

# Making Disaster Care Count: Consensus Formulation of Measures of Effectiveness for Natural Disaster Acute Phase Medical Response

Rajesh K. Daftary, MD, MPH;<sup>1,2</sup> Andrea T. Cruz, MD, MPH;<sup>1</sup> Erik J. Reaves, DO, MTM&H;<sup>3</sup> Frederick M. Burkle, Jr., MD, MPH;<sup>4</sup> Michael D. Christian, MD, MSc;<sup>5,6,7</sup> Daniel B. Fagbuyi, MD;<sup>2</sup> Andrew L. Garrett, MD;<sup>8</sup> G. Bobby Kapur, MD, MPH;<sup>9</sup> Paul E. Sirbaugh, DO, MBA<sup>1,10</sup>

1. Baylor College of Medicine, Department of Pediatrics, Section of Emergency Medicine, Houston, Texas USA
2. The George Washington University School of Medicine and Health Sciences, Department of Pediatrics and Emergency Medicine, Washington, D.C. USA
3. US Naval Medical Research Unit No. 6, Lima, Peru
4. Harvard Humanitarian Initiative, Harvard School of Public Health, Cambridge, Massachusetts USA
5. Critical Care & Infectious Diseases Mount Sinai Hospital & University Health Network, Toronto, Ontario Canada
6. Royal Canadian Air Force, National Defence, Canada
7. Faculty of Medicine and Dalla Lana School of Public Health, University of Toronto, Ontario Canada
8. Department of Health and Human Services, Office of the Assistant Secretary of Preparedness and Response, Washington, DC USA
9. Baylor College of Medicine, Department of Emergency Medicine, Houston, Texas USA
10. City of Houston, Emergency Medical Services, Texas USA

## Correspondence:

Rajesh K. Daftary, MD  
George Washington University School of Medicine  
Department of Pediatrics  
Division of Emergency Medicine  
111 Michigan Ave, NW  
Washington, DC 20010 USA  
E-mail: rajdaftary@gmail.com

**Conflicts of interest:** Two authors, (Andrew Garrett and Daniel Fagbuyi), are employed directly, or as consultants, by US federal

## Abstract

**Introduction:** No standard exists for provision of care following catastrophic natural disasters. Host nations, funders, and overseeing agencies need a method to identify the most effective interventions when allocating finite resources. Measures of effectiveness are real-time indicators that can be used to link early action with downstream impact.

**Hypothesis:** Group consensus methods can be used to develop measures of effectiveness detailing the major functions of post natural disaster acute phase medical response.

**Methods:** A review of peer-reviewed disaster response publications (2001-2011) identified potential measures describing domestic and international medical response. A steering committee comprised of six persons with publications pertaining to disaster response, and those serving in leadership capacity for a disaster response organization, was assembled. The committee determined which measures identified in the literature review had the best potential to gauge effectiveness during post-disaster acute-phase medical response. Using a modified Delphi technique, a second, larger group (Expert Panel) evaluated these measures and novel measures suggested (or “free-texted”) by participants for importance, validity, usability, and feasibility. After three iterations, the highest rated measures were selected.

**Results:** The literature review identified 397 measures. The steering committee approved 116 (29.2%) of these measures for advancement to the Delphi process. In Round 1, 25 (22%) measures attained >75% approval and, accompanied by 77 free-text measures, graduated to Round 2. There, 56 (50%) measures achieved >75% approval. In Round 3, 37 (66%) measures achieved median scores of 4 or higher (on a 5-point ordinal scale). These selected measures describe major aspects of disaster response, including: Evaluation, Treatment, Disposition, Public Health, and Team Logistics. Of participants from the Expert Panel, 24/39 (63%) completed all rounds. Thirty-three percent of these experts represented international agencies; 42% represented US government agencies.

