Development and Validation of a Comprehensive Framework for Disaster Evaluation Typologies

Diana Wong
DipAppSc(Nsg), GradDipEmergHlth, MCP (Nsg)


Monash University Accident Research Centre, Disaster Resilience Initiative, Victoria, Australia
Copyright Notice

© Diana Wong 2018

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner’s permission.
Candidate Declaration

Monash University: Declaration for thesis based or partially based on conjointly published or unpublished work.

Thesis Including Published Works Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university of equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two original papers published in peer-reviewed journals and three unpublished pieces of work. The core theme of the thesis is disaster evaluation typologies. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working with the Monash University Accident Research Centre (MUARC) under the supervision of Emeritus Professor Frank Archer (Primary Supervisor), Dr Caroline Spencer and Professor Leanne Boyd.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

Publications During Enrolment

Publications arising from my research degree enrolment include the following.

Published in Peer-Reviewed Journals:


Unpublished Publications:


<table>
<thead>
<tr>
<th>Thesis chapter</th>
<th>Publication title</th>
<th>Status (published, in press, accepted or returned for revision)</th>
<th>Nature and % of student contribution</th>
<th>Co-author name(s) and % of co-author’s contribution*</th>
<th>Co-author(s), Monash student Y/N*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 7</td>
<td>Disaster Metrics: A Comprehensive Framework for Disaster Evaluation Typologies Prehosp Disaster Med. 2017; 32(5):1–14 (1)</td>
<td>Published</td>
<td>75% Concept development, literature review, iterative process, collecting data, design of tables, writing first draft, making amendments and finalising publication, negotiating with Editor-in-Chief of Prehospital and Disaster Medicine</td>
<td>Caroline Spencer 5%  Leanne Boyd 5%  Frederick Burkle Jnr. 5%  Frank Archer 10%</td>
<td>N</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>Peer Reviewed Validation of a Comprehensive Framework for Disaster Evaluation Typologies (submitted)</td>
<td>Submitted</td>
<td>75% Concept development, literature review, iterative process, collecting data, design of tables, writing first draft, making amendments and finalising publication, negotiating with Editor-in-Chief of Prehospital and Disaster Medicine</td>
<td>Caroline Spencer 5%  Leanne Boyd 5%  Frederick Burkle Jnr. 5%  Frank Archer 10%</td>
<td>N</td>
</tr>
</tbody>
</table>

**Note: Nature of Student Contribution**

1. Prehospital and Disaster Medicine (PDM) is the international peer-reviewed journal of the World Association for Disaster and Emergency Medicine (WADEM). PDM has published the most number of disaster health articles compared with other journals (5, 6). Given that my profession is emergency nursing, this journal in particular is highly relevant and a credible resource. PDM has a Scientific Journal Ranking (SJR) of a Q1 journal for emergency nursing.
I have not renumbered sections of this unpublished paper in order to generate a consistent presentation within the thesis.

Student signature: [Blank] Date: 21 August, 2018

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student’s and co-author’s contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: [Blank] Date: 21 August, 2018
Ethics Approvals

All projects reported in this thesis were approved by the Monash University Human Research Ethics Committee (MUHREC) as detailed below:

**Ethics Approval for the Utstein Study**

Project: Disaster Health Evaluation: A Review of the Utstein Style Guidelines/Template, MUHREC Approval CF14/158 – 2014000041 (refer to Appendix 3).

**Ethics Approval for the Validation of a Comprehensive Framework for Disaster Evaluation Typologies 2017**

Abstract

The risk of disasters resulting in mass casualties and affecting the lives of many communities continues to increase as the global population grows in number and density. The scale, frequency and impact of crises that demand international humanitarian response is also increasing. There is a moral imperative to improve the way in which disaster evaluations are undertaken and reported with the aim of reducing preventable mortality and morbidity in future events. To reduce the risk of harm and improve the health response, an evidence-based approach to disaster evaluation is urgently required. One way of reducing disaster risk to the international community is by using an agreed, standardised and comprehensive framework to structure disaster evaluations.

Disaster evaluation frameworks can support accountability, improve consistency, provide common and agreed on terminology, and explain relationships between evaluation typologies and the disaster timeline. Importantly, frameworks can also reinforce a shared responsibility for delivering better disaster interventions to the community and collectively contribute to the assessment of the effectiveness of the interventions delivered in a disaster setting.

Research has found, however, that although some evaluation frameworks exist in the disaster setting, most are fragmented, uni-focused, not validated at an international level and not used to structure disaster evaluations and research. The research undertaken during this thesis confirmed this gap, which stimulated the subsequent development and validation of a comprehensive framework that would effectively inform disaster evaluations and thereby contribute to closing this gap.

This unique, unifying comprehensive framework for disaster evaluation typologies provides structure, consistent terminology and standards for reporting that will lead to better comparability of evaluations. This is expected to provide a more robust evidence-base on which to base decision making related to interventions delivered in the disaster setting, as well as inform policy and practice. It has relevance at an international level and is expected to benefit the disaster, humanitarian and development sectors. This work promotes an environment for constructive dialogue on evaluations in the disaster setting to strengthen the evidence-base for interventions across the disaster spectrum.
I could not have completed my thesis without support from my family, supervisors and mentors, friends and work colleagues. Very special gratitude to my family who have always supported me, believed in me and provided encouragement each and every day. They have inspired me to keep writing and to see the end result – a completed thesis! For this unwavering support, I will always be grateful.

Special thanks also to my supervisors Emeritus Professor Frank Archer, Dr Caroline Spencer and Professor Leanne Boyd, who have provided the guidance, wisdom and direction that is required when undertaking a PhD. Their knowledge, insight and wisdom in navigating the PhD landscape is very much appreciated. Thank you to Ms Samantha Bailey for always helping me manage the intricacies of university enrolment and mandatory education; thank goodness she was there to smooth the way.

Undertaking a PhD has offered me so many amazing opportunities and I have appreciated each and every one of them, especially working with Professor Virginia Murray (Public Health England), Professor Frederick Burkle Jnr. (Senior Fellow and Scientist, Harvard Humanitarian Institute, Harvard University and Harvard School of Public Health) and Jonathan Abrahams (Disaster Risk Management Focal Point, World Health Organization). These mentors facilitated my travelling to Sendai, Japan to attend and participate in the 3rd World Conference on Disaster Risk Reduction, where I was able to witness the historical agreement of the Sendai Framework for Disaster Risk Reduction 2015–2030, and to Geneva to present ‘Disaster Evaluation Typologies’ at the UNISDR Science and Technology Conference on the Implementation of the Sendai Framework.

To my university and hospital colleagues and friends, you made me laugh when I wanted to cry, you made me smile and keep moving forward. For all those that I have not named specifically and who have been part of the journey, you know who you are, your kind words, your support and encouragement has helped, and you have my deepest thanks.
In loving memory of my father Lodewyk Johannes Van Dam who had the courage and foresight to leave war-torn Europe and travel to Australia to give his family a better life. You never had the chance to complete your university studies as a result of World War II. I dedicate this thesis to you and to all those who have been impacted by disasters or conflict, and whose lives have been changed forever.
# Table of Contents

Copyright Notice ........................................................................................................... ii
Candidate Declaration ...................................................................................................... iii
Ethics Approvals ............................................................................................................ vii
Abstract ........................................................................................................................ viii
Acknowledgements ....................................................................................................... ix
Dedication ....................................................................................................................... x
Table of Contents .......................................................................................................... xi
List of Figures ............................................................................................................... xiv
List of Tables ............................................................................................................... xvi
Glossary ......................................................................................................................... xviii
Prolegomenon ................................................................................................................ xxii

## Chapter 1: Introduction

1.1 Setting ..................................................................................................................... 1
1.2 Problem Statement ................................................................................................. 2
1.3 Aim of the Research ............................................................................................... 2
1.4 Significance of the Research ................................................................................... 3
1.5 Structure of Thesis ................................................................................................. 4

## Chapter 2: Background and Context for the Research

2.1 Introduction ............................................................................................................. 8
2.2 Context .................................................................................................................... 8
2.3 Key Definitions ....................................................................................................... 11
   2.4.1 The Importance and Relevance of this Paper .................................................. 13
   2.4.2 3ie Scoping Paper 1: ‘What Evidence Is Available and What Is Required, In Humanitarian Assistance?’ ......................................................... 15
   2.4.3 Implications of the 3ie Scoping Paper 1 for this Thesis ..................................... 115
2.5 Research Questions for this Thesis ........................................................................ 116
   2.5.1 Overarching Research Questions .................................................................... 116
   2.5.2 Sub Questions .................................................................................................. 117
2.6 Summary ................................................................................................................ 119

## Chapter 3: Conceptual Framework, Research Design and Methodology

3.1 Introduction ............................................................................................................. 120
3.2 Conceptual Framework of this Thesis .................................................................... 120
3.3 Research Design and Methodology ....................................................................... 121
   3.3.1 Establishment of the Core Structure ............................................................... 122
   3.3.2 Development of a Comprehensive Framework .............................................. 126
   3.3.3 Validation of a Comprehensive Framework ..................................................... 127
3.4 Summary ................................................................................................................ 129

## Chapter 4: Literature Review of Evaluation Frameworks and Typologies in the Disaster Setting

4.1 Introduction ............................................................................................................. 130
Chapter 5: Pre-Establishment of the Core Structure - Two Studies

5.1 Introduction .............................................................................. 168
5.2 Study of the Utstein Guidelines .................................................. 168
  5.2.1 Research Questions .............................................................. 169
  5.2.2 Research Design and Methodology .......................................... 169
  5.2.3 Results .............................................................................. 176
  5.2.4 Discussion ......................................................................... 186
  5.2.5 Summary ........................................................................... 189
5.3 Study of Reviews of Australian Disaster Events .............................. 189
  5.3.1 Research Questions .............................................................. 190
  5.3.2 Research Design and Methodology .......................................... 191
  5.3.3 Results .............................................................................. 203
  5.3.4 Discussion ......................................................................... 206
  5.4 Summary .............................................................................. 206

Chapter 6: Establishment of the Core Structure ........................................ 207

6.1 Introduction .............................................................................. 207
6.2 Establishment of the Core Structure .............................................. 207
  6.2.1 Layer 1 ........................................................................... 209
  6.2.2 Layer 2 ........................................................................... 214
  6.2.3 Layer 3 ........................................................................... 220
6.3 Discussion .............................................................................. 223
6.4 Conclusion .............................................................................. 224

Chapter 7: Development of the Comprehensive Framework for Disaster

Evaluation Typologies 2017 ................................................................. 225

7.1 Introduction .............................................................................. 225
7.2 Research Questions ................................................................... 225
7.3 Research Design and Methodology .............................................. 226
  7.3.1 Literature Review .............................................................. 226
  7.3.2 Framework Construction .................................................... 227
7.4 Results – Comprehensive Framework for Disaster Evaluation Typologies 2017 .............................................. 241
  7.4.1 Comprehensive Framework for Disaster Evaluation Typologies 2017 .............................................. 241
7.5 Discussion .............................................................................. 243
7.6 Conclusion .............................................................................. 244

Chapter 8: Validation of the Comprehensive Framework for Disaster

Evaluation Typologies 2017 ................................................................. 245

8.1 Introduction .............................................................................. 245
8.2 Research Questions ................................................................... 246
8.3 Research Design and Methodology .............................................. 247
  8.3.1 Literature Review .............................................................. 248
  8.3.2 Action Research / Conference Presentations ............................ 249
  8.3.3 Mixed Methods Research – Online Survey: Design and Sample .... 250
List of Figures

Figure 1: Structure of Thesis ........................................................................................................... 5
Figure 2: Research Aim and Questions .......................................................................................... 118
Figure 3: Comprehensive Search Strategy – Disaster Evaluation Frameworks ....................... 133
Figure 4: Comprehensive Search Strategy – Disaster Evaluation Typologies ......................... 134
Figure 5: Disaster Evaluation Framework Timeline ...................................................................... 141
Figure 6: Key Elements as Identified in the Utstein Guidelines 2003 ........................................ 151
Figure 7: The Impacts Framework (18) ......................................................................................... 160
Figure 8: Regionalisation of Disaster Events in the Study ............................................................ 193
Figure 9: Core Structure, Layer 1 ................................................................................................ 209
Figure 10: OCHA Cluster Coordination (404) ............................................................................. 211
Figure 11: New Zealand's Holistic Framework for Disaster Recovery (405) .............................. 212
Figure 12: Core Structure, Layer 2 ............................................................................................. 215
Figure 13: Core Structure, Layer 3 ............................................................................................. 221
Figure 14: Core Structure ............................................................................................................. 224
Figure 15: Six Themes .................................................................................................................. 227
Figure 16: Core Structure and Accountability ............................................................................ 228
Figure 17: Core Structure and Evaluation Standards and Evidence ........................................... 230
Figure 18: Core Structure and Baselines ..................................................................................... 235
Figure 19: Core Structure and Consequences ............................................................................ 237
Figure 20: Core Structure and Outcomes ................................................................................... 238
Figure 21: Core Structure and Impact Evaluations ..................................................................... 240
Figure 22: Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017) .................................................................................................................. 242
Figure 23: WHO Regional Offices ............................................................................................... 252
Figure 24: CFDET 2017 Core Structure ...................................................................................... 281
Figure 25: CFDET 2017 Core Structure Layer 1 ........................................................................ 282
Figure 26: CFDET 2017 Core Structure Layer 2 ........................................................................ 283
Figure 27: CFDET 2017 Core Structure Layer 3 ........................................................................ 284
Figure 28: CFDET 2018 Core Structure ...................................................................................... 286
Figure 29: CFDET 2017: Baselines ............................................................................................. 287
Figure 30: CFDET 2018 Baseline Typologies ............................................................................. 289
Figure 31: CFDET 2017 Consequences ...................................................................................... 290
Figure 32: CFDET 2018 Consequence Typologies ...........................................291
Figure 33: CFDET 2017 Outcomes .................................................................292
Figure 34: CFDET 2018 Outcome Typologies ..............................................293
Figure 35: CFDET 2017 Impact Evaluations ..................................................294
Figure 36: CFDET 2017 Accountability .........................................................295
Figure 37: CFDET 2018 Outcome Typologies ..............................................297
Figure 38: CFDET 2017 Evaluation Standards and Evidence .........................298
Figure 39: CFDET 2018 Evaluation Standards and Evidence .........................299
Figure 40: Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018) ..........................................................................................302
List of Tables

Table 1: Literature Review Results – Disaster Evaluation Frameworks .........................137
Table 2: Usable Papers from the Literature Review of Disaster Evaluation Frameworks .................................................................................................................................138
Table 3: Expert Opinion Results - Disaster Evaluation Frameworks .................................138
Table 4: Snowballing Results – Disaster Evaluation Frameworks ................................139
Table 5: Literature Review Results - Disaster Evaluation Typologies ............................162
Table 6: Literature Review Results: Evaluation Typologies – Baselines .........................163
Table 7: Literature Review Results: Evaluation Typologies – Consequences ...............164
Table 8: Literature Review Results: Evaluation Typologies – Outcomes .......................164
Table 9: Literature Review Results: Evaluation Typologies – Impact Evaluations ..........165
Table 10: Literature Search Strategy 2003 to 2017 .........................................................171
Table 11: Usable Papers from the Literature Review of Disaster Evaluation Frameworks .................................................................................................................................172
Table 12: Snowballing Results - Disaster Evaluation Frameworks .................................173
Table 13: Initial Survey Responses ................................................................................177
Table 14: Stratification by Country ................................................................................178
Table 15: Stratification by WHO Region ........................................................................178
Table 16: Relevant Experience of Participants ................................................................179
Table 17: Question 1A – Are you aware of the Utstein Guidelines 2003? ......................179
Table 18: Question 1B – How did you gain this awareness? ...........................................180
Table 19: Question 1C – Do you use them? Yes/No .......................................................181
Table 20: Question 1D – If ‘Yes’, are they valuable? .......................................................181
Table 21: Question 1E – How are they valuable? .............................................................182
Table 22: Question 1F - Are you aware of anyone else using them? .............................182
Table 23: Thematic Analysis of Question 2 – From the literature review, we can see that the Utstein Guidelines are being referenced but not used. Do you have any suggestions to offer on why they are not being used? .................................................................183
Table 24: Question 3 – Are you aware of any other similar guidelines being used? ......184
Table 25: Question 4 – Do you have any suggestions on how we can improve the evaluation of disaster health? .................................................................185
Table 26: Key Element Tool .........................................................................................191
Table 27: Key Elements to be Included in the Establishment of the Core Structure ....204
Table 28: Summary of Survey Responses .......................................................... 254
Table 29: Initial Survey Responses .................................................................. 255
Table 30: Age Groups .................................................................................... 256
Table 31: Gender ........................................................................................... 256
Table 32: Stratification by Country ............................................................... 257
Table 33: Stratification by WHO Region ......................................................... 257
Table 34: Area of Expertise ......................................................................... 258
Table 35: Years of Full-Time Equivalent in Current Field ......................... 258
Table 36: Undertaking Evaluations ............................................................... 259
Table 37: Survey Responses to CFDET 2017 Elements .......................... 259
Table 38: Core Structure Comments – Disaster Timeline Elements Not Covered ...... 260
Table 39: Core Structure Comments – Substitution of Any Disaster Timeline Elements .................................................................................. 262
Table 40: Baseline Comments .................................................................... 263
Table 41: Consequences Comments ............................................................ 264
Table 42: Outcomes Comments ................................................................. 266
Table 43: Impact Evaluation Comments ...................................................... 268
Table 44: Accountability Comments .............................................................. 270
Table 45: Evaluation Standards and Evidence Comments ......................... 272
Table 46: Survey Responses to CFDET 2017 as a Unifying Framework .. 273
Table 47: Relationship Between Evaluation Typologies and Disaster Timeline Comments .................................................................................. 273
Table 48: Comprehensive Framework Potential Value Comments ............ 275
Table 49: Supporting and Promoting Evaluations Comments ..................... 276
Table 50: Usefulness in Teaching Evaluation Comments ......................... 277
Table 51: Additional Thoughts on Improving the Comprehensive Framework ...... 277
Table 52: Perceived Barriers to Undertaking Evaluations in the Disaster Setting ...... 278
Table 53: Perceived Enablers to Undertaking Evaluations in the Disaster Setting ...... 279
Table 54: Utility Standards Criteria Applied to the Research .................... 313
Table 55: Feasibility Standards Criteria Applied to the Research ............. 314
Table 56: Propriety Standards Criteria Applied to the Research ................ 316
Table 57: Accuracy Standards Criteria Applied to the Research .................. 317
**Glossary**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>3ie</td>
<td>International Initiative for Impact Evaluation</td>
</tr>
<tr>
<td>AADM</td>
<td>American Academy of Disaster Medicine</td>
</tr>
<tr>
<td>ACFID</td>
<td>Australian Council for International Development</td>
</tr>
<tr>
<td>AEA</td>
<td>American Evaluation Association</td>
</tr>
<tr>
<td>AEMI</td>
<td>Australian Emergency Management Institute</td>
</tr>
<tr>
<td>AES</td>
<td>Australasian Evaluation Society</td>
</tr>
<tr>
<td>AFAC</td>
<td>Australasian Fire and Emergency Service Authorities Council</td>
</tr>
<tr>
<td>AIDR</td>
<td>Australian Institute for Disaster Resilience</td>
</tr>
<tr>
<td>AJEM</td>
<td>Australian Journal of Emergency Medicine</td>
</tr>
<tr>
<td>ALNAP</td>
<td>Active Learning for Accountability and Performance</td>
</tr>
<tr>
<td>ANZBA</td>
<td>Australian and New Zealand Burn Association</td>
</tr>
<tr>
<td>APCDM</td>
<td>Asia Pacific Conference on Disaster Medicine</td>
</tr>
<tr>
<td>AR</td>
<td>Action Research</td>
</tr>
<tr>
<td>AUSAID</td>
<td>Australian Agency for International Development</td>
</tr>
<tr>
<td>AUD</td>
<td>Australian dollar</td>
</tr>
<tr>
<td>AUSMAT</td>
<td>Australian Medical Assistance Team</td>
</tr>
<tr>
<td>BBB</td>
<td>Build Back Better</td>
</tr>
<tr>
<td>BoM</td>
<td>Bureau of Meteorology</td>
</tr>
<tr>
<td>BSF</td>
<td>Basic Societal Functions</td>
</tr>
<tr>
<td>CDC</td>
<td>Centre for Disease Control and Prevention</td>
</tr>
<tr>
<td>CF</td>
<td>Comprehensive Framework</td>
</tr>
<tr>
<td>CFDET 2017</td>
<td>Comprehensive Framework for Disaster Emergency Typologies 2017</td>
</tr>
<tr>
<td>CFDET 2018</td>
<td>Comprehensive Framework for Disaster Emergency Typologies 2018</td>
</tr>
<tr>
<td>CHS</td>
<td>Core Humanitarian Standards</td>
</tr>
<tr>
<td>COD</td>
<td>Common Operational Datasets</td>
</tr>
<tr>
<td>CONFIDE</td>
<td>CONsensus Guidelines on Reports of Field Interventions in Disasters and Emergencies</td>
</tr>
<tr>
<td>COP21</td>
<td>Climate Change Conference 2015</td>
</tr>
<tr>
<td>CRED</td>
<td>Centre for Research on the Epidemiology of Disasters</td>
</tr>
<tr>
<td>DFAT</td>
<td>Department of Foreign Affairs and Trade</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>DFID</td>
<td>Department for International Development</td>
</tr>
<tr>
<td>DG SANCO</td>
<td>Directorate General for Health and Consumers</td>
</tr>
<tr>
<td>DLM</td>
<td>Disaster Logic Model</td>
</tr>
<tr>
<td>DRR</td>
<td>Disaster Risk Reduction</td>
</tr>
<tr>
<td>DRSG</td>
<td>Pittsburgh Disaster Reanimatology Study Group</td>
</tr>
<tr>
<td>DSS</td>
<td>Disaster Severity Scale</td>
</tr>
<tr>
<td>EDMD</td>
<td>European Master in Disaster Medicine</td>
</tr>
<tr>
<td>EMA</td>
<td>Emergency Management Australia</td>
</tr>
<tr>
<td>EMDM</td>
<td>Emergency Management and Disaster Medicine</td>
</tr>
<tr>
<td>EMSB</td>
<td>Emergency Management of Severe Burns</td>
</tr>
<tr>
<td>EMT</td>
<td>Emergency Medical Teams</td>
</tr>
<tr>
<td>ENP</td>
<td>European Neighbourhood Policy</td>
</tr>
<tr>
<td>ERD</td>
<td>Emergency and Research Database</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FAADM</td>
<td>Fellow of the American Academy of Disaster Medicine</td>
</tr>
<tr>
<td>FMAT</td>
<td>Foreign Medical Assistance Team</td>
</tr>
<tr>
<td>FOD</td>
<td>Fundamental Operational Datasets</td>
</tr>
<tr>
<td>GDN</td>
<td>The Global Research Capacity Building Program</td>
</tr>
<tr>
<td>GDN-3ie</td>
<td>The Global Research Capacity Building Program and International Initiative for Impact Evaluation</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross Domestic Product</td>
</tr>
<tr>
<td>GFDRR</td>
<td>Global Facility for Disaster Reduction and Recovery</td>
</tr>
<tr>
<td>GHA</td>
<td>Global Humanitarian Assistance</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information Systems</td>
</tr>
<tr>
<td>Habitat III</td>
<td>Urban Conference on Housing and Sustainable Urban Development</td>
</tr>
<tr>
<td>HDR</td>
<td>Higher Degree by Research</td>
</tr>
<tr>
<td>HELP</td>
<td>Humanitarian Evaluation, Learning and Performance Portal</td>
</tr>
<tr>
<td>HEMU</td>
<td>Health Emergency Management Unit</td>
</tr>
<tr>
<td>HIA</td>
<td>Health Impact Assessment</td>
</tr>
<tr>
<td>HITW</td>
<td>Humanitarian Interventions Thematic Window</td>
</tr>
<tr>
<td>IASC</td>
<td>Inter-Agency Standing Committee</td>
</tr>
<tr>
<td>ICN</td>
<td>International Council of Nurses</td>
</tr>
<tr>
<td>ICRC</td>
<td>International Committee of the Red Cross</td>
</tr>
<tr>
<td>IEG</td>
<td>Independent Evaluation Group</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>IFLA</td>
<td>International Federation of Library Associations and Institutions</td>
</tr>
<tr>
<td>IFRC</td>
<td>International Federation of Red Cross and Red Crescent Societies</td>
</tr>
<tr>
<td>IHR</td>
<td>International Health Regulations</td>
</tr>
<tr>
<td>IMT</td>
<td>Incident Management Training</td>
</tr>
<tr>
<td>IRA</td>
<td>Initial Rapid Assessment</td>
</tr>
<tr>
<td>JCSEE</td>
<td>Joint Committee on Standards for Educational Evaluation</td>
</tr>
<tr>
<td>KAMEDO</td>
<td>Katastrofmedicinska Organisationskommitten</td>
</tr>
<tr>
<td>M &amp; E</td>
<td>Monitoring and Evaluation</td>
</tr>
<tr>
<td>MD</td>
<td>Doctor of Medicine</td>
</tr>
<tr>
<td>MDG</td>
<td>Millennium Development Goals</td>
</tr>
<tr>
<td>MIMMS</td>
<td>Major Incident Medical Management Support</td>
</tr>
<tr>
<td>MIRA</td>
<td>Multi-Cluster/Sector Initial Rapid Needs Assessment</td>
</tr>
<tr>
<td>MIRI</td>
<td>Monash Injury Research Institute</td>
</tr>
<tr>
<td>MMR</td>
<td>Mixed Methods Research</td>
</tr>
<tr>
<td>MS</td>
<td>Master of Science</td>
</tr>
<tr>
<td>MSF</td>
<td>Medecins Sans Frontieres</td>
</tr>
<tr>
<td>MUDRI</td>
<td>Monash University Disaster Resilience Initiative</td>
</tr>
<tr>
<td>MUHREC</td>
<td>Monash University Human Research Ethics Committee</td>
</tr>
<tr>
<td>MUARC</td>
<td>Monash University Accident Research Centre</td>
</tr>
<tr>
<td>NCCTRC</td>
<td>National Critical Care and Trauma Response Centre</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-Government Organisation</td>
</tr>
<tr>
<td>NSDR</td>
<td>National Strategy for Disaster Resilience</td>
</tr>
<tr>
<td>NSW</td>
<td>New South Wales</td>
</tr>
<tr>
<td>OCHA</td>
<td>Office for the Coordination of Humanitarian Affairs</td>
</tr>
<tr>
<td>ODI</td>
<td>Overseas Development Institute</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>OECD-DAC</td>
<td>Organisation for Economic Co-operation and Development - Development Assistance Committee</td>
</tr>
<tr>
<td>OFDA</td>
<td>United States Office of Foreign Disaster Assistance</td>
</tr>
<tr>
<td>PDNA</td>
<td>Post Disaster Needs Assessment</td>
</tr>
<tr>
<td>PDM</td>
<td>Prehospital and Disaster Medicine</td>
</tr>
<tr>
<td>PR</td>
<td>Public Relations</td>
</tr>
<tr>
<td>Qld</td>
<td>Queensland</td>
</tr>
<tr>
<td>RC</td>
<td>Rockefeller Foundation 100 Resilient Cities Framework</td>
</tr>
<tr>
<td>SDG</td>
<td>Sustainable Development Goals</td>
</tr>
</tbody>
</table>
SFDRR | Sendai Framework for Disaster Risk Reduction 2015 – 2030
SHO | Samenwerkende Hulporganisaties (Cooperating Aid Organisations)
SJR | Scientific Journal Ranking
SIPRI | Stockholm International Peace Research Institute
ToR | Terms of Reference
TFQCDM | Task Force on Quality Control of Disaster Management
UK | United Kingdom
UN | United Nations
UNEG | United Nations Evaluation Group
UNHCR | United Nations High Commissioner for Refugees
UNICEF | United Nations International Children’s Emergency Fund
UNISDR | United Nations Office for Disaster Risk Reduction
URL | Uniform Resource Locator
US | United States
USA | United States of America
USAID | United States Agency for International Development
USD | United States dollar
VBRC | Victorian Bushfires Royal Commission
WA | Western Australia
WADEM | World Association for Disaster and Emergency Medicine
WASH | Water, sanitation and hygiene
WB | World Bank
WCDEM | World Congress on Disaster and Emergency Medicine
WHO | World Health Organization
WHS | World Humanitarian Summit
Self

With over 30 years’ experience in emergency and disaster health, and as the daughter of migrants impacted by World War II, I feel very strongly about making a difference to those suffering now and those that may suffer in the future as a result of disasters and conflict.

I knew from a very early age that nursing was my calling and that I would be able to make a difference in peoples’ lives. I started my nursing career at a time when training was moving out of the hospitals and into the universities. This created much debate and some uncertainty; however, my focus was unwavering. On graduation, I commenced work in the Emergency Department at Liverpool Hospital – a place that I continue in today. During my nursing career, I have worked in a variety of roles including Registered Nurse, Registered Midwife, Certified Emergency Nurse, Clinical Nurse Consultant and Clinical Nurse Specialist in the Emergency Department, Hospital Emergency Manager, Trauma Case Manager, and I am a currently qualified member of the Australian Medical Assistance Team (AUSMAT). This places me well to act as a consultant and make leadership contributions in the emergency health sector.

When considering a postgraduate research degree, my initial thoughts were related to undertaking an evaluation of the Quakers Hill nursing home fires of 2011 from a health perspective. As a consequence of preliminary research, it became apparent that there was no complete or comprehensive evaluation methodology that would support such an undertaking. Consequently, it was believed that the development of a comprehensive framework for disaster evaluation typologies would have significant impact, and this became the research priority. From a health perspective and through a quality improvement lens, I am passionate about improving practice by consolidating past experiences and improving future response efforts, and sharing knowledge for the benefit of better-quality patient care.

---

1 A prolegomenon consists of prefatory remarks; specifically, a formal essay or critical discussion serving to introduce and interpret an extended work (Merriam-Webster, online, Prolegomenon, 2018).
List of Publications

These have been listed in my candidate declaration.

Funded Research Grant

In addition to the publications mentioned earlier, I was instrumental in identifying the opportunity to participate in a funded research project through an international competitive process. With the support of my supervisors, a successful submission was forwarded for the following project that also resulted in a published technical report:


This technical paper is the first in a series of nine scoping papers published by 3ie. The complete series of scoping papers can be located at the following URL:
http://www.3ieimpact.org/en/publications/3ie-scoping-paper-series/

Peer-Reviewed Conference Presentations

The following list identifies the conference presentations I have delivered, or have been delivered on my behalf, during the course of this postgraduate study.

2018 Evidence for Global & Disaster Health Special Interest Group, The World Library and Information Congress 84th International Federation of Library Associations and Institutions (IFLA) General Conference and Assembly, 24–30 August 2018, Kuala Lumpur, Malaysia. ‘Disaster Metrics: A Comprehensive Framework for Disaster Evaluation Typologies.’ This was presented on my behalf by Emma Farrow from Public Health England.


2015 World Congress on Disaster and Emergency Medicine, 21–24 April 2015, Cape Town, South Africa. Three papers were presented at this conference as the lead author:
• ‘A Review of the History and Use of Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style’ (7)
• ‘Thematic Analysis of 7 Australian Disaster Reports or Inquiries’ (8)
• ‘A Framework for Structuring Post Disaster Evaluations’ (9).

An additional paper was also presented where I was a co-author:

• ‘What Evidence is Available and What is Required, in Humanitarian Assistance?’, presented by Claire Allen.

2014  Keynote speaker at 12th Asia Pacific Conference on Disaster Medicine, 17–19 September 2014, Tokyo, Japan, delivering the following addresses:

• ‘Disaster Metrics: Advancing Disaster Evaluation Methodologies – the Next Great Challenge for Disaster Health’
• ‘Disaster Evaluation Repositories (A Work in Progress)’.

2013  18th World Congress on Disaster and Emergency Medicine, 21–23 May 2015, Manchester, United Kingdom. This was presented on my behalf by my Primary Supervisor: ‘Collections/Repositories of Post Emergency / Disaster Evaluation Reports – a Valuable but Unknown and Untapped Resource’.

Additional Presentations Delivered at Monash University and Liverpool Hospital

In addition to presenting at the above peer-reviewed conferences, additional opportunities were taken to broaden and share my knowledge of disaster evaluations while also expanding my network. The following presentations were delivered at Monash University and Liverpool Hospital.

Seminars – Monash University

- 2016 – MUDRI Mid-Year HDR Colloquium, ‘Disaster Health Evaluation’
• 2013 – MUDRI Second Annual Monash Disaster Research Symposium, ‘Methodological Issues in Disaster Evaluation’
• 2012 – MUDRI Inaugural Monash Disaster Research Symposium, ‘Disaster Evaluation’.

Seminars – Liverpool Hospital

• 2014 – Liverpool Hospital Nursing and Midwifery Research Forum, ‘Disaster Health Evaluation’.

Attendance at Evaluation Training Programs

To support my continued professional development and interest in disaster evaluations, I attended the following evaluation training programs:

• 2014 – Introduction to Systematic Reviews of Health Interventions; School of Public Health and Preventative Medicine, Monash University; Melbourne. (2 days)
• 2014 – Next Steps in Monitoring and Evaluation; Australian Council for International Development (ACFID); Melbourne. (2 days)
• 2014 – Next Steps in Monitoring and Evaluation Workshop; Australian Evaluation Society; Sydney. (1 day)
• 2012 – Evaluation of Humanitarian Assistance – Beginner and Advanced Levels; Brussels, Belgium. (5 days).

External Industry and Professional Engagement

During my candidature I was very keen to maintain my industry links and professional engagement. This enabled me to continually integrate the learnings from my research endeavours with my professional leadership engagement at both national and international levels. Of special note were:

• Invitation to attend the University of Melbourne, Health and Disaster Risk Reduction: State of the Art and Implications for Australian Findings, expert
workshop regarding the Sendai Framework July 10, 2017 University of Melbourne, Victoria, Australia.

Other significant contributions include:

- Research Priorities for Older People in the Context of Natural Hazards, Disasters and Emergencies, Macquarie University, Sydney, 2018.
- Expert workshop held at University of Melbourne to discuss Disaster Risk Reduction for the Health Sector under the Sendai Framework for Disaster Risk Reduction (refer to Appendix 1), 2017.
- Crowded Places and Mass Gatherings Expert Forum, Australian Institute for Disaster Resilience (AIDR), Sydney, 2017. This led to the publication of ‘Safe and Healthy Crowded Places, Australian Disaster Resilience Handbook Collection, Handbook 15’ in 2018, in which I have been acknowledged as a contributor. The publication can be found at the following URL: https://knowledge.aidr.org.au/resources/handbook-15-safe-and-healthy-crowded-places/ (10).
- WHO Emergency Medical Teams (EMT), Hong Kong, 2016.
- Subject matter expert for the University of Sydney, Graduate Certificate in Emergency Nursing, ‘Hospital Disaster Preparedness and Response’, 2015–2016.
- People in Disasters; Canterbury District Health Board; Christchurch ‘Operationalising the Deployment and Coordination of Foreign Medical Teams – Supporting the World Health Organization FMT Global Registry’, and ‘The


- Australian Medical Assistance Team (AUSMAT) Training; National Critical Care and Trauma Response Centre (NCCTRC) and NSW Health Emergency Management Unit (HEMU), 2014.

- Lead author in successful submission for Australasian Emergency Department of the Year 2012, awarded to Liverpool Hospital, Emergency Department, 2012.
Chapter 1: Introduction

The thesis, ‘Development and Validation of a Comprehensive Framework for Disaster Evaluation Typologies’, focuses on developing and validating a comprehensive framework that supports a systematic and timely approach to disaster reporting and the evaluations of health interventions initiated as a response to disasters. It is anticipated that this new framework will facilitate agreement on identifying, structuring and relating the various evaluation typologies found in the disaster setting. The outputs of this thesis are relevant to the fields of disaster health and emergency preparedness at an international level. This focused approach to disaster evaluation may assist in reducing the devastating losses that communities face each year as a result of disaster events.

1.1 Setting

The risk of disasters resulting in mass casualties and affecting the lives of many communities continues to increase as the global population grows in number and density. The scale, frequency and impact of ‘disastrous’ natural hazards, compounded by climate change (11), is also increasing (12, 13). Rapid and unsustainable urbanisation (12, 14, 15) and increasing violence and conflict causing instability and mass displacement of people present long-term challenges to the global community (13). There is a moral imperative to improve the way in which disaster evaluations are undertaken and reported with the aim of reducing preventable mortality and morbidity in future events. To reduce the risk of harm and to improve the health response, an evidence-based approach to disaster evaluation is urgently required. One way of reducing disaster risk to the international community is by using a framework to structure disaster evaluations.

Disaster evaluation frameworks can support accountability, improve consistency, provide common and agreed on terminology, and explain relationships between evaluation typologies and the disaster timeline. Importantly, frameworks can also reinforce a shared responsibility for delivering better disaster interventions to the community and collectively contribute to the assessment of the effectiveness of these interventions. Currently, there is insufficient high-quality evidence available (3). Traditional evaluations, which monitor processes or assess if targets have been achieved, are inadequate. There is a paucity of theory-based, reliable evidence causally linking health interventions to relevant outcomes in the disaster and humanitarian setting. Reliable,
high-quality evidence is required to improve the effectiveness and efficiency of such interventions (16).

To reduce the risk of harm, an evidence-based approach to disaster reporting and disaster evaluations is urgently required. As part of an evidence-based approach, this thesis develops and validates a comprehensive framework for disaster evaluation typologies at an international level. This unique and unifying framework ‘could provide consistency in terminology and standards for reporting across the different phases of a disaster, with a view to providing comparability to better understand the process, outcomes and impacts of the efficacy and efficiency of interventions’ (1)(p.502). The sharing of these methodological experiences has the ability to contribute to the continued development of these standards and guidelines to build disaster science (1).

1.2 Problem Statement

Research suggests that, although some evaluation frameworks exist in the disaster setting, most are fragmented, uni-focused, not validated at an international level and not used to structure disaster evaluations and research. An agreed, comprehensive framework to structure disaster evaluation typologies has yet to be developed. This gap stimulated the theme for this thesis. The research undertaken during this thesis confirmed this gap and stimulated the subsequent development and validation of a comprehensive framework that would effectively inform disaster evaluations and thereby contribute to closing this gap.

1.3 Aim of the Research

The aim of the research is to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline. The comprehensive framework consists of current evaluation typologies structured to highlight current processes and relationships. This framework provides consistency in terminology and standards for reporting across the different phases of a disaster. This consistency will allow for improved comparability to better understand the process, outcomes and impacts of the efficacy and efficiency of interventions. It is anticipated that this new framework will facilitate an agreement on organising and describing the various evaluation typologies found in the disaster setting. This work will add structure to the current understanding and help to underpin the diversity of disaster evaluation typologies that currently exist (1).
1.4 Significance of the Research

A comprehensive framework for disaster evaluation typologies has the ability to provide consistency in terminology and standards for reporting across different phases of a disaster. It will provide comparability to better understand the process, outcomes and impacts of the efficacy and efficiency of interventions. Sharing methodological experiences would also contribute to the further development of these standards and guidelines to systematically build disaster science. This body of work facilitates a better understanding of disaster evaluations as this is a specialised area of study at an international level and may help simplify the complex nature of disaster evaluations.

The comprehensive framework that has been developed and validated as part of this thesis builds on the earlier work of disaster researchers such as the Task Force on Quality Control of Disaster Management (TFQCDM) (17), Stephenson (18), Powers and Daily (19), Fattah et al. (20-24), Birnbaum et al. (25-34) and Sundnes (35-57). This unique work not only consolidates diverse non-validated disaster frameworks into a single diagram, but provides the conceptual relationships between the various evaluation typologies and disaster timeline (1). The framework also allows for a better awareness and understanding of the different types of evaluations in the disaster setting.

Significantly, the developed framework incorporates globally significant disaster, humanitarian and development frameworks and agendas, such as the Sendai Framework for Disaster Risk Reduction 2015–2030 (2015) (58) and the Sustainable Development Goals, 17 Goals to Transform Our World (2015) (59). The recognition and inclusion of these supports and promotes the international focus of disaster risk reduction, in keeping with the Sendai Framework, and the focus on reducing disaster risk and losses in lives, livelihoods and health in affected communities (58).

The results obtained from rigorous evaluations undertaken in the disaster setting has the ability to build the science of disaster health. Understanding the causal relationships between the interventions delivered during a disaster and subsequent changes in targeted results provides evidence of what works, for whom, why, and at what cost. This can also assist in identifying any unintended consequences of interventions (3, 16). Further, the evidence gained can be used to inform policy and practice with a quality improvement theme, underpinned by the concepts of learning and accountability.
1.5 Structure of Thesis

This thesis has ‘disaster evaluation typologies’ as its unifying theme, and outlines the development and validation of a unique comprehensive framework for structuring disaster evaluation typologies. It consists of 10 chapters, which are discussed below and shown in Figure 1.

Chapter 1 establishes the background and need for this research by identifying the problem statement and the aim of the research. The significance of the study at an international level is highlighted and the structure of the thesis presented.

Chapter 2 focuses on the background and context for this research. Key definitions and concepts are clearly defined to ensure consistency in understanding the context in which they are used. This chapter introduces the 3ie Scoping Paper 1, ‘What evidence is available and what is required, in humanitarian assistance?’ (2014) (3) that the researcher was directly involved in. This internationally funded project provided an independent analysis of the evidence-base of impact evaluations in the humanitarian sector. The results demonstrated a paucity of evidence and lack of high-quality impact evaluations to inform policy and practice (3, 16). The report identified gaps and inconsistencies in methodologies and terminologies in disaster evaluations, and a clear need to develop and validate a comprehensive framework to structure disaster evaluations. The overarching research questions that underpin this thesis are clearly stated.

Chapter 3 describes the conceptual framework, research design and methodology that underpin this thesis.

Chapter 4 details the literature review undertaken of evaluation frameworks and evaluation typologies in the disaster setting. The methodology, including key terms and databases searched, are provided in detail. An analysis of the results shows that while evaluation frameworks have been developed to support consistent disaster evaluations, they are fragmented, uni-focused, not validated and not widely used to structure disaster
**INTRODUCTION & BACKGROUND TO THESIS**

| Chapter 1: Introduction | Chapter 2: Background and context for this research including 3ie Scoping Paper 1 |

---

**RESEARCH METHODOLOGY**

| Chapter 3: Research design, methodology and conceptual framework | Chapter 4: Literature review of evaluation frameworks and evaluation typologies in the disaster setting |

---

**ESTABLISHMENT OF THE CORE STRUCTURE**

| Chapter 5: Pre-establishment of the Core Structure – 2 studies | Chapter 6: Establishment of the Core Structure |

---

**DEVELOPMENT OF THE COMPREHENSIVE FRAMEWORK FOR DISASTER EVALUATION TYPOLOGIES**

| Chapter 7: Development of the Comprehensive Framework for Disaster Evaluation Typologies |

---

**VALIDATION OF THE COMPREHENSIVE FRAMEWORK FOR DISASTER EVALUATION TYPOLOGIES**

| Chapter 8: Validation of the Comprehensive Framework for Disaster Evaluation Typologies |

---

**DISCUSSION & CONCLUSION**

| Chapter 9: Discussion, research translation and future research | Chapter 10: Conclusion |

---

*Figure 1: Structure of Thesis*
evaluation and research. Consequently, there is no agreed on comprehensive framework to structure disaster evaluation typologies.

To establish the core structure for the comprehensive framework, the work has been separated into two chapters. Chapter 5 outlines the methodology used to identify the key elements that should be included in the core structure that forms the foundation for the new and unified comprehensive framework for disaster evaluation typologies. The importance of the core structure cannot be understated, as it is this element to which all other entities will be related and that demonstrates the disaster timeline with an added emphasis on strengthening resilience. Two studies were structured to first identify key elements and then test their inclusion in a sample of reviews based on recent Australian disaster events.

Chapter 6 describes the establishment of the core structure, which consists of three layers. These three layers include a preliminary and simplistic view of the disaster timeline, an expanded view of the disaster timeline and the concept of strengthening resilience as an overarching theme. Significantly, this core structure includes globally important disaster, humanitarian and development frameworks and agendas.

Having established the core structure, the continued development of the comprehensive framework is discussed in detail in Chapter 7. The resulting Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017) identifies the different disaster evaluation typologies and demonstrates key relationships in a single diagram. Interdependencies and relationships between the various evaluation typologies, along the disaster timeline within the disaster setting, are also identified.

To gain international acceptance of the finalised CFDET 2017 and to avoid making similar mistakes to previous frameworks, an international validation process was undertaken. Chapter 8 discusses the design of this study, which included a mixed methods approach using an online survey. Survey results are presented, analysed and discussed, demonstrating that the results received from the main questions were positive and based on the established acceptance criteria CFDET 2017 was deemed validated. While not necessary to the acceptance of CFDET 2017, some comments were considered valuable and modifications were made to strengthen the framework and aid comprehension. To distinguish the modified framework, it was relabelled the Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018).
Chapter 9 presents a generalised discussion of the research journey and summarises the research undertaken and its contribution to the area of disaster science while also providing a stimulus for a more structured and rigorous approach to developing the evidence-base of disaster preparedness and management. This chapter considers the research questions, unites the literature, theory and framework and explores research translation to policy, practice and education in the disaster evaluation setting, while considering limitations of this research and opportunities for future research.

Chapter 10 presents the conclusions of this research.
Chapter 2: Background and Context for the Research

2.1 Introduction

This chapter presents the background and context behind the need for the development and validation of a comprehensive framework for disaster evaluation typologies. Core definitions and concepts are presented to support consistent understanding and application to the work contained in this thesis. Further, the involvement in an international scoping study ‘What evidence is available and what is required, in humanitarian assistance?’ (3) (refer to Section 2.4.2) supports the need for the development of a comprehensive framework for disaster evaluation typologies.

2.2 Context

In 2013, I received an email from Active Learning Network for Accountability and Performance\(^2\) (ALNAP) outlining an international research opportunity funded by the Global Research Capacity Building Program (GDN) and the International Initiative for Impact Evaluation (3ie). GDN-3ie fund impact evaluations in the humanitarian sector by means of a Humanitarian Interventions Thematic Window (HITW). The overall objective of the HITW is to improve the quality of life in humanitarian crises caused by natural disasters and complex political emergencies (60), and improve recovery and resilience efforts. The 3ie project predominantly focused on the current state of evidence related to impact evaluations in the humanitarian and development sectors. There are many similarities to the disaster and emergency sectors that made participation in this study worthwhile and relevant to the development and validation of a comprehensive framework for disaster evaluation typologies.

A successful application was lodged by Diana Wong and Emeritus Professor Frank Archer on behalf of Monash University, in partnership with Professor Mike Clarke from

\(^2\) ALNAP is one of the lead organisations dedicated to understanding how to improve responses to humanitarian crises (see the ALNAP website). It is a global network of organisations and individuals that support this aim. Their mission statement is directly aligned with the motivation for this thesis.
Evidence Aid in the United Kingdom (UK) and Dr Anneli Eriksson from the Karolinska Institutet in Sweden.

The completed 3ie Scoping Paper 1 was part of a larger study supported by the UK Department for International Development (DFID) and United States Agency for International Development (USAID). The paper examined the range of evidence and the need for evidence in humanitarian assistance. A companion study was undertaken by 3ie, along with assistance from the Stockholm International Peace Research Institute (SIPRI) and humanitarian assistance experts on a methods paper. The teams that worked on the scoping and methods papers, the humanitarian assistance experts and the steering committee members can be found on page i of the 3ie Scoping Paper 1. The contribution made by the researchers specific to the 3ie Scoping Paper 1 is listed in the Candidate Declaration section specific to the 3ie Scoping Paper 1 (refer to pages iii–iv). This researcher’s contribution has been recognised as 30% of the total effort required to complete this part of the project. The detailed contribution is discussed in the Candidate Declaration section.

The overall project used various research methodologies including a literature review, identification of priority setting reviews, review of inventories of disasters / humanitarian crises evaluations (specifically identifying impact evaluations as defined by 3ie), case studies, key informant interviews and developing a ‘gap map’ of evidence needs and priorities. The complete 3ie Scoping Paper 1 is included in Section 2.4.2. The relevance of the 3ie Scoping Paper 1 to this thesis is discussed in Sections 2.4.1 and 2.4.3.

The 3ie Scoping Paper 1 validates the premise that there is an urgent requirement to increase the evidence-base of what works, for whom and at what cost, in the fields of disasters and humanitarian assistance (3).

Despite humanitarian assistance costing governments and non-government organisations (NGOs) billions of dollars in response to not only natural hazards but man-made emergencies, there is scant evidence on how best to intervene (16). 3ie believes that rigorous impact evaluations of humanitarian assistance could overcome some of the major methodological, practical and ethical challenges that face the global community when responding to such events (16). Understanding the causal relationships between the interventions delivered and changes in targeted results provides evidence of what works, for whom, why, at what cost, and can also assist in identifying any unintended consequences of interventions (3, 16, 61).
A review of the Global Humanitarian Assistance (GHA) reports (62-64) shows that while annual international contributions to humanitarian aid continue to increase, the shortfall remains approximately 40–50% of actual requirements (65). Sadly, the demand for resources overwhelmingly exceeds supply and lives are in danger on a daily basis due to natural hazards and man-made emergencies (such as conflict) and the lack of resources to manage these events at all levels (i.e., pre-event, event and post-event) (16). Effective and efficient assistance and service delivery are key factors to reducing the impact of these events and providing relevant information that can be used to inform decision making in the disaster setting (16, 66). The researcher’s desire is to foster the delivery of more appropriate, efficient and cost-effective interventions to help those that are suffering. The utilisation of rigorous evaluations, including impact evaluations, provides one way of gathering consistent and credible evidence on interventions delivered in the disaster setting. Subsequent analysis can be conducted to help increase the knowledge base of disaster science. Further, the evidence gained can be used to inform policy and practice based on a quality improvement theme, underpinned by the concepts of learning and accountability.

While the aims of development and foreign aid are not necessarily the same as those of the disaster and humanitarian sectors, there are similarities in the debate about the effectiveness of interventions, impact and funding. Three examples of poor practice within both the host country and at an international level include the responses to the Haiti earthquake in 2010, the Ebola outbreak in Western Africa in 2014 and the Nepal earthquake in 2015. In Haiti, a cholera outbreak killed thousands of people and infected more than 6% of the population over a two-year period. Additionally, to compound the crisis, the Red Cross was accused of building only six homes in the country with nearly USD half a billion in donated funds (67). The Ebola crisis in Western Africa was described as a ‘failure of international collective action’ by Medecins San Frontieres (MSF) (68). The international community was criticised for an absence of timely and effective humanitarian assistance when it was desperately required (68, 69).

The final example of poor practice can be seen in Nepal, where two years after the 2015 earthquake, a staggering 70% of the affected population were still living in temporary shelters (70, 71) and only 5% of the destroyed houses had been rebuilt despite nearly AUD 4 billion in aid (72). It is clear that some humanitarian assistance and development policies are not delivering on their promises (73). Dramatic calls at an international level seek to understand the nature and the scope of such inefficiencies, to develop tactical and
effective responses (74), and to reduce and manage future risk (75). Improperly prepared health responses in the disaster setting may undermine the long-term wellbeing of the affected community through inappropriate use of resources and the inconsistent application of both international and local standards of care. To increase accountability and quality improvement, many governing bodies have mandated that health responses, (76) and specifically disaster health responses, be evidence-based (15) in an effort to inform policy and practice and to improve outcomes. In Australia, the Department of Foreign Affairs and Trade (DFAT) now require a monitoring and evaluation component as a requirement for all humanitarian development and crises funding to demonstrate effectiveness (77, 78).

When considering definitions and terminology used in the disaster setting, it should be emphasised that disasters are complex and dynamic situations (79). They are multi-sectoral and multi-disciplinary events that require the involvement of people from multiple disciplines, agencies and organisations (80). Each ‘actor’ brings with them a unique set of skills and experience, a potentially different perspective in which they view the disaster and, definitely, a different language with which to communicate their assessment of the situation. Standardisation of definitions and terminology is crucial to ensure commonality of purpose and the ability to work towards a common goal, potentially helping to reduce the socioeconomic costs to the community affected by disasters (81).

2.3 Key Definitions

This thesis adopts the following key definitions, which are also consistent with those adopted by the 3ie project.

*Evaluation:* ‘a systematic investigation of the worth or merit of an object’ (82)(p.3). For the purpose of this thesis, evaluation involves ‘making judgements about the merit, worth, value, significance, credibility and utility’ (83)(p.185) of interventions that are delivered to a disaster-affected community. Significantly, evaluation provides a systematic method via which the effectiveness and efficiency of the interventions delivered can be measured, with the aim of contributing to continuous quality improvement and increased accountability (84, 85).

*Impact evaluation:* studies of measurable changes in outcomes and impact, and in some cases even outputs, that can be attributed in a causal way to the activities and interventions
or a program undertaken in the context of humanitarian assistance (86). This can include both intended and unintended consequences (16).

**Disaster:** ‘a serious disruption of the functioning of a community or a society at any scale due to hazardous events interacting with conditions of exposure, vulnerability and capacity, leading to one or more of the following: human, material, economic and environmental losses and impacts’ (87). UNISDR’s definition, although basic, has been utilised for this work as it is commonly used in the disaster sector, easily accessible, recently published, based on broad international collaboration and promotes a common understanding on the subject (87).

**Emergency:** is sometimes used interchangeably with the term disaster; for example, ‘in the context of biological and technological hazards or health emergencies, which, however, can also relate to hazardous events that do not result in the serious disruption of the functioning of a community or society’ (87).

**Natural hazard:** For the purpose of this thesis, the term ‘natural hazard’ is used instead of the term ‘natural disaster’, as there is a strong belief in the disaster sector that there is no such thing as a ‘natural disaster’ (88, 89). Bankoff describes hazards as natural events, occurring more or less frequently, and to a greater or lesser magnitude. Disasters are not ‘natural’; however, there are ‘natural hazards’ that lead to disasters. What makes a disaster depends on the way society is ordered; human systems place some people more at risk of harm than others (88, 89). The basis of this vulnerability is caused by inequality, poverty, political ideology, class and power relations that turn natural hazards into disasters (90).

**Humanitarian crisis:** ‘a situation in which there is an exceptional and generalized threat to human life, health or subsistence. These crises usually appear in the context of an existing situation of a lack of protection where the consequences of pre-existing factors, such as poverty, inequality, lack of access to basic services, are exacerbated by a natural disaster or armed conflict’ (91)(p.44).

**Humanitarian assistance:** ‘action designed to save lives, alleviate suffering and maintain and protect human dignity during and in the aftermath of emergencies’ (91)(p.44).

**Framework:** ‘an essential supporting structure underlying a system or concept’ (92).

**Comprehensive:** ‘of large content or scope’ (93).
Comprehensive framework: Comprehensive frameworks exist in other disciplines, however, no clear definition of a ‘comprehensive framework’ exists in the disaster literature. Considering the definitions of the words ‘comprehensive’ and ‘framework’ provided above, a comprehensive framework could be defined as ‘an essential supporting structure underlying a system or concept of large scope or magnitude’.

Typologies: ‘a system for putting things into groups according to how they are similar; a classification according to general types or categories’ (94). For the purpose of this thesis, typologies are limited to the area of evaluation and specifically the different types of evaluations relevant to the disaster or humanitarian or development sectors.


2.4.1 The Importance and Relevance of this Paper

The 3ie Scoping Paper 1 was an important part of a larger international project conducted by DFID and USAID. The overall aims of 3ie Scoping Paper 1 were to identify priority areas and evaluation questions for impact evaluations in the humanitarian sector and help inform the design of a thematic call for proposals to conduct impact evaluations of humanitarian assistance for crises arising from sudden onset disasters and protracted emergencies (91).

Importantly, the 3ie Scoping Paper 1 was part of research that would feed into an HITW. The overall objective was to improve the quality of life in humanitarian crises through humanitarian assistance, including improved recovery and resilience efforts. This would be achieved by first increasing the evidence-base of what works, why, and at what cost in the field of humanitarian aid with a focus on learning about innovative approaches; second, by generating lessons learned through the synthesis of high-quality evidence on implementation and effectiveness of interventions; third, by building capacity to produce and use evidence from high-quality impact evaluations and reviews; and finally, by ensuring that evidence was available to policy makers in policy-friendly formats (91).

It was the expectation of the 3ie group that the HITW would support the production of 10-15 new, high-quality, rigorous impact evaluations and three-five new systematic reviews through a range of financing and management mechanisms. The studies would answer priority evaluation questions determined during the preparatory phase. A parallel
study of methodological aspects of impact evaluations was discussed by the steering committee and others on 20-21 March 2014 in London. The end results led to a final report that provided a foundation for, and the launch of, 3ie’s HITW in 2014. Currently, 3ie is funding seven impact evaluations under this thematic window (95).

The 3ie Scoping Paper 1 ‘What evidence is available and what is required, in humanitarian assistance?’ explored:

- What is the state of evaluation of humanitarian assistance interventions?
- What is the knowledge gap in the sector?
- What questions are important for implementing agencies in the sector (3, 91)?

In relation to this thesis, the outcomes from this project are particularly significant as they support the assertion that there is a gap when it comes to high-quality evidence in this sector to inform policy and practice, a view supported by other researchers (5, 96-99). The lack of high-quality evaluations to underpin the evidence-base for health interventions remains an issue.
2.4.2 3ie Scoping Paper 1: ‘What Evidence Is Available and What Is Required, In Humanitarian Assistance?’

What evidence is available and what is required, in humanitarian assistance?

December 2014

Mike Clarke
Claire Allen
Frank Archer
Diana Wong
Anneli Eriksson
Jyotsna Puri
About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programmes. We are the global leader in funding and producing high-quality evidence of what works, how, why and at what cost. We believe that better and policy-relevant evidence will make development more effective and improve people’s lives.

3ie scoping papers

3ie thematic window grant programmes typically start with a consultative process that includes a scoping study that identifies the current state of impact evaluation evidence in a particular sector. Scoping studies lay out the landscape of what is known and help identify priority policy questions. Although thematic windows are conducted in response to donor demand, this consultation process gives developing country stakeholders a voice in setting key priorities and identifying research questions.

About this scoping paper

This scoping paper provides an independent analysis of the evidence base of evaluations in humanitarian assistance and identifies key gaps and priorities in need of rigorous evidence. It is part of background scoping research and consultation undertaken to assess the scope and methods for impact evaluation in the humanitarian sector. The working paper, What methods may be used in impact evaluations of humanitarian assistance? examines the extent to which impact evaluation methods can provide evidence to help improve the effectiveness and efficiency in humanitarian action. All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. Any comments or queries should be directed to the corresponding author Mike Clarke at m.clarke@qub.ac.uk


3ie Scoping Paper Series executive editor: Beryl Leach
Managing editor: Jyotsna Puri
Production manager: Kanika Jha and Omita Goyal
Assistant production manager: Pradeep Singh
Copy editor: Sarah Chatwin
Proofreader: Mathew PJ
Cover design: John F McGill
Printer: VIA Interactive
Cover photo: Sylvain Liechti/UN Photo

©International Initiative for Impact Evaluation (3ie), 2014
What evidence is available and what is required, in humanitarian assistance?

Mike Clarke
Evidence Aid

Claire Allen
Evidence Aid

Frank Archer
Monash University, Melbourne, Australia

Diana Wong
Monash University, Melbourne, Australia

Anneli Eriksson
Karolinska Institutet

Jyotsna Puri
International Initiative for Impact Evaluation (3ie)
Acknowledgements

This paper is part of a larger study supported by UKaid through the Department for International Development and USAID. The scoping paper examines the scope of evidence and need for evidence in humanitarian assistance. The scoping paper has been prepared by Evidence Aid with support from 3ie. 3ie, along with assistance from the SIPRI and humanitarian assistance experts, led the work on the methods paper.

Jyotsna Puri provided overall leadership and management of these papers with support from Deo-Gracias Houndolo and Peter Giesen. Bharat Dhody provided research assistance.

