RESEARCH REPORT

Use of Simulation in Full-Scale Exercises for Response to Disasters and Mass-Casualty Incidents: A Scoping Review

Jen Heng Pek, MSc;[1] Li Juan Joy Quah, MSc;[2] Martina Valente, PhD;[3,4] Luca Ragazzoni, PhD;[3] Francesco Della Corte[3]

1. Department of Emergency Medicine, Sengkang General Hospital, Singapore
2. Department of Emergency Medicine, Singapore General Hospital, Singapore
3. CRIMEDIM – Center for Research and Training in Disaster Medicine, Humanitarian Aid and Global Health, Università del Piemonte Orientale, Novara, Italy
4. Department for Sustainable Development and Ecological Transition, Università del Piemonte Orientale, Vercelli, Italy

Correspondence:
Jen Heng Pek
Sengkang General Hospital
110 Sengkang E Way, Singapore 544886
E-mail: pek.jen.heng@singhealth.com.sg

Conflicts of interest/funding: All authors do not have any financial disclosures or conflicts of interest.

Keywords: disaster; mass casualty; simulation

Abbreviation:
WHO: World Health Organization

Received: September 9, 2023
Revised: October 7, 2023
Accepted: October 18, 2023

doi:10.1017/S1049023X2300660X

Abstract

Disasters or mass-casualty incidents are uncommon events. The use of simulation is an ideal training modality in full-scale exercises as it immerses the participants in a replication of the actual environment where they can respond to simulated casualties in accordance with existing protocols.

The objective of this scoping review is to answer the research question: "How effective is simulation, as assessed in full-scale exercises, for response to disasters and mass-casualty incidents world-wide?" Studies on full-scale exercises, as defined in World Health Organization (WHO) simulation exercise toolbox, that were published in peer-reviewed journals using the English language from 2001 through 2021 were included. Twenty studies were included from searching PubMed, Embase, and Web of Science. Simulated casualties were the most common simulation modality. Using Kirkpatrick's levels of evaluation to synthesize the data, simulation was reported to be generally effective and mostly demonstrated at the levels of learning of individuals and/or systems, as well as reaction of individuals. Evaluations at levels of behavior and results were limited due to the uncommon nature of disasters and mass-casualty incidents. However, evaluation outcomes across the full-scale exercises were varied, leading to the inability to consolidate effectiveness of simulation into a single measure. It is recommended for best evidence-based practices for simulation to be adhered to in full-scale exercises so that the trainings could translate into better outcomes for casualties during an actual disaster or mass-casualty incident. In addition, the reporting of simulation use in full-scale exercises should be standardized using a framework, and the evaluation process should be rigorous so that effectiveness could be determined and compared across full-scale exercises.

Pek JH, Quah LJJ, Valente M, Ragazzoni L, Della Corte F. Use of simulation in full-scale exercises for response to disasters and mass-casualty incidents: a scoping review. ***Prehosp Disaster Med.*** **2023;38(6):792–806.**

Introduction

Disasters or mass-casualty incidents are uncommon events. Responders must be properly trained in order to execute protocols efficiently and treat casualties effectively. This training needs to occur at an individual level for competency of personal skills, an organizational level for capability of agency response, and a systemic level for coordination of overall efforts. According to the World Health Organization (WHO; Geneva, Switzerland) simulation exercise toolbox, there are four types of simulation exercises – tabletop exercise, drill, functional exercise, and full-scale exercise. Table-top exercise is a facilitated discussion of a scenario in a low-stress environment to identify gaps in the response protocol so that improvements can be made. Drill is a supervised activity performed as realistically as possible in order to practice and perfect a particular aspect of the response protocol. Functional exercise is an extensive assessment of the various aspects of an organization's response protocol in a realistic simulation.[1] Full-scale exercise is the most complex and realistic, as it includes the mobilization of emergency personnel, equipment, and resources across different organizations in order to evaluate the emergency management system in a highly stressful environment which replicates the actual response conditions.[1] Therefore, the most appropriate simulation exercise should be selected based on the objectives of the training, as well as the involvement of individuals, organizations, or systems taking part in it.

Use of simulation in disaster or mass-casualty incident is particularly relevant for two reasons – first, a disaster or a mass-casualty incident is uncommon and simulation is useful as it provides exposure and learning opportunities to uncommon events; and second, a simulated environment resembles actual response environment, allowing learners to be immersed in an experience which facilitates their learning.[2–4] When taking part in the simulation, learners can acquire important skills such as interpersonal communication, teamwork, leadership, decision making, prioritization, and stress management – all important when responding to a disaster or mass-casualty incident.[5] Learners have also reported perceived self-efficacy and preparedness, as well as increased self-assessed knowledge, confidence, and skills to manage a disaster or mass-casualty incident following simulation exercises.[6]

While simulation is beneficial as an educational tool, the current evidence of its application in full-scale exercises has largely been based on individual reports. Given the immense amount of time, resources, and cost required for the planning, conducting, and after-action-review of full-scale exercises, there is a need to synthesize the available evidence on the effectiveness of simulation so that recommendations can be made for best practices surrounding the use of simulation in full-scale exercises, and suggestions can be made for future works. Ultimately, when training needs are met and learners benefit from the learning experience, the preparedness of the emergency management system should translate to providing better care and achieving improved outcomes for casualties during the response to an actual disaster or mass-casualty incident.

Therefore, the objective of this study is to perform a scoping review on the effectiveness of simulation used in full-scale exercises for response to disasters and mass-casualty incidents world-wide in order to answer the following research question: "How effective is simulation, as assessed in full-scale exercises, for response to disasters and mass-casualty incidents world-wide?" Through this, the authors hope to inform readers how effectiveness has been assessed as reported by individual studies, as well as what are the levels of effectiveness and overall effectiveness as synthesized by this scoping review.

Report

Methods

A scoping review was carried out in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-SCR).[7]

Eligibility Criteria—The inclusion criteria were studies on full-scale exercises, as defined in the WHO simulation exercise toolbox, for response to disasters or mass-casualty incidents.[1] Three key components of this definition were used to assess for eligibility, and they were: (1) simulation of a real event and response conditions, (2) mobilization of personnel, equipment, and resources across multiple organizations, as well as (3) evaluation of the plan, coordination, and capability of the emergency management system. The studies had to be published in peer-reviewed journals using the English language from 2001 through 2021. The exclusion criteria were studies on tabletop exercises, drills, and functional exercises, as well as studies that mentioned full-scale exercises but they were not based on the definition in WHO simulation exercise toolbox. Studies involving mass-casualty incidents in military operations were also excluded as the combat nature demands a routine response to mass-casualty incidents in an austere environment which would be different from the non-military context.

Information Sources and Search Strategy—The studies were obtained by searching PubMed (National Center for Biotechnology Information, National Institutes of Health; Bethesda, Maryland USA), Embase (Elsevier; Amsterdam, Netherlands), and Web of Science (Clarivate Analytics; London, United Kingdom) with the help of a university librarian. The search was conducted on July 10, 2022 (Supplementary Material Table 1 for the detailed search strategy; available online only).