**Conclusion:** Experts identified response measures that reflect major functions of an acute medical response. Measures of effectiveness facilitate real-time assessment of performance and can signal where practices should be improved to better aid community preparedness and response. These measures can promote unification of medical assistance, allow for comparison of responses, and bring accountability to post-disaster acute-phase medical care. This is the first consensus-developed reporting tool constructed using objective measures to describe the functions of acute phase disaster medical response. It should be evaluated by agencies providing medical response during the next major natural disaster.

agencies dealing with disaster management. They received no compensation or benefit by their participation in this research, and their conflict of interest is minimal.

**Keywords:** disaster; disaster epidemiology; disaster medicine; disaster relief; measures of effectiveness

## Abbreviation:

MOE: measure of effectiveness

**Received:** February 21, 2014

**Revised:** June 12, 2014

**Accepted:** July 6, 2014

Online publication: September 16, 2014

doi:10.1017/S1049023X14000922

Daftary RK, Cruz AT, Reaves EJ, Burkle FM Jr, Christian MD, Fagbuyi DB, Garrett AL, Kapur GB, Sirbaugh PE. Making disaster care count: consensus formulation of measures of effectiveness for natural disaster acute phase medical response. *Prehosp Disaster Med.* 2014;29(5):461-467.

## Introduction

Over the past decade, natural disasters have claimed hundreds of thousands of lives.<sup>1</sup> As urbanization and climate change accelerate, the incidence of devastating natural disasters will likely increase.<sup>2</sup> Medical responders have struggled to provide care in the immediate aftermath of these disasters, often with widely variable results.<sup>3</sup> Improperly prepared response teams may engage in practices that undermine the long-term wellbeing of a population through inappropriate use of resources, inconsistent application of local standards of care, and alienation and ultimate disengagement of local health staff.<sup>4</sup> To increase accountability, many governing bodies have mandated that humanitarian assistance be evidence-based and be followed by after-action reports that include measures of effectiveness (MOEs).<sup>5-9</sup>

Measures of effectiveness are clearly identified and agreed-upon targets that can be used to evaluate progress towards a desired goal. They facilitate communication among stakeholders and beneficiaries as to whether essential standards are being met, and link policy to action. Further, some MOEs can be assessed in real time, thereby facilitating immediate feedback and informing responders on the need for modification of practices.<sup>10</sup> Prior efforts have been made to standardize reporting. The Sphere Project (International Council of Volunteer Agencies, Geneva Switzerland) is among the most notable collaborative efforts to produce clear targets for aid agencies working to serve displaced populations. In its latest edition of *Humanitarian Charter and Minimum Standards in Disaster Response* (2011), the group outlines standards for a wide range of services, including water and sanitation, nutrition, and health care.<sup>11</sup> Incorporating Sphere standards with measures used by other humanitarian agencies led to development of the Rapid Epidemiologic Assessment tool to identify major areas of need within displaced populations.<sup>12</sup> Recent efforts have produced more detailed after-action reporting tools.<sup>13</sup> These new tools are data intensive, however, and are not suited for real-time assessment. Additionally, the information is less applicable to individual agency contributions, relies partially on subjective rating scales, and may be more applicable to resource-rich settings.

The international disaster community is in need of a reporting tool that can describe the ongoing efforts of a disaster response team providing medical care. The ideal tool should be accessible, able to facilitate real-time decision making, easy to understand, constructed of MOEs that are widely regarded as indicative of the quality of care provided, informative as to problem areas that need review, and reasonably completed within two hours, thereby not detracting attention from provision of medical care.<sup>12</sup>

There have been prior calls for monitoring and evaluation standards for disaster medical response.<sup>5,9,12-14</sup> Measures of effectiveness have been difficult to construct because little data exists to link agency response with population outcomes and impact. Previous efforts to construct a standardized reporting template have produced disparate results, few of which have been adopted widely. In matters where no consensus exists for a difficult-to-test hypothesis, group consensus methods have

been utilized to interpret and consolidate existing data and viewpoints.<sup>15</sup>

The goal of this study was to use group consensus methods to construct an MOE-based reporting template to describe acute phase medical care provided by agencies responding to a major natural disaster.