The team that worked on the scoping and methods paper included: Anastasia Aladyshева, Claire Allen, Frank Archer, Tilman Brück, Mike Clarke, Anneli Eriksson, Yashodhan Ghorpade, Peter Giesen, Vegard Iversen, Jyotsna Puri and Diana Wong. We are grateful to Howard White, Jeannie Annan, John Mitchell, Francesca Bonino, Joanna Macrae, Christine Kolbe, Alison Girdwood, Jonathan Patrick and Joanna Macrae.

We are also grateful to members of the steering committee which included: Caroline Andresen, Jeannie Annan, Alison Girdwood, Langdon Greenhalgh, Penny Hawkins, Christine Kolbe, Joanna Macrae, John Mitchell, John Murray, Jennie Richmond and Howard White.
Executive summary

This paper investigates the current landscape of evidence, with particular emphasis on evidence from impact evaluation, in the humanitarian sector. This is in an attempt to identify areas in which actionable evidence is available and those where more evidence is needed so as to direct research to where it will be most valuable.

The study incorporates a wide array of methods to assess available evidence including an online survey of 395 participants who are knowledgeable with regards to the humanitarian sector, semi-structured interviews with 53 experts from the humanitarian sector, extensive literature reviews of repositories of humanitarian studies and strategy documents of major humanitarian organisations, and a gap map that presents the results of a thorough search for completed, ongoing and planned impact evaluations of humanitarian interventions.

We find that, with the exception of health and nutrition, most areas in the humanitarian sector suffer from a paucity of evidence. An evidence gap map included in this study provides an illustration of the complete landscape of evidence in the sector. The upshot is that there is agreement amongst policymakers that decisions should be based on research evidence and it is recognised that impact evaluations can and should have a greater role to play in building the evidence base.

Key recommendations coming out of this study beyond informing the areas and questions for impact evaluation are that:

- Humanitarians must agree upon a way of prioritising research needs. This study puts forward a framework for choosing areas for further research.

- Efforts need to be made to index and classify existing evidence and a single unifying repository or portal should be made to improve the ease of accessibility to existing evidence.

- A single set of templates and reporting guidelines should be agreed upon to aid in the indexing and classification of evaluation studies. It would also be beneficial to agree upon standards for data collection in these studies.

One of the big challenges encountered in undertaking this project was that there is a lack of databases and repositories that index these studies and assessments. We confirmed that there is a general lack of evidence of good high quality evidence in humanitarian assistance, especially of studies that show a causal relationship between assistance and changes in targeted results. Our findings show that operational research within humanitarian assistance can be very useful and can be undertaken. Criteria that may inform prioritising further impact evaluation related research areas are assessing the feasibility of undertaking impact evaluation, seeing whether these address current or anticipated knowledge gaps, synthesising currently available results in different contexts to learn better, focusing on innovation and choosing the populations to study with care. Last but not least we also discuss the various phases of humanitarian assistance to gauge when evaluation related research is most required. Methods protocols can go a long way to plan and help programmes if planned in advance and included in a prospective manner in humanitarian programming.
Contents

Acknowledgements ........................................................................................................... i
Executive summary ......................................................................................................... ii
List of tables ................................................................................................................... v
Abbreviations and acronyms ........................................................................................ vi
1. Objectives ................................................................................................................... 1
2. Methods ..................................................................................................................... 1
3. Key findings from the scoping interviews .................................................................. 2
4. Key findings from the online survey .......................................................................... 2
   Profile of respondents ................................................................................................... 3
   Experience with impact evaluations ............................................................................ 5
   Areas requiring evidence from impact evaluations .................................................... 7
   Areas recommended for further research .................................................................. 10
5. Key findings from the semi-structured interviews ...................................................... 12
6. Key findings from review of existing priorities in strategy documents and impact evaluations ........................................................................................................... 13
7. Gap map of evidence needs and priorities ................................................................. 15
8. Gap map for needs ..................................................................................................... 15
   Gap map 1: Map of evidence needs ......................................................................... 16
   Gap map 2: Availability of evidence ......................................................................... 18
9. Gap map for existing evidence .................................................................................. 20
   Discussion of the gap map for existing evidence ...................................................... 20
10. Discussion of intervention and action domains in the needs gap map ....................... 20
    Accountability ......................................................................................................... 21
    Camp coordination and management ..................................................................... 21
    Early recovery ......................................................................................................... 21
    Education ................................................................................................................ 21
    Emergency shelter ................................................................................................. 21
    Emergency telecommunications .............................................................................. 22
    Food security ......................................................................................................... 22
    Health .................................................................................................................... 22
    Humanitarian assistance as a whole ........................................................................ 22
    Logistics .................................................................................................................. 23
    Nutrition ................................................................................................................ 23
    Organisational aspects ........................................................................................... 23
    Protection ................................................................................................................. 23
    Water, sanitation and hygiene ................................................................................ 23
11. Conclusions ............................................................................................................... 24
    Quality of the current evidence base ....................................................................... 24
12. Key areas for impact evaluations ............................................................................ 26
    Importance of operational research ....................................................................... 26
    Research prioritisation ............................................................................................ 26
    Helping the planning of impact evaluations and strengthening monitoring and evaluation systems ........................................................................................................... 27
    A framework for choosing areas for further research ......................................... 27
13. Priorities for impact evaluations .............................................................................. 31
14. Amenability of suggested areas to impact evaluation studies ................................ 31
    References ............................................................................................................. 34
Appendix 1: Methods and content of the online survey ................................................. 43
Appendix 2: People who were interviewed (* indicates those who were interviewed during the first round of semi-structured interviews) ............ 51
Appendix 3a: Methods used during the first phase of scoping interviews ........ 53
Appendix 3b: Methods used during the second phase of semi-structured interviews ........................................................................... 54
Appendix 4a: Categories used for coding of suggestions for interventions or actions for impact evaluations in the context of humanitarian crises ........ 56
Appendix 4b: Categories used for coding of suggestions for outcomes for impact evaluations in the context of humanitarian crises ......................... 57
Appendix 5: Search strategies used ........................................................................ 58
Appendix 6: Methods used for review of existing impact evaluations .............. 68
Appendix 7: Repositories ...................................................................................... 70
Appendix 8: Template used to record information about each repository ........ 72
Appendix 9: Methods used to identify reports on evidence needs or priorities for research ............................................................................. 75
Appendix 10: Summary of each report of research priorities (presented in chronological order, starting with the most recent report) .................... 77
Appendix 11: Results from a card ranking exercise on priorities and actions... 89
List of tables

Table 1: Region in which survey respondents were based and worked .........................3
Table 2: Length of time working in humanitarian assistance or related areas .................4
Table 3: Areas of expertise and years of experience ..................................................5
Table 4: Costs of impact evaluations ........................................................................6
Table 5: Areas requiring evidence to inform policy or practice in humanitarian assistance ..................................................................................................................8
Table 6: Humanitarian crises phases requiring evidence to inform policy or practice in humanitarian assistance ............................................................................................9
Table 7: Type of humanitarian crises regarded as most important for study in impact evaluations .........................................................................................................................9
Table 8: Top priorities for evaluations (percentage of responses) ................................10
Table 9: Top priorities for outcomes in impact evaluations ........................................11
### Abbreviations and acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>Action Contre la Faim</td>
</tr>
<tr>
<td>ALNAP</td>
<td>Active Learning Network for Accountability and Performance in Humanitarian Action</td>
</tr>
<tr>
<td>CAC</td>
<td>comprehensive abortion care</td>
</tr>
<tr>
<td>CHNRI</td>
<td>Child Health and Nutrition Research Initiative</td>
</tr>
<tr>
<td>DFID</td>
<td>Department for International Development</td>
</tr>
<tr>
<td>DG-ECHO</td>
<td>Directorate-General for Humanitarian Aid and Civil Protection (European Commission)</td>
</tr>
<tr>
<td>EA</td>
<td>Evidence Aid</td>
</tr>
<tr>
<td>ELRHA</td>
<td>Enhancing Learning and Research for Humanitarian Assistance</td>
</tr>
<tr>
<td>ERC</td>
<td>Evaluation Resource Center</td>
</tr>
<tr>
<td>FAO</td>
<td>Food and Agriculture Organization</td>
</tr>
<tr>
<td>GBV</td>
<td>gender-based violence</td>
</tr>
<tr>
<td>IAWG</td>
<td>Inter-agency Working Group in Humanitarian Settings</td>
</tr>
<tr>
<td>ICRC</td>
<td>International Committee of the Red Cross</td>
</tr>
<tr>
<td>IFRC</td>
<td>International Federation of Red Cross and Red Crescent Societies</td>
</tr>
<tr>
<td>IOM</td>
<td>International Organization for Migration</td>
</tr>
<tr>
<td>IPC-IG</td>
<td>International Policy Centre for Inclusive Growth</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>monitoring and evaluation</td>
</tr>
<tr>
<td>MCDEM</td>
<td>Ministry of Civil Defence and Emergency Management, New Zealand</td>
</tr>
<tr>
<td>MISP</td>
<td>minimal initial service package</td>
</tr>
<tr>
<td>MNH</td>
<td>maternal and newborn health</td>
</tr>
<tr>
<td>NGO</td>
<td>non-governmental organization</td>
</tr>
<tr>
<td>OCHA</td>
<td>United Nations Office for the Coordination of Humanitarian Affairs</td>
</tr>
<tr>
<td>PDM</td>
<td>pre-hospital and disaster medicine</td>
</tr>
<tr>
<td>UKCDS</td>
<td>UK Collaborative on Development Sciences</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>UNECLAC</td>
<td>United Nations Economic Commission for Latin America and the Caribbean</td>
</tr>
<tr>
<td>UNEG</td>
<td>United Nations Evaluation Group</td>
</tr>
<tr>
<td>UNHCR</td>
<td>United Nations High Commissioner for Refugees</td>
</tr>
<tr>
<td>UNICEF</td>
<td>United Nations Children’s Fund</td>
</tr>
<tr>
<td>WADEM</td>
<td>World Association for Disaster and Emergency Medicine</td>
</tr>
<tr>
<td>WASH/WATSAN</td>
<td>water, sanitation and hygiene</td>
</tr>
<tr>
<td>WFP</td>
<td>World Food Programme</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
1. Objectives

The specific objectives of this paper are:

- to provide an independent analysis of the quality of the evidence base of evaluations of humanitarian assistance;
- to identify the priority areas in need of evidence and research questions that are suitable for impact evaluations in the humanitarian sector;
- to recommend areas of humanitarian assistance where groups of more rigorous impact evaluations may add value;
- to provide a reference point to quickly identify what ‘good-quality’ evidence exists; and
- to provide a clear, communicable tool to support wider engagement in impact evaluations of humanitarian assistance.

2. Methods

Four studies undertaken by Evidence Aid (EA), in collaboration with Monash University, Melbourne, Australia, between November 2013 and February 2014, inform this paper:

- An online survey of stakeholders across the humanitarian sector to capture emerging research and evaluation priorities and research questions. We analysed a total of 395 responses for this report. Appendix 1 contains the methods and content of the online survey.
- We conducted semi-structured interviews with 53 key informants (Appendix 2) to identify evidence needs and priorities for research to inform policy and practice. Appendixes 3a and 3b show the methods used to conduct analyses, and Appendixes 4a and 4b present the coding sheets that were used.
- We reviewed recent humanitarian sector literature and strategy documents and identified emerging trends in evidence needs and research priorities for informing policy and practice. Appendix 5 shows the search strategies that were used.
- We identified completed, on-going and planned humanitarian assistance-related impact evaluations and related systematic reviews. Appendix 5 shows the search strategies that were used; Appendix 6 describes the methods used to identify and review the studies; Appendix 7 lists the repositories we searched; Appendix 8 shows the template used to record information on each repository.
- We convened a workshop in London on the 21 March 2014 that brought together donors, researchers and implementing agencies engaged in the humanitarian sector. We received feedback on the initial results from our scoping study at this workshop and also engaged stakeholders in various activities to help ascertain the areas in which they felt evidence was needed, the results of a card ranking exercise where participants did just this is available in Appendix 11.
• We identified 20 examples of reports that set out evidence needs or priorities for future research (see Appendix 9 for the methods used). Appendix 10 provides a summary of each of these reports.

3. Key findings from the scoping interviews

We conducted nine scoping interviews from 6 to 20 December 2013 to inform the design of the other three methodologies (some interviews were with more than one person). Appendix 2 lists the 10 interviewees.

Key outcomes were:

• Prioritisation of impact evaluations in the context of humanitarian crises is uncommon;
• Policymakers and practitioners are willing to consider evidence from impact evaluations when making decisions about humanitarian assistance;
• Respondents suggested that the scoping paper should focus on humanitarian crises during and following conflict, challenges for failed or failing states and a broad range of sudden-onset disasters and protracted emergencies;
• Respondents deemed the structure of the gap map appropriate (i.e. matching interventions against outcomes);
• Interviewees found the clusters not mutually exclusive and artificial, and recommended additional sources for organising the gap map;
• Many interventions and actions cut across sectors and require coordination across a variety of actors;
• Outcomes should include unintended consequences of the policy or intervention (e.g. misuse of the humanitarian assistance);
• Issues such as gender-based violence (GBV), shelter, and water, sanitation and hygiene (WASH), and settings such as conflict and the urban environment should be prioritised in understanding evidence;
• Forward-looking impact evaluations should be prioritised over those showing how effective the interventions and actions were in the past;
• The scoping paper’s recommendations for the areas to study for impact evaluations should be broad and should avoid the granularity that might be introduced by people with a focus on technical issues;
• The scoping paper should recommend policies and interventions that are likely to be followed up; and
• The scoping paper will need to be explicit about the definitions used for various terms, including humanitarian assistance, impact, outcome, evidence and research.

4. Key findings from the online survey

This section provides the main results of an online survey, which was open for three weeks during January 2014, and advertised through a variety of means. The online survey form was tested for one week before it was disseminated widely.
Profile of respondents

We asked respondents to provide details of their current post, the region they are based in and their sectoral expertise relevant to sudden-onset disasters and protracted emergencies, including the length of their experience. Respondents came from a wide range of organisations, including national and international non-governmental organisations (NGOs) and United Nations agencies as well as independent consultants. They were based around the world and worked in a variety of low-, middle- and high-income countries and regions. The most represented regions were Europe and North America. The most represented region in which respondents usually worked was Africa followed by Asia (see Table 1). Almost half (46 per cent) of the respondents had more than 10 years of experience in the sector (Table 2). When asked about their area of expertise, 8% of the respondents did not provide this information. Of those who did, a large proportion (78 per cent) reported more than one area of expertise. Over half (52 per cent) worked in research, usually coupled with one or more other areas. The next most common area was health (41 per cent). Analysis of the ‘other’ areas revealed that monitoring and evaluation (M&E) and communication were particularly strongly represented (Table 3).

Table 1: Region in which survey respondents were based and worked

<table>
<thead>
<tr>
<th>Region</th>
<th>Area where respondents were based (percentage of respondents)</th>
<th>Area where respondents worked (percentage of respondents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>29.6</td>
<td>13.8</td>
</tr>
<tr>
<td>North America</td>
<td>28.1</td>
<td>15.1</td>
</tr>
<tr>
<td>Asia</td>
<td>15.3</td>
<td>36.7</td>
</tr>
<tr>
<td>Africa</td>
<td>13.3</td>
<td>43.1</td>
</tr>
<tr>
<td>Middle East</td>
<td>5.4</td>
<td>16.8</td>
</tr>
<tr>
<td>Oceania</td>
<td>4.6</td>
<td>5.1</td>
</tr>
<tr>
<td>South America</td>
<td>1.5</td>
<td>3.3</td>
</tr>
<tr>
<td>Multiple regions</td>
<td>1.2</td>
<td>14.3</td>
</tr>
<tr>
<td>Central America</td>
<td>1.0</td>
<td>5.9</td>
</tr>
</tbody>
</table>

Note: Of a total of 392 respondents. (Three respondents did not provide these details.)
<table>
<thead>
<tr>
<th>Length of experience</th>
<th>Percentage of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2 years</td>
<td>8.9</td>
</tr>
<tr>
<td>2–4 years</td>
<td>14.4</td>
</tr>
<tr>
<td>5–9 years</td>
<td>26.0</td>
</tr>
<tr>
<td>10–19 years</td>
<td>29.0</td>
</tr>
<tr>
<td>&gt;20 years</td>
<td>16.8</td>
</tr>
<tr>
<td>Not applicable</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Note: Of a total of 395 respondents.
### Table 3: Areas of expertise and years of experience

<table>
<thead>
<tr>
<th>Area of expertise</th>
<th>&lt;2 years</th>
<th>2–4 years</th>
<th>5–9 years</th>
<th>10–19 years</th>
<th>&gt;20 years</th>
<th>NA</th>
<th>Percentage of total respondents *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camp coordination and management</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>17</td>
<td>12</td>
<td>1</td>
<td>13.3</td>
</tr>
<tr>
<td>Early recovery</td>
<td>3</td>
<td>4</td>
<td>21</td>
<td>22</td>
<td>19</td>
<td>0</td>
<td>19.0</td>
</tr>
<tr>
<td>Education</td>
<td>12</td>
<td>14</td>
<td>25</td>
<td>28</td>
<td>20</td>
<td>4</td>
<td>28.3</td>
</tr>
<tr>
<td>Emergency shelter</td>
<td>3</td>
<td>7</td>
<td>14</td>
<td>24</td>
<td>9</td>
<td>1</td>
<td>16.0</td>
</tr>
<tr>
<td>Emergency telecommunications</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>4.1</td>
</tr>
<tr>
<td>Food security</td>
<td>7</td>
<td>11</td>
<td>23</td>
<td>41</td>
<td>22</td>
<td>2</td>
<td>29.3</td>
</tr>
<tr>
<td>Health</td>
<td>16</td>
<td>17</td>
<td>37</td>
<td>41</td>
<td>30</td>
<td>9</td>
<td>41.3</td>
</tr>
<tr>
<td>Logistics</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>14</td>
<td>2</td>
<td>11.4</td>
</tr>
<tr>
<td>Nutrition</td>
<td>3</td>
<td>9</td>
<td>18</td>
<td>26</td>
<td>14</td>
<td>1</td>
<td>19.3</td>
</tr>
<tr>
<td>Protection</td>
<td>4</td>
<td>10</td>
<td>26</td>
<td>36</td>
<td>17</td>
<td>3</td>
<td>25.8</td>
</tr>
<tr>
<td>Research</td>
<td>20</td>
<td>29</td>
<td>42</td>
<td>53</td>
<td>33</td>
<td>12</td>
<td>52.4</td>
</tr>
<tr>
<td>Water, sanitation and hygiene</td>
<td>4</td>
<td>6</td>
<td>13</td>
<td>25</td>
<td>16</td>
<td>0</td>
<td>17.7</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>23</td>
<td>51</td>
<td>54</td>
<td>32</td>
<td>5</td>
<td>48.5</td>
</tr>
</tbody>
</table>

Note: * Of a total of 363 respondents, with 181 providing more than one area of expertise.

### Experience with impact evaluations

When asked about their experience with impact evaluation, almost two fifths of the respondents (56) said that they had commissioned or conducted an impact evaluation. However, the 125 responses to the free text question asking them to provide a brief outline of the impact evaluation (including the area of work and references to any reports) showed that some of these evaluations may not have been impact evaluations in accordance with the definition used for this study. For example, some respondents highlighted work that appeared to be related to the effects of particular disasters, rather than specific interventions, with entries such as ‘impact of floods on food security’ and ‘2010 Haiti Earthquake’ and some provided very brief information only (e.g. ‘shelter,
food and NFI, ‘child protection intervention in emergency’ and ‘disaster management’). However, where specific details of an impact evaluation were provided, these were checked for the inventory of existing impact evaluations.

This variety of interpretations of the phrase impact evaluation makes it difficult to interpret the information provided on the costs of these impact evaluations (Table 4). First, only 28 per cent of all respondents provided this information. Second, it is clear that these figures might not relate to impact evaluations of the type envisaged in this study. For instance, the answer that an impact evaluation cost more than US$1,000,000 came from a respondent who did not provide her name or contact details, and the details they provided on the impact evaluation were limited to ‘impact of religion on character and economic development; impact of food on resilience building’.

### Table 4: Costs of impact evaluations

<table>
<thead>
<tr>
<th>Cost</th>
<th>Percentage of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than US$50,000</td>
<td>58.8</td>
</tr>
<tr>
<td>US$50,000–99,999</td>
<td>19.3</td>
</tr>
<tr>
<td>US$100,000–249,999</td>
<td>10.5</td>
</tr>
<tr>
<td>US$250,000–499,999</td>
<td>7.9</td>
</tr>
<tr>
<td>US$500,000–999,999</td>
<td>2.6</td>
</tr>
<tr>
<td>More than US$1,000,000</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Note: Of a total of 115 respondents.

When asked if they had ever used the findings of an impact evaluation, almost 60 per cent of respondents (227) said that they had used the findings of an impact evaluation. However, as with the above, responses to the free text question (164 responses) seeking a brief description of the impact evaluation indicate that some of these evaluations may not have been impact evaluations in accordance with the definition used for this study.

When asked if they were aware of relevant impact evaluations but had not used their results, more than one third (141) wrote something in response. However, most of these answers provided little detail. Most merely made statements such as ‘it has been used’, ‘unaware’ and ‘not aware of any’. Where a clear statement was made (35 respondents), a few wrote that they had not used any known impact evaluations because of the poor quality of impact evaluations, but most gave a reason such as a lack of time (9) or understanding (3) of how to apply the findings (see Box 1 for some reasons cited for under-utilisation of evidence from impact evaluations). These reasons are similar to those encountered in, for example, the use of research evidence by healthcare practitioners.1

1Kahveci and Meads 2009.
Thirty-seven per cent (119) respondents answered ‘yes’ when asked if they knew of studies that had been or are being undertaken to identify priorities for impact evaluations in the context of humanitarian crises. Of these respondents, 96 provided further information in their free-text response to the question that asked for brief details of this work but most of these answers were vague and identified general areas in which such work had taken place, rather than specific projects. Twenty responses related to specific projects, and 72 responses related to general activities, or were unclear or might have been the respondent’s priorities (or areas in which they would like to see prioritisation). Specific projects included this current study, activities in mental health,\(^2\) systematic reviews,\(^3\) and the interventions noted in Section 6 of this paper by Blanchet \textit{et al.},\(^4\) funded by the Department for International Development (DFID) and the Wellcome Trust.

**Areas requiring evidence from impact evaluations**

Seventy-five per cent of the respondents answered the question requiring them to select up to three areas (out of a total of 14) for which they would most like evidence to inform policy or practice in humanitarian assistance (Table 5).

Less than half of the respondents (45 per cent) selected impact measurement and 26.8 per cent selected monitoring, which appears to reflect both a general desire for impact

\(^2\)Tol \textit{et al.} 2011.  
\(^3\) Evidence Aid 2013.  
\(^4\) Blanchet \textit{et al.} 2013.
Studies and also the need to have better evidence and reliable ways to measure impact. This is in keeping with the challenges of impact measurement and evaluation in the humanitarian sector that have been highlighted recently.\(^5\) Accountability and health were the most frequently mentioned areas after impact measurement (both 39.1 per cent).

**Table 5: Areas requiring evidence to inform policy or practice in humanitarian assistance**

<table>
<thead>
<tr>
<th>Area</th>
<th>Total number of respondents indicating the category as a key area for evidence (1)</th>
<th>Column (1) as a percentage of total non-missing respondents</th>
<th>Number (percentage among those with expertise in this area)</th>
<th>Number (percentage among those without expertise in this area)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clusters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camp coordination and management</td>
<td>24</td>
<td>8.0%</td>
<td>9 (19%)</td>
<td>15 (4%)</td>
</tr>
<tr>
<td>Early recovery</td>
<td>65</td>
<td>21.7%</td>
<td>25 (36%)</td>
<td>38 (1%)</td>
</tr>
<tr>
<td>Education</td>
<td>48</td>
<td>16.1%</td>
<td>32 (31%)</td>
<td>16 (4%)</td>
</tr>
<tr>
<td>Emergency shelter</td>
<td>41</td>
<td>13.7%</td>
<td>21 (36%)</td>
<td>20 (5%)</td>
</tr>
<tr>
<td>Emergency telecommunications</td>
<td>21</td>
<td>7.0%</td>
<td>7 (47%)</td>
<td>14 (4%)</td>
</tr>
<tr>
<td>Food security</td>
<td>69</td>
<td>23.1%</td>
<td>48 (45%)</td>
<td>19 (6%)</td>
</tr>
<tr>
<td>Health</td>
<td>117</td>
<td>39.1%</td>
<td>89 (59%)</td>
<td>27 (9%)</td>
</tr>
<tr>
<td>Logistics</td>
<td>32</td>
<td>10.7%</td>
<td>11 (27%)</td>
<td>21 (6%)</td>
</tr>
<tr>
<td>Nutrition</td>
<td>38</td>
<td>12.7%</td>
<td>18 (25%)</td>
<td>20 (5%)</td>
</tr>
<tr>
<td>Protection</td>
<td>79</td>
<td>26.4%</td>
<td>42 (45%)</td>
<td>35 (10%)</td>
</tr>
<tr>
<td>Water, sanitation and hygiene</td>
<td>61</td>
<td>20.4%</td>
<td>23 (36%)</td>
<td>37 (10%)</td>
</tr>
<tr>
<td><strong>Evaluation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accountability</td>
<td>117</td>
<td>39.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact measurement</td>
<td>135</td>
<td>45.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monitoring</td>
<td>80</td>
<td>26.8%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^5\)Giesen 2013; Knox Clarke and Darcy 2014.
<table>
<thead>
<tr>
<th>Other</th>
<th>79</th>
</tr>
</thead>
</table>

Note: Of a total of 299 respondents, with 284 marking more than one response.

One might expect a strong correlation between the area selected as in need of more evidence and the person’s area of experience but it was only in health that we found that. More than half the people with a particular area of experience rated that area as a priority; this figure rose to 59 per cent (of 152 respondents) for the health sector. However, the proportions of people who prioritised an area did vary widely between those respondents who did, and did not, have personal expertise in that area. People with stated expertise in an area were much more likely to prioritise that area (Table 5).

More than three quarters (301) of the respondents answered the question about the humanitarian crises phases for which they would most like evidence to inform policy or practice in humanitarian assistance. Each of the four areas suggested were selected by 40–60 per cent of respondents (Table 6).

**Table 6: Humanitarian crises phases requiring evidence to inform policy or practice in humanitarian assistance**

<table>
<thead>
<tr>
<th>Humanitarian crises phase</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prolonged response or engagement</td>
<td>58.8</td>
</tr>
<tr>
<td>Immediate, short-term response</td>
<td>52.8</td>
</tr>
<tr>
<td>Resilience</td>
<td>51.8</td>
</tr>
<tr>
<td>Risk reduction</td>
<td>45.8</td>
</tr>
<tr>
<td>No opinion</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Note: Of a total of 301 respondents, with 196 marking more than one response.

Seventy-two per cent of the respondents answered the question on which type of humanitarian crises they considered the most important for impact evaluations. Almost two thirds of these ranked sudden-onset disasters or protracted humanitarian emergencies equally (Table 7). Respondents were also able to suggest alternatives, but there were few such suggestions.⁶

**Table 7: Type of humanitarian crises regarded as most important for study in impact evaluations**

<table>
<thead>
<tr>
<th>Type of humanitarian crises</th>
<th>Percentage</th>
</tr>
</thead>
</table>

⁶ Two respondents listed conflict and two listed extreme weather or seasonal disasters.
Areas recommended for further research

In both the interviews and the online survey, respondents recommended research at broad levels rather than highlighting specific types of intervention, action or strategy for evaluation.

In the online survey, respondents were asked to suggest up to three interventions or actions for which they would most like to see impact evaluations. Respondents made nearly 650 suggestions. These suggestions were coded to one of 67 categories (Appendix 4a), most of which related to a type of intervention but with some that prioritised interventions more generally for a specific setting, phase or population. Some categories were more finely coded than others, reflecting greater detail provided by some respondents. For some respondents, two or more of their suggestions fell within the same category, such that there is a total of 598 suggestions when the number per category is summed. The top priorities for interventions, actions and strategies to be evaluated are listed in Table 8.

Table 8: Top priorities for evaluations (percentage of responses)

<table>
<thead>
<tr>
<th>Area or intervention</th>
<th>Percentage of responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health (in particular mental health)</td>
<td>38.0 (21.2)</td>
</tr>
<tr>
<td>Cash transfer or alternative (including vouchers or food)</td>
<td>33.6</td>
</tr>
<tr>
<td>Education</td>
<td>22.4</td>
</tr>
<tr>
<td>Shelter</td>
<td>17.6</td>
</tr>
<tr>
<td>Food aid (including food or cash)</td>
<td>18.4</td>
</tr>
<tr>
<td>Organisational aspects</td>
<td>18.4</td>
</tr>
<tr>
<td>Protection</td>
<td>16.0</td>
</tr>
<tr>
<td>Water, sanitation and hygiene</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Note: Of a total of 288 respondents.

7 The question included the following to orient respondents as to what was intended by the phrase ‘interventions or action’: ‘such as cash transfer or food aid, types of emergency shelter, education programmes for children, psychotherapy following trauma, etc.’

8 173 respondents made three suggestions each, 53 two suggestions and 24 one suggestion.
<table>
<thead>
<tr>
<th>Disaster risk reduction</th>
<th>9.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Livelihood</td>
<td>8.4</td>
</tr>
<tr>
<td>Nutrition</td>
<td>8.0</td>
</tr>
</tbody>
</table>

Note: Of 250 respondents.

Outcomes recommended for impact evaluations

In the online survey, when asked to provide up to three items in response to the question ‘Which outcomes do you regard as the most important to be measured in impact evaluations in the context of humanitarian crises?’, respondents provided just over 600 suggestions. Health was the dominant outcome, with 35 respondents suggesting mortality alone, 44 morbidity alone and 66 both. Many respondents stressed the importance of sustainability when considering outcomes.

Table 9: Top priorities for outcomes in impact evaluations

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Percentage of response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health (personal health outcomes or access to healthcare)</td>
<td>68.8</td>
</tr>
<tr>
<td>Income</td>
<td>30.8</td>
</tr>
<tr>
<td>Education</td>
<td>20.3</td>
</tr>
<tr>
<td>Resilience</td>
<td>11.4</td>
</tr>
<tr>
<td>Protection</td>
<td>10.1</td>
</tr>
<tr>
<td>Food security or access</td>
<td>8.4</td>
</tr>
<tr>
<td>Livelihoods</td>
<td>7.6</td>
</tr>
<tr>
<td>Nutrition</td>
<td>5.9</td>
</tr>
<tr>
<td>Empowerment</td>
<td>5.5</td>
</tr>
<tr>
<td>Equity</td>
<td>5.1</td>
</tr>
<tr>
<td>Recover</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Note: Of 250 respondents.

---

9 The question included the following to orient respondents as to what was meant by the term ‘outcome’: ‘[for example, educational ability, mortality, morbidity, sustainable income, etc]’.
10 161 respondents made three suggestions each, 54 two suggestions and 22 one suggestion. These suggestions were coded to one of 57 categories (Appendix 4b). Some categories were more finely coded than others, to reflect the greater detail provided by some respondents. If respondents noted the importance of sustainability for a specific outcome, this information was also captured by the coding. For some respondents, two or more of their suggestions fell within the same category, such that there is a total of 521 suggestions when the number per category is summed.
Several respondents highlighted a need for research to agree meanings for some of these outcomes and ways to measure them, in particular for issues related to equity, dignity, livelihood and resilience. This seems to reflect a broader debate about whether issues such as resilience and equity are within the remit of humanitarian action. It is important to note that humanitarian action can mean different things to different people and this needs to be recognised when launching a call for humanitarian impact evaluations. The meaning of ‘impact’ and a standard set of outcomes\(^\text{11}\) that could be applied across a series of impact evaluations were also noted.

Two thirds of all respondents answered the question about the populations that they would regard as the most important for inclusion in impact evaluations in the context of humanitarian crises. Of these, 68 per cent replied that the focus should be on specific subgroups, with 17 per cent replying that all members of the population should be included; 8 per cent made the point that the choice of population should depend on the context of the intervention. Among the subgroups suggested, there was particular emphasis on children.

5. Key findings from the semi-structured interviews

Evidence Aid conducted semi-structured interviews with 53 key informants exploring issues in more depth than in the online survey (Appendix 2). The content of these interviews was analysed by coding the responses allowing quantitative analysis.\(^\text{12}\) Of the respondents, 63 per cent were employed in research and more than 40 per cent in health.

Respondents indicated the need for research about long-term impact of humanitarian assistance, humanitarian assistance as a whole, high-level coordination, participation of local communities and disaster risk reduction. Respondents stressed the importance of context.

Respondents also suggested that impact should be evaluated based on humanitarian values, such as dignity and impartiality or from a human rights perspective, to avoid a ‘too instrumental approach’, in which the focus would be on outcomes that might be easier to measure such as morbidity and mortality.

Several interviewees stressed that there is a need to be more evidence-based in the choices they make about policies, interventions, actions and strategies. The interviews also identified a need to look at the utility of impact evaluations and what effect they

\(^{11}\) Williamson et al. 2012.

\(^{12}\) Forty interviews were conducted from 9 January to 6 February 2014, by one of three interviewers (CA, MC, AE). Most interviews involved just one interviewee, but there were two interviews in which two interviewees participated, so that a total of 42 people were interviewed in this phase. The interviews typically took 30–45 minutes, with the longest lasting 90 minutes. Their findings are combined with those from the interviewees in the first phase in the analyses presented in this paper, to provide information from 53 key informants (Appendix 2).
have on policy and practice, by making findings available to people working in the field, and not just those based in an organisation’s headquarters.

Some interviewees highlighted the importance of moving from the discussions of research prioritisation and of how to do impact evaluations, to action. They noted that this will need adequate resources, and greater priority for research in sudden-onset disasters and protracted humanitarian crises, and for research to be undertaken outside of privileged or well-established universities.

6. Key findings from review of existing priorities in strategy documents and impact evaluations

We identified 20 examples of reports that set out evidence needs or priorities for future research (Appendix 10) the search strategies and methods used to identify these documents can be found in Appendixes 5 & 9. The documents on evidence needs and priorities include highly focused work on particular aspects of humanitarian assistance, such as medical rehabilitation of spinal cord injury following earthquakes,13 as well as wide-ranging priority setting documents about how an organisation’s research programme might ensure that its own work is as effective and efficient as possible, such as the Australian Aid research strategy.14

The examples span 16 years, from a meeting organised by the World Health Organization (WHO) in 1997,15 to the comprehensive evidence review of research on health interventions in humanitarian crises by Blanchet et al., published in November 2013.16 A number of organisations including CONCERN, Norwegian Refugee Council and the United Nations Children’s Fund (UNICEF) are also developing research strategies. We found recent examples17 that stressed the importance of evidence informing policy and practice as a basis for establishing priorities for research.

The following points summarise key findings from this part of our investigation:

- Most priorities in these reports match those identified by the online survey and interviews, in particular health,18,19 food aid and nutrition,20 protection,21 water, sanitation and hygiene,22 disaster risk reduction,23 and livelihood.24

---

14 Australian Aid 2012.
16 Blanchet et al. 2013.
19 The frequency of reports focusing on priorities for health research, especially in the papers from the peer-reviewed literature, may reflect the research culture in this discipline (particularly in those areas of academia where publishing in the peer-reviewed literature is common), and the stronger demand or need for this type of evidence in health care.
• Many of the reports originate from the UK or USA, in keeping with an imbalance noted by Roy et al. in the distribution of research: 'Considering that 85 per cent of disasters and 95 per cent of disaster-related deaths occur in the developing world, the overwhelming number of casualties has contributed insignificantly to the world’s peer-reviewed literature. Less than 1 per cent of all disaster-related publications are about disasters in the developing world.'

• Some of the reports specifically identified the need for improvements in the quality of research and evaluations, in keeping with the conclusions of other recent reports.

• We found few impact evaluations of high quality, the methods used to identify these are described in Appendix 6, the repositories that were searched are listed in Appendix 7 and the template used to record information on each of these repositories has been made available in Appendix 8. The majority of randomised trials were in healthcare, with many of these relating to mental health issues, such as the prevention or treatment of post-traumatic stress disorder. Better-quality impact evaluations from the grey literature are quite recent (typically published after 2009). A number of articles retrieved by our searches confirmed the lack of impact evaluations in the humanitarian sector. For example, a 2013 report on USAID evaluations noted that only 11 (3 per cent) of the 340 evaluations conducted from 2009 to 2012 and reviewed in the study, were impact evaluations that 'included a comparison group to help determine what would have occurred in the absence of USAID’s assistance'. The authors did not regard this as a surprise because 'impact evaluations of this type are new to USAID', subsequent to 2011 policy. In a similar context, a 2012 review of evaluations undertaken in the not-for-profit sector in the USA identified that 4–6 per cent of evaluations used controls or random allocation. A 2008 report following a review of 1,000 US Department of Defense after-action reports and lessons learned found that seven (0.7 per cent) reports referred to, but did not discuss, 'impact assessment or outcome-based measures of effectiveness'.

• We identified several systematic reviews relevant to the effects of humanitarian assistance, particularly in the last few years. This growth is not surprising given the considerable rise in the number of systematic reviews in recent years. We did not have resources to assess each of the included studies in these systematic reviews to determine whether or not they were impact evaluations. For example, the Timbie et al. review of strategies to manage and allocate scarce resources

23UNISDR 2005; Treadgold 2006; Roy et al. 2011; Evidence Aid 2013.
24 Roy et al. 2011; Australian Aid 2012.
26 Giesen 2013; Knox Clarke and Darcy 2014.
27 Hagebecker et al. 2013.
28 Morahu and Pankaj 2012.
29 Reaves et al. 2008.
30 Moher et al. 2007; Bastian et al. 2010.
during mass casualty events included 74 studies,\textsuperscript{32} while Blanchet et al. identified more than 700 studies related to the effects of health interventions in humanitarian crises.\textsuperscript{33}

7. Gap map of evidence needs and priorities

We produced an evidence gap map using broad domains, to reflect the generally high level of suggestions for impact evaluations.\textsuperscript{34} Based on priorities from the online survey, semi-structured interviews and existing prioritisation exercises, the rows of the gap map are interventions and actions in each area of humanitarian assistance presented in the online survey, along with the need for an assessment of humanitarian assistance as a whole, and of organisational aspects (including the coordination of humanitarian assistance). These broad domains are the rows in the gap map. We used red for things that more than 20 per cent of respondents in the online survey suggested were a priority, yellow for things that 10–20 per cent of respondents in the online survey suggested or that were common in the interviews, and blue for things that 5–10 per cent of respondents in the online survey suggested.\textsuperscript{35}

The column headings or outcomes relate to the policy, intervention, action or strategy in need of evaluations. Some of these (education, health and income) emerged as outcomes that were felt to be priorities for assessment across the range of interventions and actions in humanitarian assistance.

8. Gap map for needs

This shows the priority given to the need for evidence to inform policy and practice. Education, health and income are the highest priority. For clarity, the individual cells have been coloured to suggest the strongest associations between the intervention or action and the outcome. We provide information about each of the rows after the gap map showing the need for evidence.

\textsuperscript{32}Timbie et al. 2013.
\textsuperscript{33}Blanchet et al. 2013.
\textsuperscript{34}However, finer detail can be provided within these to include, for example, specific types of intervention, and specific types of outcome measure. These might include access to education, school attendance and educational performance for the education domain and availability, accessibility and use for the food security domain.
\textsuperscript{35}Impact measurement and ‘monitoring’ have not been included because their frequency in the answers about the need for evidence seems likely to be a reflection of the need for impact evaluations in general and for research into the methods of impact evaluations.
## Gap map 1: Map of evidence needs

<table>
<thead>
<tr>
<th>Outcomes -&gt; Interventions</th>
<th>Education</th>
<th>Empowerment</th>
<th>Equity</th>
<th>Food security</th>
<th>Health</th>
<th>Income</th>
<th>Livelihood</th>
<th>Nutrition</th>
<th>Protection</th>
<th>Recovery</th>
<th>Resilience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accountability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camp coordination and management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early recovery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency shelter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency telecommunications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food security</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental health</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humanitarian assistance as a whole</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nutrition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisational aspects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gap map 2: Availability of evidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes -&gt; Programmes</strong></td>
<td>Education</td>
<td>Empowerment</td>
<td>Equity</td>
<td>Food security</td>
<td>Health</td>
<td>Income</td>
<td>Livelihood</td>
<td>Nutrition</td>
<td>Protection</td>
<td>Recovery</td>
<td>Resilience</td>
</tr>
<tr>
<td>Accountability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camp coordination and management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early recovery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency shelter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency telecommunications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food security</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental health</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humanitarian assistance as a whole</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
9. Gap map for existing evidence

We have presented the second gap map to illustrate the amount of evidence from impact evaluations in areas identified as in need of evidence to inform policy and practice. Empty cells should be interpreted as indicating uncertainty about whether there is any evidence from impact evaluations in these areas, rather than a definite lack thereof.

Discussion of the gap map for existing evidence

Several things are important to note here:

- **Searches**: Complete identification of all relevant impact evaluations for each cell requires more comprehensive or targeted searches of the published and grey literature, as well as searches for unpublished examples, than was possible with the resources available for this study.

- **Unintended consequences and duration of follow-up**: Respondents highlighted a need to assess both intended and unintended consequences of humanitarian assistance, to measure the medium- and long-term impact (in particular to explore the sustainability of any benefits), and to examine the utility of the evaluation itself.

- **Priorities**: A majority of respondents favoured including both sudden-onset disasters and protracted emergencies. Furthermore 40-60 per cent of respondents rated as ‘important’ each of the four phases (resilience, risk reduction, immediate short-term response, and prolonged response or engagement).

- **Subgroup analysis**: Respondents favoured the study of specific population groups in future impact evaluations, in particular children and vulnerable or underrepresented people. Several respondents stressed the importance of consulting with intended beneficiaries in impact evaluation.

- **Insufficient evidence**: There was general consensus that insufficient good-quality evidence is available, confirming the need for new, high-quality research in any priority area. This suggests that there is little risk that future research will duplicate efforts or that it could not be justified on scientific, ethical or environmental grounds. New, robust and well-designed impact evaluations could make an important contribution for most, if not all, of the areas covered by the gap map and would ideally be synthesised through systematic reviews of all available evidence.

10. Discussion of intervention and action domains (rows) in the needs gap map

In this section, we discuss some important findings about the needs for new research.
Accountability

Of the online survey respondents, 39.1 per cent identified accountability to beneficiaries, authorities and donors as a high priority (making it the most common priority, along with impact measurement and health). Accountability also features in several of the documents presenting evidence needs and research priorities. It has been assessed in several recent impact evaluations. For example, Fearon et al. report a randomised trial in which a DFID-funded community-driven reconstruction programme implemented by the International Rescue Committee in post-conflict Liberia was evaluated in 42 randomly selected communities compared with 41 control communities during 2006 to 2007. 36

Camp coordination and management

Eight per cent of the respondents to the online survey identified this as a priority. Examples of impact evaluations were identified, for example in relation to specific aspects of camp management, such as the use of insecticide-treated plastic sheeting for malaria prevention. 37

Early recovery

Of the respondents to the online survey, 21.7 per cent identified interventions after the initial life-saving phase (which is usually after about six weeks of a sudden-onset disaster) as a high priority. We identified several impact evaluations that examine the effects of interventions intended to improve early recovery. 38

Education

Of the online survey respondents, 16.1 per cent identified education of children and young people as a high priority. Some respondents noted the importance of studies to assess the effects of better education for responders. We found several impact evaluations of education projects. 39

Emergency shelter

Of the online survey respondents, 13.7 per cent identified impact of emergency shelter provision as a high priority (e.g. how to ensure that emergency shelter does not become a permanent solution). We identified a 2009 systematic review of health effects of relocation following disaster. 40

37 Mittal et al. 2011; Burns et al. 2012.
40 Uscher-Pines 2009.
Emergency telecommunications

Seven per cent of the respondents to the online survey identified telecommunication between those providing humanitarian assistance and the affected population as a high priority.

Food security

Of the respondents to the online survey, 23.1 per cent identified issues of food availability, access and utilisation following a sudden-onset disaster or during a protracted emergency as a high priority, and features in several of the documents presenting evidence needs and research priorities. The United Nations High Commissioner for Refugees (UNHCR) and the World Food Programme (WFP) commissioned an impact evaluation of food assistance in protracted refugee situations (Bangladesh, Chad, Ethiopia and Rwanda) in 2011–2012.\textsuperscript{43} We also found a systematic review of the association between nutrition and food security in disasters.\textsuperscript{42}

Health

Of the online survey respondents, 39.1 per cent identified health, including illness, disease and other health-related conditions, as a high priority (making it the most common priority along with accountability and impact measurement).

Respondents drew particular attention to mental health, mainly in relation to medium- and long-term outcomes in the prevention or treatment of post-traumatic stress disorder. We identified a relatively large number of randomised trials and controlled prospective studies, as well as other impact evaluations and systematic reviews.

The recent review by Blanchet et al.\textsuperscript{43} provides a comprehensive account of research into the effects of health interventions.\textsuperscript{43} It provides details on outcomes related to accountability, coordination, nutrition, the security of healthcare workers, urbanisation, and water, sanitation and hygiene. This work also demonstrates the crucial importance of health as an outcome of nutrition and water and sanitation. For instance, the 2005 systematic review by Fewtrell et al.\textsuperscript{44} examined the effects of water, sanitation and hygiene interventions for reducing diarrhoea in low-resource settings.\textsuperscript{44}

Humanitarian assistance as a whole

Many of the interviewees identified as a high priority the lack of robust evidence on the impact of humanitarian assistance as a whole. Evaluations of humanitarian assistance as a whole are unlikely given that it is difficult to imagine a control situation of no such response, but some elements might be evaluated. For example, the synthesis by Harvey and Bailey brings together some information on evaluation of cash transfer programmes.\textsuperscript{45}

\textsuperscript{41}Cantelli et al. 2012.
\textsuperscript{42}Rivero and Cantero 2013.
\textsuperscript{43}Blanchet et al. 2013.
\textsuperscript{44}Fewtrell et al. 2005.
\textsuperscript{45}Harvey and Bailey 2011.
Logistics

Of the respondents to the online survey, 10.7 per cent identified logistics services for the humanitarian community as a whole as a high priority.

Nutrition

Of the online survey respondents, 12.7 per cent identified interventions to improve the nutritional status of the affected populations as a high priority. An example of a recent impact evaluation is the cluster randomised trial by Action Contre la Faim-France in Chad in 2010, assessing the impact of ready-to-use supplementary food on the prevention of wasting in 1,038 children aged 3 to 36 months.\textsuperscript{46} We also identified a systematic review of the association between nutrition and food security in disasters.\textsuperscript{47}

Organisational aspects

Many of the interviewees identified as a high priority the organisation of humanitarian assistance, such as coordination across the various agencies and actors providing assistance, and between them and local communities (including with local leaders and workers). We identified several impact evaluations, including a recently published protocol for a systematic review on the effects of multidisciplinary team response to support survivors of mass casualty disasters\textsuperscript{48} and a 2009 systematic review considering disaster management in general.\textsuperscript{49}

Protection

Several of the interviewees, as well as 26.4 per cent of the respondents to the online survey, identified protection to internally displaced persons and other populations as a high priority, with particular concerns around the need to protect women against gender-based violence. We identified several impact evaluations and a 2013 systematic review of studies to reduce the risk and incidence of sexual violence in armed conflict and other humanitarian crises.\textsuperscript{50}

Water, sanitation and hygiene

Many of the interviewees, and 26.4 per cent of the respondents to the online survey, identified interventions to provide clean water, dispose of human and other waste, and safeguard health through improved sanitation and hygiene as a high priority. We identified several impact evaluations as well as the afore-mentioned systematic review by Fewtrell et al., which brought together information from 46 studies of water, sanitation and hygiene interventions intended to reduce diarrhoea in low-resource settings.\textsuperscript{51}

\textsuperscript{46}Huybrechts et al. 2012.
\textsuperscript{47}Rivero and Cantero 2013.
\textsuperscript{48}Adam et al. 2013.
\textsuperscript{49}Lettieri et al. 2009.
\textsuperscript{50}Spangaro et al. 2013.
\textsuperscript{51}Fewtrell et al. 2005.
11. Conclusions

The findings from the various studies conducted by Evidence Aid and reported in this paper highlight:

- wide-ranging acceptance of the role of research evidence in decision making;
- general agreement around the areas in most need of this evidence; and
- recognition of how existing impact evaluations have not met these needs but could do so in the future.  

We recognise that richer analyses of the results of these studies could be conducted, time permitting. However, drawing on the findings presented above, we now discuss our conclusions about the quality of the current evidence base in humanitarian assistance, key areas for impact evaluation research and priorities for impact evaluations and, finally, the amenability of certain areas for impact evaluations.

Quality of the current evidence base

Lack of evidence base: Our studies confirm the general lack of a reliable and robust evidence base from studies assessing the causal relationship between a policy or intervention and outcomes or impact. These findings are further confirmed by other recent studies and systematic reviews. Only a small proportion of the many evaluations of humanitarian assistance use designs with a counterfactual, control or comparator group that allows studies to attribute a measurable change in outcomes or impact indicators to programmes or policies. However, we also found several examples of randomised trials, showing that it is possible to generate evidence for specific questions through studies of this type. However, this evidence base is limited and concentrated in certain areas, such as mental health.  

Quality and relevance of research: Many respondents in the online survey and the semi-structured interviews highlighted the need to improve the quality and relevance of research in humanitarian assistance. ‘Impact measurement’ was the most frequent selection as a priority area for evidence in the online survey. This reflects a desire for better evidence and knowledge about how to measure impact reliably, but it is worth noting that respondents might have a variety of definitions in mind for ‘impact measurement’.

---

There are numerous recent guides to the conduct of impact evaluations and critiques of their methods, prepared by and for individual agencies and NGOs\textsuperscript{55} and in the published literature.\textsuperscript{56}

\textit{Repositories of data:} Other key challenges to the quality of the evidence base relate to the difficulties of identifying studies (given the potential problems of publication or selective reporting biases) and the underlying routine data that might be used for studies in the future. The fact that different methodologies constitute impact evaluations makes it difficult to find them. There is no agreed indexing system for these studies (unlike, for example, the taxonomy used to index different types of healthcare study in bibliographic databases such as MEDLINE), and there is no single repository dedicated to them. People wishing to find studies to inform future policy or practice, or to justify and design future impact evaluations, need to search multiple places and may need to check through hundreds or thousands of irrelevant records if they wish to be comprehensive. \textbf{This problem could be improved by the indexing and classification of impact evaluations and the development of a repository that either collects all such studies in a single place or provides a portal through which they can be identified in other repositories.} Tagging papers with keywords like ‘counterfactual analysis’ or ‘randomised trial’ that help identify the methodological peculiarities of a study would considerably simplify the task of searching for studies particularly because broader terms like ‘impact evaluation’ seem to constitute a wide range of methodologies. A facility to register impact evaluations before they begin would also be beneficial for the identification of planned or ongoing unpublished or inaccessible work. These benefits are widely accepted for prospective registries of controlled trials in healthcare.\textsuperscript{57}

\textit{Quality standards:} Indexing and classification of impact evaluations would be made easier by clearer reporting through the use of \textbf{templates and reporting guidelines}, such as those brought together by the EQUATOR Network for a variety of types of research.\textsuperscript{58} Other challenges in the quality of the evidence base relate to the data that are available for studies of outcomes and impact. This includes the lack of an agreed common set of data to be collected, variations in how data are collected, and concerns about the quality of that data.\textsuperscript{59} The development and use of standardised data formats\textsuperscript{60} and an agreed set of core outcomes\textsuperscript{61} might mitigate some of these challenges. Care will have to be exercised, however. A recent systematic review of templates for reporting pre-hospital major incident medical management found 12 templates for the same purpose, and concluded ‘our findings show that more than one template exists for generating reports from the medical management of major incidents. Limitations are present in the existing templates regarding internal and external validity, and none of them have been tested for feasibility in real-life incidents.\textsuperscript{62} These templates could be used to record relevant data in the immediate aftermath of a disaster and to report these data in a standard format. This would make it easier to compare, contrast and

\textsuperscript{55}For example, ACF 2011; Giesen 2013.
\textsuperscript{56}For example, de Pee et al. 2011; Pfefferbaum et al. 2013, 2014.
\textsuperscript{57}Ghersi and Pang 2009.
\textsuperscript{58}Simera and Altman 2013.
\textsuperscript{59}Kar-Purkayastha et al. 2011.
\textsuperscript{60}Debacker et al. 2012.
\textsuperscript{61}Williamson et al. 2012.
\textsuperscript{62}Fattah et al. 2013.
combine data from different settings and to use routine data in impact evaluations. Related to this, a June 2012 editorial by the Editor-in-Chief of Prehospital and Disaster Medicine, Samuel Stratton, advised that authors of future articles reporting on disaster and acute medical response research will need to use one such template.63

12. Key areas for impact evaluations

Importance of operational research

An evaluation of the European Commission integrated approach of food security and nutrition in the humanitarian context concludes that ‘partners do not always understand why and when DG ECHO [Directorate-General for Humanitarian Aid and Civil Protection (European Commission)] supports operational research, indicating the need for greater collaboration with partners on research priorities’. The report made what it called a ‘critical’ recommendation that action was required at Headquarters to ‘identify a forum to coordinate more with other donors, particularly USAID/Food for Peace, on policy operational approaches and research into the role of specific nutritional products’.64

Research prioritisation

There are a number of recent papers about the setting of research priorities for humanitarian assistance and there is a reasonable body of literature on the setting of priorities in healthcare research.65 Among the articles specific to humanitarian assistance is a recent book chapter by Murray and Kessel, which highlights the need for agreement on the prioritisation process given that ‘undertaking health and social research to help facilitate disaster risk reduction and disaster risk management is vitally important to increase preparedness to respond to disasters, to enable the most effective action to be taken once disasters have occurred and to understand better the consequences of disasters’.66

UNICEF also stressed the need for formal methods of prioritisation in 2011:

The efficiency of knowledge generation and dissemination at both the global and country levels is diminished by a lack of coordinated, systematic planning and rigorous evaluations. Insufficient coordination among HQ [UNICEF headquarters], ROs [regional offices] and COs [country offices] in establishing research priorities and planning evaluations detracts from development of a focused research agenda in ECD [early childhood development] and results in missed opportunities to leverage resources for more rigorous, longer-term country-specific and multi-country evaluations.

Current processes at the country and global levels do not facilitate sequencing of evaluations into formative and summative stages.67

63 Stratton 2012.
64 Haveret et al. 2013.
65 Oliver and Gray 2006.
66 Murray and Kessel 2014.
67 UNICEF 2011.
Helping the planning of impact evaluations and strengthening monitoring and evaluation systems

In another example, Action Contre la Faim (ACF) have the following policy on the evaluation of their own work:

All ACF interventions should be evaluated: from single projects and multi-project programmes to country-level and regional strategies. Although this commitment to evaluations does not change, the type of evaluation used does vary according to the size and length of the intervention. Smaller interventions (<€400,000) should be evaluated once using internal self-evaluation tools, whilst larger interventions (>€1,000,000) should be evaluated twice using external evaluators.68

A framework for choosing areas for further research

The above examples help provide a set of criteria that might inform the process of prioritising research areas. We suggest that the following issues be considered:

- **Feasibility of undertaking impact evaluations**: Difficulties in undertaking impact evaluations may be methodological (it may be difficult to find comparison groups), operational (e.g. in the ability to define and deliver policies, interventions, actions or strategies that are being evaluated) or institutional (e.g. unwillingness to evaluate). All these factors should be considered when developing priorities for impact evaluations.

- **What to evaluate?** Is it better to choose areas that are likely to be particularly easy or difficult to evaluate for attributable impact? For example, an impact evaluation which examines the effects of a particular medical procedure, nutritional product or hand-washing strategy on a specific outcome might be relatively easy and could be conducted as a randomised trial in which people are randomly allocated to the intervention or an alternative. In contrast, a much more difficult impact evaluation might examine the effects of a complex intervention to improve the protection of women and children in a camp and require the assessment of a range of outcomes, some of which are difficult to measure including gender-based violence, dignity and livelihoods. Both are feasible impact evaluations. But one is easier. Another consideration might also be given to whether the impact evaluations should be undertaken in areas where there is a stronger tradition of prospective, comparative research (such as, but not limited to, those using mixed methods with valid counterfactuals). This would mean, in particular, those areas most closely associated with healthcare, where there is now an established tradition of randomised trials, with several hundred thousand controlled trials already conducted and tens of thousands ongoing.69 On the other hand, it may also be important to undertake and demonstrate impact evaluations in those areas where they are less common.

---

68ACF 2011.
69Gherzi and Pang 2009; Bastian et al. 2010.
• **Use of existing evidence when prioritising individual impact evaluations:** Should the focus be on areas in which there is the least research already or areas where a relatively large amount of research has been conducted but this is not sufficiently reliable or robust? An advantage of conducting new research where a body of studies already exists (such as in relation to the prevention of mental health problems) is that the successes and failures of that past research might help to ensure the optimum design of a new study or might allow meta-analyses in which the findings of the new study are combined with the existing studies in a systematic review. On the other hand, an advantage of embarking on a study in an area that has been under-researched in the past is that this might provide the first, and only, evidence in an area of high priority and might stimulate further such research.

• **Creating review standards:** How should existing research be reviewed to confirm the scientific, ethical and environmental justification for new research studies, and to place their findings in context? In healthcare research, it is now widely accepted that new studies should not be done without a systematic review of the existing evidence, to confirm that there is sufficient uncertainty to warrant the new study. Some funders of healthcare research now require details of such a systematic review in the application for funding. Furthermore, following the conduct of a new study, some journals now require that the findings be placed in the context of other relevant studies, to provide the reader with an up-to-date summary of the evidence base.

• **Choosing the interventions to evaluate – innovation:** Should the focus be on assessing innovative interventions, rather than those that are already in wide use? By providing evidence that is deemed to be reliable and robust and to have minimised bias, high-quality impact evaluations might be particularly important for the uptake of innovative interventions. For example, the 2013 report by the Feinstein International Center, *The Use of Evidence in Humanitarian Decision Making*, noted: ‘with most innovative ideas, the humanitarian community requires evidence to show that a new intervention is more appropriate and effective than traditional approaches (even though the same evidence isn’t necessarily required to prove that the traditional approaches are appropriate).’ It highlighted an analysis of the funding requirements for proposed food security interventions of five large donor agencies, which found that organisations are required to demonstrate a larger body of evidence to support innovative food assistance approaches compared with traditional programmes. This suggests that high-quality evidence on the benefits of innovative interventions may make risk-averse decision makers more comfortable with implementing the innovation. Of course, it is also possible that evidence from impact evaluations suggesting that widely used interventions are ineffective will have less impact, because of a reluctance to change.

---

70 Clarke and Hopewell 2013.
71 Clark and Horton 2010.
72 Darcy et al. 2013.
- **Choosing the interventions to evaluate – relationship with the development sector:** Should the focus be on interventions where there is considerable overlap with the development sector? As the humanitarian and development sectors become more intertwined this is a much debated issue, as there will be increased opportunities to conduct impact evaluations of policies and interventions that overlap these sectors. Given the wider experience with impact evaluations in the development sector, this might help to demonstrate the feasibility of impact evaluations to people in the humanitarian sector who are sceptical of this.

- **Choosing the interventions to evaluate – uncertainty, controversy and debate:** Should the focus be on comparing policies or interventions where there is already considerable uncertainty, controversy or debate about their relative effects? Where there are policies or interventions currently with a wide variation in practice, disagreement or active debate about the most appropriate choice, high-quality impact evaluations may have more opportunity to lead to change than in areas where there is general acceptance of effectiveness and reluctance to change.

- **Choosing the populations to study:** Should the evaluations focus on particular subgroups of people (e.g. the vulnerable or disadvantaged), or the population as a whole (with analyses for specific subgroups)? The respondents to the online survey strongly supported a focus on specific subgroups, with 68 per cent of those who commented on this issue favouring this approach. In contrast, 17 per cent replied that all members of the population should be included and 7 per cent that there should be a mixture of the whole population but with special attention to subgroups. Among the subgroups suggested, there was particular emphasis on children as a group that are often neglected in evaluations.

