESULTS

Thirteen local and national experts agreed to participate on the panel (Table 1). After six key informant interviews were conducted, it appeared that no new criteria were being suggested. These criteria were incorporated into a modified Delphi survey and distributed online to all thirteen panelists. Three rounds of voting were performed (Table 1). Twelve panelists participated in the first round, 12 in the second round, and 13 in the third round. After the

first two rounds, the criteria were modified according to respondent suggestions. In the final round, over 90% agreement was achieved for all but one criterion. This final criterion was further edited and a single e-mail vote was conducted to determine consensus. Agreement was obtained from 100% of the panelists for the edits to this criterion. A consensus-based gold standard definition was successfully developed for each of the mass casualty triage categories (Figure 1).

DISCUSSION

This project developed a list of EMS and hospital-based treatments that can be used to define the specific mass casualty triage category a patient should have been assigned. These gold standard definitions are not intended for use during clinical practice, but are intended to be used for quality improvement or research. That is, they define which triage category is the “right” one after everything is known about the victim’s medical needs and diagnoses. Mass casualty triage systems were developed as a tool to assist providers in recognizing patient needs based on limited information and diagnostic equipment without the benefit of hindsight. Regardless of how robust a mass casualty triage system is, it is unlikely to be 100% accurate when used in a real-world scenario. This project allows researchers and other evaluators to start making objective and reproducible assessments of mass casualty triage tools and to further the science in this field to ultimately improve patient outcomes.

The criteria in our functional gold standard definitions are discrete interventions or conditions that are likely to be well documented in a patient’s medical record. Therefore, this quality improvement/research tool can likely be used even by investigators with limited resources.

We approached this work from an all-hazards perspective, and therefore considered clinical scenarios involving chemical and radiologic contamination. However, we did not consider decontamination triage and recommend it as a topic of further study.

We also postulated a mass casualty scenario with a single, centralized scene and did not take resource limitations into account in developing our definitions. Part of the process of assigning the expectant category requires providers to account for resource limitations. Therefore, our gold standard for the expectant category only identifies those patients who would be considered expectant under any condition. The user will need to evaluate the scenario they are reviewing to determine the resource limitations and make any adjustments that might need to be made to the definition for this category based on those limitations. It seems likely that these decisions will need to be made on a case by case basis, but further work in this area may help to improve this process.

This project has several limitations. The gold standard definitions were developed by a panel of local and national experts. While this panel had a variety of backgrounds and experiences, it may not have represented a broad enough view on the topic. Ideally, this work will need to be vetted by a broader range of stakeholders and externally validated, especially since the Delphi process may suppress outlying opinions. Further, these definitions have not been operationalized and used during an actual medical record review. We speculate that the

identified criteria can be easily obtained for all patients, but this needs to be formally demonstrated along with determining the inter- and intra-rater reliability.

CONCLUSION

A consensus-based, functional gold standard definition for each mass casualty triage category was developed. These gold standard definitions can be used to evaluate the accuracy of mass casualty triage systems after an actual incident or during training; and will be a valuable tool for advancing the science of mass casualty triage.

Acknowledgments

Source of Funding: Courtney McKee was awarded an Institutional National Resource Service Award in Aging and Injury Research (National Institute on Aging Grant No. 2T35AG029793–06) to fund her work on this project during the summer of 2012 as well as a travel stipend to offset the costs of presenting this project at a national meeting in Bonita Springs, FL. No other external support was received for this project.