## Methods

### Literature Review

A literature review was conducted to provide participants with a summary of the current state of performance reporting in disaster medical response. Using the search terms listed in Table 1, a PubMed (National Center for Biotechnology Information, Bethesda, Maryland USA) search identified articles published from June 2001 through June 2011 describing medical care provided following the 20 disasters of any type with the highest casualty count during this period.<sup>1</sup> PubMed was used as a primary source of article selection, as it provided access to a large number of peer-reviewed publications. While not exhaustive, it is a recognized starting point for disaster-related literature reviews.<sup>16,17</sup>

A single reviewer (RD) screened these articles for measures describing provided medical care. Articles were excluded if they referred to care only in the nonacute phase of a response (more than two weeks following an event). An additional search using Google Scholar (Google, Mountain View, California USA) was conducted for other disaster literature that used consensus methodology to develop response standards. At the time of this review, only the work of the Sphere Project met this additional criterion.<sup>11</sup>

The selected articles were reviewed for all measures pertaining to the input, activities, outputs, outcomes, and impact of disaster medical response.<sup>5</sup> All qualitative and quantitative descriptors were extracted by a single reviewer and compiled in a database (RD). A second reviewer (AC) evaluated the database for appropriate grouping.

These measures were sorted (RD, AC) into a spreadsheet using the above domains, and then further sorted into major categories (evaluation, treatment, disposition, prevention and population [public health], and team [logistics]). The steering committee (FB, MC, DF, AG, GK, ER; Appendix A, available online) was comprised of persons with publications pertaining to disaster response, or those serving in leadership capacities for

### Search Terms Used for Literature Review

Medical response AND (((“Mass Casualty Incidents”[MAJR]) OR (“Disasters”[MAJR])) AND ((earthquake AND (Haiti OR China OR Pakistan OR India OR Iran OR Japan OR Java OR Algeria OR Afghanistan)) OR Tsunami OR (Cyclone AND (Myanmar OR Bangladesh)) OR Darfur OR “swine flu” OR (Heatwave and France) OR Katrina OR “World trade”)) NOT economics

Daftary © 2014 Prehospital and Disaster Medicine

Table 1. Search Terms Used for Literature Review

disaster response organizations. They selected measures according to the following prompt:

A single-impact natural disaster has struck a large community, resulting in mass-traumatic injury and incapacitation of existing medical infrastructure. Medical teams are deployed to assist with provision of medical care spanning 72 hours post event (T+72hours) to the next two weeks (T+14 days). What clinical measures are predictive of reduction in morbidity and mortality during this time?

Those measures receiving approval from at least four out of six (66.7%) committee members were graduated into the Delphi stage. This target was chosen because it falls within the 55%-80% commonly used to indicate consensus.<sup>18,19</sup>

#### *Delphi Process*

A subject matter expert panel (referred to as the Expert Panel) comprised of persons recognized for contributions to disaster response, agency representatives, and those in leadership positions overseeing disaster-related functions, was nominated by the steering committee using pre-established criteria.<sup>18</sup> The nominees were directed to an online survey site (Survey Monkey Inc., Palo Alto, California USA).

In Round 1, participants from the Expert Panel were presented with those measures identified by the steering committee. Participants were asked to identify which measures in each category applied to the above-mentioned prompt, and to suggest ("free-text") additional measures. At that time, the steering committee was invited to submit free-text measures through the survey site. Contributed measures were compiled with those of the Expert Panel.

Three reviewers (AC, RD, ER) independently evaluated the consolidated free-text measures according to the following predetermined criteria: can be objectively measured, describes the acute response, and not already presented in Round 1. These reviewers then discussed, via phone conference, whether each measure met the above criteria. Unanimous consensus was required to approve measures or consolidate redundant measures. Those that could be reworded into a rate (numerator/denominator/time) were, if the reworded measure received unanimous consent by the three reviewers. Those measures not meeting unanimous consensus were eliminated. Measures that were suggested by the Expert Panel, but that had already been reviewed in some form by the steering committee in an earlier round, were eliminated.