Study Selection Process—Two independent reviewers screened for relevant studies using the title and abstract. Short-listed studies were then assessed for eligibility using the full text by the same two independent reviewers. Any conflict was resolved by discussion between the two reviewers to reach a consensus. The reason for excluding studies was recorded.

Data Collection and Analysis—Data were extracted from the studies in a standardized form by one reviewer and cross-checked by a second reviewer before qualitative analysis. The following data were collected:

- Type of full-scale exercise for response to disaster or mass-casualty incidents (eg, single versus multiple agencies, natural versus man-made disaster, number of casualties, or hazardous material);
- Type of personnel involved in full-scale exercise (eg, doctors, nurses, paramedics, or students);
- Type of simulation modality used (eg, low- versus high-fidelity, standardized patients versus mannequin versus task trainers versus hybrid);
- Objectives (eg, teaching knowledge or skill) of full-scale exercise;
- Briefing and debriefing with personnel involved in full-scale exercise; and
- Evaluation for the effectiveness of full-scale exercise (eg, reaction, learning, behavior, results based on Kirkpatrick's levels of evaluation8) and reported evaluation outcome.

Results

Among 5,488 articles screened, 20 studies met the inclusion criteria and were included in the scoping review (Figure 1).

Study Characteristics—All the included studies were observational, mostly published after 2010 (n = 16; 80%); Figure 2. The studies originated from nine countries with the majority (n = 8; 40%) coming from the United States, followed by Italy (n = 3; 15%). One-quarter of the studies (n = 5; 25%) were published in the journal *Prehospital and Disaster Medicine*. Characteristics of the included studies are summarized in Table 1.[9–28]

Scenarios of Full-Scale Exercises—Five (25%) involved transportation accidents – aircraft (n = 2), vehicular (n = 2), and train (n = 1). Five (25%) involved hazardous materials – chemical (n = 3), biological (n = 1,) and one with no further details provided. Four (20%) involved natural disasters – earthquakes (n = 2), tornadoes (n = 1), and tsunamis (n = 1). Two (10%) involved terrorists and two (10%) involved active shooters. One (5%) involved a pandemic from Avian flu. While the majority (n = 16; 80%) were single-site exercises, there were four (20%) multi-site exercises, ranging from two to four sites (Supplementary Material Table 2 for details of the scenarios for full-scale exercises; available online only).

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

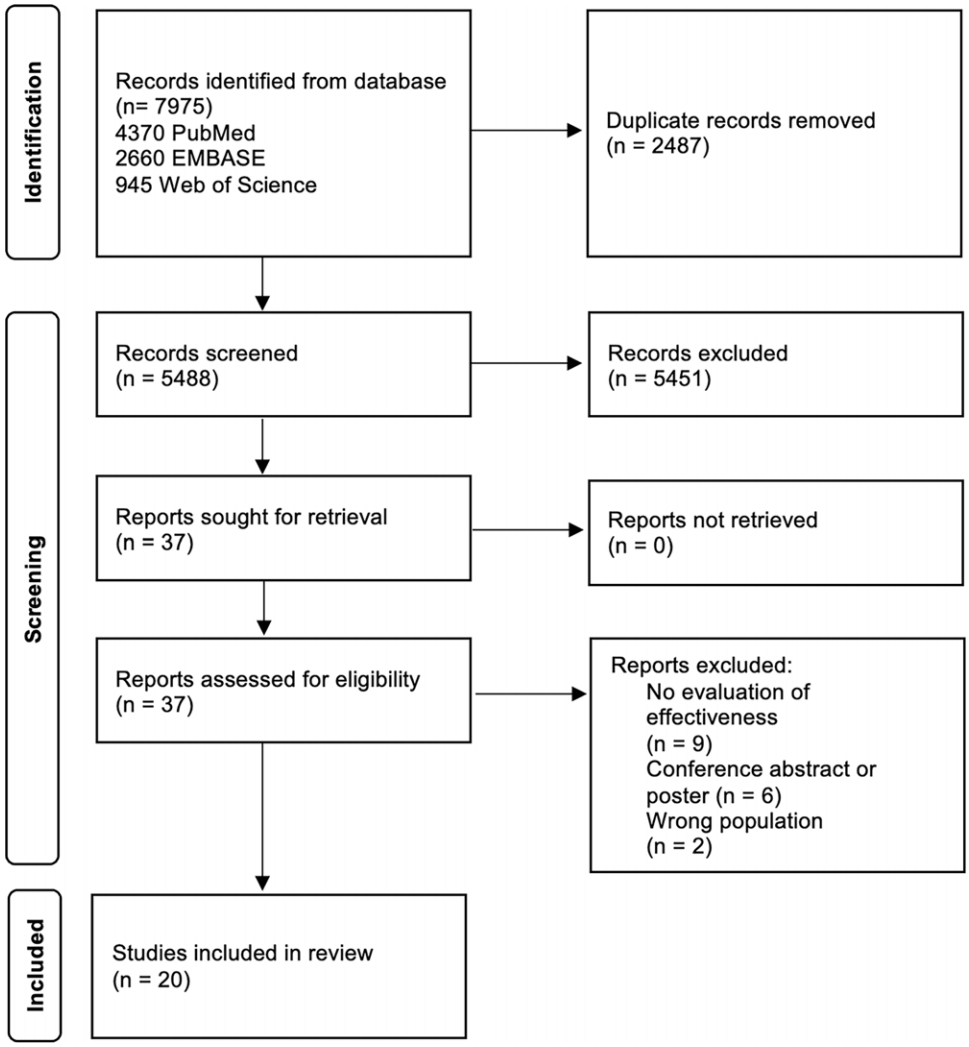

Figure 1. PRISMA Flow Diagram.

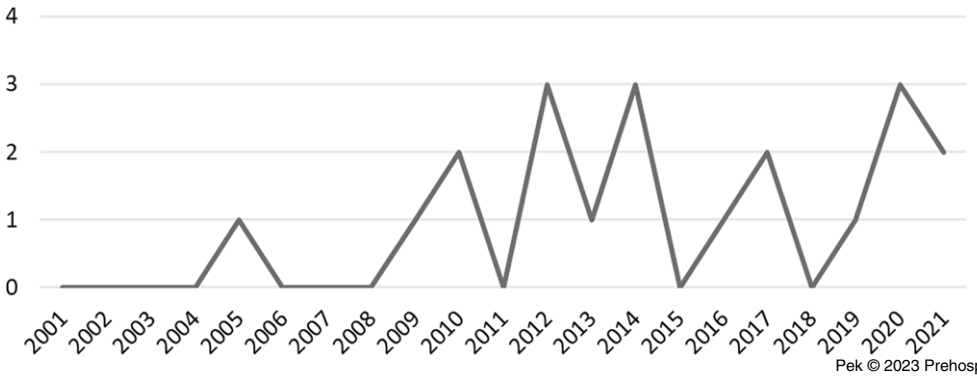

Figure 2. The Trend in the Number of Studies Over the Years.