- **Settings for the impact evaluations:** Should the focus be on sudden-onset disasters (possibly with the need to put some impact evaluations 'on the shelf' for future events, as with the UK National Institute for Health Research programme of studies for pandemic influenza or for ongoing protracted emergencies? In both the online survey and the semi-structured interviews, there was widespread support for impact evaluations in both sudden-onset disasters and protracted humanitarian emergencies. Specifically, 63.9 per cent of the respondents to this question in the online survey rated both settings as equally important, while 18.1 per cent singled out protracted humanitarian emergencies and 13.2 per cent singled out sudden-onset disasters.

- **Phases for the impact evaluations:** Should the focus be on impact evaluations in some or all of resilience, risk reduction, immediate short-term response, and prolonged response or engagement? In the online survey, each of these was

---

77 Whittall et al. 2014.
78 White 2013.
79 O’Neill et al. 2014.
70 Yong 2012.
ranked as a priority by 40–60 per cent of the respondents: resilience (51.8 per cent), risk reduction (45.8 per cent), immediate short-term response (52.8 per cent), and prolonged response or engagement (58.8 per cent).

- **Choosing the outcomes to measure**: Should a core outcome set be developed, which would allow the effects of different interventions, actions and strategies to be compared and contrasted across the sector? Information on 200 core outcome sets in healthcare have already been collated by the COMET Initiative’s database. These provide an agreed, standardised set of outcomes for measurement in all trials in a particular area of health and, in some cases, for other assessments as well, such as clinical audit. These are not intended as a closed list of outcomes, and researchers are encouraged to measure additional outcomes outside the core set, as appropriate to their study. However, if core outcome sets are used consistently, they would make it easier to compare, contrast and combine the results of separate studies on the same topic; and would also reduce the possibility that outcomes that are important to decision makers are overlooked or not reported by researchers.

- **Methodology research**: Bearing in mind the strength of the feedback on the need for evidence in the area of impact measurement, should research into the methods to be used for impact evaluations in humanitarian assistance be encouraged? Although it might be beyond the scope of the funding available for the Humanitarian Interventions Thematic Window to support such research directly, this might be something to discuss with agencies that fund such work. It could be achieved through a mixture of stand-alone research into the methodology and methodology research embedded within the impact evaluations. The potential benefits of the latter may be something to consider when evaluating future proposals for funding of impact evaluations. Furthermore, initiatives are now underway to develop simple protocols for methodology research, which could be included in prospective impact evaluations.

- **Impact evaluation of the impact evaluations**: Should each impact evaluation include a self-evaluation to be conducted either by the team commissioned to do the impact evaluation or by an independent consultant, in order to determine the impact of the impact evaluation on future policy, practice and outcomes? This was raised in a few of the semi-structured interviews which highlighted the need for evidence to show that the findings of impact evaluations do influence practice and policy in future sudden-onset disasters and in ongoing and future protracted emergencies. This needs to be referenced to all work already done on utilisation.

- **Dissemination and implementation of findings**: Should each impact evaluation include an implementation or knowledge translation plan? Alongside the desire for evaluations of the effect of impact evaluations, there was a recognition that special efforts might be needed to ensure that the findings of impact evaluations

77 Williamson et al. 2012.
79 Smith et al. 2013.
are accessible to those who might use them when making decisions and choices about practice and policy. It was noted that this should include both those based in the headquarters of agencies and those working in the field. It would also be worth considering how the findings might be made available to those who took part in the impact evaluation.

13. Priorities for impact evaluations

A list of the research questions identified for specific aspects of policy or for particular interventions through the online survey, semi-structured interviews and review of existing documents would include several hundred suggestions. Rather, in keeping with the objective to recommend areas of humanitarian assistance where groups of more rigorous impact evaluations would add value, we suggest the following broad areas for inclusion when planning and prioritising research. These were selected on the basis of the high frequency with which they were suggested as priority topics and the lack of reliable and robust evidence in existing impact evaluations:

- interventions to improve the livelihood of the affected population, in particular to explore the relative effects of providing goods (such as food), cash or vouchers;
- interventions to improve health (but with consideration to be given as to whether or not to exclude some areas of mental health because of the existence of several randomised trials and systematic reviews in this area already);
- interventions to protect vulnerable populations, in particular to prevent gender-based violence;
- interventions to improve food security and nutrition;
- interventions relating to water, sanitation and hygiene;
- specific elements within a whole programme of humanitarian assistance, which might allow a composite estimate of the overall effect of the programme to be inferred; and
- coordination of the humanitarian assistance, in particular across the different agencies and actors, and with members of the local communities (including local leaders and workers)—this might include an assessment of the effects of a programme of humanitarian assistance as a whole, and interventions to improve accountability.

For each of these areas, outcomes that are directly related to the policy, intervention, action or strategy should be measured, along with those that might identify unintended consequences. These outcomes should be measured in the short, medium and long term to explore the sustainability of any benefits. It is likely that different types of impact evaluation would be most appropriate and feasible for these topics, and that the specific policies, interventions, actions or strategies to evaluate would vary depending on the setting for the sudden-onset disaster or protracted humanitarian emergency.

14. Amenability of suggested areas to impact evaluation studies

In considering the amenability of these priority areas to impact evaluations, it is worth noting a 2013 report by the United Nations Evaluation Group of the importance of using
an **evaluability assessment** as part of the prioritisation process. They wrote that this would:

> ... include the mapping, systematization and analysis of any baseline and/or monitoring data that were produced by the managers of the intervention/body of work to be evaluated; these data will be important to inform the development of the impact evaluation tools. The main output of the evaluability assessment should be a full approach paper, including an evaluation matrix, that sets out in a detailed and explicit manner the analytical and methodological approach of the evaluation. ...For the impact evaluation of very large or complex interventions, the evaluability assessment may be a study in itself. ...By identifying what is possible to evaluate at a given point in time, highlighting those evaluation questions that are most critical, and specifying assumptions in the programme logic most in need of empirical verification, an evaluability assessment can identify priorities for impact evaluation.\(^8^0\)

The amenability of the areas in Section 12 of this paper to impact evaluations will depend in large part on the **complexity** of the question to be addressed. For instance, where it is relatively easy to isolate the intervention to be investigated, a randomised trial in which participants are randomly allocated to the intervention, or not, may be feasible. Such trials could take place with randomisation at the level of individuals, or at the level of a group of individuals, such as a family, school, village or region in a “cluster randomised trial”.\(^8^1\)

Where the impact evaluation would seek to determine the effects of policies or more complex interventions, these are still amenable to randomised trials, which could be designed using frameworks developed for the assessment of such interventions.\(^8^2\) However, impact evaluations that would seek to investigate the effects of system-wide policies or interventions might not be amenable to randomised trials, because of the difficulty of creating a sufficient number of intervention and control participants to benefit from the balancing out of confounding variables through the chance process inherent in random allocation. In such circumstances, **alternative designs** might be used. These include the assessment of individual elements of the policy or complex interventions in different geographic areas within a single disaster or protracted emergency, at different times within a single event, or in multiple events which happen frequently. Different types of disaster or protracted emergency would be more amenable to these different methods. For example, rare events that occur over a large geographic area (such as major earthquakes or tsunamis and extreme windstorms) might be suitable for the first option; protracted emergencies (such as famine and displaced populations) might be suitable for evaluating different interventions at different times; and small but repeated events (such as localised extreme weather, flooding or landslides) might be suitable for evaluating different interventions in different events.

In relation to **sudden-onset disasters**, a particular challenge will be the ability to have a prospective impact evaluation, such as a randomised trial or other comparative study

---

\(^{8^0}\)United Nations Evaluation Group 2013.  
\(^{8^1}\)Clarke 2009.  
\(^{8^2}\)Craig et al. 2008; Corry et al. 2013.
pre-prepared and ready to be activated. Without such ‘on the shelf’ studies, the question might not be amenable to an impact evaluation, for example if it takes days or weeks to design and activate the study protocol. Therefore, to overcome this challenge, the impact evaluation should have been pre-designed and be ready to initiate at the appropriate time in the disaster. This is possible and it means that questions about the effects of early-phase interventions are amenable to high-quality impact evaluations, if the necessary preparation and investment has been made. Given that the uncertainty around the occurrence of sudden-onset disasters relates more to when and where they will occur, rather than whether or not they will occur somewhere at some time, this investment is unlikely to be wasted.83

Another issue for consideration is the importance of ethical approval. Few of the studies we identified discussed ethical approval or oversight. This issue was also raised by some respondents in the online survey and interviews, and noted in some documents.84 Research ethics was the subject of a 2011 Brocher Foundation workshop, co-organised by Dublin City University and Evidence Aid.85 It included a contribution from Doris Schopper at the Center for Education and Research in Humanitarian Action in Geneva, drawing on the experience of Médecins Sans Frontières,86 which outlined the need for international guidance that would facilitate research in these emergency situations, bearing in mind the potential vulnerability of the population to be studied in the impact evaluation.87

A final concern that we would like to raise is that there seems to be little consensus in the humanitarian sector on what ‘impact evaluation’ is. In our investigation of repositories we found the term ‘impact evaluation’ used to describe a plethora of disparate studies. As such we would like to re-emphasize the need to clarify terminology associated with ‘impact evaluation’ to better organise the literature in this sector and the need for evaluators and researchers in the humanitarian sector to be exposed to training on impact evaluations.

83Yong 2012.
84For example, WHO 1997; Shaikh and Musani 2006; Roy et al. 2011.
85O’Mathúna et al. 2014.
87Schopper 2014.
References


Bastian, H, Glasziou, P and Chalmers, I, 2010. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Medicine, 7(9), e1000326.


Institute of Medicine and New York Academy of Medicine, 2012. Identifying disaster medical and public health research priorities: Data needs arising in response to Hurricane Sandy: Meeting summary. Available at:


Smith, V, Clarke, M, Devane, D, Begley, C, Shorter, G and Maguire, L, 2013. SWAT 1: what effects do site visits by the principal investigator have on recruitment in a multicentre randomized trial? Journal of Evidence-Based Medicine, 6(3), pp.136–37.


Turner, L, Shamseer, L, Altman, DG, Weeks, L, Peters, J, Kober, T, Dias, S, Schulz, KF, Plint, AC and Moher, D, 2012. Consolidated standards of reporting trials (CONSORT) and
the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database of Systematic Reviews, (11), MR000030.


Appendix 1: Methods and content of the online survey

The content of the online survey was agreed between Evidence Aid and 3ie through a series of drafts from late December 2013 to early January 2014. The final version had three main sections, with a mixture of closed (multiple choice) and open questions (see below). This was to facilitate analyses, while also giving respondents the opportunity to provide their opinions and experience in their own words. The first section asked about the respondent and their experience in humanitarian assistance and related areas. The second was designed to gather information about their experience of, and priorities for, impact evaluations, through a mixture of questions seeking either short answers or selection from multiple-choice lists. The survey concluded with an opportunity for respondents to provide additional comments. We stressed that participation in the survey was voluntary and that responses would not be linked to identifiable individuals in any reports.

The survey was made available on the internet, using SurveyMonkey. A ‘soft launch’ took place on 6 January 2014, to allow live testing by a small group of people working in the humanitarian sector or associated with Evidence Aid. The main launch took place on 8 January with distribution of an email notification to approximately 600 people connected with Evidence Aid and 3ie. We also made announcements through the Evidence Aid social media channels in Facebook, LinkedIn and Twitter, and on the homepage of the Evidence Aid website from 13 to 27 January (www.EvidenceAid.org).

We are aware that the information was retweeted from many Twitter accounts, including @Reliefweb (32,600 followers on 8 January), @hildabast (2,233 followers on 10 January), @aid_leap (1,428 followers on 11 January), @UKCDSS (1,369 followers on 15 January), @cochraneccollab (27,826 followers on 20 January) and @HEARDatUNSW (1,059 followers on 27 January). We posted information to the email distribution lists of ISCRAM, DisasterOutreach and HIFA2015 on 8 January and placed notices on various websites, including those of ELRHA (Enhancing Learning and Research for Humanitarian Assistance) and People In Aid.

The closing date for the survey was set as 27 January 2014 and we retrieved 399 responses from SurveyMonkey on 3 February and loaded them into Excel for coding and analysis, excluding those that were part of the drafting or preliminary testing of the survey. (A subsequent check of SurveyMonkey on 15 February revealed a small number of additional responses.) We checked the retrieved records individually to remove any (n=4) that appeared to have been made by people working through the survey to see which questions were being asked, rather than to provide meaningful answers; leaving 395 for the analyses presented in this report. To assist with the analyses, we read all free text responses to the survey and developed a provisional coding sheet. This was supplemented with codes identified in the analysis of the notes of the semi-structured interviews and was then used to code responses. This facilitated the identification and aggregation of common themes. The coding sheets are shown in Appendices 4a and 4b.
The content of the survey is reproduced below:

**INFORMATION SHEET**

Dear Respondent,

Thank you for your interest in this survey. We are seeking responses from a wide range of actors in the humanitarian emergency sector and related areas, to identify areas of humanitarian assistance that most need evidence to inform policy and practice.

Please share the link (www.surveymonkey.com/s/VLP8JB2) with colleagues, as appropriate, so that they can also contribute.

**OBJECTIVES**

This survey will provide valuable information for an ongoing project being undertaken by Evidence Aid on behalf of the International Initiative for Impact Evaluation (3ie). The aims of the project are to:

- Help an international consortium of donors, including DFID and USAID, to identify priority areas and evaluation questions for impact evaluations in the humanitarian sector.
- Help inform the design of a thematic call for proposals to conduct impact evaluations of humanitarian assistance to crises arising from sudden-onset disasters and protracted emergencies.

**STRUCTURE OF THE SURVEY**

The survey has three sections. The first section asks about you and your experience in this sector. The second is designed to gather information about your experience of, and priorities for, impact evaluations, through a mixture of questions seeking short answers or selection from multiple choice lists. The survey concludes with an opportunity for you to provide additional comments.

Participation in the survey is voluntary and responses will not be linked to identifiable individuals in any reports.

Questions about the survey, or the project more generally, can be sent to Professor Mike Clarke, Evidence Aid and Queen’s University Belfast (m.clarke@qub.ac.uk, +4428-90635059).

**DEFINITIONS**

To help orientate you to the issues covered in this survey, we are using the following definitions.

*Humanitarian crisis:* a situation in which there is an exceptional and generalized threat to human life, health or subsistence. These crises usually appear in the context of an
existing situation of a lack of protection where the consequences of preexisting factors
(such as poverty, inequality, lack of access to basic services) are exacerbated by a
natural disaster or armed conflict.

*Humanitarian assistance*: action designed to save lives, alleviate suffering and maintain
and protect human dignity during and in the aftermath of emergencies.

*Impact evaluations in the context of humanitarian crises*: studies assessing the causal
relationship between specific forms of humanitarian assistance and the impact on the
ultimate beneficiaries of this assistance. These usually seek to estimate the magnitude of
the impact, and also examine processes that enable and obstruct this impact. Some
impact evaluations also examine differences between subgroups.
SECTION 1: PARTICIPANT DETAILS

This first section asks about you and your experience in this sector.

1. Your name and email address (optional)
2. Organisations you have worked for in the humanitarian sector or related areas
3. Your current position
4. What are your areas of expertise (select as many as necessary)?
   Camp Coordination and Management
   Health
   Early Recovery
   Logistics
   Education
   Nutrition
   Emergency Shelter
   Protection
   Emergency Telecommunications
   Research
   Food Security
   Water, sanitation and hygiene
   Other (please specify)

5. How long have you worked in humanitarian assistance or areas related to it?
   <2 years
   2–4 years
   5–9 years
   10–19 years
   >20 years
   Not applicable
6. In which country/region are you currently based?

7. In which countries/regions do you usually work?

SECTION 2a: YOUR EXPERIENCE OF IMPACT EVALUATIONS IN THE CONTEXT OF HUMANITARIAN CRIZES

8. Have you commissioned or conducted an impact evaluation in the context of a humanitarian crisis?
   No / Yes

8a. Which intervention or project/programme did you do an impact evaluation for (please provide a brief outline of the impact evaluation, including the area of work and references to any reports)?

8b. How much did the impact evaluation cost?
   Less than US$50,000
   US$50,000–99,999
   US$100,000–249,999
   US$250,000–499,999
   US$500,000–999,999
   More than US$1,000,000
   Don’t know

9. Have you ever used the findings of an impact evaluation?
   No / Yes / Don’t know

9a. Please provide a brief description of the impact evaluation you have used, or a reference to it:

9b. Please outline how you used it:

10. If you are aware of a potentially relevant impact evaluation but have not used its findings, please provide a brief explanation for why you have not used it:
SECTION 2b: PRIORITIES FOR IMPACT EVALUATIONS IN THE CONTEXT OF HUMANITARIAN CRISSES

11. Do you know of any work to identify priorities for impact evaluations that should be conducted in the context of humanitarian crises?
No / Yes

11a. Please provide brief details of this work to identify priorities for impact evaluations, including any outputs and references to any documentation:

12. Please select up to three areas for which you would most like evidence to inform policy or practice in humanitarian assistance:
Accountability
Camp Coordination and Management
Early Recovery
Education
Emergency Shelter
Emergency Telecommunications
Food Security
Health
Impact measurement
Logistics
Monitoring
Nutrition
Protection
Water, sanitation and hygiene
Other (please specify)

12a. If you would like to be more specific, please do so:

13. Please select the humanitarian crises phases for which you would most like evidence to inform policy or practice:
Resilience
Risk reduction
Immediate, short-term response
Prolonged response or engagement
No opinion
Other (please specify)

13a. If you would like to be more specific, please do so:

14. Which interventions or actions (such as cash transfer or food aid, types of emergency shelter, education programmes for children, psychotherapy following trauma, etc.) would you most like to see impact evaluations in the context of humanitarian crises (please provide up to three):

14a. If you would like to be more specific, please do so:

15. Which outcomes (for example educational ability, mortality, morbidity, sustainable income, etc.) do you regard as the most important to be measured in impact evaluations in the context of humanitarian crises (please provide up to three):

15a. If you would like to be more specific, please do so:

16. Which populations (for example children, elderly, people with physical disability, women, etc.) do you regard as the most important to be included in impact evaluations in the context of humanitarian crises:

16a. If you would like to be more specific, please do so:

17. Which type of humanitarian crisis would you regard as the most important for study in impact evaluations?
Sudden-onset disasters
Protracted humanitarian emergencies
Both equally
No opinion
Other (please specify)

17a. If you would like to be more specific, please do so:
SECTION 3: ADDITIONAL COMMENTS

18. Please provide any additional comments you wish to make:

Thank you for taking part in this survey.

If you have any questions about the survey or the project more generally, please contact Professor Mike Clarke (E-mail m.clarke@qub.ac.uk; telephone +442890635059; post: Evidence Aid, Centre for Public Health, (ICS B), Queens University Belfast, Grosvenor Road, Belfast, BT12 6BJ, Northern Ireland).
Appendix 2: People who were interviewed (* indicates those who were interviewed during the first round of semi-structured interviews)

*Jonathan Abrahams, Emergency Risk Management and Humanitarian Response, World Health Organization, Switzerland
Myriam Ait-Aissa, Action Contre la Faim, France
Colin Armstrong, UK Collaborative on Development Sciences, UK
Veronique Barbelet, Overseas Development Institute, UK
Xavier Bosch-Capblanch, Swiss Tropical and Public Health Institute, Switzerland
*Tilman Brück, Stockholm International Peace Research Institute, Sweden
Margie Buchanan-Smith, Overseas Development Institute, UK
Skip Burkle, Harvard Humanitarian Initiative, USA
Sally Burrows, United Nations World Food Programme, USA
Laura Cartana, Swiss Tropical and Public Health Institute, Switzerland
Andy Catley, Feinstein International Center, Tufts University, USA
Jennifer Chan, Northwestern University; Harvard Humanitarian Initiative, USA
*Rudi Coninx, Policy, Practice & Evaluation Unit, World Health Organization, Switzerland
John Cosgrove, independent consultant, UK
Annie Devonport, Disasters Emergency Committee, UK
*Belinda Duff, Evaluation Adviser (OCHA-RDC)
Wendy Fenton, Overseas Development Institute, UK
*Peter Giesen, Humanitarian Strategy Consultants, Netherlands
Brendan Gormley, independent consultant, UK
*Scott Green, United Nations Office for the Co-ordination of Humanitarian Affairs, USA
Paul Harvey, Humanitarian Outcomes, UK
Samuel Hauenstein Swan, Action Against Hunger, UK
Alistair Humphrey, Medical Officer of Health (Canterbury), New Zealand
Randolph Kent, Humanitarian Futures Programme, UK
*Paul Knox Clarke, ALNAP, UK
Nick Lezama, Uniformed Services University of the Health Sciences, USA
David Loquerio, Humanitarian Accountability Partnership, Switzerland
*Joanna Macrae, Department for International Development, UK
Daniel Maxwell, Feinstein International Center, Tufts University, USA
Virginia Murray, Extreme Events and Health Protection, Public Health England, UK
Alice Obrecht, Humanitarian Futures Programme, UK
Nuala O’Brien, Irish Aid, Ireland
Rachel Pounds, Enhancing Learning and Research for Humanitarian Assistance, UK
Bernadette Peterhans, Swiss Tropical and Public Health Institute, Switzerland
Monica Ramos, Save the Children, France
Tony Redmond, Humanitarian and Conflict Response Institute, University of Manchester, UK
Matthew Reid, Médecins Sans Frontières, South Africa
*Jennie Richmond, Oxfam, UK
Elias Sagmeister, Global Public Policy Institute, Germany
David Sanderson, Centre for Development and Emergency Practice, Oxford Brookes University, UK
Kevin Savage, World Vision International, Switzerland
Andy Seal, Nutrition in Crisis Research Group, University College London, UK
Hugo Slim, Institute for Ethics, Law and Armed Conflict, University of Oxford, UK
Emanuele Sozzi, University of Brighton, UK
Huw Taylor, University of Brighton, UK
*Vivien Margaret Walden, Oxfam, UK
Cara Winters, Norwegian Refugee Council, Norway
Anthony Zwi, University of New South Wales, Australia

A further four people were interviewed but did not confirm their willingness to be listed here.
Appendix 3a: Methods used during the first phase of scoping interviews

We used a semi-structured approach to ensure that specific topics were covered, while providing the opportunity for the depth and breadth of the discussion to vary depending on the response to each question. Notes were taken during the interviews, and subsequently analysed to identify key themes. This analysis involved the initial highlighting of elements of the interview of most relevance to this study, and then a careful reading through of the notes to identify key themes. We then sought out these themes in the notes for each interview, coded them and entered them into an Excel spreadsheet to facilitate analysis. The main purpose of these interviews was to guide the subsequent conduct of the scoping study, so that the final product would help those responsible for defining the content of the Humanitarian Interventions Thematic Window. The choice of people to interview reflected this need. The intention was that this first phase of interviews would take place before 24 December 2013, so as to inform the drafting of the online survey and the questions for the subsequent semi-structured interviews. Unfortunately, this was not possible in all cases and some members of the Steering Committee could not be interviewed. We covered the following issues during these interviews:

- Are you aware of any existing efforts to identify or prioritise impact evaluations that are needed in the context of humanitarian crises?
- In what areas would you like to see evidence from impact evaluations?
- Would policymakers and practitioners use evidence from impact evaluations?
- Should the scope of the Humanitarian Interventions Thematic Window cover certain types of humanitarian crises only (such as sudden-onset disasters or protracted emergencies, or more specific settings) or certain phases for the implementation of interventions and actions (such as risk reduction, response, recovery or resilience)?
- Would a structure for the gap map that matched interventions or actions against outcomes be an appropriate way to match evidence needs with existing impact evaluations?
- Should the presentation of the information be structured in accordance with the framework used to define the clusters in the humanitarian sector (camp coordination and management, early recovery, education, emergency shelter, emergency telecommunications, food security, health, logistics, nutrition, protection, telecommunications, and water, sanitation and hygiene)?
- Who else, and what resources, should we consult for this study?
Appendix 3b: Methods used during the second phase of semi-structured interviews

We selected people to invite for these interviews by reviewing a list of 535 people from the humanitarian sector for whom Evidence Aid or 3ie had contact information (name, address and organisation). 3ie suggested that Evidence Aid should interview two of these people in particular, and others were systematically chosen on the basis of the following: ideally, there would be only one interview per organisation; the person was already interacting with Evidence Aid and was likely to respond in the limited time available for the interviews; and their role within their organisation suggested that they would be involved in identifying evidence needs or conducting research. We contacted 50 people on 10 January and sent weekly reminders up to 31 January. As the early responses suggested that the target of 40 semi-structured interviews to supplement those from the scoping phase was achievable, if someone did not reply to their invitation we did not replace them with another person from the list of potential contacts. We also asked interviewees for their suggestions for other people to invite (a process called snowballing) and additional suggestions were made in discussions between 3ie and Evidence Aid. Many of the additional suggestions came from multiple interviewees and, in order to accommodate these people, we extended the original planned date for the closure of this phase. In total, we sent 100 invitations.

We conducted the interviews in parallel with the online survey, so that both elements could be completed during January 2014, for analysis in February. The framework for the interviews is shown below. As with the initial round of semi-structured interviews, notes were taken during the interviews, and subsequently analysed to identify key themes. This analysis involved the initial highlighting of elements of the interview of most relevance to this study, and then a careful reading through of the notes to identify issues that were added to the coding sheets developed for the analysis of the online survey. The final coding sheets (Appendices 4a and 4b) were then used on the notes for each of these interviews (and on the notes from the first phase of interviews) to categorise responses. We entered the resulting data into an Excel spreadsheet to facilitate analysis.

We used the following questions during these interviews:

1. Background of respondent
   1a) In what way are you involved in humanitarian assistance?
   1b) For how long have you worked in the humanitarian sector?
   1c) What are the main areas of your work?
   1d) Are you involved in research (if so, how)?

2. Experience of impact evaluations
   2a) Have you been involved in impact evaluations in the context of humanitarian crises?
   2b) How (for example, commissioning, conducting, or implementing the findings)?
3. Priorities for impact evaluations

3a) What are the general areas for which you consider impact evaluations are most needed to inform policy and practice (consider the type of humanitarian assistance, the outcomes to be addressed and the relevant setting and population)?

3b) Do you have specific topics for impact evaluations that you would like to see within these areas?

3c) Why do you think that these areas are of high priority?

4. Further information

4a) Are there any documents that might help with this project?

4b) Who else would you recommend that we speak to?

4c) How might we promote the online survey?
Appendix 4a: Categories used for coding of suggestions for interventions or actions for impact evaluations in the context of humanitarian crises

<table>
<thead>
<tr>
<th>Category</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>Legal</td>
</tr>
<tr>
<td>Accountability</td>
<td>Livelihood</td>
</tr>
<tr>
<td>Animal health</td>
<td>Nutrition</td>
</tr>
<tr>
<td>Behaviour change</td>
<td>Organisation</td>
</tr>
<tr>
<td>Cash</td>
<td>Organisation: camps</td>
</tr>
<tr>
<td>Cash or voucher</td>
<td>Organisation: certification</td>
</tr>
<tr>
<td>Child focused</td>
<td>Organisation: community</td>
</tr>
<tr>
<td>Climate change</td>
<td>Organisation: coordination</td>
</tr>
<tr>
<td>Communication</td>
<td>Organisation: financing</td>
</tr>
<tr>
<td>Conflict mitigation</td>
<td>Organisation: government</td>
</tr>
<tr>
<td>Development</td>
<td>Organisation: health</td>
</tr>
<tr>
<td>Donation</td>
<td>Organisation: human resources</td>
</tr>
<tr>
<td>Disaster risk reduction</td>
<td>Organisation: leadership</td>
</tr>
<tr>
<td>Early phase interventions</td>
<td>Organisation: local</td>
</tr>
<tr>
<td>Early warning</td>
<td>Organisation: local government</td>
</tr>
<tr>
<td>Education</td>
<td>Organisation: local workers</td>
</tr>
<tr>
<td>Ethics</td>
<td>Organisation: private sector</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Organisation: responders</td>
</tr>
<tr>
<td>Food aid</td>
<td>Organisation: supply chain</td>
</tr>
<tr>
<td>Food or cash</td>
<td>Prioritisation</td>
</tr>
<tr>
<td>Food security</td>
<td>Protection</td>
</tr>
<tr>
<td>Gender</td>
<td>Protracted emergencies</td>
</tr>
<tr>
<td>Gender-based violence (GBV)</td>
<td>Recovery</td>
</tr>
<tr>
<td>Health</td>
<td>Refugees</td>
</tr>
<tr>
<td>Health: chronic diseases</td>
<td>Rehabilitation</td>
</tr>
<tr>
<td>Health: infectious disease</td>
<td>Rehabilitation: urban</td>
</tr>
<tr>
<td>Health: lay workers</td>
<td>Resilience</td>
</tr>
<tr>
<td>Health: mental</td>
<td>Responders</td>
</tr>
<tr>
<td>Health: routine healthcare</td>
<td>Shelter</td>
</tr>
<tr>
<td>Health: vaccine</td>
<td>Sustainability</td>
</tr>
<tr>
<td>Humanitarian assistance as a whole</td>
<td>Technology</td>
</tr>
<tr>
<td>Impact evaluation</td>
<td>Vouchers</td>
</tr>
<tr>
<td>Information</td>
<td>Water, sanitation and hygiene (WASH or WATSAN)</td>
</tr>
</tbody>
</table>
Appendix 4b: Categories used for coding of suggestions for outcomes for impact evaluations in the context of humanitarian crises

Access
Animal health
Assets
Behaviour change
Conflict
Cost
Displacement
Disaster risk reduction
Economic
economic development
Education
Education: ability
Education: access
Education: access and ability
Efficiency
Employment
Empowerment
Environment
Equity
Family
Finance
Food: access
Food security
Health: access
Health: malnutrition
Health: morbidity
Health: morbidity and mortality
Health: mortality
Income
Information
Investment
Legal
Livelihood
Mobility
Nutrition
Organisation
Organisation: community
Organisation: government
Organisation: local
Organisation: policy
Organisation: prioritisation
Preparedness
Protection
Recovery
Resilience

Response
Risk reduction
Satisfaction
Services
Shelter
Social functioning
Sustainability
Violence
Vulnerability
Water quality
Water, sanitation and hygiene: access
Appendix 5: Search strategies used

We used a variety of searches to identify evidence needs, priorities for research, examples of impact evaluations of humanitarian assistance, and emerging trends in the design and conduct of impact evaluations. There was overlap between the searches and, for completeness, we list each of them here. These include both simple searches to find potentially pivotal documents that use specific terminology, as well as highly detailed searches that had been carefully designed by Evidence Aid to identify controlled studies and systematic reviews across a wide range of disasters and humanitarian emergencies which had been conducted in 2011 and which were run again in 2013 to identify more recent papers which were added to those found previously, and tailored searches of the repositories of evaluations.

a. Ovid MEDLINE (November 2013) [controlled studies]

1  bhopal accidental release/ or chernobyl nuclear accident/ or radioactive hazard release/po or exp radiation injuries/po (10,397)

2  ((nuclear or atomic) adj3 (disaster$ or accident$ or incident$ or meltdown or melt-down or explosion$ or catastrophe$ or calamity$ or crisis or crises or leak$ or seep$ or breach$ or hazard$)).ti,ab,ot. (2,104)

3  ((nuclear or atomic or dirty or biological$) adj3 (bomb$ or weapon$ or WMD or warfare or attack$ or assault$ or strike$ or missile$ or warhead$ or warhead$)).ti,ab,ot. (4,805)

4  (((Atomic or radiation or radio-active or radioactive or nuclear) adj3 (expos$ or contaminat$ or releas$ or fallout or fall-out or disaster$ or accident$ or incident$ or explosion$ or meltdown or melt-down or explosion$ or leak$ or seep$ or breach$ or calamity$)) and (prevent$ or iodine or Radio-iodine or Radioiodine or I-131 or I131 or iodide$)).ti,ab,ot. (2,222)

5  (chernobyl or fukushima or "three mile island").ti,ab,ot. (3,980)

6  or/1-5 (20,816)

7  mass casualty incidents/ (942)

8  avalanches/ or earthquakes/ or landslides/ or tidal waves/ or tsunamis/ or volcanic eruptions/ or cyclonic storms/ or tornadoes/ (4,114)

9  (((Natural$ or technological$ or man-made or manmade or climatolog$ or geophysical$ or geo-physical$ or hydrolog$ or meteorolog$) adj3 (catastrophe$ or disaster$ or crisis or crises$ or emergency or emergencies or calamity$ or devastat$ or hazard$)).ti,ab,ot. (2,681)

---

88 Misso et al. 2012.
10 (humanitarian adj3 (catastroph$ or disaster$ or crisis or crises or aid or relief or assist$ or support$ or respon$ or emergency or emergencies or calamit$)).ti,ab,ot. (881)

11 (earthquake$ or quake$ or seismic tremor$ or seismic temblor$ or seismic eruption$ or tsunami$ or avalanche$ or flood$ or cyclon$ or hurricane$ or tornado$ or limnic eruption$ or lake overturn$ or lake over-turn$ or wildfire$ or wild-fire$ or forestfire$ or forest-fire$ or bushfire$ or bush-fire$ or tidal wave$ or twister or twisters).ti,ab,ot. (17,320)

12 (landslide$ or volcanic or volca$ or land-slide$ or mudslide$ or mud-slide$ or pyroclastic flow$ or lahar or lahars or pyroclastic density current$ or pyroclastic surge$ or debris flow$ or Jokulhlaup$).ti,ab,ot. (2,191)

13 bhopal.ti,ab,ot. (318)

14 ((armero or carabellda or Vargas) adj2 traged$).ti,ab,ot. (2)

15 or/7-14 (23,866)

16 refugees/ or disaster victims/ (6,751)

17 (displaced adj3 (population$ or people$ or person$) adj3 (site$ or shelter$ or accommodation$ or camp$ or tent$ or structure$ or settlement$ or housing$)).ti,ab,ot. (114)

18 (victim$ or survivor$) adj3 displacement).ti,ab,ot. (8)

19 (evacuee$ or refugee$) adj3 (site$ or shelter$ or accommodation$ or camp$ or tent$ or structure$ or settlement$ or housing$)).ti,ab,ot. (862)

20 (temporary or emergency or evacuat$) adj3 (site$ or shelter$ or accommodation$ or camp$ or tent$ or structure$ or settlement$ or housing$)).ti,ab,ot. (1,201)

21 or/16-20 (8,158)

22 Disaster planning/ or rescue work/ (12119)

23 (major incident$ or mass trauma$ or mass casualt$ or multiple casualt$) adj3 (incident$ or event$ or situation$ or disaster$ or emergency$ or catastrop$) adj3 (plan$ or prepar$ or respon$ or train$ or equip$ or organi$ or organiz$ or arrang$)).ti,ab,ot. (272)

24 (MDI adj3 (plan$ or prepar$ or respon$ or train$ or equip$ or organi$ or organiz$ or arrang$)).ti,ab,ot. (56)

25 ((disaster$ or catastrop$) adj3 (plan$ or prepar$ or respon$ or equip$ or organi$ or organiz$ or train$)).ti,ab,ot. (3,692)
26 ((natural or humanitarian or technological or man-made) adj3 (disaster$ or emergency$ or catastrophe$) adj3 (plan$ or prepare$ or respond$ or train$ or equip$ or organise$ or organize$ or arrange$)).ti,ab,ot. (212)

27 ((rail$ or train or trains or plane or planes or aeroplane$ or aviation or aircraft$ or jet or jets or seaplane$ or sea-plane$ or bus or buses or underground or subway$ or tube or station or station$ or airport$ or air-port$ or airterminal$ or air-terminals$) adj3 (crash or collision or accident$ or bomb$ or attack$)).ti,ab,ot. (1,352)

28 ((ferry or ferries or ship or ships) adj3 (crash or collision or accident$ or bomb$ or attack$)).ti,ab,ot. (34)

29 ((cruiseline$ or cruise-liner$) adj3 (crash or collision or accident$ or bomb$ or attack$)).ti,ab,ot. (0)

30 or/22-29 (14,978)

31 ((natural or technological or man-made or manmade or humanitarian) adj3 (catastrophe$ or disaster$ or crisis$ or aid or emergency or emergencies or catastrophe$)).ti,ab,ot. (2,865)

32 ((disaster$ or catastrophe$) adj3 setting$).ti,ab,ot. (132)

33 (quake or seismic tremor$ or earthquake$ or landslide$ or land-slide$ or tsunami$ or volcanic or avalanche$ or flood$ or cyclon$ or hurricane$ or tornado$ or typhoon$ or whirlwind$ or whirl-wind$ or wildfire$ or wild-fire$ or forestfire$ or forest-fire$ or bushfire$ or bush-fire$ or tidal wave$ or twister or twist$ or twisters).ti,ab,ot. (18,833)

34 (displaced population$ or refugee$ or displaced people$ or displaced person$ or evacuee$).ti,ab,ot. (6,220)

35 ((mass trauma$ or mass casualty$ or multiple casualty$) adj3 (event or events or episodes$ or calamity$ or incident$ or situation$ or disasters$ or emergency$ or catastrophe$ or crisis or crises$)).ti,ab,ot. (863)

36 (Mount St Helens or Pinatubo or Mount Etna or Tungurahua or Eyjafjallajokull).ti,ab,ot. (128)

37 ((terror$ or suicide) adj3 (attack$ or atrocity$ or bomb$ or campaign$)).ti,ab,ot. (1,833)

38 "September 11 Terrorist Attacks"/ or (twins towers or world trade centre).ti,ab,ot. (961)

39 or/31-38 (29,700)

40 ((infrastructure or infra-structure or communication$ or sanitation or water or transport$ or shelter$ or accommodation$ or camp$ or housing or tent or tents or
settlement$ or power or electricity or sewage or road or roads or rail$ or phone$ or twitter$ or retweet$ or internet$ or web).ti,ab,ot. (1,322,418)

41  (health$ or wellbeing or well-being or hospital$ or doctor$ or medic$ or clinic or clinics or psycholog$ or clinician$ or practitioner$ or paramedic$ or first responder$ or first aid$ or emergency services$ or rescue worker$ or ambulance$ or physician$ or surgeon$).ti,ab,ot. (3,517,652)

42  (pam or pams).ti,ab,ot. (4,496)

43  ((hospital$ or health or surg$ or medical) adj4 (personnel$ or employ$ or worker$ or team$ or staff$ or workforce$ or work force$ or manpower$)).ti,ab,ot. (100,286)

44  (nursing or nurse or nurses or matron or matrons or auxiliary or auxiliaries or midwi$).ti,ab,ot. (324,690)

45  (peacekeep$ or peace-keep$ or security or NGO or NGOs or non-governmental organi?ations$ or nongovernmental organi?ations$ or Disaster Relief Operation$ or DRO or Non-combatant evacuation operation$ or Noncombatant evacuation operation$ or NEO).ti,ab,ot. (41,427)

46  (aid agenc$ or relief agenc$).ti,ab,ot. (235)

47  or/40-46 (4,767,839)

48  39 and 47 (16,991)

49  or/6,15,21,30,48 (65,067)

50  randomized controlled trial.pt. (390,995)

51  controlled clinical trial.pt. (90,070)

52  randomized.ab. (288,395)

53  placebo.ab. (157,299)

54  clinical trials as topic.sh. (175,750)

55  randomNy.ab. (200,079)

56  trial.ti. (124,923)

57  or/50-56 (897,019)

58  animals/ not (animals/ and humans/) (3,974,347)

59  57 not 58 (826,270)

60  49 and 59 (1,534)
b. Ovid MEDLINE (November 2013) [systematic reviews]

1. bhopal accidental release/ or chernobyl nuclear accident/ or radioactive hazard release/pc or exp radiation injuries/pc (10,397)

2. ((nuclear or atomic) adj3 (disasters$ or accident$ or incident$ or meltdown or meltdown or explosion$ or catastrophe$ or calamity$ or crisis or crises or leak$ or seep$ or breach$ or hazard$)).ti,ab,ot. (2,104)

3. ((nuclear or atomic or dirty or biological$) adj3 (bomb$ or weapon$ or WMD or warfare or attack$ or assault$ or strike$ or missile$ or warhead$ or warhead$)).ti,ab,ot. (4,805)

4. (((Atomic or radiation or radio-active or radioactive or nuclear) adj3 (expos$ or contaminat$ or releas$ or fallout or fall-out or disaster$ or accident$ or incident$ or explosion$ or meltdown or melt-down or explos$ or leak$ or seep$ or breach$ or calamity$)) and (prevent$ or iodine or Radio-iodine or Radioiodine or I-131 or I131 or iodide$)).ti,ab,ot. (2,222)

5. (chernobyl or fukushima or "three mile island").ti,ab,ot. (3,980)

6. or/1-5 (20,816)

7. mass casualty incidents/ (942)

8. avalanches/ or earthquakes/ or landslides/ or tidal waves/ or tsunamis/ or volcanic eruptions/ or cyclonic storms/ or tornadoes/ (4,114)

9. (((Natural$ or technological$ or man-made or manmade or climatolog$ or geophysical$ or geo-physical$ or hydrolog$ or meteorolog$) adj3 (catastrophe$ or disaster$ or crisis or crises$ or emergency or emergencies or calamity$ or devastat$ or hazard$)).ti,ab,ot. (2,681)

10. (humanitarian adj3 (catastrophe$ or disaster$ or crisis or crises or aid or relief or assist$ or support$ or respon$ or emergency or emergencies or calamity$)).ti,ab,ot. (881)

11. (earthquake$ or quake$ or seismic tremor$ or seismic temblor$ or seismic eruption$ or tsunami$ or avalanche$ or flood$ or cyclon$ or hurricane$ or tornado$ or limnic eruption$ or lake overturn$ or lake over-turn$ or wildfire$ or wild-fire$ or forestfire$ or forest-fire$ or bushfire$ or bush-fire$ or tidal wave$ or twister or twisters).ti,ab,ot. (17,320)
12 (landslide$ or volcanic or volcano$ or land-slide$ or mudslide$ or mud-slide$ or pyroclastic flow$ or lahar or lahars or pyroclastic density current$ or pyroclastic surge$ or debris flow$ or Jokulhlaup$).ti,ab,ot. (2,191)

13 bhopal.ti,ab,ot. (318)

14 ((armero or carabela$ or Vargas) adj2 traged$).ti,ab,ot. (2)

15 or/7-14 (23,866)

16 refugees/ or disaster victims/ (6,751)

17 (displaced adj3 (population$ or people$ or person$) adj3 (site$ or shelter$ or accommodation$ or camp$ or tent$ or structure$ or settlement$ or housing$)).ti,ab,ot. (114)

18 ((victim$ or survivor$) adj3 displacement).ti,ab,ot. (8)

19 ((evacuee$ or refugee$) adj3 (site$ or shelter$ or accommodation$ or camp$ or tent$ or structure$ or settlement$ or housing$)).ti,ab,ot. (862)

20 ((temporary or emergency or evacuat$) adj3 (site$ or shelter$ or accommodation$ or camp$ or tent$ or structure$ or settlement$ or housing$)).ti,ab,ot. (1,201)

21 or/16-20 (8,158)

22 disaster planning/ or rescue work/ (12,119)

23 ((major incident$ or mass trauma$ or mass casual$ or multiple casual$) adj3 (incident$ or event$ or situation$ or disaster$ or emergenc$ or catastrop$) adj3 (plan$ or prepar$ or respons$ or train$ or equip$ or organ$ or organiz$ or arrange$)).ti,ab,ot. (272)

24 (MDI adj3 (plan$ or prepar$ or respons$ or train$ or equip$ or organ$ or organiz$ or arrange$)).ti,ab,ot. (56)

25 ((disaster$ or catastrop$) adj3 (plan$ or prepar$ or respons$ or equip$ or organ$ or organiz$ or train$)).ti,ab,ot. (3,692)

26 ((natural or humanitarian or technological or man-made) adj3 (disaster$ or emergenc$ or catastrop$) adj3 (plan$ or prepar$ or respons$ or train$ or equip$ or organ$ or organiz$ or arrange$)).ti,ab,ot. (212)

27 ((rail$ or train or trains or plane or planes or aeroplane$ or aviation or aircraft$ or jet or jets or seaplane$ or sea-plane$ or bus or buses or underground or subway$ or tube or station or stations$ or airport$ or air-port$ or airtermal$ or airtermal$) adj3 (crash or collision or accident$ or bomb$ or attack$)).ti,ab,ot. (1,352)
28 ((ferry or ferries or ship or ships) adj3 (crash or collision or accident$ or bomb$ or attack$)).ti,ab,ot. (34)

29 ((cruiseline$ or cruise-liner$) adj3 (crash or collision or accident$ or bomb$ or attack$)).ti,ab,ot. (0)

30 or/22-29 (14,978)

31 ((natural or technological or man-made or manmade or humanitarian) adj3 (catastroph$ or disaster$ or crisis$ or aid or emergency or emergencies or catastrophe$)).ti,ab,ot. (2,865)

32 ((disaster$ or catastrophe$) adj3 setting$).ti,ab,ot. (132)

33 (quake or seismic tremor$ or earthquake$ or landslide$ or land-slide$ or tsunami$ or volcanic or avalanche$ or flood$ or cyclone$ or hurricane$ or tornado$ or typhoon$ or whirlwind$ or whirl-wind$ or wildfire$ or wild-fire$ or forestfire$ or forest-fire$ or bushfire$ or bush-fire$ or tidal wave$ or twister or twisters).ti,ab,ot. (18,833)

34 (displaced population$ or refugee$ or displaced people$ or displaced person$ or evacuee$).ti,ab,ot. (6,220)

35 ((mass trauma$ or mass casualty$ or multiple casualty$) adj3 (event or events or episode$ or calamity$ or incident$ or situation$ or disaster$ or emergenc$ or catastrophe$ or crisis or crises$)).ti,ab,ot. (863)

36 (Mount St Helens or Pinatubo or Mount Etna or Tungurahua or Eyjafjallajokull).ti,ab,ot. (128)

37 ((terror$ or suicide) adj3 (attack$ or atrocity$ or bomb$ or campaign$)).ti,ab,ot. (1,833)

38 “September 11 Terrorist Attacks” or (twin towers or world trade centre).ti,ab,ot. (961)

39 or/31-38 (29,700)

40 (infrastructure or infra-structure or communication$ or sanitation or water or transport$ or shelter$ or accommodation$ or camp$ or housing or tent or tents or settlement$ or power or electricity or sewage or road or roads or rail$ or phone$ or twitter$ or retweet$ or internet$ or web).ti,ab,ot. (1,322,418)

41 (health$ or wellbeing or well-being or hospital$ or doctor$ or medic$ or clinic or clinics or psychologist$ or clinician$ or practitioner$ or paramedic$ or first responder$ or first aid$ or emergency services$ or rescue worker$ or ambulance$ or physician$ or surgeon$).ti,ab,ot. (3,517,652)

42 (pav or pams).ti,ab,ot. (4,496)
((hospital$ or health or surg$ or medical) adj4 (personnel$ or employ$ or worker$ or team$ or staff$ or workforce$ or work force$ or manpower$)),ti,ab,ot. (100,286)

(nursing or nurse or nurses or matron or matrons or auxillary or auxiliaries or midwi$),ti,ab,ot. (324,690)

(peacekeep$ or peace-keep$ or security or NGO or NGOs or non-governmental organi?ation$ or nongovernmental organi?ation$ or Disaster Relief Operation$ or DRO or Non-combatant evacuation operation$ or Noncombatant evacuation operation$ or NEO),ti,ab,ot. (41,427)

(aid agenc$, or relief agenc$),ti,ab,ot. (235)

or/40-46 (4,767,839)

39 and 47 (16,991)

or/6,15,21,30,48 (65,067)

Cochrane database of systematic reviews,jn. or search.tw. or meta-analysis,pt. or MEDLINE.tw. or systematic review.tw. (234,508)

animals/ not (animals/ and humans/) (3,974,347)

50 not 51 (218,438)

49 and 52 (861)

(201105$ or 201106$ or 201107$ or 201108$ or 201109$ or 20111$ or 2012$ or 2013$ or 2014$).ed,dc. or (2011$ or 2012$ or 2013$ or 2014$).yr. (2,385,711)

53 and 54 (241)

c. Web of Science (January 2014)

(research OR "impact evaluation"OR "impact assessment"OR "impact measurement") AND (priorities OR prioritisation OR prioritization) AND (disaster OR humanitarian) 154 records

d. EMBASE (including MEDLINE) (January 2014)

(research OR "impact evaluation" OR "impact assessment" OR "impact measurement") AND (priorities OR prioritisation OR prioritization) AND (disaster OR humanitarian) 118 records
e. Web of Science (January 2014)

("impact evaluation" OR "impact assessment" OR "impact measurement" OR "randomised" OR "randomized" OR "systematic review" OR meta-analysis OR meta-analyses) AND (disaster OR humanitarian)
372 records

f. EMBASE (including MEDLINE) (January 2014)

("impact evaluation" OR "impact assessment" OR "impact measurement" OR "randomised" OR "randomized" OR "systematic review" OR meta-analysis OR meta-analyses) AND (disaster OR humanitarian)
578 records

g. www.Google.com (January 2014)

General search of the internet using “research priorities”, “research strategy”, “priorities for impact evaluation” or “impact evaluation priorities”, linked with “disaster”, “disasters” or “humanitarian”.

h. ALNAP – Disaster preparedness, resilience and risk reduction (January 2014)

Full text search for ‘evaluation’: 170 records
Full text search for ‘impact evaluation’: 138 records
Full text search for ‘impact evaluation’ (January 2004 to February 2014): 86 records
Full text search for ‘impact evaluation’ and ‘counterfactual’ (January 2004 to February 2014): 2 records
i. ALNAP - Disasters (January 2014)

Full text search for 'evaluation': 487 records
Full text search for 'impact evaluation': 379 records
Full text search for 'impact evaluation' (January 2004 to February 2014): 312 records
Full text search for 'impact evaluation' and 'counterfactual' (January 2004 to February 2014): 1 record

j. 3ie (January 2014)

Search for 'evaluation': 2,272 records
Search for 'impact evaluation': 2,172 records
Search for 'impact evaluation disaster': 19 records
Search for "impact evaluation" (i.e. as a phrase): 2,039 records
Search for "impact evaluation" (i.e. as a phrase) and 'counterfactual': 31 records

k. 3ie - Impact evaluations (http://www.3ieimpact.org/en/evidence/impact-evaluations/) (January 2014)

Search for 'disaster': 5 records
Appendix 6: Methods used for review of existing impact evaluations

Our efforts to identify existing impact evaluations included searches of both the published and the grey literature. We sought studies in which an intervention and control group were compared (including randomised trials) and systematic reviews of the effects of interventions that might be used in humanitarian assistance. Studies were not eligible if they relied on computer-based simulation, table-top exercises or other simulations to estimate the effects that an intervention or strategy might have when used in the field.99

We searched bibliographic databases of the academic literature in health and social care (MEDLINE and EMBASE) and more generally (Web of Science). This draws on earlier work by Evidence Aid to identify research of this type across a wide variety of disasters and humanitarian emergencies.90 It also benefits from other assessments of the disaster-related literature, which were not necessarily focused on the effects of interventions. For example, Smith et al. identified more than 2,000 peer-reviewed, event-specific publications in 789 journals, following 25 individual disasters or "overwhelming crises". They found a total of 652 publications following disasters or events caused by natural hazards, 966 following human-made or technological disasters or events, and 480 following conflict or complex humanitarian events.91

We searched a series of repositories that have collected evaluations from the humanitarian and development sectors. The intent was to undertake a structured review of a number of existing inventories and repositories of evaluation reports in the humanitarian sector to identify impact evaluations, including possible exemplars of good practice. The initial phase of this study had identified 12 repositories, all largely in the humanitarian sector, containing approximately 3,000 evaluation reports. During the study, 3ie supplied further repositories, and others were identified through secondary searching of the articles reviewed. We selected two repositories as an initial trial for analysis: ALNAP (Active Learning Network for Accountability and Performance in Humanitarian Action) and the IFRC (International Federation of Red Cross and Red Crescent Societies). The IFRC analysis identified 309 evaluation reports, of which 7 included ‘impact evaluation’ in their title, but only 3 included the term in their document, of which none included a counterfactual or comparison group in the study. The ALNAP analysis identified 1,100 reports, of which 15 mentioned impact evaluation in the title, but when the repository was searched further for counterfactual, comparison or control, 46 reports were retrieved. However, when we read these 46 reports, none actually used a counterfactual, comparison or control methodology, but merely mentioned these terms in the text. Other drill-down searches or filters used included disaster, humanitarian crises, evaluation, impact evaluation, counterfactual, control, compare, comparison and comparator. Appendix 8 shows the template that was developed to capture key features of each repository, and the search strategy was also documented on this. A list of the repositories and inventories located and searched is included as Appendix 7. We also

90Misso et al. 2012.
91Smith et al. 2009.
reviewed selected journals, based on their high frequency of disaster-related articles in the review by Smith et al.\textsuperscript{92}

The searches were particularly challenging because of the lack of clarity and consistency in the description of impact evaluations, and the absence of any specific index terms or keywords. Furthermore, the classification of evaluations and their reports was not sufficiently consistent across the repositories to allow these categories to be relied upon when searching for relevant examples. The identification of randomised trials and other comparative, controlled studies in the health literature was easier because of concerted efforts to improve the indexing of such studies over the last two decades\textsuperscript{93} and the adoption of reporting standards for such studies.\textsuperscript{94}

The wide range of search strategies used across the variety of sources is shown in Appendix 5. These searches were not as comprehensive as they would be in formal systematic reviews but were intended to provide a broad overview of existing impact evaluations, in keeping with the scoping nature of this study. It was not possible to conduct a detailed examination of each individual impact evaluation or systematic review in the time available, but a fuller review of these might be particularly useful for areas chosen for the Humanitarian Interventions Thematic Window, in order to help with the selection and design of specific impact evaluations.

\textsuperscript{92}Smith et al. 2009.
\textsuperscript{93}Lefebvre et al. 2013.
\textsuperscript{94}Turner et al. 2012.
### Appendix 7: Repositories

<table>
<thead>
<tr>
<th>Agency</th>
<th>Inventory provided by Monash</th>
<th>Inventory provided by 3ie</th>
<th>Search date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Journals and/or reports</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prehospital and Disaster Medicine (PDM) – Journal</td>
<td></td>
<td>+</td>
<td>21/01/2014</td>
</tr>
<tr>
<td>Published by World Association for Disaster and Emergency Medicine (WADEM)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disaster Medicine and Public Health Preparedness – Journal</td>
<td>+</td>
<td></td>
<td>22/01/2014</td>
</tr>
<tr>
<td>Published by Society for Disaster Medicine and Public Health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kamedo – Socialstyrelsen – Swedish Disaster Medicine Study Organisation</td>
<td>+</td>
<td></td>
<td>21/01/2014</td>
</tr>
<tr>
<td>National Board of Health and Welfare – Sweden</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>United Nations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Nations Economic Commission for Latin America and the Caribbean (UNECLAC)</td>
<td>+</td>
<td></td>
<td>20/01/2014</td>
</tr>
<tr>
<td>United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA)</td>
<td>+</td>
<td></td>
<td>22/01/2014</td>
</tr>
<tr>
<td>UNOCHA Evaluation Reports</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Nations High Commissioner for Refugees (UNHCR)</td>
<td>+</td>
<td></td>
<td>20/01/2014</td>
</tr>
<tr>
<td>UNHCR Evaluation Reports</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Nations Children’s Fund (UNICEF)</td>
<td>+</td>
<td></td>
<td>21/01/2014</td>
</tr>
<tr>
<td>UNICEF Evaluation and Research Database</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Nations World Food Programme (WFP)</td>
<td>+</td>
<td>+</td>
<td>23/01/2014</td>
</tr>
<tr>
<td>WFP Evaluation Library</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Nations Development Programme</td>
<td>+</td>
<td></td>
<td>06/02/2014</td>
</tr>
<tr>
<td>International Policy Centre for Inclusive Growth (IPC-IG)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td>+</td>
<td>Date</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>-----</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>United Nations Evaluation Group (UNEG)</td>
<td></td>
<td>26/01/2014</td>
<td></td>
</tr>
<tr>
<td>UNEG Evaluation Resource Center (ERC)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food and Agriculture Organization of the United Nations (FAO)</td>
<td></td>
<td>05/02/2014</td>
<td></td>
</tr>
<tr>
<td>FAO Depository Libraries and Independent Office of Evaluation of International Fund of Agricultural Development</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>International organisations/NGOs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 8: Template used to record information about each repository

<table>
<thead>
<tr>
<th>Yes/No (if applicable)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title of the website or name of the inventory</td>
<td></td>
</tr>
<tr>
<td>Which organisation or authority owns the site?</td>
<td></td>
</tr>
<tr>
<td>Who funds the site?</td>
<td></td>
</tr>
<tr>
<td>URL</td>
<td></td>
</tr>
<tr>
<td>Hyperlink for inventory</td>
<td></td>
</tr>
<tr>
<td>What is the purpose of the inventory?</td>
<td></td>
</tr>
<tr>
<td>• Disaster/emergency or humanitarian/development</td>
<td></td>
</tr>
<tr>
<td>• What are the general themes?</td>
<td></td>
</tr>
<tr>
<td>Currency:</td>
<td></td>
</tr>
<tr>
<td>• Are dates given for when the site or repository was created?</td>
<td></td>
</tr>
<tr>
<td>• Are dates given for when the site was last updated or modified?</td>
<td></td>
</tr>
<tr>
<td>• Is the site ‘up to date’ with working links? (‘up to date’: last 3 months?)</td>
<td></td>
</tr>
<tr>
<td>Date accessed</td>
<td></td>
</tr>
<tr>
<td>Is access free/open or restricted?</td>
<td></td>
</tr>
<tr>
<td>Do you need to register as a user?</td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Answer</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>How many reports/articles can be found on the site?</td>
<td></td>
</tr>
<tr>
<td>Number of reports:</td>
<td></td>
</tr>
<tr>
<td>Filters:</td>
<td></td>
</tr>
<tr>
<td>• Disaster / humanitarian crises</td>
<td></td>
</tr>
<tr>
<td>• Evaluation</td>
<td></td>
</tr>
<tr>
<td>• Impact evaluation under Aim</td>
<td></td>
</tr>
<tr>
<td>• &quot;Impact evaluation&quot;</td>
<td></td>
</tr>
<tr>
<td>• &quot;Impact&quot; and &quot;evaluation&quot; in title</td>
<td></td>
</tr>
<tr>
<td>• Counterfactual, or control (+/- group) or compare or comparison or comparator</td>
<td></td>
</tr>
<tr>
<td>How many documents should be quality checked and reviewed (including any good exemplars)</td>
<td></td>
</tr>
<tr>
<td>Do they use standards or guidelines for the reports to be included?</td>
<td></td>
</tr>
<tr>
<td>Can you add your own report or is the article reviewed or peer reviewed? If yes, by whom?</td>
<td></td>
</tr>
<tr>
<td>Is there any advertising on the site?</td>
<td></td>
</tr>
<tr>
<td>Language – is it English only or multilingual with options to select a language other than English?</td>
<td></td>
</tr>
<tr>
<td>Are there further resources or links?</td>
<td></td>
</tr>
<tr>
<td>How did we find the site? (e.g. Google)</td>
<td></td>
</tr>
<tr>
<td>Comments:</td>
<td></td>
</tr>
<tr>
<td>• Well designed and organised</td>
<td></td>
</tr>
<tr>
<td>• Easy to read and navigate</td>
<td></td>
</tr>
<tr>
<td>• Help screens are available</td>
<td></td>
</tr>
<tr>
<td>• Search feature/site map is available</td>
<td></td>
</tr>
<tr>
<td>Search capabilities:</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>---</td>
</tr>
<tr>
<td>• Is the site searchable?</td>
<td></td>
</tr>
<tr>
<td>• How do you search (e.g. using key words)?</td>
<td></td>
</tr>
<tr>
<td>Interactive – (i.e., can you enter data or information and work with this and create your own documents)?</td>
<td></td>
</tr>
<tr>
<td>Audience – to whom is the site directed?</td>
<td></td>
</tr>
<tr>
<td>Does the website/organisation collect data on who is using their site?</td>
<td></td>
</tr>
<tr>
<td>Do they use social media? If so, which ones?</td>
<td></td>
</tr>
</tbody>
</table>

Brief outline of the organisation:
Appendix 9: Methods used to identify reports on evidence needs or priorities for research

We used a variety of methods to identify examples of where organisations, agencies or groups of stakeholders had identified evidence needs or priorities for future research into the effects of humanitarian assistance. These included drawing on the findings from the online survey, the semi-structured interviews and searches of the published and grey literature. We sought documents listing evidence needs and priorities for research into the effects of policies and interventions relevant to humanitarian assistance in sudden-onset disasters and protracted emergencies. These could be multi-sectorial, multi-agency or focused on a single agency, NGO or other actor.