Qualifying free-text measures were combined with those measures receiving >75% approval by participants in Round 1. These measures were organized into major categories for Round 2 and presented to those subjects who completed the prior round. In Round 2, participants were asked to identify which measures answered the initial prompt using dichotomous choices ("yes" or "no"). They were asked to approve only those measures meeting the National Quality Forum's Criteria for Measure Construction: important, valid, usable, and feasible.<sup>20</sup> The survey site randomized category ordering for each participant.

As with the prior round, measures receiving >75% of Expert Panel approval in Round 2 advanced to Round 3. In this final round, measures were randomly ordered on a single electronic page for each subject. Subjects rated on a 5-point, ordinal scale the value of each measure as it pertained to the initial prompt

(1 = This measure has no value and should be eliminated, 3 = This measure has value in some situations, 5 = This measure has definite value in most situations and should be kept). Those individual measures with a median value equal to or greater than the median for all measures were selected as the final measures. While there is wide variability among studies in defining consensus, the use of central tendency techniques employing the median value to indicate majority consensus has been described, especially in use of ordinal scales in the modified Delphi format.<sup>15,18,21</sup>

Traditionally, the final round of a Delphi study has respondents rank the most valuable items. Because the measures in this study had already approached a high level of consensus in the previous round (>75%), and because the remaining measures pertained to different categories, an ordinal scale was used so that an Expert Panel member could evaluate an individual measure on its own merit. Additionally, providing an ordinal scale, as opposed to the previous dichotomous choice, allowed more information to be collected regarding consensus on a measure. Surveys for all three rounds have been included (Appendices B, C, and D, available online).

#### *Analysis*

Fleiss Kappa scores were calculated (STATSTODO, Trading Pty LTD, Brisbane, Queensland Australia) to assess agreement for measure evaluation by the Steering Committee. Chi-squared test was calculated to determine statistical difference in response rate by subject type (OpenEpi, Emory University, Atlanta, Georgia USA). Median and mode scores were tabulated for all measures in Round 3.

The study was approved by the Baylor College of Medicine (Houston, Texas USA) Institutional Review Board.

#### **Results**

One hundred twenty-two participants were nominated by the steering committee (range 2-49 nominees/member). All were sent electronic invitations. Forty-nine people initiated the Delphi process (40.2%), of whom 39 (79.6%) completed Round 1. Twenty-six of thirty-nine (66.7%) completed Round 2, and 24/39 (61.5%) completed Round 3. Of those who completed all rounds, 33% represent international agencies, and 42% represent a US government agency (Table 2; Appendix E, available online). There was no statistical difference between composition of responders and those of partial responders and nonresponders ( $P = .95$ ).

Of the 220 articles initially screened in the literature review, 146 met inclusion criteria. Twelve hundred eighty-seven measures were extracted (Figure 1). These were consolidated into 397 unique measures, 116 (29.2%) of which were approved by the Steering Committee, (Fleiss  $\kappa = 0.0513$  [0.0259, 0.0767]). In the first round of the Delphi process, 25/116 (22%) measures were approved by the Expert Panel. Seventy-seven free-text measures (consolidated from 347 submissions) were added to this list of measures proceeding to the next round. In Round 2, 56/102 (55%) of measures were approved. In Round 3, each measure obtained a median score of 3 or greater, indicating a high level of approval. The median response for all measures combined was 4. Only 37/56 (66%) individual measures had a median score of 4 or higher and were included in the final reporting tool (Table 3). These measures describe team logistics (15),