Agencies and Personnel Involved in Full-Scale Exercises—At least two agencies were involved in the full-scale exercises of the included studies (Table 2). Emergency Medical Services (n = 14; 70%), fire department (n = 13; 65%), hospital (n = 13; 65%), and police department (n = 12; 60%) were most often involved. The personnel involved in the full-scale exercises included participants, observers, controllers, evaluators, simulated casualties, as well as administrative or logistical support staff. However, the exact numbers of personnel involved were often not reported, but when reported, the number ranged from 102 to 990.

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Publication Year	First Author (Last Name)	Country	Journal
2005	Alexander	Canada	Academic Emergency Medicine
2009	Lenz	Germany	Prehospital and Disaster Medicine
2010	Gryth	Sweden	Prehospital and Disaster Medicine
2010	Ingrassia	Italy	Journal of Emergency Medicine
2012	Rutty	UK	Forensic Science, Medicine, and Pathology
2012	Klima	USA	Journal of Trauma and Acute Care Surgery
2012	Radestad	Sweden	Scandinavian Journal of Trauma, Resuscitation, and Emergency Medicine
2013	Austin	USA	Journal of Pediatric Nursing
2014	Shah	USA	Prehospital and Disaster Medicine
2014	Austin	USA	Journal of Emergency Nursing
2014	Djalali	Italy	Prehospital and Disaster Medicine
2016	Daniel	USA	Prehospital and Disaster Medicine
2017	Wenham	Australia	Clinical Teacher
2017	Saber	USA	Journal of Continuing Education in Nursing
2019	McElroy	USA	Surgery
2020	Lenz	USA	American Journal of Disaster Medicine
2020	Carenzo	Italy	Disaster Medicine and Public Health Preparedness
2020	Sheikhbardsiri	Iran	Journal of Public Health Management and Practice
2021	Foo	Taiwan	Emergency Medicine International
2021	Innis	Canada	Journal of Nursing Education

Table 1. Characteristics of Included Studies

Objectives and Briefings of Full-Scale Exercises—The objectives of the full-scale exercises were mentioned in 15 (75%) studies (Table 3) and included one or more of the following domains:

- Disaster management system standards and gaps;
- Disaster knowledge, skill, and attitude (eg, triage methods, casualty management);
- Disaster roles and responsibilities;
- Disaster crisis resource management, leadership, and teamwork;
- Disaster collaboration, coordination, and communication;
- Disaster workflow and protocol (eg, registration, decontamination, or evacuation); and/or
- Benefits of exercise (eg, gain knowledge or confidence).

However, briefings to the participants were only conducted in eight (40%) studies, which covered one or more of the following aspects:

- General plan of the day;
- Safety briefing;
- Education about roles;
- Familiarization with equipment or team members; and/or
- Resources available.

In two exercises, participants only had information on the date of the full-scale exercises.

Simulation Modalities Used for Full-Scale Exercises—A simulation was used most commonly in the modality of simulated casualties who were portrayed by students, residents, and volunteers (Table 4). When reported, the number of simulated casualties ranged from 16 to 445. Moulage was used in 12 (60%) exercises. Simulated casualties were briefed on their roles or received cards containing relevant information in eight (40%) exercises. Other simulation modalities include mannequins (n = 3; 15%) and part task trainer (n = 1; 5%).

Evaluations, Outcomes, and Debriefings of Full-Scale Exercises—Evaluators of the full-scale exercises were often health care professionals or experts with experience in disasters or emergencies, representatives from agencies involved in disasters, as well as faculty members from residency programs or nursing schools (Supplementary Material Table 3 for details; available online only). When reported, the number of evaluators ranged from three to 18. Evaluators were reported to be trained in five (25%) exercises, but only two provided some details on the nature of training. As the evaluation outcomes (Table 5) differed across the included studies, the authors were unable to synthesize them into a single measurement of effectiveness. However, from the report of individual studies, simulation seemed to be generally effective.

Evaluations of full-scale exercises were performed at more than one Kirkpatrick's level of evaluation in three (15%) exercises. Evaluation of learning was most common in 15 (75%) exercises and assessed response systems (n = 8), individuals (n = 2), or both (n = 5). Learning of response systems focused on whether guidelines, protocols, and policies were enacted, while learning of individuals focused on knowledge and skill. The evaluations included whether standards, criteria, or indicators were met, assessing the accuracy and providing a score, as well as time taken for tasks to be completed. Two exercises were conducted pre- and post-exercise evaluations at this learning level. Evaluation of

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Emergency Medical Services	Fire Department	Police Department	Hospital	Military	Other Agencies	Number of Personnel
2005	Alexander	x	x	x	x			>100
2009	Lenz		x				Incident Command	102
2010	Gryth	x	x	x	x		Civil Aviation Administration, Disaster Care Unit	–
2010	Ingrassia	x	x	x	x		Civil Protection	–
2012	Rutty^	x	x	x	x			607
2012	Klima*							–
2012	Radestad^	x	x	x	x			>200 ~
2013	Austin	x		x		x	Health and Mental Hygiene Department, Red Cross, Medical Reserve Corps, University, Funeral Directors' Disaster Response Team, Emergency Management Task Force	>263 ~
2014	Shah^	x			x			–
2014	Austin	x		x		x	Health and Mental Hygiene Department, Red Cross, Medical Reserve Corps, University, Funeral Directors' Disaster Response Team, Emergency Management Task Force	>26 ~
2014	Djalali*							–
2016	Daniel^		x		x		Medical Reserve Corps	>82 ~
2017	Wenham	x	x	x	x			–
2017	Saber	x	x	x			University Volunteer Ambulance Corps, University Nursing Students	–
2019	McElroy^^	x	x	x	x		Trauma System	–
2020	Lenz	x	x	x	x		First Responders	904
2020	Carenzo	x	x	x	x	X	Coast Guard	>145 ~
2020	Sheikhbardsiri				x		Emergency Operation Center of Ministry of Health, Universities, Logistical Team, Command and Management Team, Pharmaceutical and Consumer Item Team	990
2021	Foo		x				Disaster Medical Assistance Team, Urban Search and Rescue Team, Ministry of Health, Local Emergency Management Agency	266
2021	Innis	x			x		Personal Support Worker, Pre-Service Fire and Police	>450

Table 2. Agencies and Personnel Involved in Full-Scale Exercises
^ Details of agencies mentioned in part.
*Details of agencies not mentioned at all.
~ Numbers from at least one agency involved.

participants' reaction was assessed in seven (35%) exercises, and this involved asking the participants how they felt in domains such as educational value, learning experience, building self-confidence, the realism of exercise, getting them interested, and recommending others to participate. Evaluations of participants' behaviors and the results on patient outcomes were done only in one (5%) study which examined the performance of tasks and time taken to complete them in an actual mass-casualty incident, and mortality of patients involved, respectively (Table 5).