References

1. Frykberg ER. Medical management of disasters and mass casualties from terrorist bombings: how can we cope? *J Trauma*. Aug; 2002 53(2):201–212. [PubMed: 12169923]
2. Lerner EB, Schwartz RB, Coule PL, et al. Mass casualty triage: an evaluation of the data and development of a proposed national guideline. *Disaster medicine and public health preparedness*. Sep; 2008 2(Suppl 1):S25–34. [PubMed: 18769263]
3. Model uniform core criteria for mass casualty triage. *Disaster medicine and public health preparedness*. Jun; 2011 5(2):125–128. [PubMed: 21685308]
4. Lerner EB, Cone DC, Weinstein ES, et al. Mass casualty triage: an evaluation of the science and refinement of a national guideline. *Disaster medicine and public health preparedness*. Jun; 2011 5(2):129–137. [PubMed: 21685309]
5. Lerner EB, Schwartz RB, Coule PL, Pirrallo RG. Use of SALT Triage in a Simulated Mass-Casualty Incident. *Prehosp Emerg Care*. Jan-Mar;2010 14(1):21–25. [PubMed: 19947863]
6. Deluhery MR, Lerner EB, Pirrallo RG, Schwartz RB. Paramedic accuracy using SALT triage after a brief initial training. *Prehosp Emerg Care*. Oct-Dec;2011 15(4):526–532. [PubMed: 21591923]
7. Baez AA, Sztajnkrzyer MD, Smester P, Giraldez E, Vargas LE. Effectiveness of a simple Internet-based disaster triage educational tool directed toward Latin-American EMS providers. *Prehosp Emerg Care*. Apr-Jun;2005 9(2):227–230. [PubMed: 16036851]
8. Risavi BL, Salen PN, Heller MB, Arcona S. A two-hour intervention using START improves prehospital triage of mass casualty incidents. *Prehosp Emerg Care*. Apr-Jun;2001 5(2):197–199. [PubMed: 11339732]
9. 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*. Oct-Dec;2009 13(4):536–540. [PubMed: 19731169]
10. Sanddal TL, Loyacono T, Sanddal ND. Effect of JumpSTART training on immediate and short-term pediatric triage performance. *Pediatr Emerg Care*. Nov; 2004 20(11):749–753. [PubMed: 15502656]
11. Kilner T. Triage decisions of prehospital emergency health care providers, using a multiple casualty scenario paper exercise. *Emerg Med J*. Jul; 2002 19(4):348–353. [PubMed: 12101157]
12. Garner A, Lee A, Harrison K, Schultz CH. Comparative analysis of multiple-casualty incident triage algorithms. *Annals of emergency medicine*. Nov; 2001 38(5):541–548. [PubMed: 11679866]
13. Wallis LA, Carley S. Comparison of paediatric major incident primary triage tools. *Emerg Med J*. Jun; 2006 23(6):475–478. [PubMed: 16714518]
14. Wallis LA, Carley S. Validation of the Paediatric Triage Tape. *Emerg Med J*. Jan; 2006 23(1):47–50. [PubMed: 16373803]

15. Kahn CA, Schultz CH, Miller KT, Anderson CL. Does START triage work? An outcomes assessment after a disaster. *Annals of emergency medicine*. Sep; 2009 54(3):424–430. 430, e421. [PubMed: 19195739]
16. Goodman CM. The Delphi technique: a critique. *Journal of advanced nursing*. Nov; 1987 12(6): 729–734. [PubMed: 3320139]
17. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *Journal of advanced nursing*. Oct; 2000 32(4):1008–1015. [PubMed: 11095242]

Figure 1.**Gold Standard Definitions for Each Mass Casualty Triage Category**

Note: additional criteria may need to be added to the Expectant category based on resource availability during the MCI being evaluated (real-world or simulated).

Table 1

Expert Panel Participants

Participant	Institution	Expertise	Interviewed	Voted Round 1	Voted Round 2	Voted Round 3
Charles Cady, MD	Medical College of Wisconsin	EM physician, EMS		X	X	X
M. Riccardo Colella, DO	Medical College of Wisconsin	EM physician, EMS		X	X	X
David Cone, MD	Yale University	EM physician, EMS, disaster	X	X	X	X
Arthur Cooper, MD	Columbia University	Pediatric surgeon, disaster, EMS		X	X	X
Phillip Coule, MD	Georgia Regents University	EM physician, EMS, disaster	X	X	X	X
Julio Lairer, DO	Emory University	EM physician, military, disaster, EMS		X	X	X
J. Marc Liu, MD	Medical College of Wisconsin	EM physician, EMS, disaster	X	X	X	X
Ronald Pirrallo, MD	Medical College of Wisconsin	EM physician, EMS, disaster		X		X
Scott M. Sasser, MD	Emory University	EM physician, EMS			X	X
Richard Schwartz, MD	Georgia Regents University	EM physician, disaster, military	X	X	X	X
Greene Shepherd, PharmD	University of North Carolina	Pharmacist, toxicology	X	X	X	X
Raymond Swienton, MD	University of Texas Southwestern	EM physician, disaster		X	X	X
Gamunu Wijetunge	NHTSA	EMS provider, disaster policy	X	X	X	X

-EMS – emergency medical services, EM- emergency medicine