Organization Type <sup>a</sup>	Completed (%)	Partial <sup>b</sup> (%)	No Response <sup>c</sup> (%)	Total (%)
US Government	10 (41.7)	15 (60.0)	32 (43.8)	57 (46.7)
US NGO	6 (25.0)	3 (12.0)	17 (23.3)	26 (21.3)
Non US Government	1 (4.2)	2 (8.0)	3 (4.1)	6 (4.9)
Non US NGO	5 (20.8)	5 (20.0)	11 (15.1)	21 (17.2)
IGO	2 (8.3)	0	10 (13.7)	12 (9.8)
Total by Respondent Category	24 (20%)	25 (20%)	73 (60%)	122

Daftary © 2014 Prehospital and Disaster Medicine

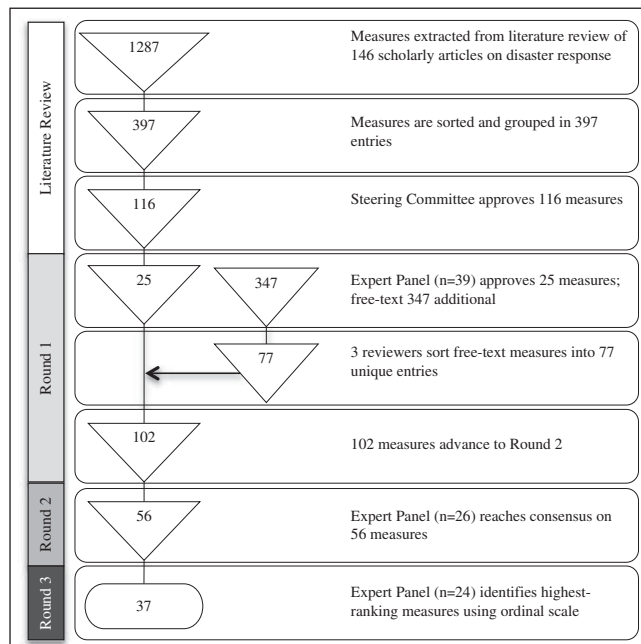
**Table 2.** Expert panel participant information. Percentages for first five data rows are within-column percentages; percentages for last row are within-row percentages. Percentages may not sum to 100% due to rounding. Chi-squared analysis of difference between responders and partial plus nonresponders:  $P = .953$ .

Abbreviations: IGO, intergovernmental organization; NGO, nongovernmental organization.

<sup>a</sup>Agency type with which respondents report affiliation. Nonresponders were identified by web search.

<sup>b</sup>Partial refers to those participants who opened the survey link but did not finish all three rounds.

<sup>c</sup>No response indicates those persons who were sent invitations but did not open the survey.



Daftary © 2014 Prehospital and Disaster Medicine

**Figure 1.** Flowchart of measure selection process. Numbers in each icon represent the number of measures that were either submitted in that round or that graduated from the prior round.

treatment (10), public health (6), disposition (4), and evaluation (2). Eleven of thirty-seven (30%) were quantitative rates.

## Discussion

Using a modified Delphi method, the authors identified 37 MOEs for disaster medical response during the acute phase following a natural disaster. Categories pertaining to treatment and team logistics had the greatest number of measures. This may reflect the focus of response teams and the source of their

greatest challenges in delivery of care. The measures with the highest median scores describe team organization (response times, incorporation of local medical staff, and establishment of clear chain of command) and assessment (number of operating health care facilities and needs assessment performed by an advanced team).

The majority of measures presented are qualitative; most of those can be answered dichotomously. Two measures (basic life support measures available and method of hand-off) are posed as open-ended questions. Only eleven rate measures graduated through all three rounds. Such measures require data intensive reporting, which may not be feasible. These measures offer detail on severity of illness (number of pediatric patients requiring ventilation and rate of transport of critical patients), treatment performance (number of patients treated and treatment of contaminated wounds with debridement and antibiotics), and outcomes (mortality rates of critical patients and cause specific mortality rates). Additional rate measures were suggested; most did not survive the consensus process.

These measures provide an opportunity for “real-time” evaluation during an acute response. It is critical that they balance the need for detailed information with the limitations of a medical response team working in austere conditions, and without the aid of data collection personnel. Such a reporting tool allows rapid assessment of effectiveness, which should be done systematically during provision of medical care.