Debriefings were carried on in five (25%) exercises and covered the following domains (Table 6):

- Participants' feelings and experience;
- Meeting objectives and identifying deficiencies in key areas of roles and responsibilities, crisis resource management, leadership and teamwork, communication, workflow and protocol, patient movement, as well as logistics; and
- Organization and conduct of full-scale exercises.

One study mentioned continued feedback for two weeks after on a web-based chat room. However, 11 (55%) of the studies did not mention debriefings and four (20%) did not provide further details about the debriefing process.

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Objectives	Briefings
2005	Alexander	Resident-specific objectives derived from emergency medicine disaster planning Core learning objectives derived from Royal College of Physicians and Surgeons of Canada core curriculum	Not mentioned
2009	Lenz	Not mentioned	Knew about exercise, but no information about time or type of hazard
2010	Gryth	To clarify if performance indicators can be used to point out weakness within the organization so as to improve the system	Know date, but no briefing
2010	Ingrassia	Not mentioned	Briefing on status of resources in the community
2012	Rutty	To consider whether or not running such an exercise increases the knowledge base of players and observers alike	Not mentioned
2012	Klima	To implement incident communications between agencies, coordinate care through incident command center, and effectively decontaminate patients after chemical spill	Not mentioned
2012	Radestad	Not mentioned	Not mentioned
2013	Austin	To introduce disaster triage methods and mass-casualty education	Safety briefings and education about roles and general plan for day
2014	Shah	To assess management of pediatric patients presenting to emergency department after a chemical exposure with primary focus on pediatric intensive care unit surge capacity	Not mentioned
2014	Austin	To introduce disaster triage methods and mass-casualty education	Safety briefings and education regarding specific roles
2014	Djalali	Not mentioned	Not mentioned
2016	Daniel	Understand process for setting up and running decontamination Understand process for setting up a registration system Understand the process of setting up a point-of-distribution site Understand how to safely don and doff hospital staff and decontaminate a victim Understand the management system positions involved Participate in and understand tasks related to crisis decision making, coordination, and communications Educate hospital staff about their roles in emergency response	Not mentioned
2017	Wenham	To demonstrate understanding of triage and triage skills within a team environment To gain an awareness of the difficulties of trauma management outside the hospital setting To demonstrate safe and effective clinical emergency management skills To demonstrate effective skills within a multidisciplinary team	Briefing mentioned, but no details
2017	Saber	Nursing students will display disaster skill competence using the Creighton Competency Evaluation Instrument Nursing students will express increased confidence in responding to disasters	Briefing mentioned, but no details other than time given to familiarize themselves with equipment and team members

Table 3. Objectives and Briefings of Full-Scale Exercises (*continued*)

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Objectives	Briefings
2019	McElroy	Participants shall identify the management structure to support effective operational coordination between all agencies and entities Facilities and agencies shall establish sustained, 2-way communications on applicable plans, policies, and procedures within 15 minutes of onset of incident Participating community health care organizations shall use the Central Ohio Healthcare Disaster Information Management System to appropriately update bed status, hospital liaison contact, and information throughout the incident Participants shall request needed resources through appropriate channels in accordance with plans, policies, and procedures Hospitals and first responders shall effectively triage patients in response to a medical surge event throughout the incident Hospitals shall effectively track patients from intake to discharge in response to a medical surge event throughout the incident Hospitals shall input victims into program used for patient tracking, verifying patient identification, and reunifying family members with patients	Not mentioned
2020	Lenz	Take patients from scene through emergency department into surgery if indicated, then appropriate floor or discharge at multiple Levels 1 and 2 trauma hospitals	Not mentioned
2020	Carenzo	Not mentioned	Not mentioned
2020	Sheikhbardsiri	To identify gaps between operation and support within mass-casualty incidents	Not mentioned
2021	Foo	To determine whether teams met the basic technical standards To identify barriers to cooperation between teams	Not mentioned
2021	Innis	To collaborate with members from different fields	Detailed information via email, 1-hour orientation session on morning of exercise

Table 3. *(continued).* Objectives and Briefings of Full-Scale Exercises

Discussion

Full-scale exercises often involve multiple agencies and individuals for the purpose of evaluating the coordination and response capabilities of the emergency management system. In this scoping review, it was found that simulation, with simulated casualties being the most common simulation modality, was generally effective as reported by the included studies. However, this effectiveness was mostly demonstrated at the levels of learning of individuals and system, as well as reaction of individuals based on Kirkpatrick's levels of evaluation.

Assessing Effectiveness of Simulation in Full-Scale Exercises

Learning of individuals and systems were assessed in full-scale exercises using published standards such as the WHO minimum technical standards for Emergency Medical Teams, or current operating procedures such as the Iranian emergency operation plan.[26,27] Evaluation tools such as the Creighton competence evaluation instrument were also used, but adaptations may be made to better suit the full-scale exercise such as the semi-quantitative performance indicators from the Center for Teaching and Research in Disaster Medicine and Traumatology (University of Linköping, Linköping, Sweden) which were modified by Center for Research and Training in Disaster Medicine, Humanitarian Aid, and Global Health (CRIMEDIM; Università del Piemonte Orientale, Novara, Italy).[19] Other measures of learning include the accuracy of tasks performed, such as the proportion of patients who were correctly triaged or the appropriateness of treatment rendered, as well as the time is taken to do so such as time for triage, treatment, or evacuation to hospital.

Reaction of individuals was assessed in full-scale exercises using completion of questionnaires, surveys, or feedback forms, as well as participation in interviews or focus group discussions. The corresponding perceptions of participants in various domains were reported mostly quantitatively rather than qualitatively. Among the domains assessed, simulation in full-scale exercises most commonly demonstrated a perception of high educational value in instilling knowledge, skill, and attitude both at an individual level and as an interprofessional team. This was followed by perceptions of simulation in full-scale exercises being an enjoyable learning experience which participants were satisfied with, and how self-

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Simulated Casualties	Preparation of Simulated Casualties	Moulage	Number of Simulated Casualties	Other Modalities
2005	Alexander	Trained volunteer paramedic students	–	Yes	80	Mannequin (pediatric and adult)
2009	Lenz	Volunteers	–	–	16	–
2010	Gryth	Trained personnel	–	–	99	Mannequin
2010	Ingrassia	Medical students	Dynamic casualty cards with vital signs and response to treatment	Yes	112	–
2012	Rutty	–	–	–	–	–
2012	Klima	Volunteers	–	Yes	281	
2012	Radestad	–	Figurant cards with injuries and pre-determined medical needs according to template from Emergo Train System victim bank	–	199*	
2013	Austin	Fraternity and sorority students, first trimester nursing students, and youth theater children	Written description of injuries and general presentation	Yes	>288	
2014	Shah	Medical students or emergency medicine residents	Patient profile card included chief complaint, physical exam findings with expected medical interventions, and final disposition	–	36	Mannequin
2014	Austin	Fraternity and sorority members, students, faculty, and volunteers from community groups	–	Yes	–	–
2014	Djalali	–	–	–	61	–
2016	Daniel	–	–	–	30	–
2017	Wenham	Local high school drama students	Briefed to have specific simulated injuries and responses to treatment	Yes	16	–
2017	Saber	–	–	Yes	28	–
2019	McElroy	–	Cards revealing injuries	Yes	445	–
2020	Lenz	–	–	Yes	126^	Partial task trainer known as the *Cut Suit* and *Hyper-Realistic* training
2020	Carenzo	Anesthesia residents	Dedicated storyboard accurately describing injuries, make-up, and evolution, as well as a set of Dynamic Casualty Cards (a series of pre-defined statuses, each including a pre-determined set of vital parameters, major complaints, and expected treatments and intervention times)	Yes	96	–
2020	Sheikhbardsiri	Participants from hospitals and a trauma center	–	–	285	–