We did not include documents that presented strategic approaches to improving the quality of research without proposing specific evidence needs or research priorities, such as the UK government’s response to the Humanitarian Emergency Response Review, which stressed the importance of the implementation and dissemination of high-quality research, without providing specific research priorities. Furthermore, if the evidence needs related to better understanding of particular outcomes, rather than the effects of interventions, these were excluded. Examples of such projects include a 2008 meeting and report by the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) on sexual violence in conflict situations, and a study of the capacity of health systems to cope with the surge in demand that arises after a disaster. We did not include work which related to research priorities for emergencies arising from terrorist action, such as the release of radiological isotopes. The recent work of Murray and Kessel to discuss different approaches to research prioritisation was also not eligible for inclusion, but provided useful insight into the issue.

In addition to the question in the online survey seeking information on work to identify priorities for evidence needs and impact evaluations and discussion in the semi-structured interviews, we sought relevant examples through the examination of websites of key organisations, and our literature searches for existing impact evaluations and for documents in which evidence needs or research priorities were described. The need for efficiency in light of the time and resource constraints of this study limited the extent of the searches and, in designing them, we were faced with the challenges arising from the lack of clarity and consistency in the description of evidence needs and impact evaluations, and the lack of specific index terms or keywords. It was not possible, therefore, to develop a highly sensitive and specific search for relevant examples in the published literature. The search strategies used for all parts of this study are shown in Appendix 5. We also undertook website-specific searches, checking the websites of key agencies (including Action Contre la Faim (ACF), ALNAP, AusAID, DFID, IFRC, J-PAL,

---

99DFID 2011.
98OCHA 2008.
97Watson et al. 2013.
96Pellmar et al. 2005.
95Murray and Kessel 2014.
 Médecins Sans Frontières (MSF), OCHA, Oxfam, UNICEF and WFP) for documents in which they presented priorities for evidence or for research that would help them to set policy relating to humanitarian assistance, or would inform policy and practice in the humanitarian sector more generally.

The information we wished to extract from each report is shown below:

- citation;
- lead organisation (where applicable);
- participants in any process to identify the needs or priorities;
- when the process took place;
- types of setting included (sudden-onset disasters / protracted humanitarian emergencies / major incident management);
- types of intervention or action included;
- method used to identify the evidence needs or priorities for research (e.g. survey, consensus, meeting, systematic review); and
- main conclusions.
Appendix 10: Summary of each report of research priorities (presented in chronological order, starting with the most recent report)


Settings: sudden-onset disasters; protracted humanitarian emergencies

This substantial review of research on health interventions was commissioned by DFID and the Wellcome Trust in 2013 and provides information on communicable disease control, WASH, nutrition, sexual and reproductive health (including gender-based violence), mental health and psychosocial support, non-communicable diseases, injury and rehabilitation, health service delivery, health systems, access to healthcare, accountability to end-users, health assessment methods, coordination, security of healthcare workers, and urbanisation. The overall aim of the project was to provide a rigorous assessment of the current quality and depth of the evidence base that informs humanitarian public health programming globally, with a specific objective to identify, through consultation with practitioners and policymakers, priority areas where further investment in the research and evidence base is most needed. The project used two main research methods: a systematic literature review on evidence on interventions of the health topics and contextual factors, and qualitative expert interviews with practitioners, policymakers and academics. Each of the sections of the report includes information on the research studies that were identified, along with recommendations for future research which include specific suggestions for interventions, actions or strategies that should be evaluated. The number of research needs listed for each area are: communicable disease control (6), WASH (5), nutrition (12), sexual and reproductive health (10), mental health and psychosocial support (14), non-communicable diseases (5), injury and rehabilitation (6), health service delivery (5), health systems (6), access to healthcare (9), accountability to end-users (8), health assessment methods (7), coordination (9), security of healthcare workers (9) and urbanisation (4).


Settings: sudden-onset disasters

The main aim of this paper was to suggest a prioritised agenda for organisational and management research on emergency planning and management relevant to healthcare in the UK, using a scoping study that was commissioned by the National Institute for Health Research, and including comparisons with the USA. The authors write ‘in general, emergency planning aims to increase the resistance and resilience of health-care supply and demand systems by implementing measures to prevent incidents, and preparing systems to respond to and recover from the incidents that do occur’ and their focus is on how an emergency planning system can have structures, processes, resources and governance that enable it to develop suitable plans, and to implement those plans effectively, and to update and revise them as necessary. They conducted a scoping study
to identify future research priorities across a wide, complex area of policy and practice, spanning different hazards, organisations and sectors; with information gathering during 2010–2011. The study used a variety of methods: a structured literature review; a survey of researchers; semi-structured interviews with 13 people from a range of UK stakeholder groups (including the ambulance and fire services, the Department of Health, a local council, voluntary and community organisations, and the Health Protection Agency); an exploration of debriefs of 20 small-scale incidents and of 2 larger case studies (the H1N1 outbreak in 2009–2010 and the Cumbria floods (2005 and 2009); and a prioritisation workshop and survey. The workshop (16 participants) and survey (a further 16 participants) rated 18 potential research topics and associated research questions, leading to 4 priority themes: public affected by health emergencies; inter- and intra-organisational collaboration; preparing responders and their organisations; and prioritisation and decision making.


Settings: sudden-onset disasters; protracted humanitarian emergencies

The objective of the Evidence Aid Priority Setting exercise was to identify approximately 30 high-priority research questions under 10 themes that could be addressed by systematic reviews in the area of planning for or response to natural disasters, humanitarian crises or other major healthcare emergencies. There was a particular focus on topics of particular relevance to low- and middle-income countries where the health impact of disasters may be greater than in high-income settings. The process started with an online survey asking humanitarian aid workers and others to provide up to 3 research questions or areas of uncertainty for which they need research evidence (101 participants). These suggestions were supplemented at 2 Evidence Aid conferences in 2011 and 2012 and from published literature, before being arranged into 43 themes. A second online survey was then used to prioritise these themes (233 participants) and the top 10 themes, along with the associated questions, were prioritised using a nominal group technique at a 2-day face-to-face workshop with 28 participants from a range of agencies and NGOs, to arrive at the top 30 priorities for systematic reviews. A full list of named participants is provided in the report. The top 10 themes, in order of priority, are water, sanitation and hygiene, disaster preparedness, disaster response, nutrition and food security, maternal and child health, coordination of humanitarian relief, quality of data/assessment tools/evaluation/impact, shelter, disaster recovery and mental health.


Settings: sudden-onset disasters; protracted humanitarian emergencies

This document presents the ACF research strategy for 2012–2015, which was compiled in 2011. It shows the main ACF axes for research and provides a list of activities needed to achieve or contribute to ACF’s overall strategy for 2010–15. These axes relate to nutrition and include specific strategic objectives and research questions, many of which would be amenable to impact evaluations. The over-arching axes are: (1) produce scientifically based operational tools and methods to improve ACF’s impact on
undernutrition; (2) identification of efficient, innovative and sustainable multi-sectoral approaches and tools to respond to disasters and build longer-term resilience to disasters; (3) produce scientific analysis on the global context in order to anticipate the main coming trends and challenges; and (4) contribute to stimulate ACF’s pre-eminence as an advocate and reference source on hunger and undernutrition.


Setting: sudden-onset disasters

This project focused on the identification of research priorities for medical rehabilitation of people who had suffered a spinal cord injury (SCI) in an earthquake. It was underpinned by a narrative literature review to identify epidemiological studies relating to the Kashmir earthquake in Pakistan (2005), the Sichuan earthquake in China (2008) and the Haiti earthquake (2010), with a follow-up review on spinal cord injury rehabilitation services provided by local and foreign providers in response to these earthquakes. This work revealed that post-disaster services were expanded by adapting local resources with international assistance to manage the significant numbers of SCI survivors but that research was limited. The authors conclude that a global disaster research agenda for spinal cord injury in earthquake settings where rehabilitation resources are scarce is needed to strengthen the evidence base for clinical management and therefore outcomes for people injured in this way. Among their suggestions for this research agenda that would be amenable to impact evaluations are response, rescue, extrication and transfer mechanisms for thoracic and thoracolumbar (rather than cervical) spinal cord injury, and treatments for clinical complications and mental health sequelae. They note that the next step will be an expansion of the review into a systematic review to identify additional research gaps and that ‘effective disaster setting data management and research collaborations of foreign and local SCI disability and rehabilitation stakeholders will be required for agenda implementation’.


Settings: sudden-onset disasters; protracted humanitarian emergencies

This evaluation was commissioned to assess the operational capacity of the European Commission’s Directorate-General for Humanitarian Aid and Civil Protection (DG-ECHO) to fund integrated food security and nutrition operations, with the aim of exploring whether DG ECHO-funded food assistance supports or hinders attention being paid to causes of acute undernutrition. It included a document review, interviews at DG-ECHO and at regional and country levels, analysis of food assistance projects and case studies in Bangladesh, Niger and South Sudan. It drew the conclusion that ‘generally there is a need for more operational research on making food assistance better suited to nutrition needs’ and highlighted three ‘important evidence gaps’ in infant and young child feeding interventions, the use of specialised foods, and blanket feeding.

Settings: sudden-onset disasters; protracted humanitarian emergencies

This report starts from the premise that the purpose of the AusAID research programme is ‘to improve the quality and effectiveness of Australian aid in developing countries’. The report notes: ‘AusAID’s research investments will be driven by the research requirements of our country and regional programs working with partner governments in line with the Comprehensive Aid Policy Framework. Research funding decisions will also be based on an assessment of poverty, national interest, capacity to make a difference and current scale and effectiveness’, with a particular emphasis on medical, agricultural and education. It also notes that ‘the Asia-Pacific region will remain the primary geographic focus of our research support, with some expansion into Africa and South Asia’. The research priorities were developed ‘through extensive consultation with AusAID program and thematic areas’ and will be revisited in 2014 or 2015 to ensure their continued relevance. The report presents the priorities under five themes: saving lives, promoting opportunities for all, sustainable economic development, effective governance, and humanitarian and disaster response, with specific items for each of these. The report also sets out how AusAID will fund, implement and use studies within its research programme.


Settings: sudden-onset disasters; protracted humanitarian emergencies

In December 2011, the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) hosted a conference on ‘Risk, Adaptation and Innovation in Humanitarian Action’ to identify priority policy research areas for the humanitarian community, and to strengthen partnerships with and among policy and research organisations. There were more than 50 participants from a global network of research institutes, universities, international NGOs and United Nations agencies. This report summarises the presentations in each of four sessions (‘Humanitarian action in a changing world’, ‘Adaptation and innovation in humanitarian action’, ‘Humanitarian action in protracted and violent conflict’ and ‘Effective humanitarian action’), and the conclusions that were drawn. It identifies four top priorities for humanitarian research: (1) evidence-driven humanitarian decision making (which would include research to understand how information can be used to ensure strategic decision making, particularly with respect to driving preventative action); (2) accountability and transparency (including research on better integration with development actors); (3) risk and agility (so that the humanitarian system can be more effective at managing risk); and (4) partnership, which was highlighted as being the most discussed issue but where the focus of the report is on the need for OCHA to build deeper and more strategic relationships with new actors in the humanitarian space, rather than the conduct of research into effective strategies to foster good partnerships.

Setting: sudden-onset disasters

This report comes from a one-day meeting in November 2012, which was hosted by the Institute of Medicine and the New York Academy of Medicine, following Hurricane Sandy. The objectives included the need to identify gaps in knowledge affecting disaster preparedness and response, and to develop a set of priorities for near-term research based on Hurricane Sandy and other recent disasters that may inform future disaster preparedness, response, and recovery plans. Participants included representatives of local and federal government agencies, healthcare providers, academia, first responders, community organisations, philanthropic organisations, and experts in disaster preparedness and response. A full list of participants is available in the report. The discussions during the meeting led to the prioritisation of more than a dozen specific research questions, which were grouped according to: healthcare institutions; community; health and response workforce; communications and the sharing of data to facilitate collaborations before, during and after emergencies.


Settings: sudden-onset disasters

This prioritisation exercise used a 3-round Delphi study with a panel of 26 people who had demonstrated their expertise through evidence of active research involvement in a literature analysis and evidence of current engagement with major incident education and training. Not all of the 26 participants contributed at each phase of the Delphi study. The first round asked participants to consider major incident research requirements broadly in 11 areas, and to identify areas where the research base was adequate, and additional areas in which research was required outside these 11 areas. Their replies were thematically reviewed and collated into a series of 221 statements, which were then rated in round 2. The 51 statements with a clear positive or negative consensus were removed after this round and the final round re-presented statements that had not reached consensus, leading to consensus being reached for a further 23 statements. In summary, the study identified 74 topics for research in the field of health service management of major incidents, which can be grouped into themes; and the statements and the themes are provided in the report. However, no prioritisation is provided, beyond this high-level identification of the 74 topics. The largest theme (10 topics) is education and training, followed by planning (9) and communications (8). Other themes include recovery, acute response, and prehospital care. The authors highlight that the panel did not identify any single topic that they agreed was well researched and understood.


Settings: sudden-onset disasters; protracted humanitarian emergencies
This conference abstract reports a study that used the Child Health and Nutrition Research Initiative (CHNRI) methodology to prioritise operational research gaps in reproductive health in crisis settings for the Inter-agency Working Group in Humanitarian Settings (IAWG). Researchers, public health and clinical practitioners working in the area compiled a list of 28 research gaps collected from prior IAWG meetings, working groups and consultations. These research gaps were reviewed by 68 researchers, public health and clinical practitioners selected from IAWG member agencies (academic, non-governmental and government), who submitted 66 additional gaps. The 94 research gaps were categorised into adolescent health, comprehensive abortion care (CAC), family planning, minimal initial service package (MISP) for reproductive health in crises, maternal and newborn health (MNH), gender-based violence (GBV), HIV/STI (sexually transmitted infections), and crosscutting issues. Sixteen members of the panel then rated the research gaps using preselected, defined criteria (need, feasibility, operationalisability, usefulness and relevance), in order to prioritise them. The highest priority gaps were in the following categories: MNH (7), family planning (7), cross-cutting (5), CAC (3), MISP (3), adolescent health (2) and GBV (n = 1). The authors conclude that, although MNH and family planning contained half of the top priority gaps, ‘there remains a need to refine these issues into specific operational research questions to better identify, measure, and improve reproductive health outcomes in crisis settings’.


Settings: sudden-onset disasters; protracted humanitarian emergencies

This project used a systematic search of the medical literature database PubMed to identify articles related to evidence on disasters in the developing world, which were then graded in accordance with the Centre for Evidence-Based Medicine levels of evidence. The authors extrapolated from the identified literature to suggest research gaps in this area. They report that the most common topics (28.2%) in the literature they identified were missions, healthcare provision and humanitarian aid during developing-world disasters, with commentaries about policies, vulnerable populations, and food, water, and nutrition being the next most commonly found in citations. They conclude that there are definite gaps in the themes covered by the literature, noting that tools, mental health, specific diseases, conflicts, ethics and epidemiology were addressed in a small proportion of articles; and they highlight that ‘mental health, in particular, lends itself to systematic research, using premeasures and comparison groups between exposed and the unexposed people’. They write that ‘future priority areas in research include long-term economic outcomes, health system recovery, occupational rehabilitation of victims, community-based disaster preparedness, resilience of communities in low-resource settings, public health interventions, monitoring and evaluation of interventions, and research tools validated for the developing world’, and conclude that ‘aid for sustaining long-term disaster research may be a more useful investment in mitigating future disasters than short-term humanitarian aid missions to the developing world’.

Settings: sudden-onset disasters; protracted humanitarian emergencies

This report describes the Mental Health and Psychosocial Support in Humanitarian Settings – Research Priority Setting (MH-SET) project which took place in 2009–2010. The project aimed to establish a consensus-based research agenda to support the prevention and treatment of mental disorders and the protection and promotion of psychosocial well-being in humanitarian settings. The authors describe it as the first systematic effort to set research priorities in this field, to their knowledge. A total of 136 advisory group members, and 114 participants in 9 focus group discussions in Peru, Uganda and Nepal, generated a set of 733 research questions. These were consolidated into a list of 74 research questions through qualitative data analysis and grouped into 4 categories, 2 of which would be most amenable to impact evaluations: mental health and psychosocial support interventions, and research and information management. An online survey rated the questions, leading to a top 10, of which 3 related to the effects of interventions relevant to mental health and psychosocial support and 2 related to issues that could be addressed in impact evaluations in research and information management. In a companion paper, the authors present their findings on the attitudes of the stakeholders towards relevance of the research questions and conclude that 'research needs to be more sensitive to questions and concerns arising from humanitarian interventions, and practitioners need to take research findings into account in designing interventions'.


Settings: sudden-onset disasters; protracted humanitarian emergencies

This report was submitted to the UNICEF WASH Steering Group in January 2009 and includes a brief review of existing evidence for WASH interventions and the identification of gaps in the existing evidence base, with proposals for research that would fill them, through consultation with key stakeholders from UN agencies, NGOs and academia. A list of the people consulted is provided in the report. The report identifies several specific topics for future research, which would be amenable to impact evaluations, and these are presented in sections, which reflect the structure used to present existing evidence, and are dedicated to: general issues; water supply; sanitation; hygiene promotion; social mobilisation; and cost-effectiveness.


\(^{100}\)Tol et al. 2012.
Settings: sudden-onset disasters; protracted humanitarian emergencies

This report arose from a request from the Centers for Disease Control and Prevention’s (CDC’s) Coordinating Office for Terrorism Preparedness and Emergency Response (COTPER) to the Institute of Medicine to convene an ad hoc committee to delineate a set of near-term research priorities for emergency preparedness and response in public health systems that are relevant to the specific expertise resident at schools of public health. An expert committee was formed and met during December 2007, in conjunction with a two-day public meeting and workshop. The committee made four recommendations for research priorities: (1) enhance the usefulness of training, through research to create best practices for the design and implementation of training (e.g. simulations, drills and exercises) and facilitate the translation of their results into improvements in public health preparedness; (2) improve communications in preparedness and response, through research to identify and develop communications in relation to preparedness and response that effectively exchange vital and accurate information in a timely manner with diverse audiences; (3) create and maintain sustainable preparedness and response systems, through research to identify the factors that affect a community’s ability to successfully respond to a crisis with public health consequences, and the systems and infrastructure needed to foster constructive responses in a sustainable manner; and (4) generate criteria and metrics to measure effectiveness and efficiency, through research to generate criteria for evaluating public health emergency preparedness, response and recovery, and metrics for measuring their efficiency and effectiveness.


Setting: sudden-onset disasters

This report arose from a breakout session at the 2006 Academic Emergency Medicine Consensus Conference on the science of the surge, and includes the statement that ‘research pertaining to surge capacity during extraordinary circumstances remains in its early stages’. There were 36 experts in disaster medicine and related fields involved in a broad discussion, and they used a structured nominal-group process to delineate 5 areas of research which they regarded as most critical from 14 potential areas of discovery identified by the group. These were: (1) defining criteria and methods for decision making regarding allocation of scarce resources; (2) determining effective triage protocols; (3) determining key decision makers for surge-capacity planning and means to evaluate response efficacy (e.g. incident command); (4) developing effective communication and information-sharing strategies (situational awareness) for public-health decision support; and (5) developing methods and evaluations for meeting workforce needs.


Setting: sudden-onset disasters

The aim of this report for the Ministry of Civil Defence and Emergency Management (MCDEM) Disaster Impact Assessment Project Team in New Zealand is to investigate and
document current best practice in New Zealand and internationally on disaster impact assessment methodologies and processes. It used a combination of four methods: review of information held by the MCEDE; semi-formal interviews with stakeholders to identify key issues and relevant documents for research at the beginning of the project; literature review of New Zealand and international documents; and sampling of three CDEM Group plans to identify the coverage given to disaster impact assessment processes. The report recommends the development of a disaster impact assessment framework for New Zealand, and highlights some areas in particular need of research, noting ‘a recurring theme when researching impact assessment was the strong focus that assessment had on response and recovery, and little to no guidance on reduction and readiness’.


**Setting: sudden-onset disasters**

These authors from the Regional Office for the Eastern Mediterranean at the World Health Organization discuss what they call the ‘research deficit’ in relation to the evidence needed to respond to a disaster in the Eastern Mediterranean region. They make recommendations about the need for research in general, including the need for changes in education and training, which would include, for example, ‘the incorporation of effective, action-oriented and user-driven approaches to research, learning and knowledge management methods into professional practices and encouraging interdisciplinary collaboration with the delivery of health care services’. They list several priority research areas for the Eastern Mediterranean region including some specific to health, the identification of barriers to the application of research results and best practices, communication and information sharing, and identification of ‘the essential elements of humanitarian response and performance indicators’.


**Setting: sudden-onset disasters**

The Hyogo Framework was adopted at the World Conference on Disaster Reduction in January 2005 in Kobe, Hyogo, Japan, following extensive consultation and meetings. One of the objectives was ‘to increase the reliability and availability of appropriate disaster-related information to the public and disaster management agencies in all regions’. The Conference adopted five priorities for action, including ‘use knowledge, innovation and education to build a culture of safety and resilience at all levels’. Within this, two key activities related to research were identified: (1) develop improved methods for predictive multi-risk assessments and socioeconomic cost–benefit analysis of risk reduction actions at all levels; incorporate these methods into decision-making processes at regional, national and local levels; and (2) strengthen the technical and scientific capacity to develop and apply methodologies, studies and models to assess
vulnerabilities to and the impact of geological, weather, water and climate-related hazards, including the improvement of regional monitoring capacities and assessments.


Settings: sudden-onset disasters; protracted humanitarian emergencies

This report followed a conference organised by WHO in October 1997, which focused on how to improve the health response to complex humanitarian emergencies, which were described as 'situations affecting large civilian populations which usually involve a combination of factors including war or civil strife, food shortages and population displacement, resulting in significant excess mortality'. There were 99 participants, representing WHO technical divisions, donor country missions, UN agencies, ICRC (International Committee of the Red Cross), IFRC, IOM (International Organization for Migration), NGOs and academic institutions; with a full list of named participants provided in the report. The methods used over the two-day conference included presentations and small group work, concluding with general discussion of the priorities. Among the objectives of the consultation was the need to 'develop research questions that are acceptable and valid to people whose dignity may have been compromised', with a focus on the early phase of the emergency, including, for the purposes of this document, applied health research to develop and deliver effective interventions. In total, the conference identified 18 priority topics for applied health research in complex emergencies, mostly related to the effects of interventions, and these were grouped under 6 themes: nutrition, reproductive health and women's health (including gender-based violence), communicable diseases (covering water supply, cholera, malaria and acute respiratory infections), mental health, health services management, and information management. The individual topics included specific suggestions for research that might now be considered impact evaluations. The report also includes the conclusions of the conference relating to the ethical aspects of research.
Appendix 11: Results from a card ranking exercise on priorities and actions

A workshop on ‘Evaluating the Impact of Humanitarian Interventions: Scope, Methods and Experiences’ was held in London, United Kingdom on 21 March 2014.

Participants were asked to place cards along a vertical and horizontal axis showing the following (items have been re-grouped under main headings to see patterns).

1. Things that are very important and very evident:

   In disasters
   - Standards for capacity building for disaster management.
   - Disaster affected people should be involved in designing recovery interventions.
   - Standards are required to improve relief and response.

   Education
   - The lack of access to early education fundamentally and irreversibly affects future earnings and social skills.

   Cash
   - The relative benefits of cash versus in-kind food aid under different conditions.

   Water, sanitation and hygiene (WASH)
   - There is a strong link between WASH programmes and health and nutrition outcomes.
   - Access to clean water is fundamental.
   - Links between WASH outcomes and health outcomes for children and women.
   - The immediate restoration of availability of clean drinking water in the aftermath of a natural disaster is a very significant (if not the most significant) life-saving measure.

   Gender-based violence
   - Gender based violence interventions are a life-saving intervention.

   Others
   - True: humanitarian action can never compensate for a lack of political action.
   - Use tranexamic acid to reduce problems of bleeding in trauma (CRASH-2).

2. Very important but not evident:

   Protection
   - What protects civilians in conflict zones?
   - Protection
   - Does ‘protection by presence’ really work?

   Gender-based violence
   - Impact of prevention of violence against women and girls programmes as a life-saving acute emergency programme.
Targeting
- How to target poor people in a way that works better than a lottery.

Evaluation
- Impact evaluation of response for management and disasters.
- Impact of livelihood programmes for women on raising household income.
- What forms of humanitarian assistance contribute to stability-related outcomes (e.g. violence)?
- How to deliver information on what works to people working in the field (which summary is most effective)?

Types of interventions
- Early childhood interventions.
- What humanitarian (or development) interventions contribute to social cohesion?
- Role of governance in resilience and sustainability.
- Hygiene promotion e.g. re: hand-washing practices as part of a community-backed approach reduce to reduce incidence of water borne diseases.

3. Moderately important, very evident
- Community involvement in decision-making for response and mitigation activities.
- Community-Based Management of Acute Malnutrition is the best way to tackle acute malnutrition.
- We do not have enough evidence. We stifle learning and innovation with a fear of failure.
- No more ‘bednets and malaria’.
- True: shelter, clean water, food distribution.
- Effects of ready-prepared supplementary foods in controlled contexts.

4. Moderately important but not evident:

Governance
- Which would be more cost effective in funding disaster response.
  (a) Donors funding disaster affected government directly.
  (b) Donors funding UN agencies.
  (c) Donors funding NGOs directly?
- Impact of different humanitarian leadership and coordination models.
- Impact of beneficiary accountability systems in improving early recovery.
- Whether third party monitoring works better than an agency monitoring itself (in terms of data quality, analysis quality and finding corruption).
- Impact of impact evaluation (including influencing factors and assumptions).
- Need to be tested: cooperation in humanitarian interventions.

Communication strategies
- Behavioural change programmes can drive down rates of stunting (chronic malnutrition).
- Incentives for learning.

Effectiveness
- Low funding for the Syrian crisis has forced NGOs and governments to work more effectively – a blessing in disguise, if lessons are learned from it.
- Supply chain cost efficiency.
- Impact of child friendly spaces on emotional wellbeing of children.
Cash versus in-kind
- Cash transfer programmes have long-term impacts on livelihoods.
- In many contexts it would be more effective to provide more people with less food or cash than the status quo.
- Tested: difference between cost-effectiveness of certain interventions, for example cash or food vouchers.
- Impact of cash versus non-food items.
- What works for prevention of malnutrition in fragile environments?
- What difference does the distribution of non-food items make to internally displaced persons?
- Sustaining impact versus cash/food interventions for better nutritional outcomes.
- What is the best way to support refugees going home? Packs of non-food items? Cash? Community support?

Disaster-risk reduction
- Models of District Disaster Response Plan and District Disaster Response Coordinators?
- What types of disaster-risk reduction programmes have demonstrable benefits?
- Linking relief and development in protracted refugee crises.

5. **Not important but very evident:**
- None

6. **Not important and not evident:**

Coordination
- Impact of humanitarian coordination on action.
- Impact of coordination.
- Ideas for impact evaluation: do the billions of dollars we invest in coordination make a difference for crisis affected people?
- Do current coordination approaches lead to more effective response?
- Inter-agency coordination in crises.

Others
- Understanding the epidemiology, effective package of services for newborn health in humanitarian settings.
- Feasibility and acceptability of long acting family planning services in humanitarian settings.
- Impact evaluations are expensive.
- Uses of (new) technologies in humanitarian action.
Most areas in the humanitarian sector suffer from a paucity of evidence. This scoping paper provides an independent analysis of the evidence base of evaluations in humanitarian assistance. It identifies key gaps and priorities in need of rigorous evidence to direct research to where it will be most valuable.

The study incorporated a wide array of methods to assess available evidence, including an online survey of participants, interviews with humanitarian sector experts, and extensive literature reviews of humanitarian studies and strategy documents of major humanitarian organisations. It also includes an evidence gap map that presents the results of a thorough search for completed, ongoing and planned impact evaluations of humanitarian interventions.

The scoping study presents conclusions about a general lack of a reliable and robust evidence base from studies assessing the causal relationship between a policy or intervention and outcomes or impact.
2.4.3 Implications of the 3ie Scoping Paper 1 for this Thesis

While the 3ie Scoping Paper 1 had a particular emphasis on evidence from impact evaluations in the humanitarian sector, it did identify three main issues that have implications for this thesis. These issues, discussed in more detail below, have been categorised under the headings of evidence, definitions and knowledge management.

2.4.3.1 Evidence

The 3ie Scoping Paper 1 confirmed that the evidence-base on which humanitarian interventions and responses are most effective is extremely lacking and frequently imperfect. This view is supported by other researchers (5, 96-99). The research results confirmed a lack of high-quality evidence in humanitarian assistance (3) and the urgent need to go beyond good intentions and find out what works. Currently, there are few impact evaluations, as defined by 3ie standards, to guide the delivery of humanitarian interventions. Evidence that is complete, robust and consistent is required to support decision making in the disaster setting.

2.4.3.2 Definitions

Clear definitions are urgently required in the disaster and evaluation sectors. Currently, there is confusion on what terms mean; greater clarity would provide better comparability and consistency of results when reviewing evaluation findings. For example, during the 3ie project, a search for ‘impact evaluations’ identified research with this in the title, but no further mention of impact evaluation or ‘causality or counterfactual’ in the document (3). This is an example of the use of the term in an incorrect context.

2.4.3.3 Knowledge Management

A lack of databases and repositories storing impact evaluations was identified by the 3ie Scoping Paper 1 (3). Access to evaluation studies could be used to compare evaluation findings, which, in turn, could inform policy and practice. Improved knowledge management is critical in the disaster setting as it provides data, information and knowledge about the event that can assist in all phases of the disaster timeline. Importantly, knowledge management can assist planning for, and responding to, disasters and complex emergencies, thereby potentially saving lives.
The primary aim of the 3ie Scoping Paper 1 was to help improve the effectiveness and efficiency of humanitarian action. This is also the underpinning philosophy driving the development and validation of a comprehensive framework for disaster evaluation typologies. As the implications show, there is a definitive need for the development of a comprehensive framework for disaster evaluations typologies, which has the potential to provide consistent disaster research and evaluation, and consistent terminology and standards for reporting across the different phases of a disaster (1). Such a framework is intended to facilitate agreement on identifying, structuring and relating the various evaluations found in the disaster setting with a view to better understand the process, outcomes and impacts of the effectiveness and efficiency of interventions.

The 3ie Scoping Project provided an opportunity to understand fully the current state of play, along with identifying the beneficiary nature of a comprehensive framework. To structure a robust research study, a further outcome of this project was the clear identification of the primary research aim and the four overarching research questions governing this thesis, as detailed below.

The findings from the 3ie project helped shape the research questions and methodologies required for the remainder of this research project.

### 2.5 Research Questions for this Thesis

The aim of this research is to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline.

#### 2.5.1 Overarching Research Questions

Existing frameworks that support consistent reporting in a disaster show a lack of validation, consistent terminology and standards for reporting across the disaster timeline (100-102). Initial research indicates that an opportunity exists for the development of a comprehensive framework to structure disaster evaluation typologies. To create such a framework to provide a structure to support disaster evaluation and research, the following four overarching research questions guide the research:

1. What does the peer-reviewed and grey literature report on comprehensive frameworks and evaluation typologies in the disaster setting?
2. What key elements would support a core structure to develop a comprehensive framework to represent evaluation typologies along the disaster timeline?

3. What comprehensive framework design could enable mapping these evaluation typologies along the disaster timeline?

4. What are the necessary steps to validate a comprehensive framework of evaluation typologies in the disaster setting?

2.5.2 Sub Questions

Each of these overarching research questions structure a series of related sub questions, as outlined in Figure 2.
RESEARCH AIM
To develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline.

SUB QUESTIONS
LITERATURE REVIEW (Chapter 4)
• What frameworks for evaluation and research exist in the disaster setting?
• What evaluation typologies exist in the disaster setting?

SUB QUESTIONS
PRE-ESTABLISHMENT OF THE CORE STRUCTURE (Chapter 5)
STUDY 1: Utstein Guidelines 2003
• How do professionals active in disaster evaluation and research view the Utstein Guidelines?
• How do professionals active in disaster evaluation and research use the Utstein Guidelines?
• What are the barriers and facilitators to professionals active in disaster evaluation and research in using the Utstein Guidelines?
STUDY 2: Reviews of Australian Disaster Events
• In what ways are the key elements of the Utstein Guidelines 2003, as amended, reflected in selected reviews of Australian disaster events?

SUB QUESTIONS
ESTABLISHING THE CORE STRUCTURE (Chapter 6)
• In what ways could Study 1: Utstein Guidelines 2003 and Study 2: Reviews of Australian Disaster Events, inform establishing a core structure for a comprehensive framework to represent evaluation typologies mapped along the disaster timeline?
• How can these key elements be mapped along the disaster timeline to support the development of a core structure?

SUB QUESTIONS
STUDY 3: DEVELOPMENT OF A COMPREHENSIVE FRAMEWORK FOR DISASTER EVALUATION TYPOLOGIES (Chapter 7)
• What process enables the development of a unifying, comprehensive framework for disaster evaluation typologies within the disaster setting?
• How can these evaluation typologies be mapped to support agreement on the identification, structure and relationships between various evaluation typologies within the disaster setting?

SUB QUESTIONS
STUDY 4: VALIDATION OF A COMPREHENSIVE FRAMEWORK FOR DISASTER EVALUATION TYPOLOGIES (Chapter 8)
• What does the peer-reviewed and grey literature report on validation processes suitable for use in validating CFDET 2017?
• In what way does the mapping between the evaluation typologies and the disaster timeline demonstrate suitability?
• What, if any, disaster evaluation typologies are missing in CFDET 2017?
• What value does CFDET 2017 have for the survey respondents’ work when undertaking evaluations in the disaster and humanitarian setting?
• In what ways would CFDET 2017 be useful for supporting and promoting evaluations in the disaster and humanitarian setting?
• In what ways would CFDET 2017 be useful for teaching evaluations in the disaster and humanitarian setting?
• What are the perceived barriers and enablers to undertaking evaluations in the disaster setting?

Figure 2: Research Aim and Questions
2.6 Summary

The background and context in this chapter provides the justification for the need to develop and validate a comprehensive framework for disaster evaluation typologies. Definitions proffered provide consistency throughout this thesis. One of the main issues encountered when reviewing work in this area was the lack of consistency across the literature; thus, the need for definitional clarity became paramount, the intent being to ensure a comprehensive and consistent understanding to promote better comparability and consistency with current reporting and future analysis.

This chapter introduced the 3ie Scoping Paper 1 (3), which supports the need for the development of a comprehensive framework and validates the premise that there is an urgent requirement to increase the evidence-base of what works, for whom and at what cost in the field of disasters and humanitarian assistance. Having identified specific gaps in current practice, the research questions underpinning the thesis were established, including the research aim – to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline.

To support addressing this research aim, Chapter 3 outlines the conceptual framework, research design and methodology that underpins this thesis.
Chapter 3: Conceptual Framework, Research Design and Methodology

3.1 Introduction

Chapter 3 outlines the research design, methodology and conceptual framework that underpins this thesis. Research methods used to address the research questions outlined in Chapter 2 are identified. The research has been undertaken in three phases. Phase 1 involves the establishment of a core structure to underpin the development of a comprehensive framework for disaster evaluations. The second phase relates to the development of a comprehensive framework for disaster evaluation typologies that meet international validation and acceptance. The final phase identifies the processes required to validate the comprehensive framework at an international level.

3.2 Conceptual Framework of this Thesis

The conceptual framework that underpins this thesis is evaluation. As previously defined in Section 2.3, evaluation has been defined as ‘a systematic investigation of the worth or merit of an object’ (82)(p.3). Evaluation involves ‘making judgements about the merit, worth, value, significance, credibility and utility’ (83)(p.185) of interventions that are delivered to a disaster-affected community. Significantly, evaluation provides a systematic method in which the effectiveness and efficiency of the interventions delivered can be measured, with the aim of contributing to continuous quality improvement and increased accountability (84, 85).

The research reported in this thesis will be measured against the following four core initial attributes of evaluation, as identified by Yarbrough (103): utility, feasibility, propriety and accuracy. Utility is defined as the usefulness of evaluations to stakeholders with the intent to increase the extent to which they find evaluation processes and products valuable in meeting their needs (103). For evaluations to be feasible, they must be efficient and cost effective (103). Propriety, including professional ethics, is adherence to the highest principles and ideals within the evaluation sector. This attribute addresses what is ‘proper, fair, legal, right, acceptable and just’ in the evaluation context (103). The final core attribute of evaluation revolves around the concept of accuracy, which is defined as ‘the
truthfulness, validity and dependability of evaluation findings and interpretations’ (103). This lends itself to a quality assurance or improvement lens.

Ethical standards exist to guide undertaking evaluations by providing a defensible set of guiding principles (103) as well as a code of professional conduct. Examples of guidelines promoting responsible and ethical conduct include, but are not limited to:


The development and validation processes undertaken during this research will be measured against the four core initial attributes mentioned previously, while ensuring the adoption of ethical standards and guidelines to support this work.

3.3 Research Design and Methodology

This is an applied thesis dealing with the role of evaluations in the disaster setting.

This thesis utilises mixed methods research (MMR) to investigate the research questions identified above. This method of research uses a combination of at least one qualitative and one quantitative component in a single research project or program. Mixed methods has experienced a tremendous increase in popularity in recent years and is believed to provide a better understanding of the research problem than either approach alone (107, 108). Utilising a mixed methods approach also supports current political currency that speaks to those who inform policy and practice (107). Combining the strengths of both approaches provides acceptance and legitimacy. One of the significant advantages and strengths of MMR is that it provides multiple ways to explore the research problem (108).

A commonly agreed on definition of MMR is the collection and analysis of data, the integration of the findings and resulting inferences that use both qualitative and quantitative approaches or methods in a single study or a program of study (108). MMR as a systematic approach to research was developed in the late 1970s, and hence, is relatively new (108, 109). Since that time, MMR has continued to grow in popularity as
researchers believe that some research problems might be best studied using both qualitative and quantitative data (108, 109).

The development and validation of a comprehensive framework to structure disaster evaluation typologies lends itself to a mixed methods approach. During the preliminary studies, that is, the pre-establishment and establishment of the core structure, the qualitative findings of the study of the Utstein Guidelines led to the study of Australian disaster events. The results from one study informed the next stage of the research. The development of the comprehensive framework utilised a thematic analysis of evaluation typologies found in the disaster setting. A mixed methods approach was used for validating the comprehensive framework because of its usefulness in addressing complex issues and the additional value of using both qualitative and quantitative methods of data collection (110). Collecting the qualitative and quantitative data concurrently allowed for the data to be collected and merged to form one interpretation of the data (108). This interpretation provided quantitative information about magnitude and frequency using a 5-point Likert scale, as well as individual participants’ qualitative information and the context in which they were commenting (108) on the comprehensive framework. Creswell describes this research design as the triangulation or concurrent mixed methods research design (108). The triangulation design in MMR is not the same as the use of the term triangulation in qualitative research, in which inquirers draw evidence from different source or different participants to develop a code or theme (108, 111). In MMR, triangulation implies that qualitative and quantitative data are merged by the researcher during the analysis stage (108).

For this thesis, the first two studies were researched using qualitative approaches. The findings were then used to inform the development of the core structure and subsequent comprehensive framework. An MMR approach utilising qualitative and quantitative methods was then used during the validation process to enrich the overall findings.

3.3.1 Establishment of the Core Structure

The literature review was structured to identify key frameworks in the disaster setting. It also helped determine the elements to be included in the establishment of a core structure. The core structure is the foundation that all other comprehensive framework elements are related to. Critical elements are included that support and underpin the development of the comprehensive framework for disaster evaluation typologies. Additionally, the necessary disaster timeline elements and important concepts on strengthening resilience,
incorporating globally significant frameworks that are an integral part of reducing disaster risk and losses (58, 59, 112-114), are demonstrated.

Two research approaches are used to establish the core structure: semi-structured interviews and a thematic analysis of Australian disaster events. Semi-structured interviews are a form of information gathering commonly used in research and applicable to many research questions (115, 116). This method provides qualitative data that generates insight and understanding into the research. Any intangible perceptions are better understood through qualitative data collection (117).

As with any interview process, the sample or target population must be identified. In defining the sample for this research, it was important to specify both inclusion and exclusion criteria (118, 119). Inclusion criteria specify an attribute or attributes that potential research participants must possess to qualify for the study, while exclusion criteria stipulate attributes that disqualify a potential research participant (120). In this research, the inclusion criteria involved participants who have experience in the fields of disasters and emergencies, the humanitarian sector and/or evaluation at a national and international level. While it can be difficult to define an ‘expert’ (121), prospective research participants were asked to self-select their chosen field as per the criteria above. Given that the field of disaster health and evaluation is an emerging field with a limited number of experienced practitioners, there was an expectation that the sample would be small (122).

By including research participants from multi-disciplinary or cross-sectoral fields (such as the three domains mentioned above), the plan was to obtain a heterogeneous sample. While these three groups have some similarities, they are distinct. The risk of using a heterogeneous sample is an extreme diversity of data that may lessen the likelihood of meaningful core themes or even sub-themes being found during analysis (120).

The sample size was decided based on theoretical and practical concerns as recommended by Robinson (120). The practical reality is that this study required a provisional decision on the sample size at the initial design stage to assist with planning the duration and required allocation of resources (120, 123). Having said that, the actual number of participants was anticipated to be 10-15 and was open to monitoring during the research (124). The actual number of prospective research participants, according to some researchers, is ‘relatively unimportant’ (123). What is important, however, is the potential for the participant to aid the researcher in developing new insight and depth into the area
of study. Once ‘saturation’ has been achieved, that is, when the research findings yield no genuinely new insights or the researcher is gathering data to the point of diminishing returns, Bowen suggests that interviews be ceased (125). Importantly, Guest and colleagues describe ‘saturation’ as the gold standard by which purposive sampling sizes are determined and noted that data saturation could be attained by as few as six interviews depending on the sample size of the population (123).

In selecting potential research participants, the sample strategy was to ensure a diversity of expertise and people with backgrounds in the areas of disaster medicine, disaster management, emergency management; humanitarian or development; and evaluation. Non-probability and purposive sampling were included as both sampling strategies were deemed appropriate due to the small target population (126, 127). Based on their a priori theoretical understanding of the topic, some participants were expected to have unique, different or important perspectives on the research. Their presence in the sample was important to allow for adequate coverage and collection of rich data (128, 129).

Prospective research participants were identified from the public domain, including from the literature review findings and those already known to the researcher. Once identified, prospective participants were contacted via email. This communication included an explanatory statement and consent form. A range of dates and times were provided and participants were required to confirm availability for a telephone or Skype conversation.

Prior to the semi-structured interviews, a theme list of questions was created using open-ended questions. Open-ended questions are an important research method that provide a variety of ideas, perceptions and thoughts from the participants, despite potential difficulty in quantifying the results (117). The theme list of questions was pilot tested prior to the commencement of the interviews.

Pilot testing, also known as pilot or feasibility studies (130), is an important and necessary step to help determine whether the questions are understandable and appropriate, if the flow of the questions is correct, if terminology is clearly defined and understood, and if there are any other unidentified issues. The instrument was tested on a convenience sample of 10 people with similar backgrounds to the prospective participants. Feedback from the pilot test was used to identify potential problems, provide an opportunity for improvement and clarify any ambiguities with wording or flow (115, 131, 132).
The theme list of questions is based on a particular subject; here, which key elements would need to be included in the development of a comprehensive framework for structuring disaster evaluation typologies. This improves consistency in the interview process, supports the collection of data in a scientific and standardised manner and adds credibility to the findings (115).

Ethics approval was required due to the involvement of human subjects in the research. This is an important part of honouring Monash University’s commitment to high ethical standards across all areas of research. Further, this ensured the research had a positive impact and supported the integrity of the researcher’s work (133).

Once approval had been obtained, the survey was run over a designated timeframe. A specific timeframe was allocated for practical reasons, including time and budget considerations. It should be noted, however, some flexibility was required due to the response rates and quality of the data received. Given the international nature of the prospective sample, interviews were conducted at a time of convenience for the participants. Interviews were recorded and transcribed with the consent of the participants. On conclusion of all interviews the data were analysed. Key themes including any relevant sub-themes or patterns within the themes were identified (115, 117).

The results from the semi-structured interviews informed the next part of the research – a thematic analysis of Australian disaster events. The sampled Australian disaster events were examined to identify whether the core structure elements identified in the previous study were evident in the chosen published reports and if any other elements not covered should be included.

Thematic analysis is a relevant qualitative research method that can be widely used across a range of epistemologies and research questions (134). Braun and Clarke describe thematic analysis as a method for identifying, analysing, organising and reporting themes found within a dataset (135). One of the limitations, however, is paucity of published articles to guide researchers on how to conduct a rigorous thematic analysis. To yield meaningful and useful results, the data needs to be collected in a precise and consistent manner and the analysis also needs to be conducted rigorously and methodologically (134).
To support the collection of precise and consistent data during the thematic analysis, a table was created that allowed mapping including inclusion or exclusion of the key elements identified in earlier research. The use of a table provides consistency, supports replicability and allows researchers to systematically review and analyse the results. This also addresses one of the criticisms of thematic analysis compared with other qualitative research methods; that is, the flexibility that thematic analysis allows and the essential tension between this flexibility, which can lead to inconsistency and a lack of coherence when developing themes from the research data (136).

The Australian disaster events were chosen based on a number of factors, including the significance, size, location and nature of the disaster or event, the impact and cost from a socioeconomic point of view, and the quality and accessibility of data, such as open access. Only events that occurred in Australia were included at this stage of the research, due to accessibility of reports and referencing events that ‘occurred in our own backyard’.

According to the guidelines published by Monash University Human Research Ethics Committee (MUHREC), ethics approval was not required as this research did not involve human subjects, animals or biohazards (133).

Results and findings from these two studies informed the establishment of a core structure that formed the foundation for the comprehensive framework for disaster evaluation typologies developed as part of this thesis.

### 3.3.2 Development of a Comprehensive Framework

The research design and methodology used to inform the development of a comprehensive framework for disaster evaluation typologies consisted of three methods: a scoping literature review, a thematic analysis of relevant evaluation typologies and an iterative process.

A scoping literature review of the peer-reviewed and grey literature identified current comprehensive frameworks and evaluation typologies or methodologies that exist in the disaster setting. Major electronic databases relevant to the health sector were used to conduct the review of the peer-reviewed literature. Additionally, the grey literature, such as generic search engines and industry-specific repositories, was searched to discover additional untapped resources that may not have been published in the peer-reviewed literature.
Scoping literature reviews, also known as scoping studies or scoping reviews, are a method used to comprehensively map evidence across a range of study designs (137). Grimshaw defines a scoping study as ‘an exploratory project that systematically maps the literature available on a topic, identifying key concepts, theories, sources of evidence and gaps in the research’ (138)(p.34). By their very nature, scoping literature reviews provide an appropriate avenue via which to search the available literature on evaluation typologies and identify any key themes and research foci.

A thematic analysis of evaluation typologies allowed for the determination of logical groupings, which were then mapped along the disaster timeline. This systematic compilation of disaster evaluation typologies, organised around the core structure, forms an integral part of the comprehensive framework.

An iterative process was used to gain feedback on the developed framework. Opportunities were taken to present and discuss this unique work at international conferences, national seminars and networking events. Feedback provided an excellent opportunity to improve the comprehensive framework prior to undertaking a more formal validation process.

3.3.3 Validation of a Comprehensive Framework

To validate the comprehensive framework at an international level, a validation process consisting of a scoping literature review and a mixed methods online survey hosted by a survey engine was used. Validation findings were reviewed and considered.

When considering the definition of ‘validation’ and how to undertake a ‘validation process’, a definition of validation was determined to provide clarity, consistency and direction. A specific literature review was conducted to identify articles that might inform the definition of validation and inform the validation process. The review investigated both peer-reviewed and grey literature.

The use of an online mixed methods questionnaire was considered appropriate as it provided a structured and consistent approach allowing for better quantification of results. Further, the mixed methods approach helped ‘triangulate’ and support the findings (117, 139). A questionnaire, including an explanatory statement and consent form, was designed seeking input on the newly developed comprehensive framework for disaster evaluation typologies.
The research questions, as outlined in Chapter 2 (refer to Sections 2.5.1 and 2.5.2), guided the questionnaire design. It was imperative that the questions were clearly written and well laid out, that the questions were not leading or confusing, and that they had been pilot tested prior to going ‘live’ with the survey (117).

One of the limitations of using a questionnaire to collect information is that they can be difficult to design (117). To address this issue, expert advice from Qualtrics staff was sought on the layout of the questionnaire and how to include the diagrams related to the comprehensive framework.

Participants were asked to use a 5-point Likert scale to identify their agreement or disagreement with a symmetric agree-disagree scale for a series of statements related to the elements and relationships in CFDET 2017 and provide responses to open-ended questions, with opportunity to comment on the disaster timeline elements and the different evaluation typologies presented. While the main intent of the survey was to gain quantifiable acceptance of the comprehensive framework elements and relationships, certain demographic information was collected to support secondary analysis.

Pilot testing was undertaken prior to the commencement of the questionnaire to test the survey platform as well as the questionnaire. The survey was hosted on the Qualtrics Insight Platform (140), which is supported by Monash University, and was run over a six-month period.

An important part of the survey process was to determine what would constitute acceptance of the comprehensive framework. As there is no single definition of consensus (121), criteria and the level used to signify validation of the comprehensive framework at an international level had to be determined. The criteria were set at number of responses equal to 30, a mean equal to 4.0 and 80% of participants ‘strongly agree’ or ‘somewhat agree’ on the Likert scale. A thematic analysis of the written comments was undertaken to create major and minor theme lists to aid the analysis. The results were reviewed and improvements made to the framework as required.

To obtain a diversity of experience and the proposed number of required responses, people with backgrounds in the areas of disaster medicine, disaster management, emergency management; humanitarian or development; and evaluation were approached via email to participate in the research. Non-probability and purposive sampling were included as both sampling strategies were deemed appropriate due to the small target
population (126, 127). The list of potential research participants was informed by the literature review process as well as from personal networks generated from attendance at academic and global events. A diversity of expertise was actively sought as it is believed to produce better quality decisions as opposed to a homogeneous one (121).

Data from completed surveys were collated and the findings compared with the acceptance criteria discussed above. Open-ended comments were reviewed and where appropriate adjustments made to the comprehensive framework to improve its structure, and ultimate acceptance and usability at the international level.

3.4 Summary

Chapter 3 detailed the conceptual framework, research design and methodology that underpins this thesis. It highlighted the three phases of the research: the establishment of a core structure that would underpin the development of a comprehensive framework for disaster evaluations, the development of a comprehensive framework for disaster evaluation typologies that would meet international validation and acceptance, and the processes required to validate the comprehensive framework at an international level. The different research methodologies utilised to support a robust and comprehensive research project were also identified, and it was shown that an evaluation lens provides the underlying conceptual framework to the entire thesis.

Having established the research foundations, the identification of the key elements to include in the core structure starts with a comprehensive literature review.
Chapter 4: Literature Review of Evaluation Frameworks and Typologies in the Disaster Setting

4.1 Introduction

An extensive literature review process undertaken as part of the 3ie Scoping Project identified repositories of humanitarian studies and strategy documents of major humanitarian organisations. An analysis template was created and used for the literature review to ensure consistency and replicability between the different repositories searched. This template was included in the appendix of 3ie Scoping Paper 1 (refer to pp.96–98 of this thesis or pp.72–74 of the 3ie Scoping Paper 1 (3)). The findings reveal that there is a lack of databases and repositories that index evaluations. Further, there is a general lack of high-quality evidence on humanitarian assistance, and especially, on evaluations that show a causal relationship between assistance and changes in targeted results (3). This continues to be an issue, as recognised by Smith et al. in their recent systematic review of disasters ‘Seven decades of disasters: A systematic review of the literature’ (5).

Following from the findings of 3ie Scoping Paper 1, a literature review was conducted to better understand the current structure of evaluations in the disaster setting. This chapter presents the comprehensive scoping literature review process to identify current evaluation frameworks and evaluation typologies in the disaster setting. The search strategy employed, databases searched and analysis of the results are presented. The findings from this chapter make a unique contribution to the knowledge base of evaluation frameworks and typologies in the disaster setting.

4.2 Literature Review: Search Strategy

Based on the definitions in Section 2.3, the comprehensive scoping literature review strategy was structured to support the two distinct themes of this research:

Theme 1: Evaluation frameworks in the disaster setting
Theme 2: Evaluation typologies in the disaster setting

The literature review was designed to provide background and understanding for the following research questions:
• What frameworks for evaluation and research exist in the disaster setting?
• What evaluation typologies or methodologies exist in the disaster setting?

A scoping literature review (141) was chosen as it allows for comprehensive mapping of the peer-reviewed and grey literature and involves additional processes such as seeking expert opinion and snowballing. Scoping reviews comprehensively map the literature in a specific area and involve multiple processes; in this case, a literature review and expert opinions were sought, including consultation with stakeholders (142). This method is particularly useful in researching complex topics (141-143) such as disasters (144-146) and undertaking evaluations (147, 148). As identified by Smith et al. in a study undertaken in 2009 (6) and updated in 2018 (5), articles related to disasters are found in more than 900 multi-disciplinary journals (5, 6), thereby making the search process challenging. To add to the complexity, there are multiple definitions for the terms ‘disaster’ (149-151) and ‘evaluation’ (152). To provide clarity and consistency, Chapter 2 identified the definitions for disaster and evaluation used in this thesis. It is with these definitions in mind that the search strategy for the peer-reviewed literature was designed and implemented.

The peer-reviewed literature was searched using major electronic databases relevant to the health sector, including CINAHL (EBSCO Information Services; Ipswich, Massachusetts, US), EMBASE (Elsevier; Amsterdam, Netherlands), Ovid MEDLINE (US National Library of Medicine, National Institutes of Health; Bethesda, Maryland US), PubMed/MEDLINE (US National Library of Medicine, National Institutes of Health; Bethesda, Maryland, US), Scopus (Elsevier; Amsterdam, Netherlands) and ScienceDirect (Elsevier; Amsterdam, Netherlands). Prior to any keyword searching, the following criteria were set:

• English language
• publication dates from 1960–2017 inclusive
• only full articles to be included in the results.

To support Theme 1 in the identification of the history and development of frameworks for evaluation in the disaster setting, the following key terms and Boolean operators were used ‘disaster’ OR ‘emergenc*’, AND ‘health’, AND ‘guideline’ OR ‘framework’ OR ‘model’ OR ‘repositor*', AND ‘evaluation’ (refer to Figure 3).
To support Theme 2 in the identification of evaluation typologies in the disaster setting, the following key terms and Boolean operators were used: ‘disaster’, AND ‘evaluation’, AND ‘typ*’ OR ‘method*’ and included typologies or methodologies rather than exemplars of specific evaluations (refer to Figure 4). Initial search terms included ‘disaster’ OR ‘emergency*’; however, the resulting combined potential records numbered in excess of 120,000. A sample of these records demonstrated a predominance of emergency department or hospital articles, which were excluded from this search strategy. Changing the search term to include ‘disaster’ only provided records of relevance.

In addition to collecting information from databases and search engines, the comprehensive scoping review was supplemented and enhanced by an expert opinion process. This process involved a convenience sample of international colleagues who commented on evaluation frameworks and evaluation typologies in the disaster setting. According to Levac, while there are no clear guidelines on how to consult with stakeholders, consultation does form an essential component of the scoping review methodology, providing an opportunity to gather additional references and insights beyond those found in more structured literature reviews (143). Seeking expert opinion increased the yield and efficiency of search efforts (153) and was deemed an important part of the research process.
Figure 3: Comprehensive Search Strategy – Disaster Evaluation Frameworks
Figure 4: Comprehensive Search Strategy – Disaster Evaluation Typologies
Further to the scoping literature review and seeking expert opinion, a technique called snowballing or secondary searching was also used. In cases where identified sources cited relevant references, the original references were also collected and, if appropriate, were included in the literature review results. This allowed for the identification of additional references by examining the bibliographies and reference lists from the most recent full-text publications and through scrutiny of the contents pages of highly relevant journals. Snowball methods can be powerful for identifying high-quality resources in obscure locations (153). Additionally, the published works of disaster experts such as Birnbaum ML, Burkle Jnr FM, Cuny FC, Fattah S, Stratton S and Sundnes KO and evaluation experts such as Quinn Patton M, Scriven M and Stufflebeam D were reviewed using author name searches in each of the databases to enhance the results.

All records initially retrieved as a result of the search strategy were reviewed by myself as the PhD candidate with a secondary review by the Primary Supervisor. Records were scanned by title and abstract. Potentially relevant records were then reviewed in full text to determine their suitability for inclusion in the scoping literature review results.

4.3 Literature Review: Key Search Results

The key search results of the comprehensive scoping literature review are presented in two general themes. Theme 1 relates to evaluation frameworks in the disaster setting. Initial findings are presented in this chapter and more specific detail provided in Chapters 5 (Pre-Establishment of the Core Structure) and 6 (Establishment of the Core Structure). Theme 2 relates to evaluation typologies in the disaster setting. Initial findings are presented in this chapter and more detail provided in Chapter 7 (Development of a Comprehensive Framework for Disaster Evaluation Typologies). The detailed search strategies and subsequent results are shown in Table 1 (Disaster Evaluation Frameworks) and Table 5 (Disaster Evaluation Typologies).

This chapter makes a unique contribution to knowledge, both from the generic literature review perspective and on evaluation frameworks and typologies in the disaster setting.

4.3.1 Theme 1: Evaluation Frameworks in the Disaster Setting

The scoping literature review yielded 31 potential records (refer to Table 1). Subsequent title and abstract review and removal of duplicate entries further refined the number of usable records to 15 (refer to Table 2). Full-text review of these articles identified that
they met the inclusion criteria for this research. Expert opinion added two records (refer to Table 3) and snowballing of the identified records provided a further 25 records of interest (refer to Table 4). Authors are listed in alphabetical order in Tables 2-4 inclusive to demonstrate multiple contributions made by certain authoritative figures in the disaster evaluation and research sector. For the purpose of this thesis, however, the significant frameworks are discussed in chronological order to better understand the historical development of disaster evaluation. A total of 42 records informed the results of this scoping literature review into evaluation frameworks in the disaster setting.

Several significant disaster evaluation frameworks, including preliminary works, were identified by the scoping literature review process. To provide structure for the following analysis, the noteworthy frameworks and bodies of work are chronologically structured around one pivotal framework and a summation of their importance in relation to disaster evaluation frameworks is provided.