Much has been written on the need for uniform reporting and establishing standards of care in humanitarian response, but little has been suggested in terms of objective measures that are indicative of impact. Group consensus methods provide a formalized process to elicit general opinion on a topic where little evidence exists and where sufficient experimental opportunities are not practical.<sup>15,22</sup> The Delphi process was selected to elicit expert opinion in a manner that would allow self-validation of findings. While the identified measures are merely the product of expert deliberation and have yet to be tested in an actual disaster, they represent the next best step towards objective indicators. By providing performance targets for response agencies, attention can be placed on important considerations, such as prearrival



Category	Measure	Median	Mean	Mode
Evaluation	Was a needs assessment performed by an advance team?	4.5	4.33	5
	What percentages of patients have been identified with moderate to severe injuries?	4	4.17	5
Treatment	Is sterilization equipment present?	4	4.17	5
	What basic life support measures are available?	4	4.04	5
	Is medical support provided to local health facilities?	4	3.92	3, 5
	What proportion of the essential drug list is not available to the prescribing provider? (no. of drug types out of stock/no. of drugs on essential drug list/day) <sup>a</sup>	4	3.88	4
	What is the healthcare provider to patient ratio in critical care areas? (no. of critical health care provider days/no. of critical care patient days/week) <sup>a</sup>	4	3.88	5
	What number of patients has been treated? (patients/day) <sup>a</sup>	4	3.79	5
	What is the rate of debridement of dirty wounds? (no. patients receiving debridement/no. of patients with contaminated wounds/day) <sup>a</sup>	4	3.79	3
	What rate of injuries requiring debridement received early antibiotics? (no. of patients with injury receiving antibiotics/no. of patients identified with injury requiring early antibiotics/day) <sup>a</sup>	4	3.75	4
	Are pediatric formulations for medications present?	4	3.75	3, 4
	What rate of children require assistive devices such as mechanical ventilation? (no. of pediatric patients days on critical care devices/no. of pediatric patient days/week) <sup>a</sup>	4	3.71	3
	Disposition	What is the cause-specific mortality rate? (deaths from "X" cause/those identified with "X" cause/day) <sup>a</sup>	4	3.96
What is the mortality rate of patients in critical care areas? (no. of deaths/no. of patients treated in a critical care area/day) <sup>a</sup>		4	3.92	3
What is the rate of transport for critically ill patients? (no. of critically ill patients transported to a higher acuity facility/no. of patients treated/day) <sup>a</sup>		4	3.83	4
What method is used to hand-off patient care on departure of team?		4	3.83	4
Team Logistics	What amount of time elapsed between recognition of surge needs and access to adequate personnel and equipment?	5	4.17	5
	Is electricity available?	4.5	4.33	5
	Are local medical staff incorporated into the response?	4.5	4.08	5
	Has a clear chain of command been in place?	4.5	4.08	5
	Has food and water been transported with the team?	4.5	4.08	3, 5
	Did search and rescue teams coordinate with medical teams?	4	4.04	5
	Are translators present?	4	3.96	4, 5
	Are critical care staff available?	4	3.88	5
	Are latrines available for patient use?	4	3.88	5
	Are your medical services being coordinated through the incident command center?	4	3.88	5
	Have medical teams been registered with a coordinating body prior to arrival?	4	3.83	5
	Have providers had prior disaster training?	4	3.79	4
	Does a disposal mechanism exist for human waste disposal?	4	3.71	4
	Is personal protective equipment available for providers?	4	3.71	3, 4, 5
Are safety measures in place for providers?	4	3.67	5	

Daftary © 2014 Prehospital and Disaster Medicine

Table 3. Measures of effectiveness for medical response to a major natural disaster (*continued*)