Table 4. Simulation Modalities Used for Full-Scale Exercises (*continued*)

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Simulated Casualties	Preparation of Simulated Casualties	Moulage	Number of Simulated Casualties	Other Modalities
2021	Foo	Disaster medical assistance team outside of duty hours	–	Yes	176	–
2021	Innis	Nursing students	Index cards with patient scenarios and instructions regarding the injuries, as well as how they were to act	Yes	–	–

Table 4. *(continued).* Simulation Modalities Used for Full-Scale Exercises
*Two separate full-scale exercises of 99 and 100 simulated casualties.
^ A total of 18 full-scale exercises over nine days.

confidence as a responder to disasters and mass-casualty incidents was improved after the full-scale exercises.

Evaluating the effectiveness of simulation in full-scale exercises at the higher Kirkpatrick levels of behavior and results would be most valuable as assessments could be carried out when participants perform their roles as responders in effective and efficient emergency management systems. Their behavior during an actual disaster or mass-casualty incident would be indicative of their readiness to translate knowledge, skill, and attitude into clinical practice when attending to casualties. The results of their actions would eventually impact outcomes of casualties ranging from immediate ones, such as mortality, to short-term ones, such as morbidity or survival to hospital discharge, and long-term ones, such as return to previous function and general well-being. However, this would be challenging to evaluate given the infrequent occurrences, and the challenge in standardizing outcomes to measure due to the varied nature of disasters and mass-casualty incidents. There was only a single report by Lenz, et al on the behavior of the responders and outcomes of the casualties during an active threat resolution at the STEM School shooting in Highlands Ranch, Colorado USA, which occurred six months after the full-scale exercise.[24]

Improving the Effectiveness of Simulation in Full-Scale Exercises

Full-scale exercises are manpower and resource intensive. They are costly in terms of finances and time required from planning to execution. In order to improve the effectiveness of simulation in full-scale exercises, the key consideration is that full-scale exercises should be integrated into a training program consisting of simulation exercises that are progressively complex – starting from tabletop exercises, moving on to drills, functional exercises, and eventually full-scale exercises. This stepwise approach can facilitate basic learning for individuals and single organizations at the earlier stages of simulation exercises and advance learning for multiple organizations and national systems at the later stages of simulation exercises. Full-scale exercise should ideally not be carried out as a stand-alone exercise without prior simulation exercises to build upon. However, in this scoping review, the authors were unable to ascertain if there were prior table-top exercises, drills, or functional exercises before these full-scale exercises were conducted.

The effectiveness of simulation in full-scale exercises could also be enhanced further by adhering to principles of learning theories and applying evidence-based practices for the conduct of simulation.[29,30] Full-scale exercises could be anchored on adult learning principles which emphasize that the learning objectives should be relevant and tailored to work experience so that participants can learn in a safe environment provided by simulation. Learning needs should be assessed in order to define learning objectives and create scenarios for full-scale exercises. The simulation modality and the level of fidelity should be determined. Participants should undergo briefing and be given time for familiarization with the simulation modality and setting. Simulated casualties should be trained, and evaluators should be standardized in their assessment. A participant-centered approach of providing feedback and debriefing should be adopted to facilitate reflective learning. In this scoping review, these best practices of the simulation were inconsistently reported across the included studies. However, the authors were unable to determine if these best practices were excluded in the conducting of full-scale exercises or in the reporting process. Therefore, a framework could be used for the conducting, which may improve effectiveness of simulation during full-scale exercises, and the same framework can be used for reporting, which may enhance documenting the effectiveness of simulation after full-scale exercises.

Increased Use of Simulation in Full-Scale Exercises

The use of simulation in full-scale exercises rose after 2010, and this could be attributed to two key drivers. Firstly, the increased frequency of disasters and mass-casualty incidents with higher morbidity and mortality rates has called to attention the need for training organizations and personnel through the conduct of full-scale exercises during the preparedness phase of disaster management.[31] The adequacy of response could also be assessed so that gaps could be identified and actions could then be instituted at national, organizational, and individual levels for improvement. The second driver is the tremendous growth of simulation as a training modality. Anchored in learning theories and guided by evidence-based practices, simulation has wide applications in health care settings and is particularly valuable for training related to disaster and mass-casualty incidents as simulation could provide realism in a safe environment for the participants.[32]

Yet, from this scoping review, despite the surge in the literature on full-scale exercises, reports on outcomes of effectiveness were limited, as well as variable in terms of what constituted effectiveness and how the effectiveness was measured. The authors were unable to determine any evolving trends across the years in participants involved, objectives identified, simulation modalities utilized, evaluation methods employed, outcomes assessed, as well as

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Evaluation Methods	Evaluation Outcomes	Kirkpatrick's Level
2005	Alexander	1) Evaluation using global assessment forms (5-point Likert scale)	A) Scenarios enjoyable mean 4.9 (SD = 0.3) B) Scenarios realistic and relevant 4.6 (SD = 0.7) C) Gained valuable knowledge toward future practice in emergency medicine 4.8 (SD = 0.4)	Reaction
2009	Lenz	1) Monitored response activities from a reference list of standard operating procedures extracted from the Fire Service Directive 5005 and the Robert Koch Institute Guideline using data collection templates and video recording	A) 20 out of 31 activities were in accordance with guidelines (10 incorrect, 1 not applicable)	Learning (System)
2010	Gryth	1) Evaluation using a template with indicators derived from concept and process modeling by the Swedish Board of Health and Welfare: - "0" meant that the standard was not met - "1" indicated that the standard was met, but not with adequate content or within specified timeframe - "2" indicated correct performance within the correct timeframe	A) Prehospital Command and Control – 3/24 points B) Strategic Command and Control – 15/22 points C) Hospital Command and Control – 17/22 points D) Staff Skills Performance – Strategic 17/22 points, Hospital 21/22 points	Learning (Individual and System)
2010	Ingrassia	1) Recorded observations of medical management on structured evaluation forms	A) Triage completed in 37 minutes, average 12.7 (SD = 8.1 minutes) B) Triage correct in 81% C) Average evacuation time for non-ambulatory 21.6 (SD = 13.1 minutes) D) Correct maneuvers in 85.2% airway, 78.7% breathing, 57.4% circulation, 65.6% others E) Total 246 radio comms, averaging 34 seconds, 29.8% with command post medical officer, 26% on patient evacuation and transportation	Learning (Individual and System)
2012	Rutty	1) Questionnaire administered post-exercise	A) 72% medium to significant knowledge gain B) 81% confirmed enforcement a lot or some of their previous knowledge C) 91% medium or great interest in learning about the processes D) 98% considered exercises as a valuable training opportunity E) 87% considered they had achieved what they had hoped for by attending the exercise F) 93% would recommend attending an exercise to another colleague	Reaction
2012	Klima	1) Assessed five areas based on specific definitions: communications, command structure, decontamination, staffing, and patient tracking	A) None of the 16 hospitals compliant in all five areas B) Mean hospital compliance in 1.9 (SD = 0.9) areas C) Compliance: patient tracking 69%, command structure 44%, staffing 37%, decontamination 25%, communication 6%	Learning (System)
2012	Radestad	1) Evaluated performance using a protocol with sets of indicators	1) Performance indicators – all results were on approved level except for prehospital command in 2008 2) Outcome indicators – preventable complications 53% and 29% in 2008 and 2010, respectively, while preventable death 29% and 41% 3) Other observations – ambulances deployed according to plan and adequate level (first victim 100 and 105 minutes after accident and last victim 235 and 273 minutes)	Learning (System)