The pivotal disaster evaluation framework identified was the ‘Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style’ (hereafter referred to as the Utstein Guidelines 2003 (17)). These guidelines were first published in 1996 (154), and then in 1999 as an executive summary (155). The final publication occurred in 2003 when Volume 1 was published, providing a structure for the conduct of research and evaluations in the disaster setting. It was anticipated that this volume would be the first of four published as supplements to Prehospital and Disaster Medicine (PDM) (17); however, with the passage of time, these did not eventuate. This work is significant as it was the first major international collaboration of disaster health experts to develop a uniform data reporting tool that could be used to aid in disaster medical response (156). The Utstein Guidelines 2003 were co-authored by TFQCDM, the World Association for Disaster and Emergency Medicine (WADEM) and the Nordic Society for Disaster Medicine (17).
### Table 1: Literature Review Results – Disaster Evaluation Frameworks

<table>
<thead>
<tr>
<th>General Search Criteria:</th>
<th>PEER REVIEWED</th>
<th>EMBASE (via Elsevier or Ovid)</th>
<th>Ovid MEDLINE</th>
<th>PubMed MEDLINE</th>
<th>Scopus</th>
<th>ScienceDirect @ Elsevier</th>
<th>GREY Google Scholar</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong> disaster or emergenc*</td>
<td>2,577</td>
<td>60,514</td>
<td>257,333</td>
<td>243,078</td>
<td>244,334</td>
<td>39</td>
<td>1,030</td>
</tr>
<tr>
<td><strong>B</strong> health</td>
<td>50,457</td>
<td>19,041</td>
<td>1,776,112</td>
<td>2,165,768</td>
<td>1,700,583</td>
<td>3,036,389</td>
<td>571,000</td>
</tr>
<tr>
<td>A and B</td>
<td>1,256</td>
<td>1,252</td>
<td>74,380</td>
<td>91,519</td>
<td>717,211</td>
<td>33</td>
<td>758</td>
</tr>
<tr>
<td><strong>C</strong> guideline or framework or model or repositor*</td>
<td>31,954</td>
<td>193,026</td>
<td>896,214</td>
<td>869,565</td>
<td>1,548,406</td>
<td>2</td>
<td>19,200</td>
</tr>
<tr>
<td>A and guideline</td>
<td>75</td>
<td>8,172</td>
<td>14,427</td>
<td>8,520</td>
<td>43,854</td>
<td>4</td>
<td>431</td>
</tr>
<tr>
<td>A and framework</td>
<td>108</td>
<td>2,585</td>
<td>13,712</td>
<td>3,150</td>
<td>14,573</td>
<td>9</td>
<td>507</td>
</tr>
<tr>
<td>A and model</td>
<td>400</td>
<td>48,059</td>
<td>184,268</td>
<td>15,374</td>
<td>357,470</td>
<td>26</td>
<td>680</td>
</tr>
<tr>
<td>A and repositor*</td>
<td>0</td>
<td>329</td>
<td>1,140</td>
<td>162</td>
<td>869</td>
<td>0</td>
<td>13,700</td>
</tr>
<tr>
<td>A and B and C</td>
<td>309</td>
<td>339</td>
<td>356,386</td>
<td>13,434</td>
<td>156,208</td>
<td>0</td>
<td>842</td>
</tr>
<tr>
<td><strong>D</strong> evaluation</td>
<td>39,123</td>
<td>136,353</td>
<td>889,009</td>
<td>742,438</td>
<td>877,267</td>
<td>6,059,533</td>
<td>585,000</td>
</tr>
<tr>
<td>A and D</td>
<td>511</td>
<td>6,491</td>
<td>24,183</td>
<td>24,690</td>
<td>251,960</td>
<td>26</td>
<td>681</td>
</tr>
<tr>
<td>A, B and D</td>
<td>267</td>
<td>147</td>
<td>8,535</td>
<td>11,358</td>
<td>1,186</td>
<td>21</td>
<td>562</td>
</tr>
<tr>
<td>A, B, C and D</td>
<td>69</td>
<td>1</td>
<td>373</td>
<td>2,180</td>
<td>944</td>
<td>0</td>
<td>15,500</td>
</tr>
<tr>
<td>Potential records (n = 31)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Usable papers (duplicates removed)</td>
<td>n = 15 (refer to Table 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expert opinion</td>
<td>n = 2 (refer to Table 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Snowballing</td>
<td>n = 25 (refer to Table 4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2: Usable Papers from the Literature Review of Disaster Evaluation Frameworks

<table>
<thead>
<tr>
<th>#</th>
<th>Author and Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Debacker et al, <em>Utstein-Style Template for Uniform Reporting of Acute Medical Response in Disasters</em>, 2012 (101)</td>
</tr>
<tr>
<td>9</td>
<td>Sundnes KO, <em>Preparedness Process</em>, 2014 (54)</td>
</tr>
<tr>
<td>10</td>
<td>Sundnes KO, <em>The Operational Processes</em>, 2014 (44)</td>
</tr>
<tr>
<td>11</td>
<td>Sundnes KO, <em>The Transactional Structure of Society: The Basic Societal Functions</em>, 2014 (40)</td>
</tr>
<tr>
<td>12</td>
<td>Sundnes KO, <em>Selection of Responses and Interventions</em>, 2014 (48)</td>
</tr>
<tr>
<td>13</td>
<td>Sundnes KO, <em>Coordination and Control</em>, 2014 (42)</td>
</tr>
</tbody>
</table>

### Table 3: Expert Opinion Results - Disaster Evaluation Frameworks

<table>
<thead>
<tr>
<th>#</th>
<th>Author and Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>KAMEDO, <em>KAMEDO Reports in English</em>, web accessed 2018 (157)</td>
</tr>
</tbody>
</table>
## Table 4: Snowballing Results – Disaster Evaluation Frameworks

<table>
<thead>
<tr>
<th>#</th>
<th>Author and Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Bradt and Aitken, <em>Disaster Medicine Reporting: The Need For New Guidelines and the CONFIDE Statement</em>, 2010 (100)</td>
</tr>
<tr>
<td>5</td>
<td>Bradt DA, <em>Evidence-Based Decision-Making (Part I): Origins and Evolution in Health Sciences</em>, 2009 (159)</td>
</tr>
<tr>
<td>6</td>
<td>Bradt DA, <em>Evidence-Based Decision-Making (Part 2): Applications in Disaster Relief Operations</em>, 2009 (160)</td>
</tr>
<tr>
<td>7</td>
<td>Cuny FC, <em>Introduction to Disaster Management: Lesson 3 - Natural Disaster Assistance and Relief Operations</em>, 1993 (161)</td>
</tr>
<tr>
<td>9</td>
<td>De Boer and Debacker, <em>Quantifying Medical Disaster Management</em>, 2006 (163)</td>
</tr>
<tr>
<td>10</td>
<td>de Boer and Debacker, <em>A More Rational Approach to Medical Disaster Management Applied Retrospectively to the Enschede Fireworks Disaster, 13 May 2000, 2003</em> (164)</td>
</tr>
<tr>
<td>11</td>
<td>Fattah et al, <em>Templates for Reporting Prehospital Major Incident Medical Management: Systematic Literature Review</em>, 2012 (20)</td>
</tr>
<tr>
<td>13</td>
<td>Fattah et al, <em>A Consensus Based Template for Reporting of Prehospital Major Incident Medical Management</em>, 2014 (22)</td>
</tr>
<tr>
<td>17</td>
<td>Leiba et al, <em>Disast-CIR: Disastrous Incidents Systematic Analysis Through Components, Interactions and Results: Application to a Large-Scale Train Accident, 2009</em> (102)</td>
</tr>
<tr>
<td>18</td>
<td>Lennquist S, <em>Protocol for Reports from Major Accidents and Disasters</em>, 2004 (168)</td>
</tr>
<tr>
<td>19</td>
<td>Powers and Daily, <em>International Disaster Nursing</em>, 2010 (19)</td>
</tr>
<tr>
<td>23</td>
<td>World Association for Disaster and Emergency Medicine, Task Force for Quality Control of Disaster Medicine, <em>Disaster Medical Response Research: A Template in the Utstein Style</em>, 1996 (154)</td>
</tr>
</tbody>
</table>
The Utstein Guidelines 2003 were a pivotal point in the evolution of disaster evaluation frameworks. Given the importance of this work, the Utstein Guidelines Volume 1, 2003 (17) is used as the central reference point to discuss evaluation frameworks in the disaster setting. Frameworks and significant bodies of work are chronologically mapped against the Utstein Guidelines 2003 (refer to Figure 5), and are classified as:

- ‘pre-Utstein’, that is, articles published before 2003
- ‘Utstein’, the Utstein Guidelines 2003 (17)
- ‘post-Utstein’, that is, articles published after 2003 and up to 2017.

Frameworks and other significant bodies of work that have guided disaster evaluation and research are discussed in more detail.
Figure 5: Disaster Evaluation Framework Timeline
4.3.1.1 Pre-Utstein

4.3.1.1.1 Katastrofmedicinska Organisationskommitten

In 1963, Katastrofmedicinska Organisationskommitten (KAMEDO) was established by the Swedish government under the National Board of Health and Welfare. The aim of KAMEDO was to analyse national and international disasters from a medical point of view, publish reports of the experience/s gained (173, 174), produce national guidelines and supervise standards in emergency and disaster medicine with the intent to standardise planning and preparedness in emergency situations (173, 175). The KAMEDO reports usually include studies on the structure of the response (that is, the rescue organisations and their activities during the disaster), as well as the medical, psychosocial and social aspects of the disaster (174). Reports were originally written in Swedish, however, there are 39 English summaries of the KAMEDO reports available. The last English report available, KAMEDO Report 98: The Disasters in Japan 2011, was published in 2016. The Swedish version was initially published in December 2013. This report was updated in May 2016 in English to include a number of extensive and important international assessments on how to improve the management of disasters in general and radiological disasters in particular (176). Notably, KAMEDO is the longest running database of disaster and emergency reports available at an international level and forms a strong basis for disaster medicine planning, education and research (177). In recognition for their international significance, the KAMEDO reports were awarded the WADEM Humanitarian Award for Excellence in Disaster Management in 2009 at the WADEM World Congress on Disaster and Emergency Medicine (178).

4.3.1.1.2 Professor Peter Safar

The next significant body of work was published by Professor Peter Safar, spanning from the 1970s through to the 1990s. Professor Safar was a consultant in disaster medicine for the US Army and a founding member of the Club of Mainz, one of the first significant groups in the field of international disaster medicine (179). He wrote more than 1,300 articles, books, book chapters and publications related to disaster medicine, critical care and resuscitation (180). WADEM honoured him by naming their premier award the ‘Peter Safar Award for Services to Prehospital and Disaster Medicine’ in recognition of his significant contributions over the years (178).
Notably, Professor Safar co-founded the World Association for Disaster and Emergency Medicine (WADEM) in 1976 and played an active role in promoting a global approach to disaster management (178). In the 1990s, he participated in a research team called the Pittsburgh Disaster Reanimatology Study Group (DRSG), consisting of notable experts including Dr Miroslav Klain (anaesthesiologist), Edmund Ricci PhD (evaluation research), Dr Ernesto A Pretto, Jr. (anaesthesiologist), Joel Abrams PhD (engineering) and Louise Comfort PhD (social science) (181). This research team, in partnership with a team of Russian and Armenian physicians, conducted the first international interdisciplinary disaster evaluation research field survey study of the earthquake in Armenia (182-184). The Armenia study led to a series of post-disaster field studies by the DRSG related to an earthquake in Costa Rica in 1991 (185), an earthquake in Turkey in 1992 (186) and an earthquake in Japan in 1994, which helped to establish the ‘Golden 24 Hours’ of emergency response in disasters and inspired Norwegian anaesthesiologist and humanitarian Dr Knut Ole Sundnes to establish, in 1995, the TFQCDM, under the auspices of the Nordic Society of Disaster Medicine and WADEM. Importantly, these three are the key organisations responsible for the creation and publication of the Utstein Guidelines 2003 (17).

4.3.1.1.3 Emeritus Professor Edmund Ricci

The 1980s and 1990s were significant due to the contributions of several disaster and emergency experts. Emeritus Professor Ricci is an expert in the technical areas of evaluation and survey research methods at the University of Pittsburgh. He designed over 300 evaluation studies and has longstanding involvement in healthcare and evaluation. Professor Ricci is one of the earliest academics to have taught courses specific to disasters and evaluations, such as the course ‘Evaluation Methods Applied to Disasters’ at the University of Pittsburgh (187). He was involved in the 2003 version of the Utstein Guidelines (17) and continues to serve on the editorial board of PDM (188).

As early as 1985, Ricci called for scientific research to be directed towards describing, understanding and evaluating disaster management as a prerequisite for quality improvement, arguing that while disaster studies existed, they lacked validity and reliability. These disaster studies would not be useful as comprehensive evaluations in the disaster setting due to their inconsistent approach and definitions. Additionally, comparative analysis was difficult because of the wide variations in data collection, data quality, data presentation and analysis. Ricci set out an approach to the evaluation of
disaster management that offered a scientific rigour anticipated to offer meaningful comparative analysis. Suggestions offered by Ricci in 1985 on how to undertake an evaluation included:

- disaster evaluations collecting consistent data
- utilising an interdisciplinary team approach including representation of persons with backgrounds in medical care, healthcare research and research methods
- the systematic collection of data to assess the quality of medical services or interventions provided to the community
- clearly stated data analysis procedures to support conclusions, allowing for comparability of results
- processes in place to address data validity and reliability concerns
- creating data banks that could host comprehensive evaluation studies. He believed that this was required to promote and allow for the comparative analysis of disaster relief activities. Such a data bank, or evaluation repository, would enable the analysis (that is, the efficiency and effectiveness) of disaster interventions from a quality improvement focus, be useful in teaching and continue to assist in the refinement and standardisation of evaluation methods in the disaster setting (189).

In the late 1980s and early 1990s, as previously mentioned, Ricci was involved in the DRSG. His motivation was to reduce avoidable morbidity and mortality in the wake of major disasters by adapting scientific evaluation concepts and methods used in other areas of healthcare at the time. Ricci suggested that the evaluation process should focus on the varying activities and processes undertaken during a disaster, including:

- structure – considering what facilities, equipment and personnel were available and how the organisation and management were structured during the response phase
- process – the interventions provided to the affected community and to the various processes, such as communication and decision making
- outcome – the degree to which avoidable morbidity and mortality were prevented
- the cost of the disaster response – including the cost of the relief effort
• adequacy of the disaster response – this included the number of disaster victims who potentially could have been saved and the extent to which the relief effort operated to its potential (190).

By evaluating the interventions and associated processes, it was believed that improved approaches to future disasters could be implemented (190).

4.3.1.1.4 Frederick Cuny

Frederick Cuny was a civil engineer who became involved in international disaster relief (191). He set out to reform the ‘flawed system’ of disaster relief from within. Cuny has been described as being one of the most dynamic individuals in disaster relief and humanitarian intervention (192, 193). Notably, Cuny leaves a legacy as one of the world’s first humanitarian engineers with a strong interest in public health (193). Cuny was ahead of his time and was the first to write a series of articles about disaster management in which he discussed the usefulness of frameworks. He also promoted the use of monitoring, evaluation and lessons learned in the disaster setting in an effort to analyse and improve disaster response (194) until his disappearance and presumed murder in war-torn Chechyna in 1995 (191). Cuny was the primary author of many of the following articles and continuing education series:

• Aim and Scope of Disaster Management: A Self Course of Instruction 1992 (195)
• Introduction to Disaster Management, 1992–1993 (194, 196-198)
• Principles of Disaster Management 1998–2001. This particular work was prepared by the Disaster Management Center at the University of Wisconsin-Madison in memory of Cuny after his disappearance in Chechyna in 1995. This work was led by Professor Birnbaum, who later went on to become a co-editor of the Utstein Guidelines 2003 (17). Joint funding was provided by the US Office of Foreign Disaster Assistance (OFDA) and the US Agency for International Development (USAID) for this work (199-212).

4.3.1.1.5 Professor Emeritus Jan de Boer

Also during the 1980s, Jan de Boer, a Professor Emeritus of Disaster Medicine at the Free University of Amsterdam, made significant contributions to disaster medicine, which was a relatively young field at the time. He published several books and many articles based on disaster medicine over a 20-year period (151, 164, 213-215). His pioneering work in
disaster health, disaster medicine and evaluation was significant at an international level. The significance of his work was in the identification that the research of disasters was hampered by a lack of uniformity and standardisation in describing events. Professor de Boer was one of the first to discuss classification and scoring systems in the disaster setting. In 1990, he published the ‘Definition and Classification of Disasters: Introduction of a Disaster Severity Scale’ (DSS); it was anticipated that this system could be utilised as a tool for evaluating the majority of disasters. The DSS included a classification, grading and scoring of the following seven parameters:

- the effect on the surrounding community and infrastructure
- the cause of the disaster
- the duration of the cause of the disaster
- the radius of the disaster area
- the number of casualties
- the nature of the injuries sustained by the living victims
- the time required by the rescue organisations for initiation of primary treatment, organisation of transport and evacuation of the injured (151).

Points were allocated per item; the higher the score, the more devastating the impact of the disaster (214). It is worth noting that Professor de Boer’s work is referenced widely (216-218); however, the DSS has only been used a few times as a framework for measuring the impact of a disaster, with mixed results (219-222).

The driving force behind the work of Professor de Boer was the desire to bring standardisation, structure and efficiency into the disaster setting, with the aim of reducing mortality, morbidity and socioeconomic costs to disaster-affected communities. He was one of the first to mention the importance of evaluating the health interventions delivered in response to disasters (164). Additionally, he suggested the establishment of a central database containing comprehensive evaluation studies, and that such a database would enable the sharing of information, including medical disaster plans, comparison of results and sharing of lessons learned. Professor de Boer strongly supported medical education based on scientific evidence and the comparative analysis of disasters (164, 223).
4.3.1.1.6 Dr Marcus Villareal

The 1990s saw significant growth in publications related to disaster medicine, including the work of Dr Villareal. Dr Villareal is an emergency medicine physician who in 1997 created a quality management tool for mass casualty emergency responses and disasters (170). Notably, he was one of the first to use the perspective of quality assurance and quality improvement when considering the emergency health response in the event of a disaster. Villarreal was keen to see efficient and timely data acquisition, and analyses to measure documentation and improve clinical performance. This could be achieved by identifying patterns and trends through monitoring and evaluation, and then using that information to improve processes. Villarreal outlined four main standards and indicators of performance in his quality management tool:

- formatting disaster medical records as data collection instruments
- developing appropriate tools that are easy to use for rapid assessments
- developing a mechanism for determination of causes of injuries
- developing methods to:
  - track patients
  - document response and recovery
  - document the circumstances associated with the event (170).

Unfortunately, Villarreal’s work has been referenced, but not used as a quality management tool in emergency or disaster responses (224, 225).

4.3.1.1.7 Emeritus Professor Marvin Birnbaum

Also emerging in the 1990s was the work undertaken by Professor Birnbaum, an Emeritus Professor of Medicine and Physiology with the University of Wisconsin. During his medical career, he specialised in critical care medicine and had a long involvement in disaster medicine at an international level. He has held the position of President of WADEM (226), was Editor-in-Chief of Prehospital and Disaster Medicine (PDM) and was co-editor of the Utstein Guidelines 2003 (17). His contribution as co-editor of the Utstein Guidelines 2003 and his ongoing work in the area of research and evaluation of the health aspects of disasters are discussed in more detail in both the Utstein and post-Utstein sections, along with a collaborative author and co-editor Sundnes (17).
4.3.1.1.8 Professor SWA Gunn

Further work in the 1990s related to definitions in the disaster setting were published by Professor Gunn, a Canadian surgeon and senior international health scientist involved in disaster management and humanitarian medicine. He is regarded as an expert in the field of disaster medicine, having served as long-time Director of the World Health Organization (WHO) Emergency Humanitarian Operations. Among his many achievements, Professor Gunn was a founding members of WADEM and a co-author of the Utstein Guidelines 2003 (17). Notably, he was the Chief of WHO’s scientific terminology section and spent a considerable amount of time defining and compiling lexical terminology for disasters and disaster medicine. Gunn argued that a common language and standardised terminology was needed to promote effective communication in the disaster setting. His significant contribution led to the creation of documents, tools and dictionaries, including multilingual versions (165), that helped promote clarity and consistency in the disaster setting (227).

4.3.1.1.9 Professor Michel Debacker

The final significant contribution in the 1990s related to definitions in the disaster setting was published by Professor Debacker, a specialist in internal medicine. He has held numerous positions linked to disaster medicine and published widely. In 1999, he co-authored a ‘Glossary of New Concepts in Disaster Medicine: A Supplement to Gunn’s Multilingual Dictionary of Disaster Medicine’ (162). Notably, he believed that the development of disaster medicine as a science and a medical specialty in its own right was dependent on clear definitions and terminology with the intent to promote effective and scientific communication (162, 228). This pre-Utstein work published by Debacker was predominantly based around disaster definitions and terminology; while extremely important, his later work published in 2011 (229) and 2012 (101, 224) discussing uniform data reporting of acute medical responses in disasters is discussed in the post-Utstein section.

These pre-Utstein papers were essentially descriptive in nature, as outlined above, and did not provide frameworks in a graphical format.
4.3.1.2 Utstein

As previously mentioned, the pivotal disaster evaluation framework identified is the Utstein Guidelines 2003 (17). These guidelines were first published in 1996 (154), and then in 1999 as an executive summary (155). The final publication was launched in 2003, at the Melbourne WADEM World Congress, when Volume 1 was published, providing a structure for the conduct of research and evaluation in the disaster setting. It was anticipated that this volume would be the first of four published as supplements to PDM (17). The significance of this work is that it was the first major international collaboration of disaster health experts to develop a uniform data reporting tool that could be used to aid in disaster medical response (156) and was thought of as the ‘gold standard’ for disaster evaluation and research. The Utstein Guidelines 2003 were the result of more than seven years of discussion and collaboration at an international level (17).

The term ‘Utstein’ has been used since the early 1990s and has become synonymous with ‘guidelines for research of different aspects of emergency medicine’ (36). The first ‘Utstein’ conference was held at Utstein Abbey, Norway’s best preserved medieval monastery, and resulted in the publication of guidelines for reporting of data from out-of-hospital cardiac arrest (230, 231). There have been subsequent specialised medical ‘Utstein style’ guidelines published, including guidelines on drowning (232), paediatric advanced life support (233) and major trauma (234, 235). What these guidelines share is the use of a consensus group of international ‘experts’ establishing uniform reporting of data with the intent to improve patient care (232) and, therefore, patient outcomes.

Initial discussions around the Utstein Guidelines 2003 commenced in 1994 (236) and were published in 1996 as a position paper in PDM by the TFQCDM and WADEM (154). TFQCDM consists of notable international disaster health experts, such as Birnbaum, Gunn, Pretto, Ricci, Sundnes and Ukai. Work on the guidelines continued for a further seven years; during this period, the Utstein Guidelines 2003 were published in peer-reviewed journals (17, 155), taught at universities and presented at international conferences (236-240). The overall aim of the development of the Utstein Guidelines 2003 was to provide sufficient structure and consistency to evaluation and research of health interventions delivered in the disaster setting to be able to maximise the efficiency and effectiveness of future responses (236, 241).

The Utstein Guidelines 2003 provided a conceptual framework in a graphical format for undertaking research and evaluation in the disaster setting. Diagrammatic representation
of the key elements in a vertical format has been reproduced from page 35 of the 2003 Guidelines (17) in Figure 6. The core of this conceptual framework has been frequently referenced in peer-reviewed papers, scholarly journals (242-246) and higher degree research theses (218, 247, 248), but rarely used as the methodological framework for undertaking disaster evaluations and research (7).

The literature review revealed one journal article (166) and two books (171, 172) that utilised the ‘conceptual framework’ and terminology of the Utstein Guidelines 2003. The three articles/books were based on the Sumatra-Andaman earthquake and subsequent Asian tsunami that occurred in December 2004.

The Utstein Guidelines 2003 were an attempt to provide a common nomenclature and uniform reporting of data that would allow comparability of findings. Importantly, these findings could influence policy and practice and it was believed that they would allow better clinical decision making (17). One of the major challenges facing the disaster setting was to overcome the limitations created by a lack of clear standards and definitions, leading to inconsistent reliability and poor operability of different data compilation initiatives (150).
Figure 6: Key Elements as Identified in the Utstein Guidelines 2003

4.3.1.3 Post Utstein

The post-Utstein era uncovered an increased interest in disaster evaluation and research as evidenced by the following authors and confirmed by Smith and colleagues in their 2018 publication (237). Significant contributions relevant to this thesis include the work...
of Bradt et al. (100, 158-160), Kulling et al. (167), Powers and Daily (19), Stephenson (18), Debacker (101), Fattah et al. (20-22, 24), Birnbaum (25-34) and Sundnes (35-57).

The following authors, selected based on their impact on the research undertaken in this thesis, are discussed in greater detail below: Birnbaum, Sundnes, Kulling, Debacker, Powers and Daily, Stephenson, Bradt and Fattah.

4.3.1.3.1 Emeritus Professor Marvin Birnbaum

Professor Marvin Birnbaum is an Emeritus Professor of Medicine and Physiology at the University of Wisconsin. He has held the positions of President and Chairman of the Board of Directors of WADEM and Editor-in-Chief of PDM, the official journal of WADEM (249), for 25 years. Professor Birnbaum is co-editor of the Utstein Guidelines 2003 and has published an evaluation on the medical aspects of the 2004 Indian Ocean earthquake and tsunami for the Southeast Asia Regional Office of the WHO (171, 172), utilising the Utstein Guidelines 2003 as a framework to structure the research.

In 2017, the American Academy of Disaster Medicine (AADM) awarded Professor Birnbaum its Distinguished Service Award for his vast contributions to the discipline of disaster medicine. According to AADM President Arthur Cooper, MD, MS, FAADM, ‘No one has done more to advance the science and practice of disaster medicine over the past dozen or more years than Marvin Birnbaum’ (250).

His current primary interest is in establishing the science of disaster health, and, together with colleagues, he has recently published Frameworks for ‘Research and Evaluation of the Health Aspects of Disasters’, Parts I–IX in PDM (26-34). All nine papers relate to research and evaluations of the health aspects of disasters with an emphasis on the creation of a Disaster Logic Model (DLM). Five frameworks have been developed to structure information and research of the health aspects of disasters (26).

Significantly, Professor Birnbaum has published more than 200 papers in the peer-reviewed literature. In particular, his publications between 2014 and 2016 related to disaster evaluation frameworks have had significant influence informing the development of the core structure that is discussed in Chapter 6.
4.3.1.3.2 Dr Knut Ole Sundnes

Dr Knut Ole Sundnes is a Norwegian physician and humanitarian who specialises in anaesthesiology and intensive care. He has been working in disaster and emergency health for over 40 years (251). On behalf of the Norwegian government, Sundnes was involved in numerous humanitarian disaster-related deployments, such as acting as the Head of Civilian Military Capacity in Afghanistan, which focused on building anaesthesia capability for civilian hospitals in Afghanistan (251). He has consulted on a variety of humanitarian publications, such as the WHO publication on ‘Emergency Preparedness and Risk Management’ (252). For his work for the international community, Sundnes has been the recipient of a variety of international medals acknowledging his commitment to improving disaster health (United Nation (UN) Interim Force Medal 1979 and United Nations Peace Forces Medal in 1995) (251). He has held a variety of important roles, such as Chairman of the TFQCDM (17), President of WADEM, editorial board of WADEM, President of Nordic Society for Disaster Medicine (17) and Head of the Office of War Surgery and Emergency Medicine and Head of Anaesthetics Services for the Norwegian Defence Forces (251).

Sundnes’ work is significant not only to this research, but the disaster and humanitarian sectors from a health perspective. In 1995, inspired by the work of Peter Safar, he established the TFQCDM under the auspices of the Nordic Society of Disaster Medicine and WADEM (181). TFQCDM subsequently became the ‘lead’ author in the publication of the Utstein Guidelines 2003 and Sundnes and Birnbaum the co-editors (17).

In 2014, Sundnes authored the ‘Health Disaster Management: Guidelines for Evaluation and Research in the ‘Utstein Style’ Structural Framework, Operational Framework and Preparedness’. This work was published as a supplement to the Scandinavian Journal of Public Health, presented 10 years after the initial publication of the Utstein Guidelines 2003 (36). It is worth noting that both original co-editors of the Utstein Guidelines 2003 published separate updates on this earlier work (25-57). Sundnes continued with his work, inspired by the belief that the need to learn and the potential to learn from disasters was vital, as more people were continuing to be exposed to disasters and increasing vulnerability became an issue (36).

The earlier 2003 work still served as a platform for understanding the deconstruction of society, and all disaster phases and the processes embedded within disaster management
Sundnes further expanded on these concepts, over 19 chapters comprised of three main parts:

1. The Structural Framework, which described the societal components generically and addressed the structures at risk of being damaged and the structures available to respond (such as relief and recovery).
2. The Operational Framework, which addressed the activities launched, including actions or interventions, to address the damage.

Sundnes deliberately chose to conclude with the beginning of the disaster timeline, as this demonstrated how society continues to move in ‘full circles’ when it deals with disasters. From a quality improvement lens, he believed that society should constantly improve its management of disasters to be better prepared for the next event (36).

Additionally, Sundnes actively promoted a common language for disasters (254). He believed that it was ‘clear that the language of disasters (or more precisely, the lack thereof) constitutes more of a barrier to research than a facilitator’ (36)(p.5). Without a ‘common language’ in the disaster setting, science would not be able to progress and the external validity of research would be practically impossible to achieve (36).

The guidelines updated by Sundnes in 2014 demonstrate an evolution in disaster evaluation frameworks. The motivation behind his work was to provide the tools necessary to enable researchers to assist in the understanding of health disasters to improve the effectiveness, efficiency, benefits/outcomes and costs of responses to health aspects of disasters and the enhancement of preparedness. This goal would be accomplished by establishing structures that facilitate the development of the science of health disasters through evaluation and research. Without structured and objective evaluation and research of the interventions and the measures to prevent or mitigate the effects of events that result in disasters, it would not be possible to learn from experience, and therefore, develop strategies to improve future responses (37) and build the science of disaster health (255). Without a uniform structure for disaster evaluation and research, important information can be lost, errors and inefficiencies can be perpetuated, and in many settings, vulnerability can continue (37).
4.3.1.3.3 Dr Per Kulling

Dr Per Kulling (1944–2013) was a physician qualified in anaesthesiology and intensive care medicine who worked for the Swedish Board of Health and Welfare as the Director of National Planning (Socialstyrelsen) (256). As an expert in the disaster and emergency medical field, he was involved in numerous publications related to disaster and emergency health (173, 257, 258) (in particular, the KAMEDO publications) and consulted with the European Commission, Health Consumers Directorate General (259-261). In 2011, he was presented with the WADEM President’s Award in Beijing for his many years of service to WADEM as a member of the Board of Directors and Chief Financial Officer (262). Kulling also worked as a seconded national expert in the Health Threats Unit of the European Union’s (EU) Directorate General for Health and Consumers (DG SANCO) in Luxembourg (261, 263).

In 2010, Dr Per Kulling and co-authors Birnbaum, Murray and Rockenschaub developed the ‘Guidelines for Reports on Health Crises and Critical Health Events’ (167, 264). This work was based on over three decades of experience with the Swedish Board of Health and Welfare in reporting on the acute health crises as per the Swedish KAMEDO project. These new guidelines refined the original work published in the Utstein Guidelines as written in 2003 and co-edited by Sundnes and Birnbaum (17). The intent was to provide specific guidelines for structured reporting of acute health events (156). By developing a common structure for reports on health, the experiences gained during a disaster response can be captured, promoting a standardised methodology for sharing results and lessons. If future reporting followed common standards, the findings would be comparable and could be used to learn and apply lessons learned within the disaster setting (264).

The development for the common reporting on health crises and critical health events was derived mainly from the following bodies of work:

- Health Disaster Management: Guidelines for Evaluation and Research in the Utstein Style
- the Swedish Disaster Medicine Study Organisation (KAMEDO)
- the Swedish Emergency Management Agency network for observers
- the WHO Regional Office for Europe (supported by the EU Health Programme) ‘Support Health Security, Preparedness Planning and Crisis Management in the
EU, EU accession and neighbouring (ENP) countries’, including expert consultations (264)(p.s16).

The guidelines suggested by Kulling et al. included the following headlines:

- Title
- Preface
- Authors
- Executive Summary
- Introduction / Material / Methodology
- Pre-Event Status (Background, Preparedness, Hazards(s) involved, Risks, Vulnerability, Resilience)
- Responses (Relief Responses, Recovery Responses)
- Development
- Discussion
- Lessons Identified and Actions Recommended
- Conclusions
- References
- Appendices
- Keywords
- Index
- Abbreviations (167, 264).

The significance of Kulling’s work is that he proposed guidelines that promoted the use of a common structure for the reporting of health crises and critical health events. These guidelines could be considered a process evaluation, and required reports to include the assessments of the pre-event status of the community, the relief and recovery responses to the events and identification of lessons learned (post-event) to inform disaster preparedness for future events (167, 264, 265). A report by Shabir (265) on the Pakistan floods of 2010 was the only report found that used the Kulling guidelines to underpin the reporting of the health impacts and response to the disaster event.

4.3.1.3.4 Professor Michel Debacker

Professor Michel Debacker is a specialist in internal medicine with additional training in intensive care and emergency medicine. He is a lecturer in disaster medicine at the Free
University Brussels and the University of Leuven. He serves as the Chairman of the Executive Committee of the European Master in Disaster Medicine and has served as the Assistant Chief of Staff of Medical Techniques and National and International Relations of the Belgian Armed Forces (266).

Prior to the publication of the Utstein Guidelines 2003, Debacker was involved in several publications related to disaster medicine (164, 267), including work with de Boer outlining disaster management definitions (162). In 2012, Debacker and colleagues published the ‘Utstein-Style Template for Uniform Data Reporting of Acute Medical Response in Disasters’ (101). Although Debacker named his paper ‘in the Utstein manner’, this was a separate enterprise to the main Utstein Guidelines 2003. The template described pre-event, event, medical response and outcome variables and their indicators relevant for evaluation and research on disaster medical response. Standardised and agreed definitions were also included in the template. The aim of this uniform reporting template was to gather empirical data on disaster medical response management. By collecting this information, it was anticipated that robust databases could be established that would allow disaster medical response investigators and researchers to collect evidence to influence response outcomes and inform policy and practice (101, 229).

According to Stratton, this work represented an important advance in disaster medical research. The document was generated by the Emergency Management and Disaster Medicine Academy (EMDM Academy) using the ‘Utstein’ style for developing medical consensus statements. The template provides the disaster and acute medical response community with standardised definitions and measurable outcome indicators. Broad support was proffered for the development of the template and funding was provided from the EMDM Academy Foundation for Acute Medicine (156).

The significance of Debacker’s work is that he defined terms and data elements necessary for uniform research, evaluation and education in the field of disaster medical science (101). Further, this work addressed a persistent challenge in disaster medical research and evaluation, the lack of adherence to defined standards for collecting and reporting of data, by proposing accurate and comparable terminology for reporting and analysis of the acute medical response phase for disasters. The intent of Debacker and colleagues was to validate and monitor the use of their medical template (156).
4.3.1.3.5 Dr Sabine Fattah

Dr Sabine Fattah is an emergency physician working for the Norwegian Air Ambulance Foundation (24). The significance of her work is that she created a structured report for reporting of the medical aspects of medical emergency retrievals (20-24). In conjunction with colleagues, Fattah made accessible a template, ‘Major Incident Reporting’, which could be used for submitting online reporting and accessing previously published reports (23). The intent was to provide a structured and uniform reporting system for use after major incidents that allows for comparative analysis and lessons learned. A recent review of the website (268) revealed 10 open access reports to date have been uploaded from five countries.

In 2013, during a systematic review of templates for reporting prehospital major incident medical management events, she identified more than 10 frameworks; however, of significance, none of these had been validated nor were commonly used to structure evaluations in the disaster setting (21). Accessing these specific frameworks was difficult as many of the links are no longer functional and others were published in Norwegian (which excluded them from the literature review as the criteria specifically limited results to English). The evidence obtained from Fattah’s systematic literature review indicated a lack of clear standards and definitions, which compounds the complexity of searching. It became evident that a core unifying and validated framework did not exist to structure disaster evaluation and research.

4.3.1.3.6 Robert Powers and Elaine Daily

In 2010, Robert Powers and Elaine Daily, both nurses, published ‘International Disaster Nursing’, the first nursing publication provided by WADEM (19). Nurses are the largest group of healthcare providers in the disaster setting (19). Importantly, Daily had previous involvement in the Utstein Guidelines 2003 (17) and would go on to publish separate updates to the Utstein Guidelines 2003 with Birnbaum et al. over 2014–2016 (25-34). Their textbook was written with the intent to help build the science of disaster health from a nursing perspective, utilising the input from international nursing experts from 14 countries. ‘International Disaster Nursing’ is the first nursing textbook to devote two chapters to the discussion on how to conduct disaster evaluation and research. These chapters discuss the importance of evaluating health interventions delivered as a result of a disaster. The conceptual and operational frameworks published in this text were based on the earlier work of the Utstein Guidelines 2003 (17). Their work is significant as it
promotes evidence-based practice and encourages nursing professionals to seek new ways to understand the health aspects of disasters and evaluate their practice.

4.3.1.3.7 Catherine Stephenson

Catherine Stephenson, while working at RMIT University and the Bushfire Co-operative Research Centre (CRC), developed ‘The Impacts Framework 2010’ (refer to Figure 7), which is based on economic principles and is a component of the Australian Natural Disasters Impacts Framework Project (18). It defines the processes used to determine the economic, social and environmental impacts and losses and benefits resulting from natural hazards. It is a guide to determine and measure the impacts resulting from an event and is not limited to any one phase of a disaster (269). This framework is significant in that it originated as an Australian contribution to disaster frameworks despite its focus on fire emergency and natural disasters.
4.3.1.3.8 Dr David Bradt

Dr David Bradt is a disaster epidemiologist working in disaster medicine and disaster public health. His professional interest is in disaster health and, as a consultant, he has worked with a variety of international humanitarian organisations such as the International Committee of the Red Cross (ICRC), WHO, UN High Commissioner for Refugees (UNHCR) and USAID’s Office of US Foreign Disaster Assistance. Dr Bradt has held faculty appointments both in the US and Australia and has served as a Rockefeller Foundation Bellagio Center Resident Scholar and a Fullbright Specialist in Public/Global Health. He has been the recipient of many prestigious awards (270).

In 2010, Bradt and Aitken co-authored a paper outlining a new approach to systemising disaster case reports. They provided specific instructions for authors and a framework,
CONsensus Guidelines on Reports of Field Interventions in Disasters and Emergencies (CONFIDE), with the intent to help authors report on the complex issues associated with disasters. The analysis of these reports could help inform policy and practice in the disaster setting (100).

A final publication, dated 2017, is the paper written by Wong et al. that discusses the development of a comprehensive framework to structure disaster evaluation typologies (refer to Appendix 8) (1). This paper is a direct result of the research undertaken and reported in this thesis.

4.3.2 Theme 2: Evaluation Typologies in the Disaster Setting

The scoping literature review initially yielded 88 potential records (refer to Table 5). Title and abstract review, removal of duplicate entries and the addition of suitable papers from expert opinion and snowballing left 47 usable papers for analysis. Full-text review of these articles identified that they met the inclusion criteria for this research.

This element of the scoping literature review proved somewhat difficult; perhaps due to the use of the term ‘typology’, which soon became apparent was not a common term used in the disaster evaluation literature. During the search process, it became evident that the inclusion of the word ‘emergenc*’ as used in the search criteria for evaluation frameworks broadened the scope too far; for example, searching for ‘disaster’ OR ‘emergenc*’ AND ‘evaluation’ AND ‘typ*’ OR ‘method*’ yielded in excess of 16,000 records. A review of ‘emergenc*’ related articles identified a large proportion of these were specific to the emergency department and therefore not relevant to this research. Further refinement of the search strategy was required and on reflection the term ‘emergenc*’ was removed from the search strategy.
### Table 5: Literature Review Results – Disaster Evaluation Typologies

<table>
<thead>
<tr>
<th>General Search Criteria:</th>
<th>PEER REVIEWED</th>
<th>GREY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language = English</td>
<td>CINAHL plus (EBSCO)</td>
<td>Google Scholar</td>
</tr>
<tr>
<td>Publication Date = 01/01/1960 to 31/12/2017</td>
<td>EMBASE (via Elsevier or Ovid)</td>
<td>Scholar</td>
</tr>
<tr>
<td>Type: Full Article</td>
<td>Ovid MEDLINE</td>
<td>Scopus</td>
</tr>
<tr>
<td></td>
<td>PubMed MEDLINE</td>
<td>@ Elsevier</td>
</tr>
<tr>
<td>A disaster or emergen*</td>
<td>2,577</td>
<td>60,541</td>
</tr>
<tr>
<td>B disaster</td>
<td>505</td>
<td>3,777</td>
</tr>
<tr>
<td>C evaluation</td>
<td>39,123</td>
<td>136,353</td>
</tr>
<tr>
<td>D typ* OR method*</td>
<td>71,907</td>
<td>815,636</td>
</tr>
<tr>
<td>A and C and D</td>
<td>282</td>
<td>4,218</td>
</tr>
<tr>
<td>B and C and D</td>
<td>51</td>
<td>137</td>
</tr>
<tr>
<td>Potential records (n = 88)</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Note: * due to large record numbers, a sample of 400 records where reviewed
Interestingly, the search strategy identified disaster evaluation frameworks, such as the Utstein Guidelines 2003 (17) and works by notable disaster experts such as Birnbaum (31, 32), Cuny (202, 209) and Sundnes (48). As these records were not specific evaluation ‘typologies’, they were disregarded from this part of the search strategy. These findings support the uniqueness of this research and highlighted the challenges faced while searching for relevant records. To accommodate this limitation, expert opinion played a significant part in obtaining records that could be used to identify disaster evaluation typologies.

To help analyse the remaining 47 papers for review, the key phases of the disaster timeline were used as an initial classification, that is, baselines (refer to Table 6), consequences (refer to Table 7), outcomes (refer to Table 8) and impact evaluations (refer to Table 9).

Table 6: Literature Review Results: Evaluation Typologies – Baselines

<table>
<thead>
<tr>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbas et al., GIS-Based Disaster Management, 2009 (271)</td>
</tr>
<tr>
<td>Bamberger M, Reconstructing Baseline Data for Impact Evaluation and Results Measurement, 2010 (272)</td>
</tr>
<tr>
<td>Burkle et al., The Changing Face of Humanitarian Crises, 2014 (12)</td>
</tr>
<tr>
<td>Cutter et al., Resilience Indicators for Benchmarking Baseline Conditions, 2010 (275)</td>
</tr>
<tr>
<td>Davies R, Evaluability Assessment, 2015 (276)</td>
</tr>
<tr>
<td>Fowler J, Appliance of Science Key to Disaster Risk Reduction, 2015 (277)</td>
</tr>
<tr>
<td>Fung V, Using GIS For Disaster Risk Reduction, 2012 (278)</td>
</tr>
<tr>
<td>HAP International et al., Core Humanitarian Standard on Quality and Accountability, 2014 (279)</td>
</tr>
<tr>
<td>Renger et al., Geographical Information Systems (GIS) as an Evaluation Tool, 2002 (281)</td>
</tr>
<tr>
<td>Singh-Peterson et al., Translation and Evaluation of the Baseline Resilience Indicators for Communities on the Sunshine Coast, Queensland Australia, 2014 (282)</td>
</tr>
<tr>
<td>WHO Regional Office for Europe and European Centre for Health Policy (ECHP), Health Impact Assessment. Main Concepts and Suggested Approach, 1999 (284)</td>
</tr>
</tbody>
</table>
### Table 7: Literature Review Results: Evaluation Typologies – Consequences

<table>
<thead>
<tr>
<th>CONSEQUENCES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck and Buchanan-Smith, <em>Joint Evaluations Coming of Age? The Quality and Future Scope of Joint Evaluations</em>, 2008 (285)</td>
<td></td>
</tr>
<tr>
<td>Cosgrave et al., <em>Real Time Evaluations of Humanitarian Action</em>, 2009 (287)</td>
<td></td>
</tr>
<tr>
<td>Global Facility for Disaster Reduction and Recovery (GFDRR), <em>Post-Disaster Needs Assessment, Guidelines, Volume A</em>, 2013 (288)</td>
<td></td>
</tr>
<tr>
<td>Inter-Agency Standing Committee (IASC), Multi-Cluster/Sector Initial Rapid Assessment (MIRA), 2012 (289)</td>
<td></td>
</tr>
<tr>
<td>Pan American Health Organization (PAHO) and WHO Regional Office for the Americas, <em>Rapid Needs Assessment</em>, 2016 (291)</td>
<td></td>
</tr>
<tr>
<td>PreventionWeb, <em>Centre for Research on the Epidemiology of Disasters (CRED)</em>, 2016 (292)</td>
<td></td>
</tr>
<tr>
<td>United Nations Office for Disaster Risk Reduction (UNISDR), <em>Disaster Statistics</em>, 2015 (294)</td>
<td></td>
</tr>
</tbody>
</table>

### Table 8: Literature Review Results: Evaluation Typologies – Outcomes

<table>
<thead>
<tr>
<th>OUTCOMES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Centers for Disease Control and Prevention (CDC), <em>A Framework for Program Evaluation</em>, 2016 (297)</td>
<td></td>
</tr>
<tr>
<td>Fattah et al., <em>Templates for Reporting Prehospital Major Incident Medical Management: Systematic Literature Review</em>, 2012 (20)</td>
<td></td>
</tr>
<tr>
<td>Fattah et al., <em>Systematic Literature Review of Templates for Reporting Prehospital Major Incident Medical Management</em>, 2013 (8)</td>
<td></td>
</tr>
<tr>
<td>Fattah et al., <em>A Consensus Based Template for Reporting of Prehospital Major Incident Medical Management</em>, 2014 (22)</td>
<td></td>
</tr>
<tr>
<td>Kulling et al., <em>Guidelines for Reports on Health Crises and Critical Health Events</em>, 2010 (167)</td>
<td></td>
</tr>
<tr>
<td>Moore et al., <em>Process Evaluation of Complex Interventions: Medical Research Council Guidance</em>, 2015 (300)</td>
<td></td>
</tr>
<tr>
<td>WHO et al., <em>Process Evaluations</em>, 2000 (304)</td>
<td></td>
</tr>
</tbody>
</table>
This analysis into sub-themes contributed to the development of the comprehensive framework of disaster evaluation typologies, as discussed in further detail in Chapter 7.

4.4 Discussion

The scoping literature review identified that there was no single, core, unifying framework to structure disaster evaluation and research in common usage. Likewise, the few frameworks that were identified were not commonly used to structure disaster evaluations. Further, more information could be found in the grey literature and humanitarian sector than in the peer-reviewed literature or in the disaster sector. There are very few impact evaluations of interventions during disasters reported in the literature (3). Most evaluation reports are descriptive, process focused and lack a core conceptual framework (313, 314). In a recent study, it was noted that less than 1% of publications specific to disaster evaluations discussed the development of, or use of, evaluations such as the Utstein Guidelines 2003 or KAMEDO (5). Research undertaken by Stratton in 2014 identified that the majority of papers submitted to PDM were surveys or descriptive in nature (122). The published reports did not demonstrate a consistent and structured approach to evaluations of interventions and the impact of interventions on the affected population was rarely mentioned (26). Many NGOs, such as the International Federation of Red Cross (IFRC) and various UN agencies, have their own internal standards for evaluations (315, 316). As evidenced by the work of ALNAP (317) and the Inter-Agency Standing Committee (IASC) (318), attempts continue to be made to consolidate standards and guidelines across the sector. This theme is expanded in Chapter 7.
National government-led disaster reviews or evaluations, however, are noticeably absent from this literature review, which probably reflects that they are either not in the public domain or were not identified by the search strategy. Further, in the Australian context, most, if not all, disaster-related organisations undertake operational debriefs and self-review following major events, however, these are commonly not made public. Two particular evaluation typologies are evident in the Australian context but likewise were not picked up by the search strategy for this review; namely, ‘lessons learned’ (319-321) and ‘quality assurance’ (322, 323), which are gaining momentum.

This literature review had a number of limitations. Searching was difficult because of the relatively small published literature on disaster evaluation frameworks and the lack of clarity and classification of disaster evaluation typologies available. As noted above, the search strategy itself, in hindsight, perhaps could have been improved. It is not clear if most or all of the frameworks or evaluation typologies existing in the disaster setting have been identified. Confusing and inconsistent terminology remained an issue and this could be explained by the number of multiple stakeholders involved in the disaster setting. Each of the stakeholders brings to the table a different perspective and approach, including different terminology and definitions with differing points of view. While there are limitations to the scoping literature review, the results support the 3ie findings from Chapter 2 that there is a lack of a high-quality, reliable and robust evaluation framework to support decision making in the disaster and humanitarian sector (3).

These findings helped frame the research questions and subsequent development of a comprehensive framework for disaster evaluation typologies, developed further in Chapter 7.

4.5 Summary

A scoping literature review was conducted to provide answers to the research questions regarding what disaster evaluation frameworks and evaluation typologies exist in the disaster setting. The literature review identified several disaster evaluation frameworks. The Utstein Guidelines 2003 (17) were identified as the pivotal disaster evaluation framework. Other evaluation frameworks were placed along a chronological timeline in relation to their publication date relative to the Utstein Guidelines 2003. This diagrammatic representation demonstrated the evolution of disaster evaluation frameworks. It is important to note that frameworks were in their infancy during the pre-
Utstein phase, they were not validated, and not used to structure evaluations and research in the disaster setting. The Utstein Guidelines 2003 addressed issues of inconsistency in definitions and terminology, aimed to improve the evidence-base of disaster health and improve healthcare delivered as a result of a disaster. Unfortunately, the Utstein Guidelines 2003 were referenced but not used to structure evaluation and research in the disaster setting. There appears to be, however, an increased interest and number of publications on disaster evaluations appearing in the literature post-2003.

The second theme of this literature review was structured to identify contemporary evaluation typologies that were, or could be, used in the disaster setting. This review identified that the context was confusing and that disaster evaluation typologies are poorly defined and lacked consistent definitions and terminology. Evaluation typologies were categorised into four sub-themes that were positioned along the disaster timeline.

Attempts at structuring evaluations in the disaster setting have been progressing for over 30 years, as demonstrated in the scoping literature review. It became evident that a core unifying, internationally validated framework to structure disaster did not exist. This provides a unique opportunity for the researcher to address this gap by developing a comprehensive framework to structure disaster evaluation typologies. Such a framework would provide structure to support disaster evaluation and research and offer consistent terminology and standards for reporting across the disaster timeline. Further, this work promotes an environment for constructive dialogue on evaluation in the disaster setting and adds to the evidence-base of disaster health.

Results from the scoping literature review strengthen the initial findings of the 3ie Scoping Paper 1 highlighting the lack of quality and comparable evidence. This chapter makes a unique contribution to knowledge, from both the generic literature review and ‘Utstein’ perspectives. Results are used to guide the research design to identify the key elements to be included in the core structure discussed in Chapter 5.
Chapter 5: Pre-Establishment of the Core Structure - Two Studies

5.1 Introduction

When starting to build a comprehensive framework to structure disaster evaluation typologies, it is important to establish a strong foundation to which all other entities will be related. This chapter details the procedures undertaken to support the establishment of the core structure by presenting the structure and outcome of two studies that inform this research. The study of the Utstein Guidelines 2003 includes the identification of key elements. This is followed by the study of reviews of Australian disaster events, which used a thematic analysis to test the inclusion or exclusion of not only the key ‘Utstein’ elements but contemporary elements.

Two research approaches were used to assist in the establishment of the core structure. The first included the use of semi-structured interviews seeking input from experienced practitioners in the disaster/emergency and humanitarian sectors regarding the use and usefulness of the Utstein Guidelines 2003, previously regarded as the ‘gold standard’ (refer to Chapter 4). Key elements from the Utstein Guidelines 2003 (17)(p.35) that could be included in the establishment of the core structure were identified.

The second research study involves a thematic analysis of Australian disaster events to test whether the key elements from the Utstein Guidelines 2003, as amended, are reflected in selected reviews of Australian disaster events. Additionally, the thematic analysis assists in the identification of other contemporary elements that might be significant enough for inclusion in the core structure.

5.2 Study of the Utstein Guidelines

The Utstein Guidelines referred to in this chapter are the 2003 ‘Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style Volume 1 Conceptual Framework of Disasters’ (17). Subsequently, separate updates have been published by both co-editors: Sundnes in 2014 (35-57) and Birnbaum et al. in 2014–2016 (25-34). However, the timeline of this thesis research meant that these updates were not available when undertaking this study.
5.2.1 Research Questions

The research questions relevant to this study are:

- How do professionals active in disaster evaluation and research view the Utstein Guidelines?
- How do professionals active in disaster evaluation and research use the Utstein Guidelines?
- What are the barriers and facilitators to professionals active in disaster evaluation and research using the Utstein Guidelines?
- What do professionals active in disaster evaluation and research identify as important key elements in the Utstein Guidelines to include in the development of a comprehensive framework for evaluations in the disaster setting?

5.2.2 Research Design and Methodology

To understand how the Utstein Guidelines 2003 are used, a literature review specific to the Utstein Guidelines 2003 in the disaster setting was undertaken. The results show that while these guidelines were being referenced, they were not widely used to structure disaster evaluations and research. Curious to know why this was the case, semi-structured interviews were undertaken to understand the reasons for the lack of uptake and use of the Utstein Guidelines 2003 to structure disaster evaluations and research. Results obtained from the semi-structured interviews assisted in identifying key elements to include in the establishment of the core structure. Both research methods are discussed in detail below.

The literature review was designed to provide background and understanding on how the Utstein Guidelines 2003 were being used in the disaster setting. A scoping literature review was chosen as this would allow comprehensive mapping of the peer-reviewed and grey literature and would also involve additional processes such as seeking expert opinion and snowballing. The justification for using this research strategy was discussed in Section 4.2.

The peer-reviewed literature was searched using major electronic databases relevant to the health sector, including CINAHL (EBSCO Information Services; Ipswich, Massachusetts, US), EMBASE (Elsevier; Amsterdam, Netherlands), Ovid MEDLINE (US National Library of Medicine, National Institutes of Health; Bethesda, Maryland,
US), PubMed/MEDLINE (US National Library of Medicine, National Institutes of Health; Bethesda, Maryland, US), Scopus (Elsevier; Amsterdam, Netherlands) and ScienceDirect (Elsevier; Amsterdam, Netherlands). Prior to keyword searching, the following criteria were set:

- English language
- publication dates from 2003-2017 inclusive. The commencement date of 2003 was chosen to coincide with the release of the Utstein Guidelines 2003 (17)
- that the study referred to the Utstein Guidelines 2003
- only full articles to be included in the results.

The following key terms and Boolean operators were used ‘Utstein’ AND ‘guideline’ OR ‘evaluation’ OR ‘framework’ OR ‘methodology’ OR ‘research’ OR ‘style’ OR ‘template’, AND ‘disaster’ OR ‘emergenc*’ OR ‘humanitarian’. The results of the literature search are shown in Table 10.

The scoping literature review yielded 18 potential records (refer to Table 10). Subsequent title and abstract review and removal of duplicate entries further refined the number of usable records to nine. These are listed in alphabetical order in Table 11 to demonstrate the multiple contributions made by the authors. Full-text review of these articles identified that they met the inclusion criteria for this research.
### Table 10: Literature Search Strategy 2003 to 2017

<table>
<thead>
<tr>
<th>General Search Criteria:</th>
<th>PEER REVIEWED</th>
<th>CINAHL plus (EBSCO)</th>
<th>EMBASE (via Elsevier or Ovid)</th>
<th>Ovid MEDLINE</th>
<th>PubMed MEDLINE</th>
<th>Scopus</th>
<th>ScienceDirect @ Elsevier</th>
<th>GREY Google Scholar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language = English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication Date = 01/01/2003 to 31/12/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type: Full Article</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>Utstein</td>
<td>220</td>
<td>261</td>
<td>292</td>
<td>434</td>
<td>416</td>
<td>1,742</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>Guideline or evaluation or framework or methodology or research or style or template</td>
<td>1,336,232</td>
<td>374,457</td>
<td>1,347,618</td>
<td>8,364,277</td>
<td>2,221,275</td>
<td>391</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>A and B</td>
<td>158</td>
<td>142</td>
<td>242</td>
<td>423</td>
<td>330</td>
<td>6</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>Disaster or emergency or humanitarian</td>
<td>177,512</td>
<td>53,469</td>
<td>65,787</td>
<td>210,917</td>
<td>208,297</td>
<td>2,435</td>
</tr>
<tr>
<td>E</td>
<td></td>
<td>A and D (disaster or emergency or humanitarian)</td>
<td>141</td>
<td>152</td>
<td>193</td>
<td>341</td>
<td>302</td>
<td>8</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td>C and D</td>
<td>104</td>
<td>86</td>
<td>158</td>
<td>336</td>
<td>236</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Potential records (n = 18)</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Usable papers (duplicates removed)</td>
<td>n = 9 (refer to Table 11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Snowballing</td>
<td>n = 19 (refer to Table 12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 11: Usable Papers from the Literature Review of Disaster Evaluation Frameworks

<table>
<thead>
<tr>
<th>Author and Title</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Castren et al., <em>Improving the Science and Evidence for the Medical Management of Disasters: Utstein Style</em>, 2012 (324)</td>
<td></td>
</tr>
<tr>
<td>Debacker et al., <em>Utstein-Style Template for Uniform Data Reporting of Acute Medical Response In Disasters</em>, 2012 (101)</td>
<td></td>
</tr>
<tr>
<td>Fattah et al., <em>Systematic Literature Review of Templates for Reporting Prehospital Major Incident Medical Management</em>, 2013 (8)</td>
<td></td>
</tr>
<tr>
<td>Kohl et al., <em>The Sumatra-Andaman Earthquake and Tsunami of 2004: The Hazards, Events, and Damage</em>, 2005 (166)</td>
<td></td>
</tr>
<tr>
<td>Radestad et al., <em>Essential Key Indicators for Disaster Medical Response Suggested to be Included in a National Uniform Protocol for Documentation of Major Incidents: A Delphi Study</em>, 2013 (325)</td>
<td></td>
</tr>
<tr>
<td>Stratton SJ, <em>The Utstein-Style Template for Uniform Data Reporting of Acute Medical Response in Disasters</em>, 2012 (156)</td>
<td></td>
</tr>
<tr>
<td>Task Force on Quality Control of Disaster Management (TFQCDM) et al., <em>Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style</em>, 2003 (17)</td>
<td></td>
</tr>
</tbody>
</table>

Snowballing of the identified records provided a further 19 records of interest. These are shown in Table 12 in alphabetical order. A total of 28 records informed the results of this scoping literature review of how the Utstein Guidelines 2003 were being used in the disaster setting.

Many of the findings have already been covered in the literature review conducted in Chapter 4 related to disaster evaluation frameworks. From an ‘Utstein’ specific point of view, it is worth noting the increased publications regarding the medical management of disasters. This health focus, and not specifically a disaster focus, is evident in publications by Debacker (101, 224, 229), Castren (324), Radestad (325), Fattah (20, 21, 23, 24) and Lennquist (168, 326-330), to name a few. These authors have identified the urgent need for a consistent approach to disaster evaluation and reporting from a medical perspective.
Of these 28 papers, each of which referenced the Utstein Guidelines 2003, only three used these guidelines to structure their evaluation study (166, 171, 172). Most commonly, the Utstein Guidelines 2003 were mentioned in the literature review and/or discussions within these papers. This is consistent with the previous findings in Chapter 4.