Category	Measure	Median	Mean	Mode
Public Health	How many healthcare facilities are open and operating? (functional centers/total centers/region) <sup>a</sup>	5	4.25	5
	What is the volume of clean water available per person per treatment facility? (liters of clean water/no. of persons treated/day) <sup>a</sup>	4.5	4.25	5
	Is communicable disease monitoring being performed?	4	4.04	4
	Are infection control procedures in place?	4	3.83	4
	Are public health staff incorporated into response?	4	3.88	5
	What percentage of infectious outbreaks are reported to an overseeing epidemiology agency within 24 hours of detection?	4	3.71	3

Daftary © 2014 Prehospital and Disaster Medicine

**Table 3** (continued). Measures of effectiveness for medical response to a major natural disaster.

Measures graduating through all three rounds of the group consensus process. Rate units are suggested in parentheses. Measures are sorted into major categories. Within each category, measures are presented in their rank-order by median score. Mean values are used only to sort within common median scores.

<sup>a</sup>Quantitative rate.

coordination with overseeing agencies and transition of care upon completion of mission.

Without such targets, the community of disaster responders risks repeating the challenges of prior disaster efforts: limited coordination of responders, suboptimal utilization of aid, and impacted populations unable to access life-saving care in the critical period following a major event. Evaluation of outcomes should be used to compare organizational performance, thereby allowing funding to be directed to the most effective organizations. It could also enable governments of impacted countries to select organizations with better performance to participate in the response. The Sphere Project made a monumental step when it elaborated on Minimum Standards for disaster response. The present work seeks to extend those efforts and provide greater accountability in the acute phase of a disaster medical response.

Validation is required. The list of measures should be evaluated by agencies providing acute medical care in the days and weeks following the next major natural disaster. Additionally, these acute response measures should be linked to subsequent recovery efforts in an attempt to identify which measures are predictive of reductions in morbidity and mortality, and establishment of a sustainable and effective humanitarian effort.

### Limitations

There were limitations to this study. Participants in the Expert Panel were nominated by the steering committee, potentially introducing selection bias. A larger panel of subjects may have also been beneficial to achieve thematic saturation; however, most studies of this design utilize 15-20 participants.<sup>18</sup> United States experts were disproportionately represented, despite a large number of international experts being approached. Response rate was not correlated with international status. It is unclear why so few people completed all rounds of the study. This may have been secondary to the complexity or repetitiveness of an iterative survey process. Further investigation into this might be helpful if the study is to be replicated. A single reviewer extracted measures during a review of PubMed articles and a search through online grey and unpublished literature.

This may have introduced classification bias and would reflect publication bias. However, the impact of potential measure omissions may have been negated in part by the ability of both groups to add free-text measures. Key words were not defined using a common standard and respondents may have interpreted items differently. The consolidation of redundant measures and classification of free-text measures could also have introduced classification bias. The authors attempted to decrease this risk by having three authors review measures using a predetermined process.

All measures in Round 3 received a median rating of 3 or greater. This indicates that all measures were perceived as having “value in some situations.” An overly long list of evaluation measures can be cumbersome and a criterion was put in place so as to refine this group of 56 to a more manageable 37. In some cases, a single respondent rating a measure one point higher or lower could have changed the median score, thereby qualifying or eliminating a measure from inclusion. Selected measures reflect the views of the subject matter expert panel; a different panel may have selected others. For this reason, data on all measures from Round 3 have been included (Table 3; Appendix F, available online).

### Conclusion

Group consensus methods were used to develop measures describing the functions of acute medical response to major natural disasters. These measures can facilitate detailed description of agency contributions and allow real-time assessment of performance. This is a crucial step in linking early actions and outputs to long-term outcomes and impact. Development of standards helps ensure that disaster care counts when it is most needed; this work is a step towards developing those standards.

**Note:** The views expressed in the paper are solely those of the authors and do not necessarily reflect the views, policies, or official position of the Royal Canadian Air Force, the Canadian Department of National Defence, the US Government, the US

Department of Defense, or the US Department of Health and Human Services.