Table 5. Evaluations of Full-Scale Exercises (*continued*)

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Evaluation Methods	Evaluation Outcomes	Kirkpatrick's Level
2013	Austin	1) Questionnaire administered post-exercise	A) Assessment (35%), triage (16%), and intervention (15%) were identified as the top three nursing roles during a disaster B) 52% reported some confidence, 21% very confident, 19% slightly more confident than prior to exercise	Reaction
2014	Shah	1) Completed specific evaluation forms designed to assess the appropriateness of response 2) Rated subjectively how closely the surge plan was followed	A) Appropriateness of treatment – respiratory complaints 11/11 received albuterol and steroids, 3/3 intubated; eye complaints 5/5 irrigated but none ocular pH checked, all referred to eye consult; all other key treatments performed B) Timeliness of treatment – all timely C) Patient education – 15/22 patients given explanation of treatment received, 4 told of effects of chlorine over-exposure, discharged patients received appropriate discharge instructions D) Specialty consult – appropriate referral E) Disposition/PICU surge – responded as per emergency operations plan of hospital F) Electronic medical record – all patients, medications, lab, imaging recorded G) Communication – several options for communication and interpersonal communication facilitated, need to communicate with greater hospital community H) Command and control – no single individual commanded the scene, poor utilization of incident command system, updates not always timely and accurate, not all relevant external agencies contacted	Learning (System)
2014	Austin	1) Questionnaire administered post-exercise (6-point Likert scale) 2) Pre- and post-test using a case study	A) Positive experience that increased their knowledge about disaster response B) Able to incorporate concepts from coursework into the mass-casualty environment C) Able to respond to patient questions and recognize patients with special needs D) Teams worked well together E) Worked well with other volunteers F) Pre- and post-test showed improved scores on case study which was statistically significant (P <.01)	Reaction and Learning (Individual)
2014	Djalali	1) Evaluated performance using nine semi-quantitative performance indicators from the Center for Teaching and Research in Disaster Medicine and Traumatology at the University of Linkoping, Sweden, but the scoring method was modified by CRIMEDIM	A) Preparedness score 3 days before exercise: 59% B) Response score during exercise: 70% C) START triage 90% correct for yellow, 100% for green D) Average time for first triage of yellow and green – 2 and 8 minutes, respectively E) Average time for bed allocation in ED for yellow and green – 1 and 5 minutes, respectively	Learning (System)
2016	Daniel	1) Assessed attainment of knowledge for the ED residents using pre- and post-test 2) Evaluated teams on the percentage of critical actions met	A) Pre- and post-test scores – mean post test scores higher than pre (62% vs 53%; P = .002) B) Team performance of critical actions ranged from 48% to 63% of objectives completed correctly C) Hospital performance of critical actions ranged from 50% to 100% of objectives completed correctly	Learning (Individual and System)

Table 5. Evaluations of Full-Scale Exercises (*continued*)

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Evaluation Methods	Evaluation Outcomes	Kirkpatrick's Level
2017	Wenham	1) Survey with 10-point Likert and qualitative using semi-structured interview within focus group	A) 94% felt all 4 learning outcomes met B) Domain score >75% – prepare for day-to-day practice, effective teaching strategy, adequate resources, should run it again C) 88% felt briefing before event adequate D) Qualitative – value of good teamwork and communication, increased understanding of stressful work environment on performance, willingness to engage in reflection of the value of interprofessional learning	Reaction
2017	Saber	1) Evaluated using Creighton Competence Evaluation Instrument	A) All nursing students scored 100% indicating disaster skill competence met B) Almost all participants reported increased in confidence (100% nursing, 100% agencies, 87.5% victim actors)	Reaction and Learning (Individual)
2019	McElroy	1) Evaluated objectives as performed without challenges, performed with some challenges, performed with major challenges, or unable to be performed	A) All 7 objectives performed with some challenges B) Communication major weakness - little coordination and communication	Learning (Individual and System)
2020	Lenz	1) Evaluated adequacy and accuracy of interventions 2) Evaluated actual response six months later during a school shooting incident	A) 265 procedures performed in prehospital, 202 in ED, 12 in OR B) Tourniquet application – 93.5% success rate C) Time for EMS to move into the MCI facility, locate casualties, extract first victim, move to casualty collection point, transport to safety – reduced from 42 minutes to 12 minutes over 9 days D) Six months later: active threat resolution	Learning (Individual and System), Behavior and Results
2020	Carenzo	1) Evaluated performance	A) Triage accuracy 85% prehospital, 84% in-hospital B) Evacuation flow respected triage priority C) Mean incident to definitive care times of 121 minutes for immediate, 163 minutes for delayed care, 130 minutes for minor wounds D) All casualties triaged and assessed but 16 not evacuated	Learning (System)
2020	Sheikhbardsiri	1) Completed evaluation checklist of 13 functional dimensions and 72 items developed by reviewing related textbooks, interviewing experts, and reading the Iranian emergency operation plan	A) Preparedness of health and treatment departments score 79.5% B) Management administrative scores higher than 80% C) Security score 94% D) Incident action plan 67% E) Periodic assessment and comprehensive information management 67% F) Request for personnel 68% G) Establishment of incident command post 69% H) Mean arrival of team 4 hours 35 minutes	Learning (System)
2021	Foo	1) Conducted assessments based on the minimum technical standards for type I Fixed Emergency Medical Teams (EMT) set forth in the EMT coordination handbook by participating in regular team meetings, monitoring them and the radio channels, as well as observing the operations 2) Health monitoring on Behavior self-reported health assessment information	A) Overall compliance rate with technical standards of type 1 fixed EMT 70.4% – gaps in local anesthesia, laboratory test, pharmacy and drug supply, sterilization B) Health monitoring showed 52.9% abnormal results	Learning (System)