Further investigation was required to understand why the Utstein Guidelines 2003 were referenced, but not used to structure disaster evaluations and research. To understand how

<table>
<thead>
<tr>
<th>Author and Title</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bradt and Aitken, Disaster Medicine Reporting: The Need for New Guidelines and the CONFIDE Statement, 2010</td>
<td>(100)</td>
</tr>
<tr>
<td>De Boer and Debacker, Quantifying Medical Disaster Management, 2006</td>
<td>(163)</td>
</tr>
<tr>
<td>Fattah et al., Templates for Reporting Pre-Hospital Major Incident Medical Management: Systematic Literature Review, 2012</td>
<td>(20)</td>
</tr>
<tr>
<td>Fattah et al., A Consensus Based Template for Reporting of Prehospital Major Incident Medical Management 2014</td>
<td>(22)</td>
</tr>
<tr>
<td>Green et al., Generic Evaluation Methods for Disaster Drills in Developing Countries, 2003</td>
<td>(331)</td>
</tr>
<tr>
<td>Kar-Purkayastha et al., Dealing with Disaster Databases – What Can We Learn From Health and Systematic Reviews? 2011</td>
<td>(332)</td>
</tr>
<tr>
<td>Kulling et al., Guidelines for Reports on Health Crises and Critical Health Events, 2010</td>
<td>(167)</td>
</tr>
<tr>
<td>Langhelle et al., Recommended Guidelines for Reviewing, Reporting and Conducting Research on Post-Resuscitation Care: The Utstein Style, 2005</td>
<td>(333)</td>
</tr>
<tr>
<td>Legemaate et al., The Evaluation of Research Methods During Disaster Exercises: Applicability for Improving Disaster Health Management, 2012</td>
<td>(334)</td>
</tr>
<tr>
<td>Leiba et al., Disast-CIR: Disastrous Incidents Systematic Analysis Through Components, Interactions and Results: Application to a Large-Scale Train Accident, 2009</td>
<td>(102)</td>
</tr>
<tr>
<td>Leiba et al., Lessons Learned from Cross-Border Medical Response to the Terrorist Bombings in Taba and Ras-el-Satan, Egypt, on 07 October 2004, 2005</td>
<td>(335)</td>
</tr>
<tr>
<td>Lennquist S, Protocol for Reports from Major Accidents and Disasters in the International Journal of Disaster Medicine, 2008</td>
<td>(327)</td>
</tr>
<tr>
<td>Lennquist S, Protocol for Reports from Major Accidents and Disasters, 2004</td>
<td>(168)</td>
</tr>
<tr>
<td>Lockey DJ, The Shootings in Oslo and Utoya Island July 22, 2011: Lessons for the International EMS Community, 2012</td>
<td>(336)</td>
</tr>
<tr>
<td>Ringdal et al., The Utstein Template for Uniform Reporting of Data Following Major Trauma: A Joint revision by SCANTEM, TARN, DGU-TR and RITG, 2008</td>
<td>(234)</td>
</tr>
<tr>
<td>Rüter et al., Performance Indicators for Major Incident Medical Management – A Possible Tool for Quality Control? 2004</td>
<td>(337)</td>
</tr>
<tr>
<td>Subbarao et al., A Consensus-Based Educational Framework and Competency Set for the Discipline of Disaster Medicine and Public Health Preparedness, 2008</td>
<td>(338)</td>
</tr>
</tbody>
</table>

Table 12: Snowballing Results - Disaster Evaluation Frameworks
professionals viewed and used the Utstein Guidelines 2003, semi-structured interviews were conducted to answer the following questions:

- Are professionals active in disaster evaluation and research aware of the Utstein Guidelines?
- How do professionals active in disaster evaluation and research view and use the Utstein Guidelines?
- What are the barriers to professionals active in disaster evaluation and research using the Utstein Guidelines?
- Are professionals active in disaster evaluation and research aware of similar guidelines, and how can disaster evaluations be improved?

To address these research questions, semi-structured interviews were undertaken; the key elements of how this was conducted are discussed in greater detail below. Semi-structured interviews were deemed an appropriate method of ascertaining the perceptions of experienced practitioners (115, 116) in the disaster/emergency and humanitarian sectors regarding the use and usefulness of the Utstein Guidelines 2003. The results obtained assist in identifying key elements to include in the establishment of the core structure.

The research process for undertaking the semi-structured interviews included identifying the sample population, including the selection criteria and sample size; writing a theme list of questions and undertaking a pilot test; writing an explanatory and consent statement; obtaining ethics approval; hosting the semi-structured interviews; collecting the data; undertaking a thematic analysis of the results; and disseminating the findings.

The sample or ‘target’ population was selected based on their experience in undertaking evaluation studies in the fields of disaster medicine, disaster management, emergency management or the humanitarian arena. This international, multi-disciplinary, cross-sectoral approach provided a depth and richness of responses due to the wide range of experience that the participants had. These skills and attributes were sought, as it was believed that experience in undertaking evaluation studies/evaluations in the disaster, emergency or humanitarian arena would provide a depth of knowledge/understanding and insight into the strengths, weaknesses, opportunities and threats (339, 340) facing evaluation in the disaster setting.

A purposive sampling strategy was used to create a target list of potential research participants (n = 30). Participants who met the selection criteria and whose contact details
were available in the public domain, known to the researcher and/or identified through the literature review were contacted via email. They were asked to read an Explanatory Statement (Appendix 5), provide consent (Appendix 6) and confirm a date and time convenient for them to participate in the semi-structured interview via telephone or Skype.

The next step was to create a theme list of questions that asked the participants to comment on the use and usefulness of the Utstein Guidelines. The following questions, were a combination of open-ended and closed questions (Appendix 7), and asked:

**Question One:**

- Are you familiar with the 2003 version of ‘Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style’? This was specified as having been authored by the Task Force of Quality Management of Disaster Management, Health Disaster Management Guidelines, Volume I, 2003 and the reference provided (17) to avoid any confusion.
- If yes, how did you gain this information/awareness?
- If yes, do you use them?
- If yes, are they valuable?
- How have you found them to be valuable?
- Are you aware of other colleagues using them?

**Question Two:**

- From the literature review, we can see that the Utstein Guidelines are being referenced, but not used. Do you have any suggestions to offer on why they are not being used?

**Question Three:**

- Are you aware of any other similar guidelines being used?
- Are you aware of any recent updates to the Utstein Guidelines?

**Question Four:**

- Do you have any suggestions on how we can improve the evaluation of disaster health?
The Explanatory Statement, consent form and theme list of initial questions related to the Utstein Guidelines 2003 can be found in Appendices 5, 6 and 7 respectively. These documents were pilot tested (n = 5) prior to the commencement of the semi-structured interviews. The feedback received provided an opportunity for improvement; only minor edits were required.

Once the necessary documents were created, pilot tested and modified as necessary, Ethics Committee approval was sought via MUHREC. The ethics application for low-risk research was prepared and accepted under project: Disaster Health Evaluation: A Review of the Utstein Style Guidelines/Template, MUHREC Approval CF14/158 – 2014000041. Please refer to Appendix 3. A 12-week period was allocated in late 2014 for the commencement of the interviews.

The semi-structured interviews (n = 18) were conducted using telephone or Skype over a 9-month period in 2014–2015. It was necessary to consider multiple time zones because of the international nature of the research participants. Another consideration was the impact of international disaster events, such as the outbreak of the Ebola crisis in Western Africa (68, 341), which precluded some of the invited research participants from engaging in the survey.

The interviews were recorded with participant consent and the results transcribed. A copy of the final interview notes was sent to the research participant for verification and to ensure the accuracy of their responses. Participants were given 14 days to respond. No edits were required and a thematic analysis was undertaken. Key themes and relevant sub-themes were identified and are presented in Section 5.2.3

5.2.3 Results

Some 30 participants in total were approached to participate in the semi-structured interviews. Of these, 18/30 (60%) completed the survey and 12/30 (40%) did not respond to the email request. For the results to be considered significant, a participation rate of 50% or higher was set. Despite the extended length of time that the surveys took (i.e., a 9-month period in 2014–2015), the ‘required number’ of responses were obtained. The data were collated and analysed according to the following:

- consent and demographic information about the participants
• survey questions 1–3, based on the use and usefulness of the Utstein Guidelines 2003
• responses related to suggestions on how to improve the evaluation of disaster health.

5.2.3.1 Consent and Demographics

Participants were asked to acknowledge that they had read and understood the explanatory statement and consented to participating in the semi-structured interviews. All participants (100%) agreed (refer to Table 13).

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory Statement</td>
<td>18/18</td>
<td>100.0%</td>
</tr>
<tr>
<td>Copyright</td>
<td>18/18</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Participants were invited from three WHO regions in an attempt to facilitate an international point of view. Table 14 shows the country breakdown, while Table 15 shows stratification by WHO region. The largest number of responses were received from the Western Pacific region.
Table 14: Stratification by Country

<table>
<thead>
<tr>
<th>STRATIFICATION (By Country)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>6</td>
<td>33.3%</td>
</tr>
<tr>
<td>Belgium</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>Canada</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>Japan</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>New Zealand</td>
<td>2</td>
<td>11.1%</td>
</tr>
<tr>
<td>Sweden</td>
<td>2</td>
<td>11.1%</td>
</tr>
<tr>
<td>Switzerland</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>UK</td>
<td>1</td>
<td>5.6%</td>
</tr>
<tr>
<td>USA</td>
<td>2</td>
<td>11.1%</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Table 15: Stratification by WHO Region

<table>
<thead>
<tr>
<th>STRATIFICATION (By WHO Region)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>African</td>
<td>0/18</td>
<td>0.0%</td>
</tr>
<tr>
<td>Americas</td>
<td>3/18</td>
<td>16.7%</td>
</tr>
<tr>
<td>Eastern Mediterranean</td>
<td>0/18</td>
<td>0.0%</td>
</tr>
<tr>
<td>European</td>
<td>5/18</td>
<td>27.8%</td>
</tr>
<tr>
<td>South East Asia</td>
<td>0/18</td>
<td>0.0%</td>
</tr>
<tr>
<td>Western Pacific</td>
<td>10/18</td>
<td>55.6%</td>
</tr>
<tr>
<td></td>
<td>39/39</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Participants were asked to choose their area(s) of expertise (academic, emergency management, healthcare and humanitarian). Participants were able to choose more than one response. The predominant areas identified by participants were healthcare – medical (11/18, 29.7%) and academics (10/18, 27.0%) (refer to Table 16).
The survey did not ask participants for their age group, gender or years of experience. This information was readily available in the public domain or known to the researcher; in any case, these variables were not used for sub-analysis.

5.2.3.2 Use and Usefulness of the Utstein Guidelines

When asked whether they were aware of the Utstein Guidelines 2003, the majority 17/18 (94.4%) agreed that they were (refer to Table 17, Question 1A).

Table 17: Question 1A – Are you aware of the Utstein Guidelines 2003?

<table>
<thead>
<tr>
<th>RESPONSE</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>17/18</td>
<td>94.4%</td>
</tr>
<tr>
<td>No</td>
<td>1/18</td>
<td>5.6%</td>
</tr>
<tr>
<td></td>
<td>18/18</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Regarding how the participants had gained awareness or knowledge about the Utstein Guidelines 2003, 7/18 (38.9%) had learned about them through promotion via the WADEM or the World Congress on Disaster and Emergency Medicine (WCDEM), the biennial world congress hosted by WADEM. Currently, the Utstein Guidelines 2003 are available on the WADEM website (342) (refer to Table 18, Question 1B). Almost one-third of participants 5/18 (27.8%) stated that they taught the Utstein Guidelines 2003 at university. Three of the participants (16.7%) were part of the initial Inception Group/Task
Force for the Utstein Guidelines, dating back as far as 1996 when the initial Utstein Guidelines were published (154). The inclusion of three participants from the initial Inception Group is significant, as the Utstein Guidelines 2003 were the first major attempt to provide structure for the conduct of evaluation and research in the disaster setting (17).

Of the remaining participants, one stated that they had found out about the Utstein Guidelines 2003 via a literature search, another chose to not answer the question and the final participant stated that they ‘were not aware of the Utstein Guidelines’.

Table 18: Question 1B – How did you gain this awareness?

<table>
<thead>
<tr>
<th>QUESTION</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCDEM (including training)</td>
<td>7/18</td>
<td>38.9%</td>
</tr>
<tr>
<td>Inception Group/Task Force</td>
<td>3/18</td>
<td>16.7%</td>
</tr>
<tr>
<td>Taught at University</td>
<td>5/18</td>
<td>27.8%</td>
</tr>
<tr>
<td>Literature Search</td>
<td>1/18</td>
<td>5.6%</td>
</tr>
<tr>
<td>No reply</td>
<td>1/18</td>
<td>5.6%</td>
</tr>
<tr>
<td>Was not aware of the Utstein</td>
<td>1/18</td>
<td>5.6%</td>
</tr>
<tr>
<td>Guidelines</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18/18</td>
<td>100.0%*</td>
</tr>
</tbody>
</table>

Note: * Total % adds to 100.2% due to rounding of the final three responses.

When asked about whether they used the Utstein Guidelines 2003, results were mixed (refer to Table 19, Question 1C). Participants were given the opportunity to answer: yes, no, yes/no or unaware. Almost half of the participants 8/18 (44.4%) specified that they did not use the Utstein Guidelines 2003. Surprisingly, one-third (6/18, 33.3%) confirmed that they did use the Utstein Guidelines 2003. Only one participant was not aware of the Utstein Guidelines 2003 and the remaining participants 3/18 (16.7%) provided both a positive and negative response. The verbal comments provided some clarification to the yes/no responses. From a ‘yes’ perspective, participants commented that they used the Utstein Guidelines 2003 to ‘teach at university’, ‘in a modified way’, and ‘for research’. From a ‘no’ perspective, participants commented that they did not use the Utstein Guidelines 2003 to structure an ‘evaluation’ and that they were ‘not operational’ and therefore not applied in the disaster setting. One of the benefits of undertaking semi-structured interviews is the ability to delve into the participants’ responses in more detail while seeking further clarification of any ambiguous comments.
Table 19: Question 1C – Do you use them? Yes/No

<table>
<thead>
<tr>
<th>RESPONSE</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6/18</td>
<td>33.3%</td>
</tr>
<tr>
<td>No</td>
<td>8/18</td>
<td>44.4%</td>
</tr>
<tr>
<td>Yes and No</td>
<td>3/18</td>
<td>16.7%</td>
</tr>
<tr>
<td>Was not aware of the Utstein Guidelines</td>
<td>1/18</td>
<td>5.6%</td>
</tr>
<tr>
<td></td>
<td>18/18</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

When asked if they thought the Utstein Guidelines 2003 were valuable, 8/18 (44.4%) agreed that they did have value, 6/18 (33.3%) expressed mixed feelings, 2/18 (11.15%) stated that the Utstein Guidelines 2003 did not have value and 2/18 (11.51%) chose not to answer (refer to Table 20, Question 1D).

Table 20: Question 1D – If ‘Yes’, are they valuable?

<table>
<thead>
<tr>
<th>RESPONSE</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>8/18</td>
<td>44.4%</td>
</tr>
<tr>
<td>No</td>
<td>2/18</td>
<td>11.15%</td>
</tr>
<tr>
<td>Yes and No</td>
<td>6/18</td>
<td>33.3%</td>
</tr>
<tr>
<td>Unable to comment</td>
<td>2/18</td>
<td>11.15%</td>
</tr>
<tr>
<td></td>
<td>18/18</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Comments related to the perceived value of the Utstein Guidelines 2003 provided further insight into how participants viewed these. Table 21 Question 1E provides a summary of the comments provided, ranked in order. While five participants provided no answer, the remaining comments were all positive.
Table 21: Question 1E – How are they valuable?

<table>
<thead>
<tr>
<th>COMMENTS</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teach them at university</td>
<td>5</td>
</tr>
<tr>
<td>No answer</td>
<td>5</td>
</tr>
<tr>
<td>Provides conceptual framework (theory and concept)</td>
<td>4</td>
</tr>
<tr>
<td>Promote standard use of terms / definitions</td>
<td>3</td>
</tr>
<tr>
<td>Unique, valuable, brilliant, wonderful tool, nothing like them</td>
<td>3</td>
</tr>
<tr>
<td>Idea /intent is good</td>
<td>2</td>
</tr>
<tr>
<td>Thorough, comprehensive</td>
<td>1</td>
</tr>
<tr>
<td>Provide structure for disaster research</td>
<td>1</td>
</tr>
<tr>
<td>Helps understand disasters</td>
<td>1</td>
</tr>
<tr>
<td>Evaluation (projected objective of evaluation)</td>
<td>1</td>
</tr>
</tbody>
</table>

When asked about whether their colleagues used the Utstein Guidelines 2003, half the participants did not know; one participant agreed that they were being referenced but not used to structure evaluation and research in the disaster setting (refer to Table 22, Question 1F).

Table 22: Question 1F - Are you aware of anyone else using them?

<table>
<thead>
<tr>
<th>RESPONSE</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>5/18</td>
<td>27.8%</td>
</tr>
<tr>
<td>No</td>
<td>9/18</td>
<td>50.0%</td>
</tr>
<tr>
<td>Aware they are referenced but not used</td>
<td>1/18</td>
<td>5.6%</td>
</tr>
<tr>
<td>No answer</td>
<td>3/18</td>
<td>16.6%</td>
</tr>
<tr>
<td></td>
<td>18/18</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Participants were asked to offer suggestions on why the Utstein Guidelines 2003 were referenced but not used to undertake evaluations in disaster health. A significant number of participants identified that the Utstein Guidelines 2003 were ‘conceptual, not practical or operational’ and ‘too difficult, complex and not easy to use’. A thematic analysis of the comments was conducted (refer to Table 23, Question 2). Interestingly, two participants mentioned the lack of a validation process.
Table 23: Thematic Analysis of Question 2 – From the literature review, we can see that the Utstein Guidelines are being referenced but not used. Do you have any suggestions to offer on why they are not being used?

<table>
<thead>
<tr>
<th>THEMATIC ANALYSIS OF PARTICIPANT COMMENTS</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual not practical or operational</td>
<td>7</td>
</tr>
<tr>
<td>Too difficult, complex, not easy to use</td>
<td>6</td>
</tr>
<tr>
<td>Need a simplified version</td>
<td>4</td>
</tr>
<tr>
<td>Book difficult to read</td>
<td>3</td>
</tr>
<tr>
<td>Questioning the usefulness of Utstein Guidelines</td>
<td>3</td>
</tr>
<tr>
<td>No underpinning theory/framework</td>
<td>2</td>
</tr>
<tr>
<td>Not validated</td>
<td>2</td>
</tr>
<tr>
<td>No consensus with definitions</td>
<td>2</td>
</tr>
<tr>
<td>Narrowly marketed, low level of awareness</td>
<td>2</td>
</tr>
<tr>
<td>Health focus and not emergency management</td>
<td>2</td>
</tr>
<tr>
<td>Authors are not considered ‘opinion leaders’, but marginal</td>
<td>1</td>
</tr>
<tr>
<td>Content has major flaws and analytical errors</td>
<td>1</td>
</tr>
<tr>
<td>Don’t know where to find them</td>
<td>1</td>
</tr>
<tr>
<td>Not a good research tool</td>
<td>1</td>
</tr>
<tr>
<td>No guidelines in evaluation and research</td>
<td>1</td>
</tr>
<tr>
<td>No comment</td>
<td>1</td>
</tr>
</tbody>
</table>

When asked if they were aware of other guidelines being used (other than the Utstein Guidelines 2003), participants offered a list of alternatives consisting of 23 options. The most recognised guidelines were the IFRC, which were identified by 3/18 (16.7%). Over 20 different organisational operational frameworks were mentioned, but there was very little consensus (refer to Table 24, Question 3).
Table 24: Question 3 – Are you aware of any other similar guidelines being used?

<table>
<thead>
<tr>
<th>COMMENTS</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>No answer</td>
<td>2</td>
</tr>
<tr>
<td>No, not aware of any other guidelines</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organisational Operational Frameworks</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>All Hazards Approach – EM</em></td>
</tr>
<tr>
<td><em>ALNAP</em></td>
</tr>
<tr>
<td><em>Bradt, D</em></td>
</tr>
<tr>
<td><em>Case reports</em></td>
</tr>
<tr>
<td><em>CRED</em></td>
</tr>
<tr>
<td><em>EDMD, Debacker, M</em></td>
</tr>
<tr>
<td><em>EMA or AEMI</em></td>
</tr>
<tr>
<td><em>Evidence Aid</em></td>
</tr>
<tr>
<td><em>Fattah, S</em></td>
</tr>
<tr>
<td><em>FMAT – deployment guidelines</em></td>
</tr>
<tr>
<td><em>ICN</em></td>
</tr>
<tr>
<td><em>IFRC</em></td>
</tr>
<tr>
<td><em>KAMEDO</em></td>
</tr>
<tr>
<td><em>Kulling et al.</em></td>
</tr>
<tr>
<td><em>Lennquist, S</em></td>
</tr>
<tr>
<td><em>MSF</em></td>
</tr>
<tr>
<td><em>National government standards</em></td>
</tr>
<tr>
<td><em>ODI</em></td>
</tr>
<tr>
<td><em>OECD-DAC</em></td>
</tr>
<tr>
<td><em>Powers, R and Daily, E</em></td>
</tr>
<tr>
<td><em>SPHERE Standards</em></td>
</tr>
<tr>
<td><em>UN</em></td>
</tr>
<tr>
<td><em>WHO</em></td>
</tr>
</tbody>
</table>

Most of these additional similar guidelines offered by respondents have been identified in the previous literature searches. The remaining suggestions did not meet the search criteria.

When asked how to improve the evaluation of disaster health, a variety of suggestions were offered by participants and these diverse comments made undertaking a thematic analysis difficult. Support was offered for the involvement of ‘think tanks’ on evaluation and communities of practice, such as ALNAP (343). Additional comments were made
specific to the complex nature of undertaking evaluations and the need to increase training in evaluation (refer to Table 25, Question 4).

**Table 25: Question 4 – Do you have any suggestions on how we can improve the evaluation of disaster health?**

<table>
<thead>
<tr>
<th>COMMENTS</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Involve think tanks on evaluation or Communities of Practice, such as ALNAP, UN, WHO, IFRC/ICRC and WADEM</td>
<td>7</td>
</tr>
<tr>
<td>Evaluation needs to be deliverable / operational / easier to use</td>
<td>4</td>
</tr>
<tr>
<td>Need evidence-base, science base, need to get research into action</td>
<td>3</td>
</tr>
<tr>
<td>Need for evaluations, contribute to evaluations</td>
<td>2</td>
</tr>
<tr>
<td>Increase training in evaluation, improve consistency, possible certification in evaluation / certified evaluator</td>
<td>2</td>
</tr>
<tr>
<td>All hazards approach</td>
<td>2</td>
</tr>
<tr>
<td>Need to revise the 2003 Guidelines and lack of completion Vol 2-4</td>
<td>2</td>
</tr>
<tr>
<td>Explain the evaluation – for better outcomes, lessons learned, evaluation template</td>
<td>2</td>
</tr>
<tr>
<td>Common language / terminology a good start</td>
<td>2</td>
</tr>
<tr>
<td>Need broader uptake of Guidelines and involvement in disaster evaluation</td>
<td>2</td>
</tr>
<tr>
<td>Any evaluation tool must be accessible and usable, people need to be able to deliver it/use it, needs to be operational and relevant to the people who will administer the intervention / project</td>
<td>2</td>
</tr>
<tr>
<td>Good to have pre-event assessment</td>
<td>2</td>
</tr>
<tr>
<td>National database – up and running with AusMAT</td>
<td>2</td>
</tr>
<tr>
<td>Discussion and collaboration, takes time</td>
<td>2</td>
</tr>
<tr>
<td>Utstein Guidelines are useful and we should think of ways to make it more user friendly and easy to grasp</td>
<td>2</td>
</tr>
<tr>
<td>Extraordinarily difficult</td>
<td>2</td>
</tr>
<tr>
<td>“Good question, when you find out, let me know”</td>
<td>2</td>
</tr>
<tr>
<td>Need guidelines for health in regard to disaster emergency management</td>
<td>1</td>
</tr>
<tr>
<td>Need top down approach</td>
<td>1</td>
</tr>
<tr>
<td>Need bottoms up approach</td>
<td>1</td>
</tr>
<tr>
<td>Create guidelines for evaluation</td>
<td>1</td>
</tr>
<tr>
<td>Review Utstein Guidelines against other evaluation frameworks (is there an opportunity for convergence?)</td>
<td>1</td>
</tr>
<tr>
<td>Change culture, have quality improvement focus (i.e. overcome barriers to evaluation – future study)</td>
<td>1</td>
</tr>
<tr>
<td>Need to simplify the Guidelines</td>
<td>1</td>
</tr>
<tr>
<td>Aim to improve accountability with evaluation</td>
<td>1</td>
</tr>
<tr>
<td>Data sharing</td>
<td>1</td>
</tr>
<tr>
<td>Standardised template for reporting and sharing information</td>
<td>1</td>
</tr>
<tr>
<td>Provide a theory that underpins an evaluation framework</td>
<td>1</td>
</tr>
</tbody>
</table>
Methodology and structure of Guidelines need to be broader
Agree that evaluation needs to be undertaken, however, do not have an answer
Check with the community or the patient – what do they want?
Need PR person to ‘sell’ the Guidelines
Sam Stratton 2012 – no framework – no publication – have not seen anything to date
Disasters are multi-disciplinary and not just ‘health based, solutions will not come if you just take a ‘health approach’, the methodology and structure needs to be broader – include social, economic, law, etc., such as ALNAP, MSF, IFRC, engage with broader disaster community, work with evaluators, international approach
Is there an ABC approach – such as lifesaving – airway, breathing, circulation?
Requirement for key performance indicators / metrics

5.2.4 Discussion

Four significant findings from this study identified:

1. Health professionals active in disaster evaluation and research are aware of the Utstein Guidelines 2003 and deemed them to have value even though they were not being used to structure disaster health evaluation and research. Three respondents offered that the Utstein Guidelines 2003 were ‘unique, valuable, brilliant, wonderful tool, nothing like them’.

2. The Utstein Guidelines 2003 were used in a variety of formats other than structuring evaluation studies; for example, teaching and as a source of conceptual framework and definitions.

3. While many comments were offered on improving disaster health evaluations, there were no consistent themes identifiable within the comments.

4. A number of similar guidelines were named by respondents, however, most of these were already identified earlier in this thesis.

Comments provided by participants provided depth and insight into their thoughts and perceptions regarding the Utstein Guidelines 2003. These results were surprising, as the literature review identified that the Utstein Guidelines 2003 were being referenced but not used to structure evaluation and research in the disaster setting. There was an expectation that the findings would also support ‘their lack of usefulness’ and this was contrary (and exciting) to the research findings. This encouraged the researcher to review in more detail what key elements of the Utstein Guidelines 2003 would be beneficial in the establishment of the Core Structure for Comprehensive Framework for Disaster
Evaluation Typologies (CFDET 2017). Noting that interview respondents identified the Utstein Guidelines 2003 as ‘too complex’, and ‘too conceptual’ and commented that ‘the book is difficult to read’ and ‘a simplified version would be more useful’ led to a difficult decision point.

As identified by the research findings, participants commented on improvements and areas of clarification that could be made to the Utstein Guidelines 2003 to improve uptake and utilisation in disaster evaluation and research. The intent of the Utstein Guidelines 2003 was admirable and it was disappointing that greater uptake and utilisation had not occurred over the years. Prominent and noteworthy disaster and emergency professionals had spent years collaborating on the Utstein Guidelines 2003 (17, 154, 155). Revised sets of Utstein Guidelines 2003 have been published individually by both the original co-editors (Birnbaum (26-34) and Sundnes (35-57)). If pre-existing flaws remain unidentified, it will not be possible to improve the Utstein Guidelines 2003 and increase their implementation. Further, if existing flaws are not identified and improved, any proposed future frameworks (such as the proposed Comprehensive Framework for Disaster Evaluation Typologies) may potentially repeat the same system errors and also be under-utilised.

No participant offered any strong response on how to improve the evaluation of disaster health. It is recognised that this a difficult question to answer as both ‘disasters’ and ‘evaluation’ are extremely complex areas to navigate (79). Multiple actors and stakeholders are involved and each brings a different philosophy, mindset, approach and terminology. There is ongoing debate on how to define ‘disasters’ (149) and ‘evaluation’ (66). There are multiple definitions for the word disaster (344). If we do not use the same language and do not have a universally accepted, organised methodology for the conduct and reporting of evaluation in the disaster setting, how can we comment on the efficiency, effectiveness and efficacy of the interventions/disaster medical response and relief efforts that are delivered? Are we then doomed to repeating the same failures in disaster management over and over again? Or, if a framework exists but is not utilised or operationalised, how can we move forward and reduce disaster risk and improve disaster management?

While there continues to be many different definitions and interpretations for both disaster and evaluation, it is encouraging to note that there is progress at an international level. This is evidenced by the synchronous adoption of landmark UN agreements such as the
Sendai Framework for Disaster Risk Reduction 2015–2030 (58), the first major agreement of the post-2015 development agenda; the Sustainable Development Goals (SDGs) (59); and the World Humanitarian Summit (WHS) (345), to name a few (346). These published agreements signify a major turning point in the international community’s efforts to tackle disaster challenges, both existing and future (347).

Given these considerations, and the findings that the Utstein Guidelines 2003 had value, even though as a total presentation they were considered ‘too complex’, what was deemed ‘useful’ for this research was to include the ‘Diagrammatic Representation of Definitions’ as outlined on page 35 (refer to Figure 6) and not the entire book ‘Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style’. The core framework of the Utstein Guidelines 2003 (17)(p.35) identified six key elements (hazard, risk, event, impact, damage and disaster), with associated enabling elements of prevention, modification, vulnerability and resilience. These six key elements provided the framework for the next study on the review of Australian disaster events, which is covered in Section 5.3.

5.2.4.1 Limitations

The limitations for this study included delays and small sample size. Delays in obtaining the ‘appropriate’ number of participants arose due to a variety of reasons including the international timeline differences, participants’ work schedules and the impact of humanitarian crises, such as the Ebola crisis in Western Africa (341).

Another delay was caused when the original Utstein Guidelines 2003 were updated and published separately by the two main co-editors, Birnbaum (25-34) and Sundnes (253). When they both concurrently published their respective updated versions in late 2014, international legal action ensued and both versions were temporarily withdrawn by the respective journals. This caused further delay to the researcher as the updated versions were inaccessible for many months.

The sample size was small; this was expected as the field of disaster evaluation is a complex and an emerging field of expertise (79, 148). By extending the timeframe of the semi-structured interviews, an adequate sample was achieved. It is worth noting that the limitations identified do not detract from the research findings.
5.2.5 Summary

The results from Study 1: The Utstein Guidelines 2003 revealed four significant findings. First, that the Utstein Guidelines 2003 were deemed to have value as they were being referenced in peer-reviewed papers and taught at universities, but were not being used to structure disaster evaluation. Second, the Utstein Guidelines 2003 were used for teaching but not to structure disaster evaluations. Third, participants could not offer any suggestions on how to improve disaster health evaluations. Finally, the evaluation frameworks suggested for inclusion had already been identified in the literature review. The findings from this study added weight to the literature review results and the urgent need to develop a comprehensive framework to help structure evaluation typologies in the disaster setting.

These results informed the structure of the next part of the research – a thematic analysis of reviews of Australian disaster events. The sampled Australian reviews were examined to identify whether the core elements from the Utstein Guidelines 2003 (17) were evident and what other key elements should be included in the core structure.

5.3 Study of Reviews of Australian Disaster Events

The Utstein Guidelines 2003 (refer to Figure 6) identified the following six key elements:

- Hazard
- Risk, to include prevention, modification and vulnerability
- Event
- Impact
- Damage
- Disaster (17).

The Utstein Guidelines 2003 provided the framework for this study. As previously identified, the researcher was not able to access the updated 2014 (25, 253) version of the Utstein Guidelines at the time that this study was undertaken.

The literature review analysis and the contemporary evolution of the disaster environment, augmented by an iterative process during the analysis of these reviews, suggested amendments to these key elements as follows:

1. Pre-event status of the society was added as a baseline perspective.
2. Impact and damage from the Utstein Guidelines 2003 were merged into a single element.

3. Rather than the use of ‘disaster’ from the Utstein Guidelines 2003, ‘disruptions / disturbances / changes in functions / consequences’ were substituted, consistent with Birnbaum’s conceptual framework republished in 2014 (25).

4. Response was added as this is a standard phase in the disaster process.

5. Recovery was added as this is also a standard phase in the disaster process.

6. Resilience was added as a strong contemporary theme emerging during the period of this study.

7. Post-event status of the society was added as an outcome perspective.

This study tests if these amended elements were included in a thematic analysis of recent Australian disaster reviews, to evaluate if they had merit or worth for inclusion in the core structure used to underpin CFDET 2017.

A selection of seven Australian disaster events dating from 2006 through to 2014 were chosen to test the validity of the key elements identified. These documents included inquiries (including a commission of inquiry and special inquiry) and royal commissions. Each Australian state and territory is governed by its own legislation and this may be why the documents used different descriptors in the title of their ‘reviews’. For the purpose of this thesis, the term ‘review’ will include any inquiry, report, review or royal commission. For the purpose of this thesis, the term ‘review’ as defined in the Oxford dictionary as ‘a formal assessment of something with the intention of instituting change if necessary’ (348) is used.

It is important to note that these reviews are formal government-sponsored inquiries supervised by experienced authorities who do not necessarily have a background in disasters per se and may not have a background in evaluations. The chairpersons consisted of high-ranking retired military and law enforcement officers and lawyers.

5.3.1 Research Questions

The research sub question relevant to this study is:

- In what ways are the key elements of the Utstein Guidelines 2003, as amended, reflected in selected reviews of Australian disaster events?
The results from this study provide an answer to the second overarching research question: ‘What key elements would support a core structure to develop a comprehensive framework to represent evaluation typologies along the disaster timeline?’

5.3.2 Research Design and Methodology

A thematic analysis of seven reviews of Australian disaster events over 2006-2014 was undertaken to test the inclusion or exclusion of the pre-identified key elements from the Utstein Guidelines 2003, as amended. The key elements, if appropriate, would be considered for inclusion in the establishment of the core structure used to underpin the development of a comprehensive framework of disaster evaluation typologies.

To ensure consistency when mapping key elements from the reviews, a table was created (refer to Table 26). The populated tables are presented and discussed under the results section.

<table>
<thead>
<tr>
<th>Table 26: Key Element Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disaster Event and Year</td>
</tr>
<tr>
<td>Report/Inquiry Name</td>
</tr>
<tr>
<td>Type of Event</td>
</tr>
<tr>
<td>Chairperson</td>
</tr>
<tr>
<td>Pre-Event Status Society</td>
</tr>
<tr>
<td>Hazard</td>
</tr>
<tr>
<td>Risk including prevention, modification, vulnerability and exposure</td>
</tr>
<tr>
<td>Event</td>
</tr>
<tr>
<td>Impact / Damage</td>
</tr>
<tr>
<td>Disruptions / Disturbances / Changes in Function / Consequences</td>
</tr>
<tr>
<td>Response</td>
</tr>
<tr>
<td>Recovery</td>
</tr>
<tr>
<td>Resilience</td>
</tr>
<tr>
<td>Post-Event Status Society</td>
</tr>
</tbody>
</table>

To highlight the iterative nature of this stage of the research, a secondary review of the disaster reviews was undertaken to identify any other key elements included in these reviews considered significant enough to require inclusion in the developing core structure. These additional key elements were complemented by the work of other notable
authors from the disaster and emergency sector such as Kulling et al. (167), Powers and Daily (19), Stephenson (18), Sundnes (35-57) and Birnbaum et al. (26-34).

The intended outcome of this study was a list of key elements structured to provide consistency for future disaster evaluations and research by informing the development of a core structure used to underpin the comprehensive framework.

For this study of reviews of disasters, Australia was chosen, as it faces many extreme natural hazards. There is an upward trend in the incidence and impact of natural hazards in Australia, and this is evidenced by a corresponding upward trend in the costs of overall losses. It is expected that flood and bushfires will increase in frequency, intensified by socioeconomic development and population growth in at risk areas (293, 349).

Australia is the world’s sixth largest country by area, the smallest continental land mass and the world’s largest island with a population of 24.7 million (350). It faces many extreme natural hazards, including bushfires, floods, severe storms, earthquakes and landslides. All states and territories are affected, and these natural hazards have major societal and economic impacts on communities (351). In recent years, there has been an ‘unprecedented number of floods, storms and bushfires that have devastated life and property across Australia’ (352)(p.i); this finding makes studying Australian events relevant and appropriate. This decision is reinforced by the recent establishment by the Australian government of the National Resilience Task Force in the Department of Home Affairs, now headed by Mr Mark Crosweller (353), the former Director General of Emergency Management Australia.

In 2012, the total economic cost of natural hazards in Australia was estimated to have exceeded AUD 6 billion. In 2015, when both financial and social costs are included, it is estimated the total cost of natural hazards exceeded AUD 9 billion, or 0.6% of Gross Domestic Product (GDP). These figures are expected to double by 2030, and to reach an average of AUD 33 billion per year by 2050 (352). Importantly, it is expected that there will be proportionately increasing costs in relation to increased mental health issues, impacts on health and wellbeing and increased disadvantage and vulnerability for those affected by disasters (354).
A recent report by Deloitte Access Economics (2016) commissioned on behalf of Australian Business Roundtable for Disaster Resilience & Safer Communities, argued that the costs of ‘natural disasters’ (for the purpose of this thesis, there is ‘no such thing as a natural disaster’, only ‘natural hazards’ (89, 355)) are underestimated by at least 50%. They identify both tangible and intangible impacts (previous reports only assess tangible impacts), such as the social impacts and subsequent ongoing impacts and associated costs, such as the impact and cost to health, wellbeing, education, employment and community networks (352).

The disasters chosen as illustrative cases in this study occurred in four Australian states over 2006-2014 and covered four types of events, including cyclone, bushfires/wildfires, floods and a public health emergency (refer to Figure 8).

Figure 8: Regionalisation of Disaster Events in the Study

---

3 The Australian Business Roundtable for Disaster Resilience & Safer Communities was formed in 2012 to influence public policy via evidence-based reporting on the unsustainable cost of disasters on life, property and the economy. The roundtable members are the Chief Executive Officers of the Australian Red Cross, IAG, Investa Office, Munich Re, Optus and Westpac. This roundtable has been influential in the Australian government establishing the new National Resilience Roundtable in the Department of Home Affairs, which has as its initial mandate developing a national disaster mitigation strategy, primarily to reduce the costs of disasters in Australia.
Seven Australian events/reviews were selected, governed by different terms of reference (ToR) and chaired by six different people:


All the reviews listed above were open access and included sufficient data for comparison. If we had chosen to look at the Newcastle earthquake that occurred in 1989, there would be insufficient data available in the public domain (352). Closed access or insufficient data makes comparison of findings difficult and results incomplete.

Utilising an ‘all-hazards’ approach to the disaster reviews was important. Studying different hazards helps in the comparison and validation of the modified framework elements. The ‘all-hazards’ approach is a concept in emergency management that acknowledges that while hazards vary in source, such as natural, technological or man-made (363), they often challenge emergency management and/or health systems in similar ways (364).
These events had a significant socioeconomic impact on Australia; in fact, the 2009 Black Saturday Bushfires and the 2011 Queensland Floods have been described as ‘extreme disasters’ (365) and ‘catastrophic’ events (349).


This review, *The Final Report of the Operation Recovery Task Force: Severe Tropical Cyclone Larry (2007)*, was written with a focus on the ‘recovery’ process after Cyclone Larry and was intended to be helpful to people with responsibility for organising recovery in future events/disasters. The first part of the document provides an overview of how the recovery was organised and what was done. Attachments are provided that contain detailed reports from individual agencies of government who assisted in ‘Operation Recovery’. Analysis and views of the recovery process in each specialised area is outlined in the attachments, and as such, provide ready information to counterpart organisations facing a similar situation (356).

Tropical Cyclone Larry struck Far North Queensland on 20 March 2006. This cyclone was classified as a category five cyclone with winds of up to 290 km/h and was one of the strongest cyclones in Australia for nearly a century. An area of over 17,000 square kilometres, an area about one-quarter the size of Tasmania, was significantly damaged. Fortunately, the Bureau of Meteorology (BoM) had accurately predicted its path and gathering power over the several days of its build up. This early warning enabled a great deal of preparation and public notification to take place. Major damage was caused in just a few hours by the fast-moving cyclone, devastating the regional city of Innisfail and nearby townships, leaving thousands of people homeless and damaging 50% of the houses in the region. Approximately 30 were injured; fortunately, there were no deaths reported. This may be partly attributable to the accurate early warning from the BoM. A state of emergency was declared as a result of the damage caused by Cyclone Larry (356).

The economic impact of Cyclone Larry and the damage to critical infrastructure and agriculture were intensive. The economic losses were estimated at nearly AUD 1 billion, with 80–90% of Australia’s banana plantation destroyed, causing a phenomenal 400–500% rise in banana prices across the country (366).
Cyclone Larry caused major damage to flora and fauna, a significant loss of habitat and damage to the Great Barrier Reef. Major coral damage followed as a result of rubbish and debris ending up in Great Barrier Reef (367). Scientists calculated that around 30% of the Wet Tropics World Heritage Area was affected to some extent by Cyclone Larry (368).


The review, 2009 Victorian Bushfires Royal Commission: Final Report Summary (2010), was written with a focus on learning/lessons learned and improving the preparedness and response to bushfires. It consists of the following sections:

- **Final Report – Summary** (357).
- **Volume I – The Fires and the Fire-Related Deaths.** This volume focuses on the progress of the fires that burned in Victoria in January and February 2009 and summarises the results of the commission’s inquiries into the 173 associated deaths (369).
- **Volume II – Fire Preparation, Response and Recovery.** This volume provides the commission’s conclusions and recommendations, focusing on fire preparation, response and recovery. Significantly, it also summarises the lessons learned from such events and how to reduce the risk and impacts of fire with a view to minimising fire-related loss of life in future (370).
- **Volume III – Establishment and Operation of the Commission.** This volume describes administration and processes that the commission adopted during the review (371).
- **Volume IV – The Statements of Lay Witnesses.** This volume contains statements and other material provided by the 100 lay witnesses who shared their experiences of the event (372).

For the purpose of this section, the key review referred to is the 2009 Victorian Bushfires Royal Commission, Final Report Summary published in 2010 (357).

As a result of the bushfires and the significant loss of life, the Victorian Bushfires Royal Commission (VBRC) was established. This was seen to be an important part of ensuring
that lessons were clearly identified and improved practices implemented. The VBRC conducted extensive investigations into the causes of the bushfires, the preparation and responses to the bushfires, and the impact of the bushfires (357).

The 7 February 2009 has become widely referred to as ‘Black Saturday’. The Black Saturday bushfires were a series of bushfires that ignited and burned across Victoria on and around Saturday 7 February 2009, in Australia’s worst ever bushfire disaster (357). The fires occurred during extreme bushfire-weather conditions and resulted in Australia’s highest ever loss of life from a bushfire: 173 people died and 414 people were injured as a result of the fires (357, 369). The lifetime costs of deaths and injuries are estimated to be around AUD 933 million with ongoing mental health issues, increased alcohol consumption in the risky/high-risk levels, increased chronic diseases and non-communicable diseases, and increased family violence and environmental damage expected (352).

The scale of the damage associated with the 2009 Victorian Bushfires is enormous and it is clearly one of the most significant disaster events in Australia’s history. Not least the lives lost, other associated damages included damaged or destroyed infrastructure at a cost of AUD 77 million, lost timber production at a cost of AUD 658 million, over 430,000 hectares burnt and damage to telecommunication assets valued at AUD 25 million (349). Considerable areas of scenic forests and woodlands that form part of Victoria’s natural heritage were also devastated by the fires (357).


The review, Queensland Floods Commission of Inquiry, Final Report, 2012, was written with a focus on all aspects of the response to the 2010/2011 flood events, and in particular, the preparation and planning of all levels of government, emergency services and the community (373).

As a result of widespread and unprecedented flooding across the state of Queensland in 2010/2011, Premier Anna Bligh announced an independent commission of inquiry to examine the flood disaster on Monday 17
January 2011. The terms of reference provided for an independent and thorough examination of the chain of events leading to the floods, and all aspects of the response and the subsequent recovery. The commission was tasked to deliver an interim report in August 2011 (373), covering matters associated with flood preparedness to enable early recommendations to be implemented before the next summer’s wet season. The final report was completed in March 2012 (358).

Australia was subject to widespread and unprecedented rainfall during the summer of 2010–2011. Major flooding occurred in north-west Western Australia, north-west Victoria, northern Tasmania, New South Wales (NSW) and especially, in the north-eastern state of Queensland (374). This particular event was unprecedented due to the extent of the inundated surface area, along with the number of record water levels. Nearly 75% of Queensland was affected by major flooding between November 2010 and January 2011. Staggeringly, this represented a combined area of nearly $1.3 \times 10^6$ km$^2$, comparable to the combined areas of France, Germany, Netherlands, Belgium, Denmark and Norway. Both large-scale flooding and flash floods took place, causing substantial loss of life (375).

Thousands of people required evacuation from towns and cities, including the capital city Brisbane. The flooding affected 70 towns, over 2.5 million people with 36 recorded fatalities directly attributed to the event. It was estimated that 100 people suffered severe injuries, with an additional 200 suffering minor injuries. The lifetime cost of deaths and injuries is estimated to be AUD 321 million, with ongoing mental health issues, increased alcohol consumption in the levels of risky to high risk, increased use of tobacco and medication, increased chronic diseases/non-communicable diseases and increased family violence (352, 376). Damage initially was estimated at around AUD 1 billion and the estimated reduction in Australia’s GDP was approximately AUD 30 billion. Three-quarters of the state of Queensland was declared a disaster zone (358).

The review, Review of the 2010–11 Flood Warnings and Response (2011), was written with a focus on conducting a detailed examination of emergency management arrangements in Victoria in the context of the 2010–2011 floods. Additionally, it provided the opportunity to revisit several of the findings and recommendations of the 2009 Victorian Bushfires Royal Commission. The Floods Review 2011 and the Royal Commission 2010 both revealed significant shortcomings in Victoria’s emergency management arrangements (359, 377).

This review was significant, as it added support to a major reform program announced in September 2011 with the release of the government’s green paper Towards a More Disaster Resilient and Safer Victoria (378) and the companion document Victoria Prepared: An Action Plan (359). Subsequently, the government’s white paper led to the Emergency Management Reform Strategy for Victoria, headed by the newly created Emergency Management Victoria (379). Review of the 2010–11 Flood Warnings and Response identified where improvements were needed to assist local government and community recovery and made 93 recommendations across a broad range of disaster timeline elements, including preparation, response and recovery matters (377).

The floods of late 2010 and early 2011 were among the worst flooding events to occur in Victoria. The state received significant rainfall events between September 2010 and March 2011 and the BoM issued more than 1500 flood watches and warnings. During that period, Victoria recorded the fifth wettest year on record and its wettest year since 1974. The subsequent flooding caused significant damage to both urban and rural communities in regional Victoria. Several communities experienced flooding two or three times in less than four months (380). Significantly, these floods affected the majority of Victorian local government areas, with 16 very severely affected. Major damage was caused to whole communities, roads, rail networks, community infrastructure and the natural environment. Some communities were inundated several times, with repeat damage to homes, businesses and the agricultural sector. As at 31 January 2013, the total cost estimate for relief and recovery was AUD 971 million (377).

The review, *A Shared Responsibility. The Report of the Perth Hills Bushfire February 2011 Review* (2011), was written with a focus on ‘shared responsibility’ to reflect mutual obligations to build community resilience between local and state agencies and the community in the event of an emergency or disaster. The purpose of the special inquiry was to examine all aspects of the bushfire risk management in the Perth Hills area and provide a report on findings and recommendations (360).

*Perth Hills Bushfire Review 2011* was described as a report and a special inquiry. There had been previous reviews on bushfires in Western Australia (WA), including a parliamentary review. Significant power was afforded the special inquiry in documenting its review, and while this event may be ‘smaller’ than previous events, it had a significant impact on WA (360).

On the weekend of 5 and 6 February 2011, two major bushfires affected metropolitan Perth. The fires raged from two directions: the first occurred in the suburbs of Roleystone and Kelmscott, south-east of Perth, while the second occurred in the north-east, affecting on the suburbs of Red Hill, Herne Hill, Millendon, Baskerville and Gidgegannup. Over 500 families required evacuation and 12 people were hospitalised for injuries related to the fires; fortunately, there were no deaths. The fires destroyed 71 homes and damaged a further 39, with the preliminary cost of damage estimated to be AUD 35 million as per the Insurance Council of Australia (381). The significance and impact of this fire was that it caused the single biggest house loss in WA due to a single bushfire event (382). On the following Monday, 7 February 2011, the Western Australian Premier declared the bushfire-ravaged areas of Roleystone, Kelmscott and Red Hill natural disaster zones (360).
The review, *2013 Tasmanian Bushfires Inquiry, Volume One (2013)*, was written with a focus on the immediate causes and circumstances of the fires, all aspects of the emergency response, preparation and planning of all levels of governments, agencies and emergency services, and the transitioning from response to recovery. No special powers were given to the special investigator during the inquiry (361).

The bushfires active on and following 4 January 2013 were the most significant bushfire emergency in Tasmania for many years, probably since the devastating fires in 1967. Over 40 fires were burning, only a handful were classified as ‘contained’ and the danger rating reached ‘catastrophic’ in many areas (361, 383).

The fires forced the evacuation of hundreds of people from the Tasman Peninsula (384). Fortunately, no person was killed in the fires; however, the physical, economic, social, psychological and environmental damage was substantial (361).

When reviewing the losses incurred as a result of the bushfires, it became apparent that much of the damage was not quantifiable even though over AUD 89 million was recorded in insured losses and the estimated overall financial cost is in excess of AUD 100 million (361, 384). Overall, structural damage consisted of 431 properties damaged or destroyed, and of these, 203 residential properties were destroyed (361).

From an agricultural aspect, it was estimated that approximately 10,000 head of livestock, (mainly sheep) were lost and over 650 km of commercial fencing was destroyed (361). The environmental impact of bushfires in Tasmania can be catastrophic, as nearly 20% of the state (approximately 1,500,000 hectares) is classified as a World Heritage Area in recognition of the state’s unique flora and fauna (385). Bushfires in Tasmania are extremely destructive as they damage vegetation unique to Tasmania and damage the soils on which the unique Tasmanian vegetation depends and is unlikely to recover from. It is thought that the fire-sensitive vegetation could take hundreds or thousands of years to recover following bushfires (386).
When reviewing the losses incurred as a result of the bushfires, it became apparent that much of the damage was not quantifiable even though over AUD 89 million was recorded in insured losses and the estimated overall financial cost is in excess of AUD100 million (361, 384).


The review, *Hazelwood Mine Fire Report (2014)*, was written with a focus on the immediate causes and circumstances of the fire and how it spread into the Hazelwood Coal Mine, the adequacy and effectiveness of measures to prevent the outbreak of fire, the adequacy and effectiveness of the response to the fire at Hazelwood Coal Mine by the owner/operator/licensee of the mine, emergency services and other relevant government agencies, and lastly, the impact on the health and wellbeing of the local community (362).

The review consisted of one volume, including an executive summary and 18 recommendations collectively for improvement directed at the Victorian government and the operator of the mine (362, 387). Subsequently, the Victorian government initiated further reviews of the consequences of this event (387-390).

The review of this event brought about significant changes in Victoria in the emergency management space, including the implementation of a new overarching body for emergency management in Victoria and the establishment of an Inspector-General Emergency Management as the assurance/authority for Victoria’s emergency management arrangements (362). The Victorian government continues to engage with the local community regarding health and environmental issues, such as monitoring air quality and the health of the community.

During the summer of 2014, Victoria experienced one of its hottest and driest summers on record and there were multiple significant fires across the state, making response difficult.
In early February 2014, a fire ignited as a result of embers spotting from a nearby bushfire, and took hold in the Hazelwood Coal Mine located in the Latrobe Valley. It is worth noting that open-cut brown coal mines are particularly vulnerable to fire, and to fire that spreads quickly and is difficult to extinguish. Coal mine fires have unique properties that differentiate them for bushfires – they typically burn slowly over a protracted period of time due to the presence of deep-seated, compacted fuel. Bushfires, on the other hand, tend to burn quickly and unpredictably (387). Additionally, the Hazelwood Mine Fire was not just one fire – it was a complex of fires that started as a series of smaller fires that ignited within the mine (362).

The Hazelwood Coal Mine fire burned for 45 days and was the largest and longest burning mine fire to occur in the Latrobe Valley. It required significant resources to bring it under control. The fire sent smoke and ash over the town of Morwell and surrounding areas for much of that time, significantly affecting air quality. This event, which started as a bushfire and became a fire in a mine, then evolved into a public health emergency for the local communities (362).

The Hazelwood Mine Fire had no direct impacts on the local community and, unlike many disasters, there were no deaths as a direct result of the fire. The health impact, however, was significant to the local communities as a result of six weeks of smoke and ash surrounding the area. The fire is believed to be responsible for a spike in deaths and the premature deaths of at least 11 people, and will require further monitoring (391). The Department of Health and the Victorian government were criticised in the Hazelwood Mine Fire Inquiry (2014) for their failings in the way the event was managed (362). The Hazelwood Health Study was commissioned to identify potential long-term health outcomes for people who may have been exposed to smoke from the mine fire; this work is expected to continue for several years (392).

5.3.3 Results

The analysis of these reviews is provided in Table 27.

All chosen elements of the Utstein Guidelines 2003, as amended, were present in the seven reviews as demonstrated in Table 27 where source = 1. This table also shows six additional key elements with a source = 2 that were identified and deemed important in the establishment of the core structure.
Table 27: Key Elements to be Included in the Establishment of the Core Structure

<table>
<thead>
<tr>
<th>Source*</th>
<th>Cyclone Larry</th>
<th>Victorian Bushfires</th>
<th>Victorian Floods</th>
<th>Queensland Floods</th>
<th>Perth Hills Fire</th>
<th>Tasmanian Fires</th>
<th>Hazelwood Mine, Victoria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Event</td>
<td>Cyclone</td>
<td>Fire</td>
<td>Floods</td>
<td>Floods</td>
<td>Fire</td>
<td>Fire</td>
<td>Fire / Public Health Emergency</td>
</tr>
<tr>
<td>Chairperson</td>
<td>Gen P Cosgrove AC, MC (Retired)</td>
<td>Hon B Teague AO R McLeod AM S Pascoe AM</td>
<td>N Comrie AO, APM</td>
<td>Hon Justice C E Holmes</td>
<td>M J Keelty AO, APM</td>
<td>Special Investigator M Hyde</td>
<td>Hon B Teague AO Prof J Catford S Petering</td>
</tr>
<tr>
<td>Pre-Event Status Society</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hazard</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Risk including prevention, modification, vulnerability and exposure</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Event</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------</td>
<td>--------------------</td>
<td>--------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Impact / Damage</td>
<td>1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Disruptions / Disturbances / Changes in Function / Consequences</td>
<td>2</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Response</td>
<td>2</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Recovery</td>
<td>2</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Resilience</td>
<td>2</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Post-Event Status Society</td>
<td>2</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note: * Source: 1 Utstein, 2 Wong et al.
5.3.4 Discussion

The reviews of seven Australian disaster events cover four states, four types of events and the period 2006–2014. The study was aimed to test in what ways the key elements of the Utstein Guidelines 2003, as amended, were reflected in the selected reviews. These reviews did not follow any published or recognised framework to structure their findings. There was also no apparent consistency in the reviews as to the order of the identified elements.

The findings demonstrate preliminary validation of the key elements to be included in the establishment of the core structure to support consistency in the evaluation of disaster health. This was exciting, as it added weight to the inclusion of these elements in the core structure. This study is unique in that it provides an analysis of how the key elements reflect reality.

5.4 Summary

The findings from both Study 1 (Utstein Guidelines 2003) and Study 2 (Reviews of Australian Disaster Events) address the second overarching research question by providing a list of 10 key elements that will be used to support the establishment of a core structure to underpin a comprehensive framework that maps evaluation typologies along the disaster timeline.
Chapter 6: Establishment of the Core Structure

6.1 Introduction

The studies in the preceding chapter identified the following key elements that will be included in the core structure to underpin the comprehensive framework:

- Pre-event status of society
- Hazard
- Risk including prevention, modification, vulnerability and exposure
- Event
- Impact / damage
- Disruptions / disturbances / changes in function / consequences
- Response
- Recovery
- Resilience
- Post-event status of society

Each of these elements are discussed in detail in the forthcoming sections. These elements are important as they provide structure, consistency in terminology and aid understanding of the disaster timeline. This provides guidance to those responsible for implementing disaster management or emergency management arrangements (393).

6.2 Establishment of the Core Structure

The conceptual framework that underpins the core structure is the disaster timeline. For this research, the terms disaster timeline, disaster phases and disaster cycle can be used interchangeably; however, the preferred term moving forward is ‘disaster timeline’.

All disasters occur in phases through time (394). There are many ways of expressing the disaster timeline. At its most basic level, disaster experts have identified three main phases (pre-event – warning, event – impact and post-event – consequences) (394, 395). Other experts prefer to use four main phases (preparedness/mitigation, planning, response and recovery) to describe the disaster timeline (396).
Importantly, the concept of phases has been used for over 80 years and evolved in an attempt to describe and promote the study of disasters (344, 395, 397, 398). Given that disasters occur and continue to occur, the concept of phases (while simplistic) helps to organise the practice of disaster and emergency management and provides a coherence to an analysis not otherwise possible (344, 398, 399). Phases are identified by their properties, rather than a time component (19), and each of the phases requires a different form of intervention/s (395). For example, disaster management actions or interventions could include pre-event prevention or mitigation strategies, emergency responses delivered during the event, and rehabilitation and development strategies post-event (145).

Regardless of how many phases are identified, the concept that each of these phases varies in length of time depending on the nature of the disaster and the nature of the response remains the same. All the divisions of phases are artificial, as one phase of a disaster may merge with another (399). The names applied to these phases may differ from country to country, organisation to organisation or community to community (394, 398). Importantly, the disaster timeline was developed with the intent of providing a coordinated and consistent approach to the management of disasters (400) and reducing the associated losses.

Other frameworks or guidelines influential in the building the core structure included work by Stephenson (18), Powers and Daily (19), Kulling et al. (167), Debacker (101), Fattah (20-22), Sundnes (35-57) and Birnbaum et al. (26-34). While Fattah identified more than 10 frameworks, she also discovered that none had been validated and they were not commonly used to structure evaluations in the disaster setting (21).

The core structure consists of three important layers:

1. Layer 1 (refer to Figure 9) – provides a preliminary and simplistic view of the disaster timeline
2. Layer 2 (refer to Figure 12) – provides an expanded view of the disaster timeline
3. Layer 3 (refer to Figure 13) – introduces the concept of strengthening resilience as an overarching theme.

It is important to note that ‘the disaster timeline must be viewed as a continuum and not as a linkage of any of the semi-discrete’ elements (401)(p.279) within the layers as described in the following sections.
6.2.1 Layer 1

At the most basic level, this layer has three core phases represented by a pre-event phase, an event phase and the post-event phase of a disaster or emergency (395). Twigg describes this sequence as ‘a linear operational model’ that divides the disaster cycle into three phases (395). The concept of phases has been used for over 80 years and helps to describe and promote the study of disasters (344, 395, 397). Importantly, phases help to organise the practice of disaster and emergency management and provide a coherence to an analysis not otherwise possible (344). Phases are identified by their properties, rather than a time component (19), and each of the phases requires a different form of intervention/s (395). For example, disaster management actions or interventions could include pre-event prevention or mitigation strategies, emergency responses delivered during the event, and rehabilitation and development strategies post-event (145).

\[\text{Figure 9: Core Structure, Layer 1}\]

6.2.1.1 Pre-Event Phase

The pre-event phase acts as a baseline for determining the status of preparedness and other baselines to enable an understanding of the damage and disturbances in function
that occurs as a result of an event (19, 40, 45, 54). Knowing the pre-event status of the affected community is useful in:

- understanding the level of preparedness
- assessing the severity of the change after the event
- monitoring and evaluating the response delivered as a consequence of the damage/event (402).

For the purpose of this research, the terms ‘society’ and ‘community’ are used interchangeably.

Therefore, it makes sense that the pre-event status of the community and/or any of the basic societal functions (BSFs) is known, as this information is essential in all disaster management and evaluations. Although this information is often fragmented, it should be ideally compiled prior to an event or a disaster occurring (19).

BSFs (17) are the major functional components of a society that may be affected either directly or indirectly by an event resulting in a disaster. They are useful in structuring the functional status of the community at any stage of the disaster timeline and are important in societies as they fulfil fundamental (key) roles, including the following 14 elements:

1. Medical
2. Public health
3. Sanitation and water supplies
4. Shelter and clothing
5. Food
6. Energy supplies
7. Search and rescue
8. Public works and engineering
9. Environment
10. Logistics and support
11. Security
12. Communications
13. Economy
14. Education (17).