### Acknowledgements

Funding by the US Department of Health and Human Services, Emergency Medical Services for Children Grant. The authors

thank Susan Torrey, Deborah Hsu, Manish Shah, and Thomas Kirsch for their mentorship and support.

### Supplementary materials

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S1049023X14000922>

### References

- Poyzner M. World Disasters Timeline. 2013. [www.mapreport.com/subtopics/d.html](http://www.mapreport.com/subtopics/d.html). Accessed August 21, 2012.
- James JJ, Subbarao I, Lanier WL. Improving the art and science of disaster medicine and public health preparedness. *Mayo Clin Proc.* 2008;83(5):559-562.
- Rosborough S. A tale of two responses: Haiti earthquake highlights the need for training in international disaster response. *Disaster Med Public Health Prep.* 2010;4(1):21-23.
- Pavignani E, Colombo S. Analysing Disrupted Health Sectors: A Modular Manual. France: World Health Organization: Department of Recovery and Transition Programmes, Health Action in Crisis; 2009.
- Reaves EJ, Schor KW, Burkle FM, Jr. Implementation of evidence-based humanitarian programs in military-led missions: part II. The impact assessment model. *Disaster Med Public Health Prep.* 2008;2(4):237-244.
- Bradt DA. Evidence-based decision-making (part II): applications in disaster relief operations. *Prehosp Disaster Med.* 2009;24(6):479-492.
- Waller SG, Ward JB, Montalvo M, Cunliffe C, Beadling C, Riley K. A new paradigm for military humanitarian medical operations: mission-generic metrics. *Mil Med.* 2011;176(8):845-851.
- Stratton SJ. The Utstein-style template for uniform data reporting of acute medical response in disasters. *Prehosp Disaster Med.* 2012;27(3):219.
- Bradt DA, Aitken P. Disaster medicine reporting: the need for new guidelines and the CONFIDE statement. *Emerg Med Australas.* 2010;22(6):483-487.
- Ciotton G (ed). *Disaster Medicine*. 3rd ed. Philadelphia, Pennsylvania USA: Elsevier; 2006.
- The Sphere Project. *Humanitarian Charter and Minimum Standards in Humanitarian Response*. 3rd ed. Northampton, United Kingdom: The Sphere Project; 2011.
- Bradt DA, Drummond CM. Rapid epidemiological assessment of health status in displaced populations—an evolution toward standardized minimum, essential data sets. *Prehosp Disaster Med.* 2003;18(1):178-185.
- Debacker M, Hubloue I, Dhondt E, et al. Utstein-style template for uniform data reporting of acute medical response in disasters. *PLoS Curr.* 2012;4:e4f6cf3e8df15a.
- Leiba A, Schwartz D, Eran T, et al. DISAST-CIR: disastrous incidents systematic analysis through components, interactions and results: application to a large-scale train accident. *J Emerg Med.* 2009;37(1):46-50.
- Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health.* 1984;74(9):979-983.
- Chan JL, Burkle FM, Jr. A framework and methodology for navigating disaster and global health in crisis literature. *PLoS Curr.* 2013;5:10.1371/currents.dis.9af6948e381dafdd3e877c441527cba0.
- Kelen G, Sauer LM. Trend analysis of disaster health articles in peer-reviewed publications pre- and post-9/11. *Am J Disaster Med.* 2008;3(6):369-376.
- Hsu C, Sandford B. The Delphi technique: making sense of consensus. *PARE.* 2007;12(10):1.
- Powell C. The Delphi technique: myths and realities. *J Adv Nurs.* 2003;41(4):376-382.
- National Quality Forum. Measure Evaluation Criteria. 2012. [http://www.qualityforum.org/docs/measure\\_evaluation\\_criteria.aspx](http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx). Accessed May 28, 2013.
- Williams PL, Webb C. The Delphi technique: a methodological discussion. *J Adv Nurs.* 1994;19(1):180-186.
- Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ.* 1995;311(7001):376-380.