Table 5. Evaluations of Full-Scale Exercises (*continued*)

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

Year Published	First Author (Last Name)	Evaluation Methods	Evaluation Outcomes	Kirkpatrick's Level
2021	Innis	1) Questionnaire administered post-exercise	A) 64% prepared for role B) 77% contributed C) 87% debriefing beneficial D) 90% help understand role of other health care professions E) 72% satisfied with experience F) 77% rated event good to excellent G) 76% provided insight to patient perspective H) 56% practice professional skills within current scope of practice I) 85% demonstrate importance of interprofessional practice J) 71% demonstrated profession's role during disaster K) Qualitative themes – communicating with patients, collaborating with health care and emergency management providers	Reaction

Table 5. *(continued).* Evaluations of Full-Scale Exercises
Abbreviations: PICU, pediatric intensive care unit; START, Simple Triage and Rapid Treatment; ED, emergency department; EMS, Emergency Medical Services; MCI, mass-casualty incident; EMT, Emergency Medical Team.

briefings and debriefings conducted, as these were inconsistently reported in the included

Year Published	First Author (Last Name)	Details of Debriefings
2005	Alexander	Feedback about the organization of the event and how they felt as physicians responsible for carrying out the plan; summary of experience; two weeks following the event, web-based chat room to elaborate thoughts and lessons learned
2009	Lenz	Not mentioned
2010	Gryth	Not mentioned
2010	Ingrassia	Not mentioned
2012	Rutty	Not mentioned
2012	Klima	An after-action review was performed and reviewed for deficiencies
2012	Radestad	Not mentioned
2013	Austin	Not mentioned
2014	Shah	Not mentioned
2014	Austin	Not mentioned
2014	Djalali	Not mentioned
2016	Daniel	Used data from assessment forms during the full-scale exercise and questionnaire after; reviewed objectives and assessed if objectives were met; safety and communication errors
2017	Wenham	Mentioned, but no details
2017	Saber	Not mentioned
2019	McElroy	Debriefing mentioned, but no details
2020	Lenz	On identifying and correcting deficiencies observed; logistics, roles and responsibilities, workflows, patient movement, communication
2020	Carenzo	Simulation data (triage, morbidity, and mortality), feedback, highlight situations which need improvement - role of leader, importance of protocols and communications
2020	Sheikhbardsiri	HOTWASH mentioned, but no details
2021	Foo	Not mentioned about debriefing, but evaluators met and conducted internal discussions prior to issuing formal report
2021	Innis	Mentioned debriefing, but no details

Table 6. Debriefings of Full-Scale Exercises

https://doi.org/10.1017/S1049023X2300660X Published online by Cambridge University Press

studies.

Recommendations for Use of Simulation in Full-Scale Exercises

Simulation in full-scale exercises should be well-conducted before its effectiveness could be accurately measured. Likewise, reports of simulation in full-scale exercises should be of good quality.[33] Future works on the use of simulation in full-scale exercises should first include standardized reporting of methodologies used in the conduct of simulation. This should be done using a framework which would allow for assessment of the quality of simulation based on six key domains which reflect best evidence-based practices – participants involved, objectives identified, simulation modalities utilized, evaluation methods employed, outcomes assessed, as well as briefings and debriefings conducted. Using this framework, key aspects of simulation use for full-scale exercises could be consistently reported to develop and advance knowledge in this field – important details may be less likely to be left out, and this may allow for comparisons between individual studies. Novel or modified methodologies could be highlighted so that evidence-based practices specific to the context of full-scale exercise could be augmented within the community.

Focusing on effectiveness of simulation in full-scale exercises, future works should clearly identify Kirkpatrick's levels of evaluation on effectiveness, and explicitly define how evaluation would be carried out for the particular level, such as self-reporting using survey or feedback through focus group discussion for reaction of individuals, benchmarking against published standards or specific protocols for learning of system or individuals, responding to casualties competently and confidently for behavior of system or individuals, as well as reporting patient-centered outcomes of casualties for results. It would also be important to provide information about how these evaluation tools are selected, how evaluation outcomes for measuring effectiveness are selected, whether these tools and outcomes have been previously used, and whether they have been validated for accuracy and precision in evaluating the effectiveness of simulation in full-scale exercises.

Limitations

Firstly, there was no universal gold standard definition for full-scale exercise, so the definition in WHO simulation exercise toolbox was used in this scoping review. In the study selection process, there were articles which were reported as full-scale exercises but did not demonstrate all three key components of the definition in WHO simulation exercise toolbox. This observation called for a need for a universal gold standard definition for full-scale exercise to be agreed upon in the community and then be used consistently for reporting in the literature.

Next, non-English articles were excluded, and this may have led to useful data being missed by this scoping review. Nonetheless, the authors were able to understand the body of evidence available to address the research question, and share the findings, gaps, and recommendations so that subsequent works in this field can benefit from this. Also, many methods exist to evaluate the effectiveness of training. In addition to Kirkpatrick's levels of evaluation, other methods such as the Context, Input, Process, and Product model and Kaufman's model of learning evaluation exist, all with their pros and cons.[34,35] While Kirkpatrick's levels may be overtly focused on outcomes instead processes of learning, and the linear causality implied across the levels may be too simplistic, it is still widely recognized and accepted in simulation and health care education, thus making it a good standard to base this scoping review on.[36] Full-scale exercises may involve learning of both individuals and response systems, therefore, the authors had to modify Kirkpatrick's level of learning and delineate whether individuals or response systems were assessed in the evaluation.

Finally, while the simulation was shown to be overall effective, the respective domains which reflect the best evidence-based practice, such as briefing and debriefing of participants, or choice of simulation modality and fidelity, could not be assessed individually to determine which aspects of the conduct of simulation contributed most to its effectiveness in full-scale exercises.

Conclusion

Full-scale exercises provide training opportunities for disasters and mass-casualty incidents, which are rare events. In the simulated setting, which is created to replicate the actual environment, participants can attend to simulated casualties in accordance with existing protocols. However, these training are labor and resource intensive, and they often require significant amounts of time for planning, coordination, and execution. This scoping review has shown that the use of simulation in full-scale exercises, based on reports by individual studies, is generally effective but has mostly been reported at lower Kirkpatrick's levels of reaction of individuals, as well as learning of individuals and/or systems. Best evidence-based practices for simulation should be adhered to in full-scale exercises so that the training would adequately prepare the participants and eventually translate into better care and outcomes for casualties during an actual disaster or mass-casualty incident. In addition, the reporting of simulation use in full-scale exercises should be standardized using a framework, and the evaluation process should be rigorous so that effectiveness could be determined and compared across full-scale exercises.