Each of these elements comprises many interdependent components. Assessments of the functional status of all, some or one of the BSFs provides a snapshot of a society’s
situation at a given point in time (37, 40, 42). BSFs are linked by a coordination and control function (ideally) provided by the respective governments (17, 42). As no function operates in isolation from the others, information about the contemporaneous status of BSFs is important to gain a better understanding of functional losses and the impact of interventions (positive or negative; intended or unintended) (37). Standards for many of these BSFs have evolved from a variety of sources such as the SPHERE Guidelines (403) and Core Humanitarian Standards (CHS) (279). Comparable frameworks to BSFs include the UN Office for the Coordination of Humanitarian Affairs (OCHA) Cluster Coordination (refer to Figure 10) (404) and New Zealand’s Holistic Framework for Disaster Recovery (refer to Figure 11), which provides general principles and concepts of recovery management (405).

Figure 10: OCHA Cluster Coordination (404)
6.2.1.2 Event Phase

The event phase is the actualisation of a hazard (19) that has the potential to adversely affect people, their community and their environment (17, 172). Events may be caused by natural or human-made hazards, or a combination of the two (26). Each event has the following characteristics:

- type of onset (slow onset or sudden onset) (87)
- complex humanitarian emergency (406-408)
- intensity (the integral of the amplitudes over a given period of time) (17)
- duration (brief – only seconds to minutes; short – hours to days; intermediate – days to weeks; prolonged – months to years; some events cross over into more than one category (17))
- scale (small scale or large scale) (87)
- magnitude (17).

Further, events can be considered primary or secondary in nature; for example, an earthquake (primary event) causing a tsunami (secondary event). In this illustration, the secondary event may be responsible for more damage than the primary event (26).
It is important to note that confusion has been caused by the incorrect use of the term ‘event’ to describe a disaster and vice versa. Even though there is no international consensus regarding the definition of a ‘disaster’ (149, 409), for the purpose of this research, the UNISDR definition of disaster is used given its widespread use and community-based approach – ‘a serious disruption of the functioning of a community or society due to hazardous events interacting with conditions of vulnerability and exposure, leading to widespread human, material, economic and environmental losses and impacts’ (410)(p.11). The definition is further expanded by identifying that disasters are a type of hazardous event in which there is significant disruption of the function of all or part of the community/society. The impact may test or exceed the capacity of a community or society to cope using its own resources, and therefore, may require assistance from external sources (410).

It is also important to note that the event does not necessarily lead to functional change or a disaster. An example of this could include a bushfire occurring in a remote area where there is limited damage and which does not lead to a ‘disaster’ for the local community. The local community is able to manage the disruptions, disturbances, change in functions or consequences and does not require outside assistance.

Identifying the event or events that cause a disaster is important and will allow for better comparisons between events. It will also assist in the discussion on the commonalities and differences that may exist, and the information gathered can be used to prepare for subsequent events. While each of the phases can be identified individually and they are a linear representation of time, their timing is not necessarily sequential and the phases can overlap (19, 39).

6.2.1.3 Post-Event Phase

While it is relatively easy to find definitions for ‘pre-event’ and ‘event’, it is more difficult to find a definition for the ‘post-event’ phase. For the purpose of this research, the ‘post-event phase’ is defined as the time period that occurs directly after an event manifests. The post-event status of society includes:

- impact of the event
- subsequent damage
- disruption to the community as a result of the damage
- change in functions as a result of the damage (consequences)
• response phase
• recovery phase.

The damage, changes in function and disaster that may occur after an event are relevant only in terms of their impact on society (19). Impact has been defined as ‘the actual process of contact between an event and a society or society’s immediate perimeter; an effect of influence’ (17)(p.153). In a contemporary setting, the impact of an event is now spoken of in terms of consequences and the associated consequence management (411).

6.2.2 Layer 2

Layer 2 provides a more detailed representation of the pre-event, event and post-event phases described in Layer 1, enhanced by the addition of extra phases (refer to Figure 12). This layer consists of:

• pre-event status of society
• hazard
• risk reduction including prevention, modification, vulnerability and exposure
• event
• impact / damage
• disruptions / disturbances / changes in function / consequences
• response
• recovery
• resilience
• post-event status of society.

It is based on a modified representation of the Utstein Guidelines 2003, as amended, and can be found in ‘Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style’ (17), as previously discussed.
6.2.2.1 Pre-Event Status

The pre-event status in this layer is reflective of the pre-event phase discussed in Layer 1.

6.2.2.2 Hazard

UNISDR 2017 defines hazard as a ‘process, phenomenon or human activity that may cause loss of life, injury or other health impacts, property damage, social and economic disruption or environmental degradation’. Hazards can be:

- natural (associated with natural processes and phenomena)
- anthropogenic (human-induced hazards)
- socio-natural (combination of the above, and including environmental degradation and climate change) (87).

The Sendai Framework (58) also recognises that hazards can be:
Hazards can be further described as single, sequential or combined in their origin and effects. Each hazard is characterised by location, intensity, probability and likely frequency (412).

It is worth noting the ongoing discussion around the term ‘natural disaster’. The phrase ‘natural disaster’, although widely used, often causes confusion and has been the subject of much debate. Strictly speaking, there is no such thing as a natural disaster – there are only natural hazards (413, 414) – and the difference between a hazard and a disaster is an important one. At its most basic, a disaster is an event that overwhelms a community’s capacity to cope. The impact of the disaster is heavily influenced by the degree of the community’s socially created vulnerabilities to the hazard, exposure and response capability (89, 415, 416), as evident in Hurricane Katrina (414, 417).

6.2.2.3 Risk Reduction

This thesis was emerging concurrently with the Sendai Framework and it is therefore appropriate to consider the broad issue of risk, which includes hazard recognition, risk management, prevention and mitigation under the general theme of disaster risk reduction (DRR) in the setting of the updated Sendai Framework (58).

UNISDR defines DRR as the ‘concept and practice of reducing disaster risks through systematic efforts to analyse and manage the causal factors of disasters, including through reduced exposure to hazards, lessened vulnerability of people and property, wise management of land and the environment, and improved preparedness for adverse events’ (418).

DRR is aimed at understanding disaster risk, preventing new and reducing existing disaster risk, and managing residual risk. All of these contribute to strengthening resilience and therefore to the achievement of sustainable development (87), in keeping with the Sendai Framework (58, 419) and the SDGs (59).
Other important risk reduction considerations include the priorities of disaster risk governance and disaster risk management. Based on the Sendai Framework, Basabe defines disaster risk governance as ‘the system of institutions, mechanisms, policy and legal frameworks and other arrangements to guide, coordinate and oversee disaster risk reduction and related areas of policy’ (p.13), while disaster risk management is the ‘application of disaster risk reduction policies, processes and actions to prevent new risk, reduce existing disaster risk and manage residual risk contributing to the strengthening of resilience’ (p.10). There is a need for risk reduction practices to inform policy and practice with the intent to reduce avoidable socioeconomic costs to communities.

6.2.2.4 Event

The event status in this layer is reflective of the event phase discussed in Layer 1. The size of the event phase shown in Layer 2 has been made smaller for inclusion in the diagram and is not intended to reflect the size of the event.

6.2.2.5 Damage Consequence Management

The previously mentioned BSFs provide a useful framework to undertake damage assessments. From an operational perspective, the OCHA Clusters provide a companion framework to not only structure response, but to measure the damage and functional consequences within and across clusters. Once the event has occurred, damage can be assessed. Damage is defined as the negative results from the impact of an event. This can affect the community and environment and may manifest in multiple ways and forms. Events may produce damage that may or may not be of sufficient magnitude to result in a disaster. It is the amount and characteristic of the damage that results from an event, tempered by the place of occurrence, society and culture, level of development, and degree of preparedness that determines whether an event results in a disaster. Damage may involve humans, animals and/or the environment.

Damage from an event may be ‘direct’ or ‘indirect’. Direct damage is damage that occurs as a result of the energy released by the event, such as building collapses and deaths as a result of an earthquake. Indirect damage is damage that is not due directly to the forces imparted by the event, and could include the occurrence of cardiac-related deaths after a disaster, increased levels of domestic violence or suicide of victims. Tragically, in some disasters, this has the potential to be more severe than that due directly to the forces exerted by the event.
6.2.2.6 Disruptions, Disturbance, Changes in Function and Consequences

Disruption and disturbances are hard to quantify even though they are a common element in most disasters (422). The nature of disruption can vary from disaster to disaster and affect individuals and communities differently. To be able to measure disruption, pre-event or baseline information is required as a point of reference to quantify the changes in functions that have occurred.

Changes in functions can be summarised as the ‘change in the ability of a major functional component of a society to operate due to damage caused by an event’ (171)(p.16). Change can also be described as ‘disturbances’ in function (172). The consequences of the damage may also be hard to quantify and may be cascading in nature, such as the Japanese earthquake and tsunami (423).

6.2.2.7 Response

Response (respond and adapt) is defined as those ‘actions taken directly before, during or immediately after a disaster in order to save lives, reduce health impacts, ensure public safety and meet the basic subsistence needs of the people affected’ (87). Disaster response is primarily focused on the immediate (rescue, surge) and short-term (relief) needs of the community and is sometimes called disaster relief. Effective, efficient and timely response relies on disaster-risk-informed preparedness measures, including the development of the response capacities of individuals, communities, organisations, countries and the international community (87).

It should be noted that the division between the response phase and the subsequent recovery phase is not clear-cut. Some response actions may extend well into the recovery phase (87), and, as an example, could include the supply of temporary housing, and water, sanitation and hygiene (WASH).

6.2.2.8 Recovery

The core structure shows response and recovery occurring at the same time. While there is some debate around this timing, it is important to note that response and recovery are not truly linear processes (424) and that recovery commences at the same time as response, as supported by the standards of Emergency Management Victoria (425).
Recovery in the disaster setting is a complicated and complex problem (424) and is probably the most least understood of all the phases in the disaster management cycle (400). Recovery is also regarded as the least researched phase of the disaster timeline (314, 426, 427).

There are many definitions for the term ‘recovery’; however, the definition that has been accepted for this thesis is that offered by the UNISDR, which defines recovery as ‘the restoring or improving of livelihoods and health, as well as economic, physical, social, cultural and environmental assets, systems and activities, of a disaster-affected community or society, aligning with the principles of sustainable development and “Build Back Better” (as emphasised in the Sendai Framework (58)), to avoid or reduce future disaster risk’ (87). Build Back Better (BBB) is a guiding principle that aims to take advantage of the recovery, rehabilitation and reconstruction phases, with an intent to increase resilience of communities and improve living and environmental conditions by integrating disaster risk reduction measures throughout the entire process (87). Conceptually, BBB means that there should be improvement in both the process and results of post-disaster humanitarian interventions as compared with the pre-event status. Improved conditions include stronger and safer buildings and settlements, a better social and ‘political context with strengthened capacities, greater equity, and enhanced economic and social development’ (424)(p.74). Importantly, the concept of BBB as a goal implies doing more than just restoring what existed in the pre-event status of society. This reconstruction approach aims to reduce vulnerability and improve living conditions while promoting more effective reconstruction. Following disasters, BBB addresses the importance of enhancing community resilience and identifies what is considered successful recovery (424).

Norman further expands on the definition of ‘recovery’ and discusses the holistic regeneration of a community following a disaster. This perspective is also supported by the 2012 work of Helen Clark in her role as the United Nations Development Programme (UNDP) Administrator. Clark believed that people, infrastructure, communities and institutions must be equipped to withstand external shocks and that recovery included community development (428, 429).

A final concept that is part of the recovery phase is reconstruction. This has been defined by the UNISDR as ‘the medium- and long-term rebuilding and sustainable restoration of resilient critical infrastructures, services, housing, facilities and livelihoods required for
the full functioning of a community or a society affected by a disaster, aligning with the principles of sustainable development and BBB, to avoid or reduce future disaster risk’ (87).

6.2.2.9 Early Recovery and Persistent Recovery

The recovery process is best described as a sequence of interdependent and often concurrent activities that progressively advance a community towards a successful recovery including sustainable resilience-based development (430). It begins with pre-disaster preparedness and includes a wide range of planning activities (431). As noted previously, recovery occurs concurrently with the relief and response phases and is often referred to as early recovery (432). Recovery continues for variable periods following an event but often extends over years; in a contemporary sense, the latter timing is referred to as persistent or sustainable recovery (433).

6.2.2.10 Post-Event Status

The post-event status in this layer is reflective of the community status ultimately achieved following the event. The current trend to measure this element of recovery includes outcomes, indicators and standards of successful recovery. A recent publication, *A Monitoring and Evaluation Framework for Disaster Recovery Programs, Version 2 – May 2018*, supported by the Australian government, provides a useful framework for the evaluation of disaster recovery frameworks (434).

6.2.3 Layer 3

The third layer of the core structure introduces the concept of ‘strengthening resilience’ as an overarching theme (refer to Figure 13). The inclusion of this emerging, international imperative embraced within the Sendai Framework (58) enhances and enriches the core structure of CFDET 2017. Contemporary to the evolution of this thesis has been the growth of resilience in the disaster and humanitarian settings (435). Key elements of strengthening resilience include actionable areas, as well as internationally identified and significant frameworks (346).
The actionable areas that provide the building blocks to strengthening resilience include:

- system structure, governance, coordination and leadership (436-438)
- a ‘culture of safety’, with consideration of risk reduction, prevention and mitigation (58, 428, 429, 439)
- reducing exposures and vulnerabilities, building anticipative, absorbing and adaptive capacities, and promoting community development (440).

The strengthening resilience element identifies the following five contemporary and influential international frameworks available prior to 2016:

1. **Sendai Framework for Disaster Risk Reduction 2015–2030 (58).** The Sendai Framework is a 15-year, voluntary, non-binding agreement that focuses on the adoption of measures that address three dimensions of disaster risk – exposure to hazards, vulnerability and capacity, and the characteristics of hazards – to prevent the creation of new risk, reduce existing risk and increase resilience to withstand residual risk. It builds on achievements and elements established under its predecessor agreement, the Hyogo Framework for Action: Building the
Resilience of Nations and Communities 2005–2015 (419). It introduces a number of important innovations, including a stronger emphasis on disaster risk management as opposed to disaster management; it supports risk-informed policy decision making, and the allocation of adequate resources to prevent the creation of new disaster risks (441). The Sendai Framework underscores that disaster risk reduction is essential to achieve sustainable development (58).

2. Sustainable Development Goals (SDGs) (2015) (442). SDGs form the 2030 Agenda for Sustainable Development and represent the latest global targets in pursuit of poverty reduction, sustainable development and peace. Resilience features in two goals and eight targets linked to poverty, built infrastructure and human settlements, agricultural production and vulnerability to climate extremes and disasters. This is a marked shift from the goals’ predecessor, the Millennium Development Goals (MDGs) (443), in which resilience did not appear all (444).

3. Climate Change Conference (COP21) (2015) (112). COP21 featured resilience as an integral component of climate change adaptation. It gave resilience a more prominent role than in previous climate change agreements and placed an emphasis on the resilience of, and links between, socioeconomic and ecological systems (444).

4. Global Facility for Disaster Reduction and Recovery (GFDRR) Framework, also known as the Post Disaster Needs Assessment (PDNA) Recovery Framework (445). To manage and reduce disaster risk, GFDRR believes that meaningful risk reduction can only be achieved by influencing development by identifying and managing risks, and designing more resilient plans and projects (445, 446).


The international frameworks mentioned above have global significance (346). In 2015 and 2016, the world’s governments agreed on the Sendai Framework for Disaster Risk Reduction, the SDGs and the Paris Agreement on Climate Change (COP21) (444). These frameworks are internationally important as they identify a set of goals and targets that, if achieved, will create a future where significant progress will have been made on disaster, sustainable development, climate and humanitarian challenges (346, 444). Further, there are additional emerging measures of resilience, such as:
• the City Resilience Framework and Index, which is an international collaboration that consists of comprehensive tools to help cities assess and measure their resilience (448-450)

• the Australian Resilience Index, which was created to identify an index of disaster resilience from an Australian perspective (451-453) and in keeping with the Australian government’s *National Strategy for Disaster Resilience* (NSDR) (436)

• the new Australian Government National Resilience Taskforce (454), which is tasked with improving resilience in the disaster setting, with a particular focus on natural hazards (353, 455).

6.3 Discussion

The complete core structure (refer to Figure 14) consists of the three layers described in section 6.2. Significantly, this core structure takes the key elements identified by the Utstein study and Australian disaster events study, as described in Chapter 5, and maps them along the disaster timeline. The addition of strengthening resilience highlights the contemporary global initiatives to reduce disaster risk. This foundation will be used to support the additional constructs discussed in Chapter 7.
When starting to build a comprehensive framework, it is important to establish a strong foundation to which all other entities will be related. This chapter detailed the procedures undertaken to support the establishment of the core structure based on the results from Studies 1 and 2 (which were covered in Chapter 5). The results of these studies provided a list of 10 key elements to be included in the core structure. The finalised core structure consists of three layers including a preliminary and simplistic view of the disaster timeline, an expanded view of the disaster timeline and the concept of strengthening resilience as an overarching theme.

The establishment of the core structure is part of the solution to the second research question in that it provides a diagrammatic representation of the key elements that support a core structure to develop a comprehensive framework to represent evaluation typologies. This core structure forms the foundation for the development of the comprehensive framework for disaster evaluation typologies, discussed in Chapter 7.
Chapter 7: Development of the Comprehensive Framework for Disaster Evaluation Typologies 2017

7.1 Introduction

There is currently no single, internationally accepted, comprehensive framework to structure disaster evaluation typologies. Chapters 5 and 6 established the core structure for such a comprehensive framework (refer to Figure 14), which includes the disaster timeline and the concept of strengthening resilience. Chapter 7 builds on the core structure by detailing the identification of various evaluation typologies and overarching themes that result in the development of a comprehensive framework for disaster evaluation typologies.

The resulting comprehensive framework has been developed to improve the approach to undertaking and reporting disaster evaluations, with the aim of reducing preventable mortality and morbidity in future events (253, 422). This also supports the improvement of disaster health science by sharing information about past experiences and improving future response efforts (85). To mitigate some of the previous issues identified, this new framework will define and provide consistent terminology along with standards for reporting across the different phases of a disaster. This, in turn, will provide comparability to better understand the process, outcomes and impacts of the efficacy and efficiency of interventions.

The Comprehensive Framework for Disaster Evaluation Typologies 2017 (hereafter referred to as CFDET 2017) is intended to be used in the disaster and humanitarian settings. Importantly, CFDET 2017 is not an evaluation typology in itself, nor is it a list of evaluation studies; rather, it is a framework that outlines different disaster evaluation typologies that can be mapped along the disaster timeline.

7.2 Research Questions

Existing frameworks that support consistent reporting in a disaster lack validation, consistent terminology and standards for reporting across the disaster timeline. The research presented in Chapters 2 and 4 indicated that there was an opportunity for the development of a comprehensive framework to structure disaster evaluation typologies. Chapter 4 addressed the first research question ‘What does the peer-reviewed and grey
literature report on comprehensive frameworks and evaluation typologies in the disaster setting?” The results of the literature review informed Chapter 5. This identified the key elements to support a core structure to develop a comprehensive framework to represent evaluation typologies. Chapter 6 continued this work by taking the key elements and establishing the core structure. This chapter builds on this work by mapping evaluation typologies along the disaster timeline.

Specific questions related to the development of the comprehensive framework are:

- What process enables the development of a unifying, comprehensive framework for disaster evaluation typologies within the disaster setting?
- How can these evaluation typologies be mapped to support agreement on the identification, structure and relationships between various evaluation typologies within the disaster setting?

7.3 Research Design and Methodology

Action research underpins this part of the study, as it has a focus on quality improvement, professional engagement and is an iterative process. When combined with mixed methods research, action research has the ability to produce sounder scientific research as it integrates professional engagement (456, 457), which is imperative to the success and utilisation of the final framework.

7.3.1 Literature Review

An analysis of the literature review findings (refer to Chapter 4) allowed the identification of relevant evaluation typologies to include in the development of a single, unifying framework. The findings of the literature review were categorised into different evaluation typologies that were currently being used, or could be used, in the disaster setting. The range of evaluation typologies were mapped along the disaster timeline. The focus of the literature review was on the typology of the evaluation methodology rather than the disaster event itself. A wide scatter of different evaluation typologies exists and there was no fundamental classification found to assist during the search process. Disaster evaluation typologies were poorly defined, lacked consistent definitions and terminology, lacked structure and were more often than not found in the grey literature. To provide structure the evaluation typologies were categorised into six themes (refer to Figure 15).
7.3.2 Framework Construction

Evaluation typologies identified during the literature review were categorised into six themes. These were further classified into two overarching themes and four evaluation typologies, each of which are discussed below and will be mapped along the disaster timeline, demonstrating relationships and interdependencies.

7.3.2.1 Overarching Themes

Two overarching themes were identified after reviewing the results from the literature review discussed in Chapter 4. It is important to note that these overarching themes are not in themselves evaluation typologies but have significant influence on the development of the comprehensive framework. They highlight the standards that should be applied when undertaking any disaster evaluation, regardless of where it is positioned along the disaster timeline. The two overarching themes are:

- accountability
- evaluation standards and evidence:
  - evaluation standards and guidelines
  - evidence-based reviews and registries
  - knowledge management.

These themes are extremely important in the disaster and humanitarian sectors and support the evaluation standards that underpin the conceptual framework of this thesis.
7.3.2.1.1 Accountability

Accountability to donors, stakeholders and beneficiaries is a cross-cutting theme across all phases of the disaster timeline and is applicable to every evaluation undertaken in the disaster setting (458). Figure 16 shows this theme as situated at the top of the framework, spanning the entire disaster timeline.

The inclusion of accountability recognises the call at an international level for greater accountability in disaster and humanitarian settings. More recently, there has been mounting pressure to strengthen quality, accountability and learning practices, while also ensuring transparency (15, 459-463). The lack of an accepted definition for accountability in the humanitarian context remains a challenge (459). The term ‘accountability’ represents a whole range of different concepts and principles including ensuring public institutions and other organisations fulfil their duties to those affected by disasters (395).

ALNAP is an example of an international organisation dedicated to improving humanitarian performance through accountability and increased learning (464).

![Figure 16: Core Structure and Accountability](image-url)
For the purpose of this research, the term ‘accountability’ is defined as the means by which power is used responsibly. This includes consideration of the views of all interested parties, including donors, stakeholders and beneficiaries (279).

Examples of accountability evaluations include, but are not limited to:


7.3.2.1.2 Evaluation Standards and Evidence

The second overarching theme, evaluation standards and evidence, is also not in itself an evaluation typology. Evaluation standards and evidence is a collective term covering:

- evaluation standards and guidelines
- evidence-based reviews and registries
- knowledge management.

Each of these are important cross-cutting themes that are relevant through the entire disaster timeline (refer to Figure 17). They provide an important opportunity to strengthen data gathering, inform decision making during disasters and build disaster science (160). Importantly, Gossip and colleagues identified that dissemination of evaluation findings was one of the weakest points in the monitoring and evaluation chain in the disaster context (85). There is an urgent need for a shared culture of learning within the disaster sector, including the sharing of lessons learned from interventions delivered during a disaster, and standardised format and terminology to improve disaster response efforts (85) while reflecting the priority for action embraced by the Sendai Framework (58).

Research methods as a paradigm have been excluded from the framework. It is recognised from a theoretical perspective that ‘research’ is on a continuum with ‘evaluation’, but the focus of this work is on evaluation typologies. Research may identify issues, answer questions or add to disaster health knowledge, but this is a conceptual approach in its own right and may cut across evaluations (467). The focus of this research is on evaluations (and evaluation typologies) and therefore includes a systematic approach to inform
decision making regarding the value of a program or intervention and includes some form of judgement.

Figure 17: Core Structure and Evaluation Standards and Evidence

Evaluation Standards and Guidelines

Evaluation standards and guidelines are a roadmap that provide guidance on the manner in which evaluations should be undertaken. They assist in the delivery of high-quality evaluation reports that have a credible basis and are suitable for use in making important decisions (468). Importantly, they play an integral part in the provision of quality humanitarian assistance in the disaster setting (469).

Evaluation standards and guidelines include generic evaluation standards such as:

- general evaluation theorists, including but not limited to Patton (119), Scriven (470) and Stufflebeam (471)
• American Evaluation Association (AEA), *Program Evaluation Standards, Summary Form*, 2016 (473), which are based on earlier work by Yarbrough (103)
• Organisation for Economic Co-Operation and Development (OECD), Development Assistance Committee (DAC), *Principles for Evaluation of Development Assistance*, originally published in 1991 (474) and updated in 2008 (475)

Additionally, there are disaster-specific evaluation standards and guidelines published by many NGOs, that include but are not limited to:

• International Federation of Red Cross and Red Crescent Societies (IFRC), *Project/Programme Monitoring and Evaluation (M & E) Guide*, 2011 (477)
• Save the Children, *Evaluation Handbook*, 2012 (478)
• International Initiative for Impact Evaluation (3ie), *Principles for Impact Evaluation*, 2016 (480)
• Centers for Disease Control and Prevention (CDC), *Program Performance and Evaluation Office – Program Evaluation*, 2016 (482)
• Australian Aid, Department of Foreign Affairs and Trade (DFAT), *Monitoring, Evaluation and Learning Framework (MELF)*, 2016 (78).

Evaluation standards and guidelines also include guidelines for responsible and ethical conduct in undertaking evaluations, which include but are not limited to:

• Australian Council for International Development (ACFID), *Guidelines for Ethical Research and Evaluation in Development*, 2015 (105)
Evidence-Based Reviews and Registries

The evidence to support informed decision making in disasters is limited and tends to accumulate slowly post-event (96, 483). Evidence collection and objective analysis is important in the disaster setting as it can be used to inform and strengthen policies, programs and practice. Evidence-based reviews and registries can promulgate best practice in real time and aid in and improve disaster response, thereby saving lives and reducing socioeconomics costs to affected communities (484).

Evidence-based reviews and registries include meta-evaluations, systematic reviews, other types of literature review typologies and registries of evaluation reports. The level and quality of evidence in this setting has recently been reviewed by Clarke and Darcy in Insufficient Evidence? The Quality and Use of Evidence in Humanitarian Action – ALNAP Study (98). Despite improvements over the last 20 years, they identified that there remains room for further development in the quality and use of evidence in the humanitarian setting. The authors also suggest that ‘evidence matters: the use of good quality evidence improves the effectiveness and accountability of humanitarian action, and is in accordance with humanitarian ethics and principles’ (98)(p.6). This is also in keeping with the evaluation standards.

Meta-evaluations are systematic and formal evaluations of evaluations (485) and provide a high level of evidence; however, they are uncommon in the disaster setting.

Examples of meta-evaluations include but are not limited to:


Systematic reviews are structured, comprehensive literature reviews that utilise a rigorous and published search strategy, with the aim of minimising selection bias (488-490).

Examples of systematic reviews in this discipline include but are not limited to:

- Blanchet et al., An Evidence Review of Research on Health Interventions in Humanitarian Crises, 2015 (491)
• Clarke et al., *What Evidence is Available and What is Required, in Humanitarian Assistance? 3ie Scoping Paper 1*, 2014 (3)

• Gallardo et al., *Core Competencies in Disaster Management and Humanitarian Assistance: A Systematic Review*, 2015 (492)


Other literature review methods include scoping reviews (141, 494), gap analyses (3) and priority settings (495).

The final subset of this overarching theme for evidence-based reviews and registries relates to registries of evaluation reports. For the purpose of this research, registries (sometimes called repositories) are defined as publicly available, free access collations of evaluation reports that have been undertaken in the disaster setting. Registries aim to help build capacity and strengthen disaster risk reduction and resilience (496). A separate review of such registries undertaken by the PhD candidate as part of their direct contribution to the 3ie Scoping Paper 1 suggests that these are not well known in the disaster sector but contain a large number of evaluation reports that might be of use to aid decision making and improve practice (3).

Examples of disaster evaluation registries include, but are not limited to:

• Independent Evaluation Group (IEG) hosted by the World Bank (WB) (497)

• Humanitarian Evaluation, Learning and Performance (HELP) hosted by ALNAP (498)

• International Federation of Red Cross and Red Crescent Societies (IFRC) (315)

• Evaluation and Research Database (ERD) hosted by United Nations International Children’s Emergency Fund (UNICEF) (316)

• Impact Evaluations hosted by 3ie (499).

**Knowledge Management**

Knowledge management includes cross-sectoral research, collaboration and dissemination of information to improve disaster risk management and the evidence-base of disaster science. Sharing knowledge and effective learning enables informed decision making regarding disaster risk reduction and disaster risk management. Ensuring the availability and accessibility of accurate and reliable disaster risk information should
reduce the risk of future disasters affecting communities through well-informed mitigation and preparedness planning (500, 501).

Evaluation standards and guidelines, evidence-based reviews and registries and knowledge management are relevant in advancing the science of disaster evaluations by providing scientific rigour, common terminology and the ability to replicate various methodologies.

7.3.2.2 Evaluation Typologies

In addition to the two overarching themes mentioned above, four additional evaluation typologies were identified and are discussed in detail below.

7.3.2.2.1 Baselines

Baselines are a series of evaluations or assessments that occur during the pre-event phase of disasters and provide information about the current state of the community (refer to Figure 18). The baseline status is useful in assessing the severity of the change after the event and monitoring and evaluating the response delivered as a result of the damage/event (253, 402).

Baseline evaluations include any information or data that have been collected prior to an event or disaster occurring. It covers both the pre-event status of a community and the actual hazard itself. Obtaining baseline information in the pre-event phase is critical in understanding the state of the community, how it has been affected by a disaster (238, 275, 502) and the subsequent damage that has occurred. This information assists in identifying community strengths, weaknesses and vulnerabilities to disasters. Additionally, this information assists in developing appropriate disaster management and disaster risk reduction strategies.
Innovations in science and technology have made it easier in recent years to collect information that helps reduce disaster risk, and therefore, plan for the future (278). The Sendai Framework, ratified in Sendai Japan in March 2015 by 187 UN Member States, acknowledges that there is a growing demand for science and technology to play a more prominent and effective role in providing evidence for policy and decision making (58). Knowledge and knowledge management is essential to the process. A strengthened evidence-base to support the implementation of disaster risk reduction strategies is also required (274, 277). Further, Priority 4 of the Sendai Framework supports the ‘further development and dissemination of instruments, such as standards, codes, operational guides and other guidance instruments to support coordinated action in disaster preparedness and response to facilitate information sharing on lessons learned and best practices for policy practice and post-disaster reconstruction programmes’ (58).

Examples of baseline evaluations include but are not limited to:

- baseline evaluations (272, 282) and evaluability (276)
- demographics and infrastructure, such as geographic information systems (271, 281)
• epidemiology and emerging threats (12, 273)
• defining minimum standards/criteria: baselines, targets and indicators (279, 283)
• preparedness, resilience capacity with a special focus on the Sendai Framework for Disaster Risk Reduction 2015–2030 (58)
• risk management, surveillance and early warning (280)
• health impact assessment (HIA) predictive (284).

7.3.2.2.2 Consequences

Consequences are a series of evaluations and/or assessments that occur after the event or disaster has occurred and include assessment of damage and changes in function (refer to Figure 19). It covers both the event and post-event phase of the disaster timeline. Systematic data collection and assessment is required to inform disaster needs analysis after an event. It is used in monitoring the effectiveness of response and recovery interventions and to aid decision making.

Examples of consequence evaluations include but are not limited to:

• rapid needs assessments (damage), which usually occur on Day 1 after the event as per the Inter-Agency Standing Committee document titled Multi-Cluster/Sector Initial Rapid Assessment Manual (MIRA) (289, 291, 503)
• detailed needs assessments (functional), which usually occur on Days 2–3 as per the Inter-Agency Standing Committee document titled Multi-Cluster/Sector Initial Rapid Assessment Manual (MIRA) (289, 503) and may include PDNA (288)
• continual assessments include monitoring and surveillance and can occur on multiple occasions after the event (289)
• independent real time evaluations (287) and collaborative joint evaluations, both of which are contemporary evaluation types (285).
The information received from these evaluations will ideally be compared with previous baseline studies and incorporated into helping to plan response and recovery for the current event, provide feedback into planning and preparing for subsequent events and assist in disaster risk reduction (294). Currently, damage and loss trends are difficult to monitor over time. This is partly due to inconsistent methodologies and the fact that very few countries keep national disaster databases. Even then, only one in five countries consistently record economic losses using validated tools and data collection methods (286). PDNAs aim to provide a common approach to post-crisis needs assessments and recovery planning (288).

The Centre for Research on the Epidemiology of Disasters (CRED) promotes research, training and information dissemination on disasters (292). In the Australian context, the Australian Business Roundtable for Disaster Resilience and Community Safety provides a first-time overview of disaster data with the aim of making Australian communities safer and more resilient to natural disasters (293, 504).
7.3.2.2.3 Outcomes

Outcomes are a series of evaluations and/or assessments that occur towards the end of the post-event phase of a disaster (refer to Figure 20).

**Figure 20: Core Structure and Outcomes**

Outcome evaluations reflect information or data that have been collected after an event or disaster has occurred. These evaluations include summative reviews of processes used in managing the event and outcomes related to the post-event status of the community. This information will ideally be incorporated into planning and preparing for the next event or disaster.

Examples of outcome evaluations include but are not limited to:

- operational/strategic/institutional reviews that are internal include debriefs, after action reviews (296, 299) and lessons learned (295, 320)
- quality reviews that are external include audits, key performance indicators and quality improvement (298, 299)
• government inquiries, such as the Federal Response to Hurricane Katrina (303) and the Victorian Bushfires Royal Commission (357)
• multi-disciplinary event reports and reviews, include the work by Kulling and the recent work by Fattah (20-22, 24)
• process and outcome evaluations (297, 304, 505) including logic maps or models, theories of change and causal links and attribution used to guide process evaluations. Each intervention and evaluation requires its own logic map that describes the sequence of actions to be undertaken and communicates what the program is and will do (301). Process evaluations are extremely important as they:
  o investigate how well an intervention or programme works
  o determine whether an intervention is a good investment
  o can be used to explain discrepancies between expected and observed outcomes
  o can provide insights to aid delivery of interventions (505, 506)
• HIA evaluations are important as they focus on sharing their evaluations findings and lesson learned (284, 507).

7.3.2.2.4 Impact Evaluations

Impact evaluations of programs, projects and interventions are evaluations that include a measure of causality or attribution (309); they can occur during any phase of the disaster timeline (refer to Figure 21). Impact evaluations are shown using the standard decision symbol of a diamond. While there are two diamonds shown in Figure 21, they should be read as being able to happen anywhere along the disaster timeline and can number greater than two.
In the disaster setting, impact evaluations have gained popularity for identifying causal links between specific interventions and outcomes. This is a result of the international community demanding accountability and improved evidence-based interventions (3, 16, 98, 508). Although there remains ongoing debate about the exact definition of impact evaluations (306, 509), they are particularly well suited to answer important questions, such as whether:

- interventions do or do not work
- interventions make a positive or negative impact
- any intended or unintended consequences
- how cost effective they are (305, 310).

It is believed impact evaluations greatly improve the effectiveness of interventions delivered in the disaster setting by identifying what works for whom, why and at what cost (16, 312).

Examples of impact evaluations include but are not limited to:
• Rogers P, *Introduction to Impact Evaluation*, 2012 (311)
• Puri et al., *What Methods May Be Used in Impact Evaluations of Humanitarian Assistance?* 2015 (310)
• Chambers et al., *Designing Impact Evaluations: Different Perspectives*, 2011 (308).

7.4 Results – Comprehensive Framework for Disaster Evaluation Typologies 2017

Individually, each of the elements from the core structure through to impact evaluations (refer to Figures 14 and 16–21) have relevance in their own right in the disaster evaluation setting. Combining them into a single, unifying comprehensive framework, however, provides a diagrammatic solution to the overarching Study 3 research question, which identifies what comprehensive framework design could enable mapping evaluation typologies along the disaster timeline. Further, this supports the overall research aim. This section amalgamates the previous seven diagrams and introduces CFDET 2017.

7.4.1 Comprehensive Framework for Disaster Evaluation Typologies 2017

CFDET 2017 (refer to Figure 22) identifies the different typologies of disaster evaluations and demonstrates key relationships in a single diagram. Importantly, it also suggests the interdependencies and relationships that exist between various evaluation typologies along the disaster timeline and within the disaster setting. This consolidates the previous Figures 14, 16-21, but more importantly, includes feedback loops from consequences and outcomes to improve baselines, reduce disaster risk and strengthen resilience. The feedback loop helps to strengthen the concept that the disaster timeline is a continuum and should be viewed in this light (401).
Figure 22: Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017)
The framework is not limited to any one phase of the disaster timeline and can be used for responding to disasters (regardless of the cause, that is, an ‘all-hazards approach’), humanitarian crises or in the development sector. It is relevant to any sized disaster.


The development of CFDET 2017 was published in PDM in 2017 (refer to Appendix 8) (1). This paper brings together in a single framework the different typologies of disaster evaluations. It suggests interdependencies and relationships that exist between various evaluation typologies found within the disaster setting. It is a unique unifying framework published for the first time in this domain.

7.5 Discussion

Disasters, whether caused by natural hazards or man-made, are complex events (79, 144, 510, 511), and undertaking structured evaluations in this setting is also a complex activity. This research created a classification of disaster evaluation typologies that provides structure, encourages common terminology and advances the evidence-base of disaster science. The role of CFDET 2017 is to support the ability to measure and evaluate the effectiveness of interventions provided in the disaster setting and thereby reduce the increasing human and economic costs associated with disasters.

CFDET 2017 is the first framework of this type globally and makes a unique contribution to current knowledge. No previous reference has been located that identifies such a wide range of evaluation typologies used in the disaster setting and further provides conceptual relationships in a single comprehensive framework. There may be evaluation typologies that have not been identified due to the difficulty in searching the grey and peer-reviewed literature. The aim of CFDET 2017 is consistent with that of James, who concluded that ‘a common Disaster Medicine and Global Health taxonomy will form the foundation of a safer, more resilient world, through more effective preparedness and response; but we must first come together for the public good’ (512)(p.184).

The development of CFDET 2017 has been an iterative process, in that it has been presented at both international and national forums and modified based on feedback. The framework was initially presented at the 19th World Conference on Disaster and Emergency Medicine (WCDEM), Cape Town, South Africa, 21–24 April 2015, and the

7.6 Conclusion

Disasters are complex events and undertaking disaster evaluations is a specialised area of study (79). While some frameworks have been developed over the years to support consistent disaster research and evaluation, they lacked validation, consistent terminology and standards. Consequently, there was no single, internationally accepted, comprehensive framework to structure disaster evaluation typologies. A strong evaluation framework for disaster settings is extremely important given the significant impact of disasters. It must utilise agreed definitions and be able to measure the impact and effectiveness of interventions. It is anticipated that CFDET 2017 creates a useful and usable framework and promotes an environment for constructive dialogue at an international level.

This chapter discussed the development process undertaken to create a comprehensive framework for disaster evaluation typologies. The resulting CFDET 2017 provides a diagrammatic answer to the overarching Study 3 research question of ‘What comprehensive design could enable mapping evaluation typologies along the disaster timeline?’ This work adds to the current knowledge base and understanding of disaster evaluations.

This unique and unifying framework has relevance at an international level and is expected to benefit the disaster, humanitarian and development sectors. This work promotes an environment for constructive dialogue on evaluation in the disaster setting and adds to the evidence-base of disaster evaluation and research. To gain international acceptance and credibility, a validation process is needed and is discussed in more detail in Chapter 8.
Chapter 8: Validation of the Comprehensive Framework for Disaster Evaluation Typologies 2017

8.1 Introduction

Chapter 7 presented the single, unifying CFDET 2017 (1), which includes six additional layers that overlaid the core structure developed in Chapter 6. The six layers were determined based on the earlier literature review that categorised the evaluation typologies into two overarching themes and four evaluation typologies. The two overarching themes were categorised as accountability and evaluation standards and evidence, while the four evaluation types included baselines, consequences, outcomes and impact evaluations. Each of these were mapped along the disaster timeline, which also highlighted their relationships and interdependencies with each other. The purpose of this chapter is to discuss the peer-review validation process, analysis of results and discussion of outcomes. This aids in rounding out the overall research aim of this thesis, which is to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline.

Previous frameworks (17-19, 25-34, 101, 167, 253), while recognised in the literature, have not been validated and are rarely used as a methodological framework to structure disaster evaluation and research (1). The importance of this thesis is not only in the development of a unifying framework, but in undertaking the additional steps of validating the developed framework and gaining international acceptance.

CFDET 2017 promotes an environment for constructive dialogue on evaluations currently in use, or that could be used, in the disaster and humanitarian setting. It improves the validity of disaster evaluation and research by clarifying the various evaluation typologies with the intent to strengthen the evidence-base of health interventions delivered in response to disasters by identifying what works for whom and why. Despite the evolution of disaster evaluation and research and concerted efforts since 2003 to develop consensus on disaster research and terminology (513), there has been little agreement from experts regarding a comprehensive and validated framework that structures disaster evaluation typologies (1). To add credibility and global relevancy to CFDET 2017, an international peer-reviewed validation process was necessary (514). Thus, with an element of global acceptance, this process would ensure that CFDET 2017
was not just the perception of the researchers, but was indeed an internationally collaborative view of evaluation typologies suitable for use in the disaster setting.

When considering the definition of ‘validation’ and how to undertake a ‘validation process’, a review of the literature highlighted multiple perspectives and a confusing array of terms and diverse typologies (515). For the purpose of this thesis, ‘validity’ is defined as the ability of the instrument to measure the attributes of the construct under study (516). In this case, the instrument under study is the comprehensive framework and the ‘attributes’ of CFDET 2017 (1) are the concepts and elements; that is, the disaster phases and evaluation typologies found in the disaster setting, and their relationships.

Validity can be further divided into ‘external validity’, which measures the generalisability of the findings, and ‘internal validity’, which refers to the confidence placed in the cause and effect relationship; that is, is there another reason or cause that can explain the results (517)? For a framework to be useful and used, it needs to be both operationally credible and feasible (21). An online survey was created to administer a questionnaire that used a mixed methods approach to ascertain the developed framework’s acceptance and validity at an international level.

8.2 Research Questions

The development of CFDET 2017 (1) provides a solution to the overarching research aim. This stage of the research extends on the development of the comprehensive framework to include an international peer-reviewed validation process to gain consensus on disaster evaluation typologies currently in use, or that could be used, in the disaster setting. This provides the solution for the overarching Study 4 research question as identified in Chapter 2 of this thesis.

On a more detailed level, the validation process was designed to answer the following sub questions:

- In what way does the mapping between the evaluation typologies and the disaster timeline demonstrate suitability?
- What, if any, disaster evaluation typologies are missing in CFDET 2017?
- What value does CFDET 2017 have for the survey respondents’ work when undertaking evaluations in the disaster and humanitarian setting?
• In what ways would CFDET 2017 be useful for supporting and promoting evaluation in the disaster and humanitarian setting?
• In what ways would CFDET 2017 be useful for teaching evaluations in the disaster and humanitarian setting?
• What are the perceived barriers and enablers to undertaking evaluations in the disaster setting?

8.3 Research Design and Methodology

To validate CFDET 2017 at an international level, a peer-reviewed validation process consisting of the following research methods was used: a literature review, presentations of the comprehensive framework seeking professional input, and a mixed methods online survey hosted by a survey engine. The use of a mixed methods online questionnaire was appropriate as it provided a structured and consistent approach allowing for better quantification of results. Further, the mixed methods approach helped ‘triangulate’ and support the findings (117, 139). Triangulation employs more than one method for corroborating findings and helps validate qualitative research through the convergence of information from different sources (518).

A questionnaire, including an explanatory statement and consent form, was designed seeking input on the newly developed comprehensive framework. Subsequent validation findings were analysed and reviewed.

To undertake a robust validation of CFDET 2017, and incorporating a quantitative and qualitative approach (514), the following phases were undertaken:

• literature review to identify any existing validation processes that could be used for validating the developed framework
• as an iterative process, presentations of CFDET 2017 were made at international and national events to support the iterative development process and seek professional engagement with an aim to continually improve the framework through consultation and reflection
• a mixed methods research online survey to collect and analyse comments and feedback on CFDET 2017.
8.3.1 Literature Review

The aim of the literature review was to investigate how to undertake a validation process of a comprehensive framework. The review was based on investigating the peer-reviewed and grey literature. First, the peer-reviewed literature was searched using major electronic databases, including PubMed/Medline (US National Library of Medicine, National Institutes of Health; Bethesda, Maryland, US), CINAHL (EBSCO Information Services; Ipswich, Massachusetts, US), EMBASE (Elsevier; Amsterdam, Netherlands), ProQuest (Ann Arbor, Michigan, US) and Science Web (Thomson Reuters; New York, New York, US). These databases were searched to identify contributions to the validation of comprehensive frameworks. The following key terms and Boolean operators were used: ‘validation’ OR ‘valid*’ AND ‘comprehensive framework’ OR ‘framework’ OR ‘model’. Additional search strategies included the words ‘disaster*’ OR ‘evaluation’ OR ‘evaluat*’.

Inclusion criteria consisted of articles in English and included several validation publications of frameworks from other areas of specialty, such as the health sector (519-523) and several Delphi validation studies (121, 243, 325). This was done to gain further insight as the number of articles related to the validation of a comprehensive framework in the disaster or evaluation setting were limited. Additional references were identified through examination of the bibliographies from the most recent publications (snowballing) (524) and through the scrutiny of the contents pages of highly relevant databases and journals (525, 526).

Second, a review of the grey literature was also undertaken, including similar keywords, using Google and Google Scholar (Google Inc.; Mountain View, California, US) and supplemented by the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP; London, UK) (343).

The results from the literature review on identifying evaluation typologies were reported in Chapter 4 and the findings integrated into the comprehensive framework. The results from the literature review for identifying the validation process for comprehensive frameworks, such as in this thesis, highlighted that no articles that referred to a validation process in the disaster setting could be replicated or modified for the purposes of this research. The results, however, did highlight once again a lack of consistency in terminology and application.
8.3.2 Action Research / Conference Presentations

Presenting CFDET 2017 at international conferences, national seminars and other academic meetings provided an excellent opportunity to discuss and receive initial feedback prior to undertaking the formal validation process. This also embraced the underlying principle of action research as an iterative process seeking input from end users (527). Evolving versions were presented and discussed at various international conferences, including:

- 18th World Congress on Disaster and Emergency Medicine (WCDEM), Manchester, UK, May 2013
- Monash University Disaster Resilience Initiative Forum, Melbourne, Australia, November 2013
- 12th Asia Pacific Conference on Disaster Medicine (APCDM), Tokyo, Japan, September 2014
- 19th World Conference on Disaster and Emergency Medicine (WCDEM), Cape Town, South Africa, April 2015
- Monash University Disaster Resilience Initiative Forum, Melbourne, Australia, November 2015
- Monash University Disaster Resilience Initiative Forum, Higher Degree by Research (HDR) Colloquium, Melbourne, Australia, July 2016
- Monash University Accident Research (MUARC), Departmental Research Seminar, Melbourne, Australia, October 2017.

Feedback received was used to make minor modifications to the framework; for example, the addition of an arrow indicating the progression of ‘time’. Other comments highlighted an opportunity to further explain the theory underpinning the structure. These interactive events helped shape the development of the framework and were a key component of the progressive peer-reviewed validation process.
8.3.3 Mixed Methods Research – Online Survey: Design and Sample

An online survey was an appropriate tool for collecting comments on the evolving and contemporary framework CFDET 2017 due to sampling size and the international nature of the participants (121). Jorm identified expert consensus as a fundamental underpinning of science, and the Delphi method as a useful way of determining expert consensus. There is no apparent ‘gold standard’ for undertaking a Delphi study to guide the validation process (121). An alternative was to consider the concept of the ‘wisdom of crowds’, as highlighted by Surowiecki (528). The term ‘crowd’ represents a collection of individuals with some experience and consists of four substantiating criteria. For a crowd to be wise, the crowd must have diversity of expertise, make decisions independently, be decentralised and autonomous, and there must be a way of coordinating and aggregating their experience (528). This study met all of these criteria. In light of Surowiecki’s comments and the inability to find a Delphi study to guide the validation of CFDET 2017, a heterogeneous crowd with experience in the different fields of disasters/emergencies, humanitarian/development and/or evaluation was chosen. They could act independently of one another and the majority of participants did not work together. It is believed that the ‘wisdom of crowds’ effect in this research is robust (529).

An online survey was an appropriate tool for collecting comments on the evolving and contemporary framework CFDET 2017 due to sampling size and the international nature of the participants (121).

Further, the use of this structured and consistent approach allowed for better quantification of results. The literature review showed there was no pre-existing survey or questionnaire appropriate to adopt or adapt for this particular validation process. A questionnaire was designed and developed (refer to Appendix 9). A pilot test was conducted (n = 9) to test the survey platform as well as the questionnaire. No functional changes or modifications were required and ethics approval was obtained from MUHREC for Project: Validation of ‘A Comprehensive Framework for Disaster Evaluation Typologies’, Project Number 7737 (refer to Appendix 4). Participation was voluntary and anonymity was preserved. The survey was hosted on the Qualtrics Insight Platform (140) and ran over a 6-month period in 2017; 140 participants were initially approached.

A conscious decision was made to not make answering questions compulsory, nor were participants required to answer questions to progress through the survey. This was deemed appropriate to prevent any stalling that may have led to a small sample size and
the need to run the survey additional times. This was also necessary given the estimated time commitment required to complete the survey (30–45 minutes); it was acknowledged that participant attrition might occur due to the length of the survey (121).

While the main intent of the survey was to gain quantifiable acceptance of CFDET 2017 elements and relationships, certain demographic information was also identified as helpful for secondary analysis. To support this, the online survey consisted of two parts:

1. A demographics section that sought basic information about the participant, such as:
   a. Age.
   b. Gender.
   c. Country of residence – geographical stratification was included as per the WHO regions of the world to assess regional representation and to add a diversity of participant expertise. WHO member states are grouped into six WHO regions: African region, Americas region, Eastern Mediterranean region, European region, South-East Asia region and Western Pacific region (530) (refer to Figure 23). The international nature of disaster risk reduction dictates global involvement and as this thesis is predominantly focused on health interventions in the disaster setting, it was deemed appropriate to continue with a ‘health theme’ and follow the WHO regional stratification.
   d. Current field of experience (disaster medicine, disaster management or emergency management; humanitarian or development sector; and evaluation – participants self-identified their current field of experience).
   e. Years of experience in their current field.
   f. Whether they undertook or supervised evaluations.
2. Mixed methods questionnaire specifically targeting CFDET 2017. Participants were provided with an extensive explanatory statement outlining the background behind the development of CFDET 2017 as well as the individual elements diagrammatically represented in the survey (refer to Appendix 9). It was highlighted that this survey was based on a publication, ‘Disaster Metrics: A Comprehensive Framework for Disaster Evaluation Typologies’ (1), that was in press at the time, and there was a requirement to honour the copyright agreement. Participants were asked to:

   a. Use a 5-point Likert scale to identify their agreement or disagreement on a symmetric agree-disagree scale for a series of statements related to the elements and relationships found within CFDET 2017.

   b. Provide responses to open-ended questions with an opportunity to comment on the disaster timeline elements and the different evaluation typologies presented. Participants were asked to comment on the relationships and linkages between the disaster timeline and evaluation typologies as demonstrated in CFDET 2017, whether they deemed CFDET 2017 to have value, if they thought CFDET 2017 supported and promoted evaluation, if they thought CFDET 2017 was useful for teaching evaluations and any general comments.
The next stage of the survey process aimed to determine what would constitute acceptance of CFDET 2017. As there is no single definition of consensus (121), a determination was made that the following criteria would be used to validate CFDET 2017:

1. \( n = 30 \) responses
2. mean of 4.0
3. 80% of participants who ‘strongly agree’ or ‘somewhat agree’ on the Likert scale.

An agreed sample size of 30 responses is based on the small target population and is supported by Jorm’s work, where he identified that there is little firm guidance in this area (121). Research undertaken by Morgan and colleagues in 2002 identified that no new concepts were identified after 20 interviews when they investigated how many qualitative data collection events they needed to study (531, 532). Further, setting the agreement of 80% of participants who ‘strongly agree’ or ‘somewhat agree’ on the Likert scale is a benchmark identified in other research (121, 533).

To obtain the required number of responses, people with backgrounds in the areas of disaster medicine, disaster management and emergency management; humanitarian or development; and evaluation were approached via email. The sampling strategy included both non-probability sampling (that is, the participants were selected by the researchers (127)) and purposive sampling. Both sampling strategies were deemed appropriate as the target population was small (126, 127). The actual list of potential participants was created from the literature review process as well as from personal networks generated from attendance at academic and global events. This diversity of expertise was actively sought as it was thought it would produce better quality decisions as opposed to homogeneity (121). In the Qualtrics Insight Platform, contact lists had to be created outside the platform and copied in, ensuring correct and current email addresses were entered.

All comments were reviewed and discussed by a team of four experienced researchers. Positive and reassuring comments of the framework were received. Some comments suggesting change were incorporated in amendments to the framework. Others were rejected based on a lack of feasibility, utility or accuracy.
8.4 Results

A total of 140 participants were approached to undertake the survey. Five people chose to opt-out with 52 starting the survey of whom 30 completed the survey (refer to Table 28).

To reach the required sample size of 30, three separate groups of participants needed to be approached. Participants were given a three-week window to complete the survey and a reminder email was automatically generated at two weeks. Group A consisted of 37 participants with 10 complete responses obtained (n = 10), Group B consisted of 48 participants with four complete responses (n = 4), and Group C consisted of 50 participants with 16 complete responses (n = 16). In total, 30 complete responses were obtained (n = 30). While Table 28 shows that 52 participants started the survey, the detailed discussion of the results reflect where participants chose to exit the survey.

Table 28: Summary of Survey Responses

<table>
<thead>
<tr>
<th></th>
<th>Invitations</th>
<th>Opt Outs</th>
<th>Started Survey</th>
<th>Completed Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>40</td>
<td>3</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Group B</td>
<td>50</td>
<td>2</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Group C</td>
<td>50</td>
<td>0</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>140</td>
<td>5</td>
<td>52</td>
<td>30</td>
</tr>
</tbody>
</table>

Having obtained the required number of responses, the data were collated and analysed according to the following areas:

1. consent, copyright and demographic information about the participants
2. results obtained regarding the individual elements that comprised CFDET 2017
3. responses related to the relationships between the evaluation typologies and the disaster timeline as a unifying framework, as well as identifying whether CFDET 2017 has value when undertaking evaluations in the disaster setting.

Means and values based on a 95% confidence interval were calculated on Likert-scale responses. The results are discussed in greater detail below.
8.4.1 Consent, Copyright and Demographics

Participants were asked to acknowledge that they had read and understood the explanatory statement and consented to participating in the online survey. They also agreed to abide by international copyright regulations as the survey was based on a publication in print (1) at the time of the survey being run. All participants (n = 52) agreed (refer to Table 29).

Table 29: Initial Survey Responses

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understand</td>
<td>52</td>
<td>100.0%</td>
</tr>
<tr>
<td>Copyright</td>
<td>52</td>
<td>100.0%</td>
</tr>
<tr>
<td>Completed Demographics Section</td>
<td>39</td>
<td>75.00%</td>
</tr>
</tbody>
</table>

The next six questions related to demographics and asked about the participants’ age group, gender, country of residence at the time of completing the online survey, current field of experience (disaster medicine, disaster management or emergency management; humanitarian or development sector; and evaluation), which they self-identified, the number of years full-time equivalent they had worked in their current field and whether they undertook or supervised evaluations in the disaster/emergency and/or humanitarian/development sector.

Some 92.3% of participants were aged 41 years and above, with 33.3% older than 60 years (refer to Table 30).
As shown in Table 31, the ratio of female-to-male participation was almost equal, coming in at 51.3% and 48.7% respectively.

Participants were invited from all six WHO regions in an attempt to facilitate an international peer-review of CFDET 2017. Table 32 shows the country breakdown, while Table 33 shows stratification by WHO region. While participants were invited from the African and South-East Asia regions, no responses were received. The largest number of responses were received from the Western Pacific region.
Table 32: Stratification by Country

<table>
<thead>
<tr>
<th>STRATIFICATION (By Country)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>17</td>
<td>43.6%</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>1</td>
<td>2.6%</td>
</tr>
<tr>
<td>Ireland</td>
<td>2</td>
<td>5.1%</td>
</tr>
<tr>
<td>Italy</td>
<td>1</td>
<td>2.6%</td>
</tr>
<tr>
<td>Japan</td>
<td>2</td>
<td>5.1%</td>
</tr>
<tr>
<td>Netherlands</td>
<td>2</td>
<td>5.1%</td>
</tr>
<tr>
<td>New Zealand</td>
<td>2</td>
<td>5.1%</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>1</td>
<td>2.6%</td>
</tr>
<tr>
<td>Switzerland</td>
<td>3</td>
<td>7.7%</td>
</tr>
<tr>
<td>UK</td>
<td>3</td>
<td>7.7%</td>
</tr>
<tr>
<td>USA</td>
<td>5</td>
<td>12.8%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>39</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Table 33: Stratification by WHO Region

<table>
<thead>
<tr>
<th>STRATIFICATION (By WHO Region)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>African</td>
<td>0/39</td>
<td>0.0%</td>
</tr>
<tr>
<td>Americas</td>
<td>5/39</td>
<td>12.8%</td>
</tr>
<tr>
<td>Eastern Mediterranean</td>
<td>1/39</td>
<td>2.6%</td>
</tr>
<tr>
<td>European</td>
<td>11/39</td>
<td>28.2%</td>
</tr>
<tr>
<td>South East Asia</td>
<td>0/39</td>
<td>0.0%</td>
</tr>
<tr>
<td>Western Pacific</td>
<td>22/39</td>
<td>56.4%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>39/39</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

With regards to areas of expertise, more than half of the participants (64.1%) identified their area of expertise as disaster medicine, disaster management or emergency management. Evaluation was the second most common response at 28.2%, with the remaining participants (7.7%) categorising themselves as belonging to the humanitarian and development sector (refer to Table 34).
After identification of area of expertise, participants were asked to identify the years of full-time equivalent in their current field. Some 66.67% of participants had greater than 11 years or more full-time experience in their self-selected area of expertise. Participants with 16 or more years of experience came in at 46.1% (refer to Table 35).

Table 34: Area of Expertise

<table>
<thead>
<tr>
<th>AREA OF EXPERTISE</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disaster medicine / management, emergency</td>
<td>25</td>
<td>64.1%</td>
</tr>
<tr>
<td>management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humanitarian or development sector</td>
<td>3</td>
<td>7.7%</td>
</tr>
<tr>
<td>Evaluation</td>
<td>11</td>
<td>28.2%</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

The majority of participants, 64.1%, indicated that they undertook evaluations or supervised evaluations as part of their normal work (refer to Table 36).
8.4.2 Survey Responses Related to CFDET 2017 Elements

The results shown in Table 37 are related to questions directed at the seven individual elements, reflecting the inclusion of the core structure (established in Chapter 6), which are illustrated in each unique layer of CFDET 2017, along with relationships and interdependencies. Six of these elements were determined from the outcome of Study 3. In Chapter 7, these were identified as the two overarching themes of accountability and evaluation standards and evidence, and the four evaluation typologies of baselines, consequences, outcomes and impact evaluations. The survey, however, was structured in a different order, and the results discussed below reflect the survey structure.

Table 36: Undertaking Evaluations

<table>
<thead>
<tr>
<th>UNDERTAKE EVALUATIONS</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>25</td>
<td>64.1%</td>
</tr>
<tr>
<td>No</td>
<td>14</td>
<td>35.9%</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Table 37: Survey Responses to CFDET 2017 Elements

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>MEAN +/- 95% CI</th>
<th>ACCEPTANCE THRESHOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Structure</td>
<td>33</td>
<td>4.34, 4.15 – 4.54</td>
<td>96.88</td>
</tr>
<tr>
<td>Baselines</td>
<td>33</td>
<td>4.09, 3.69 – 4.49</td>
<td>75.76</td>
</tr>
<tr>
<td>Consequences</td>
<td>32</td>
<td>4.34, 4.06 – 4.63</td>
<td>87.50</td>
</tr>
<tr>
<td>Outcomes</td>
<td>32</td>
<td>4.13, 3.82 – 4.43</td>
<td>84.38</td>
</tr>
<tr>
<td>Impact Evaluations</td>
<td>30</td>
<td>3.80, 3.42 – 4.18</td>
<td>63.33</td>
</tr>
<tr>
<td>Accountability</td>
<td>29</td>
<td>3.93, 3.49 – 4.37</td>
<td>65.52</td>
</tr>
<tr>
<td>Evaluation Standards</td>
<td>29</td>
<td>3.97, 3.57 – 4.40</td>
<td>82.76</td>
</tr>
</tbody>
</table>
1. The core structure outlines the fundamental framework of the disaster timeline to which all typologies are related and consists of three important layers. Layer 1, at the bottom of the diagram, provides a preliminary and simplistic view of the disaster continuum, represented by a pre-event phase, an event phase and the post-event phase of an emergency or disaster. Layer 2 expands on Layer 1 and is based on a modified representation of the Utstein Guidelines 2003 (17), as amended. Layer 3 introduces the contemporary concept of ‘strengthening resilience’ as an important and overarching theme. It underpins and enriches the core structure and is an emerging, international imperative embraced by the Sendai Framework (58).