Supplementary Materials

To view supplementary material for this article, please visit https://doi.org/10.1017/S1049023X2300660X

References

1. World Health Organization. Simulation exercises. https://www.who.int/emergencies/operations/simulation-exercises. Accessed October 10, 2022.
2. Al-Elq AH. Simulation-based medical teaching and learning. *J Family Community Med.* 2010;17(1):35–40.
3. Lateef F. Simulation-based learning: just like the real thing. *J Emerg Trauma Shock.* 2010;3(4):348–352.
4. Cannon-Bowers JA. Recent advances in scenario-based training for medical education. *Curr Opin Anaesthesiol.* 2008;21(6):784–789.
5. Flanagan B, Nestel D, Joseph M. Making patient safety the focus: crisis resource management in the undergraduate curriculum. *Med Educ.* 2004;38(1):56–66.
6. Sweeney AL, Lui YL, Watkins N, et al. Simulation exercises increase staff confidence, knowledge, and skills in managing mass-casualty incidents: a pre-test-post-test study. *Tasman Medical Journal.* 2021;3(4):112–119.
7. Preferred Reporting Items for Systematic Reviews and Meta-Analyses. PRISMA for scoping reviews. https://prisma-statement.org/Extensions/ScopingReviews. Accessed October 10, 2022.
8. Kirkpatrick DL. *Evaluating Training Programs: The Four Levels*. Oakland, California USA: Berrett-Koehler; 1994.
9. Alexander AJ, Bandiera GW, Mazurik L. A multiphase disaster training exercise for emergency medicine residents: opportunity knocks. *Acad Emerg Med.* 2005;12(5):404–409.

10. Lenz M, Richter T. Disaster response to the release of biohazardous agents: instrument development and evaluation of a firefighter's exercise. *Prehosp Disaster Med.* 2009;24(3):197–203.
11. Gryth D, Rådestad M, Nilsson H, et al. Evaluation of medical command and control using performance indicators in a full-scale, major aircraft accident exercise. *Prehosp Disaster Med.* 2010;25(2):118–123.
12. Ingrassia PL, Prato F, Geddo A, et al. Evaluation of medical management during a mass-casualty incident exercise: an objective assessment tool to enhance direct observation. *J Emerg Med.* 2010;39(5):629–636.
13. Rutty GN, Rutty JE. Did the participants of the mass fatality exercise Operation Torch learn anything? *Forensic Sci Med Pathol.* 2012;8(2):88–93.
14. Klima DA, Seiler SH, Peterson JB, et al. Full-scale regional exercises: closing the gaps in disaster preparedness. *J Trauma Acute Care Surg.* 2012;73(3):592–598.
15. Rådestad M, Nilsson H, Castrén M, Svensson L, Rüter A, Gryth D. Combining performance and outcome indicators can be used in a standardized way: a pilot study of two multidisciplinary, full-scale major aircraft exercises. *Scand J Trauma Resusc Emerg Med.* 2012;20:58.
16. Austin EN, Hannafin NM, Nelson HW. Pediatric disaster simulation in graduate and undergraduate nursing education. *J Pediatr Nurs.* 2013;28(4):393–399.
17. Shah VS, Pierce LC, Roblin P, Walker S, Sergio MN, Arquilla B. Waterworks, a full-scale chemical exposure exercise: interrogating pediatric critical care surge capacity in an inner-city tertiary care medical center. *Prehosp Disaster Med.* 2014;29(1):100–106.
18. Austin EN, Bastepe-Gray SE, Nelson HW, et al. Pediatric mass-casualty education: experiential learning through university-sponsored disaster simulation. *J Emerg Nurs.* 2014;40(5):428–433.
19. Djalali A, Carenzo L, Ragazzoni L, et al. Does hospital disaster preparedness predict response performance during a full-scale exercise? A pilot study. *Prehosp Disaster Med.* 2014;29(5):441–447.
20. Daniel P, Gist R, Grock A, Kohlhoff S, Roblin P, Arquilla B. Disaster Olympics: a model for resident education. *Prehosp Disaster Med.* 2016;31(3):237–241.
21. Wenham J, Bennett P, Gleeson W. Crash simulation: an immersive learning model. *Clin Teach.* 2018;15(6):467–471.
22. Saber DA, Strout K, Caruso LS, Ingwell-Spolan C, Koplovsky A. An interprofessional approach to continuing education with mass-casualty simulation: planning and execution. *J Contin Educ Nurs.* 2017;48(10):447–453.
23. McElroy JA, Steinberg S, Keller J, et al. Operation continued care: a large mass-casualty, full-scale exercise as a test of regional preparedness. *Surgery.* 2019;166(4):587–592.
24. Lenz A, Shelton R, Ryznar R, et al. The next nine minutes: lessons learned from the large-scale active shooter training prior to the STEM school shooting. *Am J Disaster Med.* 2020;15(4):241–249.
25. Carenzo L, Bazurro S, Colombo D, Petrini F, Ingrassia PL. An island-wide disaster drill to train the next generation of anesthesiologists: the SIAARTI Academy Experience. *Disaster Med Public Health Prep.* 2021;15(2):151–154.
26. Sheikhbardsiri H, Yarmohammadian MH, Khankeh H, et al. An operational exercise for disaster assessment and emergency preparedness in South of Iran. *J Public Health Manag Pract.* 2020;26(5):451–456.
27. Foo NP, So EC, Lu NC, et al. A 36-hour unplugged full-scale exercise: closing the gaps in interagency collaboration between the disaster medical assistance team and urban search and rescue team in disaster preparedness in Taiwan. *Emerg Med Int.* 2021;2021:5571009.
28. Innis J, Mack K. Evaluation of nursing students' experiences in an interprofessional simulated disaster exercise. *J Nurs Educ.* 2021;60(8):445–448.
29. Wang EE. Simulation and adult learning. *Dis Mon.* 2011;57(11):664–678.
30. Sittner BJ, Aebersold ML, Paige JB, et al. INACSL standards of best practice for simulation: past, present, and future. *Nurs Educ Perspect.* 2015;36(5):294–298.
31. Hodgson L. How Violent Attacks Are Changing The Demands of Mass-casualty Incidents: A Review of The Challenges Associated with Intentional Mass-casualty Incidents. www.hsaj.org/articles/16880. Accessed November 1, 2022.
32. Herrera-Aliaga E, Estrada LD. Trends and innovations of simulation for twenty first century medical education. *Front Public Health.* 2022;10:619769.
33. Cheng A, Kessler D, Mackinnon R, et al; International Network for Simulation-based Pediatric Innovation, Research, and Education (INSPIRE) Reporting Guidelines Investigators. Reporting Guidelines for Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements. *Simul Healthc.* 2016;11(4):238–248.
34. Toosi M, Modarres M, Amini M, Geranmayeh M. Context, input, process, and product evaluation model in medical education: a systematic review. *J Educ Health Promot.* 2021;10(1):199.
35. Kaufman R, Keller JM. Levels of evaluation: beyond Kirkpatrick. *Hum Resour Dev Q.* 1994;5(4):371–380.
36. Johnston S, Coyer FM, Nash R. Kirkpatrick's evaluation of simulation and debriefing in health care education: a systematic review. *J Nurs Educ.* 2018;57(7):393–398.