The majority of participants strongly agreed that the elements illustrated in the core structure reflected contemporary thinking on disaster timeline elements (n = 33, mean 4.34 (4.15-4.54) and an acceptance rate of 96.88%) (refer to Table 37).

The survey asked participants to identify any disaster timeline elements that were not covered in the core structure. Responses are provided in Table 38.

Table 38: Core Structure Comments – Disaster Timeline Elements Not Covered

<table>
<thead>
<tr>
<th>Responses to the request for comments related to whether any disaster timeline elements were not covered in the core structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>May need to consider adding Social, Built, Economic, Environmental Factors into the Core Structure. Some of the boxes I believe would need to change i.e., Hazard, then Pre-Event Status. Also, Disruption, Change in Functions Consequences should span across all boxes</td>
</tr>
<tr>
<td>My reason for saying ‘neither nor’ above is that I don’t know what contemporary thinking on disaster timeline events is</td>
</tr>
<tr>
<td>The Core Structure is well done but I am afraid that the concept of ‘disaster cycle’ might move to a second plan in this way. My concern is that the idea of continuum may get lost</td>
</tr>
<tr>
<td>Risk reduction is a confusing term. In the new Sendai Framework, it is starting to mean the policy objective of risk management. My preference is that risk reduction and risk management should be used across the spectrum of emergency and disaster risk management – i.e., prevention, preparedness, response and recovery. Rather than use the term risk reduction, I would suggest pre-event hazard reduction, vulnerability reduction, exposure reduction and emergency preparedness - response &amp; emergency preparedness - recovery. Pre-event prevention and pre-event mitigation might also work, but I prefer the former. Also, disaster tends to refer to large scale events associated with natural hazards, rather than events of any scale associated with all types of hazards. Thus, you might consider using terms ‘emergency and disaster’ or ‘emergency and disaster risk management’ which might offer more inclusive approach. Colleagues working on epidemics rarely use the term ‘disaster’</td>
</tr>
<tr>
<td>1. Normal development process continue in the background in intertwine with recovery and response</td>
</tr>
<tr>
<td>2. Disaster events may be slow-onset. or may only slowly develop after a sudden-onset disaster</td>
</tr>
<tr>
<td>The linearity of the pre-event/hazard/risk reduction part of the timeline does not adequately reflect both stochasticity and risk distribution</td>
</tr>
<tr>
<td>I would say that you can add the evaluation as part of the post event</td>
</tr>
</tbody>
</table>
Is it worth adding mention of workforce specifically? Covered as part of system but impact on an organisations people often the determinant of outcome. May be as simple as access to work (transport / security) but also personal impacts (mental health / family / personal death and injury)

Whilst it may not be relevant, I am unsure of what ‘pre-event status’ actually means, I am also conscious that most disasters are preceded by a previous event be it major or minor. Not sure if it is covered in the literature but the recovery area is very broad. I think that in practical sense we have, relief, ‘functional recovery’, then a review period to identify future options, before persistent recovery and then resilient activity. Whilst it is a timeline it is a cycle of actions

Bounce back better is not included. Event seems to be limited to flash event and e.g., no pandemic. Response starts with event and doesn’t wait for damage and functional losses to appear. Order of structural damage and functional losses is counter intuitive from a health perspective, so should be explained as damage to structures and functional losses of communities etc. In relation to the safety culture that you learn best from near misses and that you need a culture of admitting mistakes in order to be able to do that, as is promoted for airplane personnel.

I am not sure that ‘pre-event and event are clearly separated in all types of disasters, in particular slow onset disasters (droughts – when is the specific event beginning?) or cascading disasters or public health emergencies – the complexities of these are challenging. Intuitively it makes sense that it should be, however, it is a real expectation that this would happen in the event phase? Perhaps resilience peaks and troughs across the timeline?

As a core structure, it appears to have the essential elements

Response preparation – that is when an event is expected there is a series of responses initiated prior to the event itself, e.g., evacuation prior to a major flood or storm

I feel I need to read your recent article first but I have not accessed it yet so I apologies if I am not well informed in my comments

These frameworks do not address individual differences, and the fact that some affected individuals may move back and forth across a restricted range from disruption to recovery to disruption again ….. Also, thinking of the Canterbury data showing an increase in mental health requests/visits over time

The survey then asked participants to comment on whether there were any contemporary disaster timeline elements that could be substituted in the core structure. Responses are provided in Table 39.
Table 39: Core Structure Comments – Substitution of Any Disaster Timeline Elements

Responses to the request for comments related to whether there were any contemporary disaster timeline elements which could be substituted in the core structure?

I would suggest that the Recovery Phase begins at the same time as the Post-Event phase starts

Keep it simple as much as you can. Most new things are not new but simply rewording of existing work or new ‘buzz phrases’

Only question is how you include ‘development’? I would argue that development could be in response to a hazard and part of risk management but probably contrary to the views of most. Bottom line is that unless we do this we will forever be providing external assistance when we should be supporting development of local capability and capacity supported by regional arrangements

I think that recovery planning can begin in the planning/preparation/pre-event stage. This isn’t a new contemporary timeline element – just a suggestion on timeline structure

There could, however, be a greater emphasis on adapt and transform. The small text used for these terms suggest that those elements aren’t as important (and are quite hard to read)

I like the way the framework is structured and think that you could add many different perspectives. It is not easy to know what is implicit in some of the stage. In my research, I have been thinking about the phases from a particularly health specific perspective with a person-centred trajectory through disasters and from pre-existing health to post-event health but I would believe that is presumed to be part of the resilience timeline. You could add may timeline elements to the framework you propose which is its strength .... This could be a ‘living framework’ where each group has their own explicit perspective or lens to the most appropriate layer of the timeline

See above comment ‘These frameworks do not address individual differences, and the fact that some affected individuals may move back and forth across a restricted range from disruption to recovery to disruption again ..... Also, thinking of the Canterbury data showing an increase in mental health requests/visits over time’

2. The baseline evaluation typologies section comprised a series of evaluations and/or assessments that occur during the pre-event phase of disasters. They provide important information about the existing state of the community prior to a disaster. Having adequate baseline information is critical in understanding the community and how it has been affected by a disaster (275, 502) and the subsequent damage that has occurred.

When asked whether evaluation typologies included in the baseline section reflected current baseline evaluation methods, the results were borderline with a mean of 4.09 (3.69-4.49) and an acceptance rate of 75.76% (refer to Table 37).

The survey asked participants to provide comments related to whether evaluation typologies, as illustrated in baselines, reflected current baseline evaluation methods and whether there were any baseline evaluation typologies not covered. Responses are provided in Table 40.
Table 40: Baseline Comments

Responses to the request for comments related to:

- whether evaluation typologies illustrated in baselines reflected current baseline evaluation methods
- if there were any baseline evaluation typologies not covered

Only that well-being and coping capacity would be included, perhaps the latter as part of the Resilience capacity typology

Sorry, unsure. However, I was a bit surprised to see ‘evaluability’ in there. I normally think of that as an assessment of how possible it would be to evaluate something. Is that possible to do, when you do not know what that ‘something’ even is?

It looks good, but maybe too focused on health and not on other sectors
- there is not reference to Common Operational Datasets and to Fundamental Operational Datasets (COD and FOD)
- there is not reference to capacity situation
- there is not reference to pre-existing vulnerabilities

Risk management should be used about all prevention, preparedness, response and recovery measures, and not limited to the pre-event management of risks.

I have made comments above to replace risk reduction with prevention and mitigation, or preferably hazard, vulnerability and risk reduction.

If you use risk as a reflection of hazard, vulnerability, exposure and capacity you might have a simpler way to reflect the baseline status.

Capacity is also universal across systems, policy and legislation, human and financial resources, information management, planning and coordination, (health) service capacity, infrastructure, logistics, community capacity … which support/result in prevention, preparedness, response and recovery.

It is also important to have an assessment of the context – political, financial, social, health, economic and environmental.

Conflict analysis?

Reliable and accurate baseline data is often not available in developing countries. Even for developed countries, key baseline indicators may be contested (for example the US President’s characterisation of the economy under his predecessor)

The list of elements in the grey box are a mix of actions, data types, disciplines, methods, approaches and contexts. Not sure what this list is meant to comprise

If we look at this figure as a cycle where the post-event would be considered as, or followed by, the pre-event of the next disaster then you can add post evaluation (of previous events) to the baseline evaluation typologies

I think there are some clues in the NERAG (National Emergency Risk Assessment Guidelines – saved) on baseline themes, people, economy, environment, social setting and public administration. Risk assessment is baseline, but risk management includes the acts of response and recovery, (to manage the risk to your objectives of reducing incidents, containing extent, and minimising consequence)

Resilience capacity is not well specified, neither is predictive HIA

To be able to evaluate, an evaluation infrastructure should be in place before the event. Such as infrastructure includes people (expertise), products (evaluation instruments that may be used) and processes (pre-planned and validated evaluation instruments).

Another element that is missing is the information on the pre-event level of preparedness of those involved (i.e., what is the level, quality and outcome of education, training and disaster exercises)

Could baseline assessment occur in the event or post event phase- perhaps in practice they are? Does the current criteria include geography (i.e., small island states/climate change/sea level rise?)

I guess there are trajectories in life and disasters. A baseline measurement is only one point in time and the community might already be on a downward trajectory for other reasons unrelated to the disaster so would it be useful to have a ‘trajectory’ evaluation? as one of the baseline measurements?
Pre-event status presumably includes community strengths from a psychological as well as other perspectives

3. The consequence evaluation typologies section comprised a series of evaluations and/or assessments that take place after an event or disaster. Importantly, they include assessment of the damage and changes in function that have occurred as a consequence of the disaster. It covers both the event and post-event phase of the disaster timeline. Systematic data collection and assessment is required to inform disaster needs analysis after an event. It is used in monitoring the effectiveness of response and recovery interventions and to aid decision making (1).

When asked about the evaluation typologies included in the consequence section reflecting current consequence evaluation methods, the results received were positive (n = 32, mean 4.34 (4.06-4.63) and acceptance rate of 87.50%) (refer to Table 37).

The survey asked participants to provide comments related to whether evaluation typologies, as illustrated in consequences, reflected current consequence evaluation methods and whether there were any consequence evaluation typologies not covered. Responses are provided in Table 41.

**Table 41: Consequences Comments**

<table>
<thead>
<tr>
<th>Responses to the request for comments related to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• whether evaluation typologies illustrated in consequences reflected current consequence evaluation methods</td>
</tr>
<tr>
<td>• if there were any consequence evaluation typologies not covered</td>
</tr>
</tbody>
</table>

*Midterm evaluations and other forms of evaluation.*

Some of the methods/exercise seem more assessment exercises than evaluation itself.

May want to specifically identify complaints and feedback mechanisms

As previous, unsure due to lack of familiarity. However, I have never heard the language of ‘consequence evaluations’, is that a term of your own making? A form of outcome evaluation?

Consequences of what – the disaster? Or the recovery? Or both?

I went to the next screen to see what was called Outcome Evaluation. So …. Is this another name for ‘process evaluation’? If yes, then why not just call it that – be out and proud!

1. I am not sure that the timeline proposed is always realistic.

- Rapid assessment comes often on Day 2/3 (72 hours) and detailed needs assessment could also take 2 weeks.

- What is done on Day 1 is normally a Situation Report/or a Briefing Note (check ACAPS – [www.acaps.org](http://www.acaps.org)) based on Secondary Data Review

- PDNA (unfortunately) could also take 2 months it if is referred to the one run by World Bank. At least, this is what I witnessed from my personal experiences during famine/droughts/floods in Southern Africa
2. What is the timeline proposed for the continual/monitoring assessment? Is there any reference to a cycle (i.e., this could be useful for chronic emergencies: such as conflict in Somalia, South Sudan, Syria; drought in Sahel Region; floods in South Eastern Asia)

The monitoring should also include risk monitoring as risks evolve (e.g., in outbreaks), emerge and change during events.

Also, there are cascading events.

Should include the term – rehabilitation as part of recovery. Arguably rehabilitation and reconstruction fall under recovery.

In recovery, there are opportunities to ‘build back better’ and reduce risks thus presenting this framework as a continuum

1. Often lacking accurate baseline information

2. Extent of needs assessment depends on
   a) prior knowledge of a disaster type and typical consequences;
   b) existing capacities and vulnerabilities of the affected community;
   c) geographical extent and intensity of the disaster;
   d) the response actions of other actions

C and D cannot be known in advance, but if an agency has prior knowledge of A and B, they can respond quickly based on C and D

Real-Time Evaluation and Collaborative Joint Evaluation are two different issues. Joint Evaluation normally has a longer lead time than for true RTE

The ‘Day 1, 2, 3’ designations seem arbitrary. These will happen at different times in different settings. This typology leaves out longitudinal and iterative approaches. Some of the best impact evaluations in post-disaster settings take advantage of existing baseline data

While it may be hard to include in this framework I would be keen to see something that emphasises the importance of local input and local review of needs rather than only trusting the opinion of the external needs assessment

See previous comment on post functional review

I miss EWS: early warning systems

I miss RRA: rapid risk assessment, as in needed in e.g., environmental disasters to avert further damage and function loss (BSF).

I miss rapid intervention teams. Of course, these are not part of the consequences, but of the response, but the response phase is not further elaborated in this framework, so I pose this omission here

Does this include situation reports?

Are you also considering information community led evaluations?

I think the framework is great! I like the way you have the consequences crossing through the other lines. If you expand it from post event to the next pre-event to cover the next disaster you could add some extra consequences related to preparedness or behavioural change for the second experience

4. The outcome evaluation typologies section comprised a series of evaluations and/or assessments that occur towards the end of the post-event disaster phase. They reflect information or data that have been collected after an event or disaster has occurred. These evaluations include summative reviews of processes used in managing the event and outcomes related to the post-event status of the community (1).
When asked about evaluation typologies included in the outcomes section reflecting current outcome evaluation methods, the results achieved were positive (n = 32, mean 4.13 (3.82-4.43) and acceptance rate of 84.38%) (refer to Table 37).

The survey asked participants to provide comments related to whether evaluation typologies, as illustrated in outcomes, reflected current outcome evaluation methods and whether there were any outcome evaluation typologies not covered. Responses are provided in Table 42.

**Table 42: Outcomes Comments**

<table>
<thead>
<tr>
<th>Responses to the request for comments related to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• whether evaluation typologies illustrated in outcomes reflected current outcome evaluation methods</td>
</tr>
<tr>
<td>• if there were any outcome evaluation typologies not covered</td>
</tr>
</tbody>
</table>

*Impact on things other than health (e.g., education)*

So, this is confusing. You are calling it outcomes, but then you have process and outcome evaluation in there. In my world (education), we try to keep it simple.

- Process evaluation = what did we do and how well we do it
- Outcome evaluation = what happened as a consequence
- Economic evaluation = was that a good return on investment
- That way, we can talk about three types of evidence (in the main)
  - Evidence of activity – shows us what we did
  - Evidence of process quality – shows us how well we did it
  - Evidence of impact – shows us what happened as a consequence (short, medium and long term)

*It would be useful to clarify the difference between After Action Reviews and Lessons Learned; Peer assessment should be added as a look back facilitation session Retrospective should be assessed Before Action Review could be added here on in the pre-event part. If we consider ‘logistics’ as a component of the response, most of its assets should be evaluated in what was prepared before. It would be great to refer to the creation of an Incident Registry*

*Many agencies are not undertaking Real Time Evaluations during the response stage of an event – perhaps this should be included in the model*

*It is important to define outcomes as the result of the entire process – pre-event and post-event. Outcomes are not limited to post-event, because outcomes can also be achieved during pre-event and event stages*

*Humanitarian evaluations normally focus on formative rather than summative aspects. The reasons for this are complex but include the large number of independent actors seen in a typical response, and the difficulty of attributing responsibility for any outcome between them. Even contribution analysis is at best an educated guess in most circumstances. For example, if mortality falls after the beginning of assistance is it due to: food – and whose food, water – quality and quantity, medical assistance – preventive or curative, support from relatives, cash from agencies, government action, increased economic activity, normal season patterns, the demographic changes wrought by the event (the death of the weakest), movement of populations etc.*

*Same comment as the first box – not sure what this list is. Is this just a laundry list of possible evaluation types? Is it meant to be comprehensive (if so, it is not). Is there a hierarchy or prioritisation implied here?*

*Again, if we consider this framework as a cycle then gov. or agencies should have set some goals in dealing with the next disaster (the current one), hence, goal-based evaluation can be added to the list*
Increasingly one of the measures of outcomes is ‘public opinion’. While not a formal measure the picture painted in the media (and the social media) will determine the political and community view of the outcome. There are medial measures that could be used as a surrogate tool but not available everywhere.

Post event new risk assessment

In my experience, there is a big difference between evaluations a. informal and formal; b. mono vs multidisciplinary; c responder vs sector

I have no expectations of evaluations anymore. Lesson 1 in research is that you must have an undisturbed setting in order to do research, and this can be defined (sic definitely) not be fulfilled in disaster evaluations.

Research based on a priory-tested and validated research instruments

Again, thinking about information evaluations as well as formal

Relating things like building codes to damage assessments post disaster, to build shelter in a more resilient way

There are often community, special interest, cultural evaluations that also take place

Probably in this phase, we will evaluate indicators of the minimum standards and modify if necessary

Overall the evaluation that drives change in policy is the evaluation made by the public and this is largely driven by public media. I recommend that you somehow work this into the framework. Also, public sentiment and cultural drivers influence how evaluations are conducted and interpreted. The tint the lens that influences how results are interpreted

The outcomes are very broad, therefore one can argue anything to be included in those

Don’t know if you have covered basic demographic/population changes – there may not only be health outcome impacts but population change impacts to consider. e.g., post Katrina where a great proportion of the population moved on and the new population differed considerably.

The multi-disciplinary review strategy is quite important

5. Impact evaluations of programs, projects and interventions comprise evaluations that include a measure of causality or attribution and can occur during any phase of the disaster timeline (1). In the disaster setting, impact evaluations have gained popularity for identifying causal links between specific interventions and outcomes. This is a result of the international community demanding accountability and improved evidence-based interventions (3, 16, 98).

When asked about the impacts evaluation section reflecting current impact evaluations methods, the results were below the established level of agreement as previously outlined (n = 30, mean 3.80 (3.42-4.18) and acceptance rate of 63.33%) (refer to Table 37). While the mean was lower than the acceptance criteria, the use of a 95% confidence interval showed the results were acceptable.

It is possible that these results reflect the current, ongoing debates, such as the difficult methodological and theoretical challenges faced related to impact evaluations and the many diverse views held (312, 534-536).
The survey asked participants to provide comments related to whether evaluation typologies, as illustrated in impact evaluations, reflected current impact evaluation methods and whether evaluation typologies could be portrayed in a different manner. Responses are provided in Table 43.

Table 43: Impact Evaluation Comments

<table>
<thead>
<tr>
<th>Responses to the request for comments related to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• whether the evaluation typology of impact evaluations reflected current impact evaluation methods</td>
</tr>
<tr>
<td>• if impact evaluation typologies could be portrayed in a different manner</td>
</tr>
</tbody>
</table>

Baselines/endline studies can be linked with this as a source of information for impact measurement. Generally, my understanding of impact evaluations is it is summative and towards if not at the end of the intervention, as it takes time for outputs and outcomes to realise to assess.

I get frustrated by the idea of impact evaluation being something different from outcome evaluation. What is an ‘outcome’, if it is not something that ‘comes out’? Ergo, the whole idea of outcome evaluation NOT wrestling with the question of causality/contribution is a bit of nonsense.

I am thinking you could have separate timelines for:
- The risk – e.g., hazard, vulnerability, exposure, event …. Damage, disruption
- The capacity to manage the risk – e.g., prevention, preparedness, response and recovery

Otherwise it is confusing to put response and recovery in the same timeline as the risk, damage, disruption …. etc.

Most of the available impact evaluations of humanitarian action are deeply flawed:
- first, there is the problem of attribution
- second is the problem of mechanism

We cannot say, for example what the impact of increasing a ration from 1,500 Kcal per head has on mortality. All the normal measures of impact are the results of multiple inputs (mortality depends on climates, seasons, disease, medical care, food, shelter, water quality, water quantity, access to resources, knowledge of hygiene, sanitation and a host of other factors).

Not sure why you place them where you do but sat ‘they can be placed anywhere’. The difficulty in placing them is driving in part by the forced linear/temporal organisation of the figure. Also, why are they presented completely separately from other outcome evaluation typologies? I agree that they are a different type of evaluation but not a different beast altogether ....

As per the presented timeline, the pre-event period doesn’t include interventions which make IE a not applicable exercise.

Some progressive donors may actually be appropriately humble enough to seek information about contribution, not just attribution, recognising that many other factors and donors may be contributing to any changes which have occurred.

Concerned that having 2 ‘example’ impact evaluation diamonds will lock concrete thinkers into believing that impact evaluations can only occur in these 2 spaces
b) impact evaluations have to occur here each time
Could impact evaluations be included as a footnote or umbrella / bar going across the whole timeline to show that can occur anywhere

From a risk perspective, all interventions come with a declared or undeclared assumption as to the impact they will have. This is the element to be ‘monitored’ as to it having that assumed impact

As a separate horizontal line, in between the pre/post event status and the pre/post event timeline

Maybe this aspect should not be included in the illustration at all
Apart from this: the diamond could not only be between the layers above each other but also in a horizontal manner between the phases

I think causal link assessment is closely linked with inquiry style assessments, in the pre-event stage this would probably be referred to as a risk assessment

The description given must support the diagram – alone, the diagram tells us very little

Potentially as a dotted line across the continuum

I believe the impact evaluation in the post phase has importance as the different timing so I would suggest to place another diamond in the post-event status

Diamonds are often used for decision making, it seems appropriate for this

Evaluation method, type?

Possibly using arrows to where it moves the population concerned, but I like the way you have portrayed it above

6. Accountability was included as an individual element in CFDET 2017, even though it is not generally classified as an evaluation typology.

When asked about accountability and whether it reflected current accountability methods, the results obtained did not meet acceptance criteria (n = 29, mean 3.93 (3.49-4.37) and an acceptance rate of 65.52%) (refer to Table 37). While the mean was lower than the acceptance criteria, the use of a 95% confidence interval showed the results were acceptable. These results may reflect the current debates surrounding accountability and the many diverse views held (459).

The survey asked participants to provide comments related to whether accountability reflected current accountability evaluation methods, and if accountability typologies could be portrayed in a different manner. Responses are provided in Table 44.
Table 44: Accountability Comments

Responses to the request for comments related to:

- whether the evaluation typology of accountability reflected current accountability evaluation methods
- if accountability typologies could be portrayed in a different manner

Include accountability to those you are delivering the aid

SPHERE, Core Humanitarian Standards, Complaints and feedback mechanisms, donor requirements etc.

It would be worth linking accountability to both monitoring and evaluation typologies, i.e., you do them for both learning as well as accountability. You could also reference the ‘Accountability to Affected Populations Initiative’ of the IASC

There is no portrayal of accountability typologies at all, unless you are referring to the ‘up’, ‘down’, ‘across’, ‘down’ typology of donors, stakeholders and beneficiaries. If that is what you mean, I think that there is a fourth one, which is within the organisation itself. It is kind of like the version of the ‘across’, but with rather than between. Also, I am not sure why you single out accountability as a product of evaluation, without also calling out LEARNING

For further deepening, please check ‘Core Humanitarian Standard’ Alliance

What kind of assessment tools are outlined about accountability? Why is there no reference to the Sphere Projects related tools?

Accountability described above seems to be focused on the international situation, rather than accountability being seen in the context of responsibilities in any setting. The reference to donors could be expanded to financial institutions; accountability should be reflected in the accountability of emergency management actors to governments and parliaments; and for the international response community to governing bodies such as the World Health Assembly and the United Nations General Assembly. Ultimately, for example, politicians are accountable to those who put them in power, including the electors, and inquiries can be political as well as operational, resulting in resignations and sackings

Donors, stakeholders, and the affected population have different priorities and accountability to these may be different. (I avoid the term beneficiaries as is assumed that the affected population have benefited from the assistance provided – a leap of faith which is not always justified). Accountabilities may conflict – for example – donors’ accountability demands information on who received assistance, but such information may put those assisted at risk in some circumstances

Are ‘donors, stakeholders, beneficiaries’ meant to be the ‘typologies’ of accountability assessments?

I don’t think people in a context should only be portrayed as ‘beneficiaries’ although I appreciate donors do use this language. To me, there are simply citizens of a country of residents of a location. Rarely, do they see themselves as ‘beneficiaries’ primarily

I think you will need to define very clearly what you mean by accountability. Accountability vs responsibility is often confused

If we go back to the original element of risk management, you identify accountability there.

I have no expertise in this area

Please also read my first remark; this has to do with political, economical and PR context. Maybe this should be designed as a circle around the layers?

On the other hand, this is very much related to the impact evaluation which may lead to some brain cracking how to integrate the event as it happens (and why it happened – backward looking) with the interactions and the evaluation of the interactions (on-line looking)

I agree with the definition provided but would recommend transparency as a component of good accountable practice

Beneficiaries, stakeholders, related organisations and donors to emphasise human-centred approach not money-centred

Include host nation (national, regional, and local levels) as key entities in accountability
... Seems disconnected

Government, public (tax payer)

Donors, stakeholders and beneficiaries – I hate to be negative here but I think beneficiaries implies a benefit and I think 'recipients' or even better maybe 'affected locals' as that doesn’t imply the assistance is helpful. As we know it isn’t always ..... donors I would use ‘foreign responders’ or ‘outside responders’ as again I think there is an implication there. I think the line across the top is good although I think it should be at the bottom in a more supporting location/role. 

Throughout, by ‘current’ I am referring to the framework proposed in this survey.

7. Evaluation standards and evidence included evaluation standards and guidelines, evidence-based reviews and registries, and knowledge management (henceforth referred to as evaluation standards and evidence). It was included as an individual element in CFDET 2017, even though it is not an evaluation typology. Evaluation standards and evidence are important as a cross-cutting theme that covers the entire disaster timeline. They are relevant in advancing the science of disaster evaluations by providing scientific rigour, common terminology and the ability to replicate various methodologies (1).

When asked about evaluation standards and evidence and whether it reflected current evaluation standards and evidence methods, the results were borderline (n = 29; mean 3.97 (3.57-4.40) and acceptance rate of 82.76%) (refer to Table 37). While the mean was lower than the acceptance criteria, the use of a 95% confidence interval showed the results were acceptable.

The survey asked participants to provide comments related to whether the evaluation typologies illustrated in evaluation standards and evidence reflected current evaluation standards and evidence methods, and if there were any evaluation standards and evidence typologies not covered in evaluation standards and evidence. Responses are provided in Table 45.
### Table 45: Evaluation Standards and Evidence Comments

Responses to the request for comments related to:
- whether the evaluation typologies illustrated in evaluation standards and evidence reflected current evaluation standards and evidence methods
- if there were any evaluation standards and evidence typologies not covered in evaluation standards and evidence

<table>
<thead>
<tr>
<th>Separate reviews from databases, show the individual studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sphere, CHS, ALNAP etc.</strong></td>
</tr>
</tbody>
</table>

*I think this is the ideal and what many organisations are striving towards, but I don’t believe it’s been widely adapted. For those agencies attempting this there are many challenges to evaluation – internal and external. As you use the term Knowledge Management, then it is more ‘research’. Research could be a separate box, or Knowledge Management should take a broader strategic approach, e.g., refer to WHO Knowledge Management Strategy*

<table>
<thead>
<tr>
<th>Are these boxes meant to be MECA? Why are they located where they are in the framework?</th>
</tr>
</thead>
<tbody>
<tr>
<td>I believe that some disasters have already evaluation standards or management systems which include some sort of standards, e.g., health disaster management. So, it might be better to list the existing sets (if applicable)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Not sure where monitoring fits into this</th>
</tr>
</thead>
<tbody>
<tr>
<td>The typologies seem to generic. There are many specific and focussed research questions and null hypotheses that can be studies. They will only refer to one element. As I see it, much disaster research wants too much at the same time while also a lot can be understood by just focusing on one aspect</td>
</tr>
</tbody>
</table>

| Unclear what the expected scope is here – appears to be encompassing elements of professional evaluation, research practice and ethical action, each of which are multifaceted and have additional potential elements. For example, does the reference to Evaluation and Guidelines assume that this encompasses all practice which contributes to the development of such material, that is, professional practice of individuals, organisations and services as well as development and application of the process of evaluation? Should / does this consider elements specific to generic disaster research, which may or may not ultimately contribute to such documents? |

<table>
<thead>
<tr>
<th>Ethical consideration should be placed in this domain</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Question the relevance of Knowledge Management?</em></td>
</tr>
</tbody>
</table>

### 8.4.3 CFDET 2017

The results in Table 46 are related to four specific questions regarding CFDET 2017 as a complete unifying framework.
In the first question, participants were asked to rate and comment on whether the relationships between the evaluation typologies and the disaster timeline were appropriately demonstrated in CFDET 2017. The results received were positive (n = 30, mean 4.10 (3.80-4.40) and an acceptance rate of 86.67%) (refer to Table 46).

The survey asked participants to provide comments related to whether the relationships between the evaluation typologies and the disaster timeline were appropriately demonstrated in the comprehensive framework. Responses are provided in Table 47.

Table 46: Survey Responses to CFDET 2017 as a Unifying Framework

<table>
<thead>
<tr>
<th>COMPREHENSIVE FRAMEWORK FOR DISASTER EVALUATION TYPOLOGIES 2017</th>
<th>n</th>
<th>MEAN +/- 95% CI</th>
<th>ACCEPTANCE THRESHOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationships</td>
<td>30</td>
<td>4.10</td>
<td>86.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.80 – 4.40</td>
<td></td>
</tr>
<tr>
<td>Value</td>
<td>29</td>
<td>4.28</td>
<td>86.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.97 – 4.58</td>
<td></td>
</tr>
<tr>
<td>Support evaluations</td>
<td>29</td>
<td>4.28</td>
<td>82.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.96 – 4.60</td>
<td></td>
</tr>
<tr>
<td>Teaching</td>
<td>30</td>
<td>4.33</td>
<td>86.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.03 – 4.63</td>
<td></td>
</tr>
</tbody>
</table>

Table 47: Relationship Between Evaluation Typologies and Disaster Timeline

Comments

Responses to the request for comments related to whether the relationships between the evaluation typologies and the disaster timeline are appropriately demonstrated in the comprehensive framework

‘Appropriate’ can be relative to the context and stakeholders

Related, the conceptual framework is rather dense – does it help users make sense of the typologies (e.g., a snapshot that summarises key concepts), or is it confusing? I guess, related to this is who is the audience – if it is academics or technical specialists, the framework will likely have more currency that a program manager or other user

It makes sense for the way it is illustrated. The only concern is that it is quite linear in its approach and does not take into account the vast complexities of this work, both from a timeline perspective and a phase perspective. There are lot of assumptions behind the diagram, including a single organisation involved in a single event. It might also be worth highlighting somewhere that this framework has not said anything about how the typologies are implemented or approached

Great work putting it altogether into one diagram. It’s pretty intuitive. Well done!
I still don’t understand the difference between consequences and outcomes thought.
Maybe you could have LEARNING at the bottom (as distinct from accountability at the top) – maybe the feedback look could pass through that?

The only thing that got lost is the sense of ‘cycle’ which is so important in Disaster Risk Reduction
If you use a risk based approach then the baselines refer to underlying risk – does this change over time, including the effects of prevention and mitigation measures on controlling or reducing risk. A risk based approach would also indicate measures to ‘prevent and mitigate’ risk are also contributors to outcomes. At the moment it looks like the focus on outcomes is the result of the response and recovery efforts. National and sub-national polices, legislation and frameworks (and regional frameworks) should be included along the global frameworks. The former has more influence that international frameworks. The International Health Regulations should also be included among the international frameworks.

Yikes! I am sorry to be so negative. I am struggling to glean how this framework is useful? Does it help me decide what kind of evaluation I need to do? When to do it? How to prioritize?

It is indeed a comprehensive. In evaluation, Impact comes after outcome. Similarly, IE is aiming to assess and measure the impact that resulted from the (inputs – outputs – outcomes). So, pushing IE toward the end of the framework and link it back to the baseline in forms of feedback loop might be a good idea.

I don’t have enough expertise in this area to really understand the complexity of the details you’ve included here, but it looks roughly okay to me. I’m not into extreme rigour (because my preference is towards providing opportunities for people in countries and NGOs to learn, reflect and consider ideas, rather than demonstrate accountability) and don’t think that everything in this area needs to be tightly defined. To me, it’s just an area of work, with some appropriately fuzzy edges, to reflect the complexity of power relationships, cultural value differences (between cultures that think accountability is important and others that think the relationships are more important) and various alliances and connections.

Not sure how to do it but think the link between evaluation standards / evidence / knowledge management and the rest of the framework is loose. Almost looks like these only occur at pre-event / event state the way that diagram is structured. What if these were all in one box above accountability to that reinforces the concept that can and should be used at any stage as appropriate.

This is generally what we do, true, but is the very similar to a risk management framework that would be applied and hasn’t been properly articulated, i.e., all of these elements/actions are part of managing the risk to the resilience aim.

See all remarks above. This is a good start but might miss some elements. On the other hand, one can endless work on this without ever being perfect so at some stage it has to be decided ‘this is it’ (for the time being).

All the main points are covered, I am unsure that in reality there is always clearly defined boundaries between disaster timeline elements and therefore evaluation typologies.

I agree that the relationships are appropriately demonstrated. However, my perception may be based on the information in each element being explained in individual detail prior to. Although I don’t think this is the case. I would like to not again that the corresponding detail needs to be linked to impact evaluations for those symbols to have a greater meaning.

The arrows showing the relationship between consequences and outcomes and the feedback loop to baselines should be double ended, i.e., indicating that the impact goes both ways.

It explains key elements to consider, and how they relate to each other to produce outcomes.

Clean and clear. Layers show the complexities.

Great work – a comprehensive overview allows all actors to see where their work fits and how it is related to other sections or how it crosses.

Perhaps I was not rating these various sections correctly. My ‘framework’ was to take each of the figures along the way to the final Figure Eight, which put it all together. I think this final framework is excellent.

The second question asked was based on whether CFDET 2017 potentially held value for the participant when undertaking evaluations in the disaster or humanitarian setting.
While the number of responses was slightly lower than preferred (n = 29), the overall results and comments received were positive (mean 4.28 (3.97-4.58) and an acceptance rate of 86.21%) (refer to Table 46).

The survey asked participants to provide comments related to whether the comprehensive framework potentially held value for their work when undertaking evaluations in this setting. Responses are provided in Table 48.

**Table 48: Comprehensive Framework Potential Value Comments**

<table>
<thead>
<tr>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses to the request for comments related to whether the Comprehensive Framework potentially holds value for their work when undertaking evaluations in this setting</td>
</tr>
<tr>
<td><em>I think this can inform different understandings of the range of not just evaluation, but sources of data to analyse and inform better service deliver in humanitarian contexts</em></td>
</tr>
<tr>
<td><em>I don’t work in this setting</em></td>
</tr>
<tr>
<td><em>It’s an ideal to aspire to – not sure how many agencies will have the resources to undertake this level of evaluation</em></td>
</tr>
<tr>
<td><em>There is a need to show evaluation as a ‘continuum’ – which your framework shows. The results of evaluation will review each aspect of ‘risk management’ practice, address accountability, make recommendations and providing evidence for improving practice and strengthening capacity etc.</em></td>
</tr>
<tr>
<td><em>In development sector, someone can look at poverty as a disaster. Hence, applying this framework would result in positive results</em></td>
</tr>
<tr>
<td><em>It could be valuable, but it isn’t the only thing that would frame any evaluation in this area</em></td>
</tr>
<tr>
<td><em>Would prefer to see in context of risk management</em></td>
</tr>
<tr>
<td><em>It helps to understand the wider context and provides insight how current and future studies related to the overall picture. It is also helpful for educational purposes</em></td>
</tr>
<tr>
<td><em>It adds value for typology selection when conducting evaluations – some testing (by evaluating) to validate would be great</em></td>
</tr>
</tbody>
</table>

The third question was based on whether CFDET 2017 would be useful for supporting and promoting evaluations in the disaster or humanitarian setting. While the number of responses was slightly lower than preferred (n = 29), the overall results received were positive (mean 4.28 (3.96-4.60) and an acceptance rate of 82.76%) (refer to Table 46).

The survey asked participants to provide comments related to usefulness of the comprehensive framework for supporting and promoting evaluations in the disaster or humanitarian setting. Responses are provided in Table 49.
Table 49: Supporting and Promoting Evaluations Comments

<table>
<thead>
<tr>
<th>Responses to the request for comments related to the usefulness of the comprehensive framework for supporting and promoting evaluations in the disaster or humanitarian setting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>As clearly explained in the introduction of this survey, there is not clear standard and framework for disaster evaluation. Such a well-developed systematic framework would definitely help in supporting and promoting sound disaster evaluation</strong></td>
</tr>
<tr>
<td><strong>It depends on who is being asked – donors and NGOs all have their own ideas about what makes an evaluation effective and useful and how it’s used etc.</strong></td>
</tr>
<tr>
<td><strong>It synthesizes the disaster evaluation process</strong></td>
</tr>
<tr>
<td><strong>The framework is useful visual representation that can be used to support and promote evaluations. This could work at all government levels, and at the community level – households, businesses, clubs etc.</strong></td>
</tr>
<tr>
<td><strong>I can come back to this framework if I miss to evaluate during the field work when I am supposed to engage multitasking</strong></td>
</tr>
<tr>
<td><strong>Identifies the factors to evaluate / measure</strong></td>
</tr>
<tr>
<td><strong>Evaluators, particularly workers who are researching as a result of an event and are not academics would benefit from a clear framework to base their evolution on</strong></td>
</tr>
<tr>
<td><strong>I would question the value of this. It is good as a reference point but is not aiding in evaluation by providing a methodology</strong></td>
</tr>
<tr>
<td><strong>Again, by providing an internationally agreed upon framework it has the benefit of providing guidance on the research and clearer understanding of responsibilities in this field</strong></td>
</tr>
</tbody>
</table>

The final question put forward was related to the complete and re-constructed CFDET 2017. It was based on whether the comprehensive framework would be useful for teaching evaluation in the disaster or humanitarian setting. The results received were positive (n = 30, mean 4.33 (4.03-4.63) and an acceptance rate of 86.67%) (refer to Table 46).

The survey asked participants to provide comments related to whether the comprehensive framework would be useful for teaching evaluation in the disaster or humanitarian setting. Responses are provided in Table 50.
Table 50: Usefulness in Teaching Evaluation Comments

Responses to the request for comments related to whether the comprehensive framework would be useful for teaching evaluation in the disaster or humanitarian setting

<table>
<thead>
<tr>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>It could be useful, alongside a whole lot of other material on different aspects of evaluation. Disaster evaluation shouldn’t really be much different from other focus areas of evaluation really</td>
</tr>
<tr>
<td>As mentioned above, each element is clearly linked to the various phases of disaster and provided a very useful visual representation, that is not only useful for garnering support but also for pulling apart in a teaching environment</td>
</tr>
<tr>
<td>Again, useful as overview, starting point and to consider the conceptual basis</td>
</tr>
<tr>
<td>The framework graphically aids students learning the complexity and importance of relations among all elements</td>
</tr>
<tr>
<td>Breaks down the system to simpler parts</td>
</tr>
<tr>
<td>Clean and organised</td>
</tr>
<tr>
<td>Helps to understand the process</td>
</tr>
<tr>
<td>As above, ‘it would particularly be useful in explaining to others where the research fits, how it can be applied, what research could follow etc. It provides a clear visual representation of any work in the area’</td>
</tr>
<tr>
<td>This is a very logical step by step framework</td>
</tr>
</tbody>
</table>

Participants were asked to provide generalised comments about the comprehensive framework as a whole from an improvement perspective. Responses are provided in Table 51.

Table 51: Additional Thoughts on Improving the Comprehensive Framework

<table>
<thead>
<tr>
<th>Thoughts on improving the comprehensive framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>At this point, I would say the best way for improvement is by testing the framework on the ground</td>
</tr>
<tr>
<td>My suggestion is not to suggest it replaced all other forms of organising ideas in this area of work. Just suggest it adds to and pulls together many other elements</td>
</tr>
<tr>
<td>See comments through survey</td>
</tr>
<tr>
<td>Prepare it as a risk management framework. Maybe 8r’s typology, based on Ellis 5 R’s – research, risk reduction, readiness, response, relief, recovery (functional), review, recovery (resilient)</td>
</tr>
<tr>
<td>Depends on the feed-back on this questionnaire. If the response is confusing and very varied, each step in the process that has led to this framework should be reconsidered. If there is much unanimity, this can be adapted straight forward</td>
</tr>
<tr>
<td>Consider cross over / merging at the intersection of timeline elements</td>
</tr>
<tr>
<td>Those that I have already mentioned</td>
</tr>
<tr>
<td>Please be bold enough to revise once this work is out and be utilised for a certain period of time</td>
</tr>
<tr>
<td>I have always thought that the linear approach is really does not accurately represent the situation. I have always thought that the model should be represented as a circle, or as a line with feedback loops. It does not illustrate how the end outcome actually impacts pre-event status for future events. (The factors are accurate, but the way the model is depicted should be revised)</td>
</tr>
<tr>
<td>Only to add the element of preparation for an expected disaster, e.g., a storm</td>
</tr>
</tbody>
</table>
I guess the only comment now is that you could make many subnotes for each section of the diagram to clarify what each section means. I think that the diagrams should be easily understood by those with no knowledge of the field.

Participants were asked to provide comments on any barrier(s) that they could perceive in undertaking evaluations in the disaster setting. Responses are provided in Table 52.

**Table 52: Perceived Barriers to Undertaking Evaluations in the Disaster Setting**

<table>
<thead>
<tr>
<th>Responses to the request for comments on any barriers perceived in undertaking evaluations in this setting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The disaster setting is complex and sophisticated by nature and that is the first barrier (or challenge) in conducting these types of evaluations</strong></td>
</tr>
<tr>
<td><strong>The main barrier in my experience is the expectation of donors that their tiny contribution was the key contributor to change. Evaluations are largely flawed when taken from the donor perspective, especially when they struggle to reflect the information and reflection priorities of citizens in countries affected by disasters</strong></td>
</tr>
<tr>
<td><strong>Creating awareness let alone change. I can see resistance coming for adherents to the existing WADEM Research Framework which is overly complicated and hardly usable as a result</strong></td>
</tr>
<tr>
<td><strong>The usual suspects: short term finances due to the last disaster, losing expertise and manpower when money dries up, and then starting all over again with new people with limited expertise in the next disaster.</strong></td>
</tr>
<tr>
<td><strong>Poorly prepared researchers who have to start at once, and research that has to be invented in a few hours</strong></td>
</tr>
<tr>
<td><strong>Poor understanding of evaluation</strong></td>
</tr>
<tr>
<td><strong>I am not sure if I could fully understand the concepts you have written in this article. Figures do not clearly illustrate the concept of evaluation typology</strong></td>
</tr>
<tr>
<td><strong>Accessibility and consistency of data; use of terminology (e.g., varied by state; can be overcome)</strong></td>
</tr>
<tr>
<td><strong>Barriers are related to the actual processes needed to make the concepts operational, and as such are variable within different settings</strong></td>
</tr>
<tr>
<td><strong>Need defined operational definitions for each factor, and validated variables to measure each factor. From this we could develop validated scales. That way we are all measuring the same things and can more accurately measure intensity</strong></td>
</tr>
<tr>
<td><strong>Lack of experience</strong></td>
</tr>
<tr>
<td><strong>Evaluations should have a means of deciding any aspect is good, bad or worse. There is not mechanism for that in this framework, not it points out what methodologies are available</strong></td>
</tr>
<tr>
<td><strong>I presume you mean barriers to evaluations in disasters? Ethical barriers, issues of RCTs, population dispersal etc.</strong></td>
</tr>
<tr>
<td><strong>The ability to develop practical, on-the-ground tools and guidelines for those who will be doing the evaluations – in many cases, relatively unsophisticated.</strong></td>
</tr>
<tr>
<td><strong>The diagram itself is, by necessity, so complex and difficult to easily grasp as a visual. This will be off-putting to many practitioners</strong></td>
</tr>
<tr>
<td><strong>By definition, evaluations are not seen as priorities among EM/humanitarian workers. Further research should be undertaken to identify/overcome this cultural barrier</strong></td>
</tr>
</tbody>
</table>
Participants were asked to provide comments on any enabler(s) that they could perceive in undertaking evaluations in the disaster setting. Responses are provided in Table 53.

### Table 53: Perceived Enablers to Undertaking Evaluations in the Disaster Setting

<table>
<thead>
<tr>
<th>Responses to the request for comments on any enablers perceived in undertaking evaluations in this setting</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity building, increasing awareness and sharing knowledge are key elements for better application of these evaluations</td>
<td></td>
</tr>
<tr>
<td>Early publication of evaluations linked to the framework</td>
<td></td>
</tr>
<tr>
<td>Have a number of evaluation issues prepared as ‘half baked bread’ and put them in the political fire when there is a right moment. Most of all this refers to policy making. Maintain good relationships, and build up confidence with the right people at ministries and sponsor organisations, be pro-active</td>
<td></td>
</tr>
<tr>
<td>Clear guidance for evaluators</td>
<td></td>
</tr>
<tr>
<td>The use of a national framework will help promote the various elements and enhance understanding of the need to undertake evaluations in the various settings suggested by the framework</td>
<td></td>
</tr>
<tr>
<td>Recognition of generic utility and potential for comparisons that are apparent with a process such as this</td>
<td></td>
</tr>
<tr>
<td>If we all use the same framework we can compare outcomes. Thank you for undertaking this project!</td>
<td></td>
</tr>
<tr>
<td>Enthusiasm</td>
<td></td>
</tr>
</tbody>
</table>

Validation results were received and collated. The number of results, the mean and the acceptance rates were calculated and presented in this chapter. Overall, the results received from the 11 main questions were positive, with the majority of responses resulting in a mean > 4.00 and an acceptance rate of > 80%. The responses with a mean < 4.00 were still within range with the application of a 95% CI.

A review of the results indicated that improvements to the comprehensive framework could be made. These changes, including the reason(s) why they were adopted, are discussed in detail in the next section.

### 8.5 Discussion

One of the major limitations of previous frameworks is their lack of validation and subsequent use to structure disaster evaluation and research (1). To overcome this particular shortfall, an international peer-reviewed validation process was undertaken to seek acceptance and validation of CFDET 2017.

It is important to involve end users in not only the development process but in the validation of the comprehensive framework, to facilitate the identification of issues and
additional ideas that could enhance the framework and increase its usability and uptake. The involvement of end users also creates ownership of the process and the subsequent outcome, which enhances and creates a stronger framework (537).

While every attempt was made to ensure the validation process was as robust as possible, one limitation was identified, relating to obtaining a suitable sample size. To overcome this, the survey period was extended and additional experts were invited to participate until the acceptance threshold of responses was received.

Based on the comments received from the survey, minor modifications were adopted to enhance the readability, usability and understanding of CFDET 2017. In addition, the opportunity was taken to reflect on recent updates based on an environmental scan of the literature. Each of the changes will be discussed based on the diagram they relate to and supporting rationale for their adoption presented. To alleviate any bias and enhance rigour, a mixed methods approach was used to assist with data collection, thereby enabling triangulation of the findings.

Overall, the majority of results from the validation survey supported CFDET 2017. Some of the responses, however, indicated that changes to CFDET 2017 would enhance the unifying framework. Significantly, no participant suggested the inclusion of any other additional evaluation typologies. While this was reassuring for the validation of CFDET 2017, as is standard practice in finalising surveys, projects or theses, a quick environmental convenience scan was undertaken to see if there were any typologies not identified in the previous literature review or the international survey and that may have emerged recently.

These changes are discussed in the following sub-sections specific to the core structure, four evaluation typologies and two overarching themes. The adoption of these changes has led to the development of a modified framework; to avoid any confusion, this will be known as the Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018).

8.5.1 Core Structure

The original CFDET 2017 core structure (refer to Figure 24) consisted of the following three layers:
Layer 1. Simplified disaster timeline including stages of pre-event, event and post-event (refer to Figure 25).

Layer 2. A more detailed look at the disaster timeline identifying pre-event status, hazard, risk reduction, event, damage, disruption, change in functions, consequences, response, recovery and post-event status (refer to Figure 26).

Layer 3. Strengthening resilience (refer to Figure 27).

Figure 24: CFDET 2017 Core Structure

While the core structure as a whole was a single section in the survey, a total of 26 comments were received (refer to Tables 38 and 39). The comments are discussed based on the three distinct layers that make up the core structure.
8.5.1.1 Core Structure Layer 1

Participants did not provide any comments that could be directly related to this layer; therefore, no modifications were required.
8.5.1.2 Core Structure Layer 2

The comments from survey participants were reviewed; the following comment was considered extremely relevant and led to improvements in Layer 2:

Should include the term – rehabilitation as part of recovery. Arguably rehabilitation and reconstruction fall under recovery. In recovery, there are opportunities to ‘build back better’ and reduce risks thus presenting this framework as a continuum.

Reflecting on the above comment, the term ‘rehabilitation’ was added to the recovery section in Layer 2. In the Sendai Framework, rehabilitation is mentioned as a major priority (refer to Priority 4: Enhancing disaster preparedness for effective response) and is also highlighted in the section ‘Build Back Better’ (58). For the purpose of this thesis, rehabilitation is defined as ‘a set of measures aimed at restoring normal living conditions through the repair and reestablishment of vital services interrupted or degraded by a disaster or emergency’ (446)(p.76). It is important to note that the included recovery actions can occur concurrently and are not mutually exclusive. The recovery section, which originally read as ‘(recover grow/thrive/transform), Reconstruction, Development, Renewal, Regeneration’ in CFDET 2017 now reads as ‘Reconstruction, Renewal,
Regeneration, Rehabilitation, Development’. This is reflected in Figure 28, which shows the revised core structure for CFDET 2018.

8.5.1.3 Core Structure Layer 3

Survey participants provided two comments in relation to Layer 3 of the core structure. Of these, the following comment was considered extremely relevant and led to improvements in Layer 3:

The International Health Regulations should also be included among the international frameworks.

On reflection, this comment was deemed valid, and a change to Layer 3, known as ‘strengthening resilience’, was made to include the addition of the International Health Regulations (IHR) 2005 (538). Further, the researcher took the opportunity to add two additional international frameworks deemed globally significant: The World Humanitarian Summit (WHS) 2016 (345) and Habitat III 2016 (539).

The IHR 2005 is an international legal instrument legally binding on the 196 signatory countries, including the member states of the WHO (538). IHR 2005 provides the
governing framework for global health security (540) and aims to help international communities prevent and respond to acute public health risks that have the potential to cross borders and threaten the health of people worldwide (538, 541).

The WHS 2016 was first convened in May 2016 at the behest of the former UN Secretary-General Ban Ki-Moon (345). The aim of the summit was to generate international commitment to prevent the suffering of millions of people around the world caught up in humanitarian crises, instead of just managing them (542). Further, it was an attempt to support the new Agenda for Humanity and bring meaningful change to the world’s most vulnerable people (539). It was believed that the ‘world could better respond to, and prevent, humanitarian crises’ (345). The commitments made by the UN Member States in relation to WHS 2016 are not legally binding, and no independent accountability mechanism or framework to track progress has been announced (543).

The UN Conference on Housing and Sustainable Urban Development, better known as Habitat III 2016 (539), is another framework added to ‘strengthening resilience’. Habitat conferences take place once every 20 years and provide an opportunity to develop a sustainable plan for a rapidly urbanising world. Currently, 54% of the world’s population live in urban areas, a percentage that is expected to increase to 68% by 2050 (544). At the foundation of the conference is the New Urban Agenda, a declaration of intent to make cities more liveable and consider and plan for the impacts of housing, migration, infrastructure, climate change and inclusion issues (545).

Additionally, resilience in CFDET 2017 core structure already includes ‘Rockefeller 100 RC Framework’. Evolving from this international framework has been a companion typology ‘The City Resilience Index’ designed to measure resilience in the setting of the Rockefeller City Resilience Framework (449). Although influential in guiding the 100 cities within the Rockefeller 100 Resilient Cities Project, it may be too early to determine its use or its influence on resilience in those cities. In Australia, an Australian Natural Disaster Resilience Index is also evolving (451-453, 546), but again it is too early to determine its use or its influence on resilience in Australia.

The modified strengthening resilience section now reads:

- System structure: governance, coordination and leadership.
- ‘Culture of safety’: risk reduction, prevention and mitigation.
• Reduce exposures and vulnerabilities, build anticipative, absorbing and adaptive capacities, and promote community development (Bahadur, 2015).


All guidelines and frameworks included in the core structure are significant at an international level and their inclusion demonstrates the contemporary nature and living status of the comprehensive framework. Although not all are legally binding, these international frameworks have shaped policy and practice globally (547) and add rigour and strength to the comprehensive framework. It is unlikely that further international frameworks of this significance will be held before 2027, as many are due for review within the next 10-15 years, such as the Sendai Framework (58) and the SDGs (548).

In summary, two changes were made to the core structure. These are now reflected in Figure 28, which shows the modified core structure.

Figure 28: CFDET 2018 Core Structure
8.5.2 Baselines

The original CFDET 2017 baselines are shown in Figure 29 and consist of:

- baseline evaluations/evaluability
- demographics and infrastructure (GIS)
- epidemiology and emerging threats
- define minimum standards/criteria: baselines, targets and indicators
- preparedness, resilience capacity (Sendai Framework for Disaster Risk Reduction 2015–2030)
- risk management, surveillance and early warning
- HIA predictive.

Survey participants provided a total of 13 comments in relation to baselines (refer to Table 40). A review of these by four researchers identified comments that led to improvements in baselines, including:

I was a bit surprised ‘evaluability’ is there. I normally think of that as an assessment of how possible it would be to evaluate something. Is that possible to do, when you do not know what that ‘something’ even is?

Figure 29: CFDET 2017: Baselines
It looks good, but maybe too focused on health and not on other sectors. There is no reference to Common Operational Datasets and to Fundamental Operational Datasets (COD and FOD), there is not reference to capacity situation, there is not reference to pre-existing vulnerabilities.

A review of these comments led to the following modifications to baselines:

- ‘Evaluability’ was moved from baselines to the overarching theme of evaluation standards and guidelines.
- The term ‘assessment’ has been included and now reads ‘Baseline Evaluations/Assessments’. Baseline evaluations and baseline assessments remain an integral part of disaster risk reduction as they provide a critical reference point for assessing changes and impacts (549). Further, they establish a basis for comparing the community situation before and after a disaster. The inclusion of the word ‘assessments’ is intended to provide further clarity in the baselines section.
- Changes to demographics and infrastructure were made and now include ‘demographics, infrastructure (GIS) and existing common statistical and operational databases’. ‘Common statistical and operational databases’ include common operational datasets (COD), functional operational datasets (FOD) and relevant databases. CODs and FODs are fundamental operational datasets used in humanitarian emergencies. They are important in improving the effectiveness of humanitarian assistance, in supporting technical standards, improving the quality of the data and strengthening interoperability and harmonisation. Both datasets are developed and endorsed by the IASC (550, 551) and maintained by OCHA (552).
- The final change to baselines was the inclusion of the title ‘baseline typologies’ as a descriptor to provide clarification and direction to the items included in the baselines.

In summary, four changes were made to baselines as a result of the feedback received.

Additionally, a recent environmental scan of the literature identified the expansion of the use of Targets and Indicators in this domain. This typology is already included in CFDET 2017 Baselines but their use has now evolved into ‘Resilience’ and ‘Outcome’ phases of the disaster timeline and has also been consolidated in the Baseline Typologies reflected in CFDET 2018.
CFDET 2017 Baselines already included ‘Define Minimum Standards Criteria: Baselines, Targets and Indicators’, however, to reinforce the importance of this typology and in recognition of the evolution of the Targets and Indicators developed and endorsed by the UN after the Sendai Conference, an opportunity was taken to make an additional change to the comprehensive framework. The emergence of Targets and Indicators is a sign of the maturing of the disaster evaluation approach. Significantly, Targets and Indicators will now be used as the framework for all countries to report on their progress towards disaster risk reduction goals within the Sendai Framework every two years (58).

Point four of CFDET 2018 Baseline Typologies will now read as:

- Define Minimum Standards Criteria: Baselines, Targets and Indicators, Disaster Risk Reduction Targets and Indicators (Sendai).

The UN Targets and Indicators, and additional support resources, are available online at the following URL: https://www.preventionweb.net/drr-framework/sendai-framework-monitor/ (441). Australia has already submitted its first progress report Australia: Sendai Framework Data Readiness Review, 2018 (553).

These changes are reflected in Figure 30, which shows the modified baseline typologies.

Figure 30: CFDET 2018 Baseline Typologies
8.5.3 Consequences

The original CFDET 2017 consequences (shown in Figure 31) consisted of the following:

- rapid needs assessment (damage), Day 1
- detailed needs assessment (functional), Days 2–3, PDNA
- continual assessments: monitoring and surveillance
- independent real time and collaborative joint evaluations.

![Figure 31: CFDET 2017 Consequences](image)

Survey participants provided 13 comments in relation to consequences (refer to Table 41). Of these, the following comments led to improvements in consequences:

The ‘Day 1, 2, 3’ designations seem arbitrary. These will happen at different times in different settings. This typology leaves out longitudinal and iterative approaches.

I am not sure that the timeline proposed is always realistic. Rapid Assessment comes often on Day 2/3 (72 hours) and Detailed Needs Assessment could also take 2 weeks. What is done on Day 1 is normally a Situation Report/or a Briefing Note based on Secondary Data Review. PDNA (unfortunately) could also take 2 months.
A review of these comments led to the following modifications being made to the consequences section:

- The term ‘for example’ (e.g.) was added to the time frames mentioned under rapid needs assessment (damage), e.g., Day 1 and detailed needs assessment (functional), e.g., Days 2–3 as per the IASC Initial Rapid Assessment (IRA) documents (554).
- The final change to consequences was the inclusion of the title ‘consequence typologies’ as a descriptor, which provides further clarification and direction to the items included in consequences.

In summary, two changes were made to consequences. These are reflected in Figure 32, which shows the modified consequence typologies.

![Figure 32: CFDET 2018 Consequence Typologies](image)

8.5.4 Outcomes

The original CFDET 2017 outcomes (shown in Figure 33) consisted of the following:

- operational/strategic/institutional reviews (internal): debrief, after action reviews, lessons learned
• quality reviews (external): audit, key performance indicators, quality improvement  
• government inquiries  
• multi-disciplinary event reports/reviews  
• process and outcome evaluations  
• HIA evaluations.

Survey participants provided 21 comments in relation to outcomes. While all comments provide some value, one comment in particular was considered significant and led to the most important change to the comprehensive framework:

There is no portrayal of accountability typologies at all.

This particular comment led to the following two changes to the framework:

• The addition of accountability as an outcomes evaluation method. This will be discussed in more detail in the accountability section (refer to Section 8.5.6).
• The title outcome typologies was included as a descriptor to provide further clarification and direction to the items within ‘outcomes’.

Figure 33: CFDET 2017 Outcomes
In addition, the opportunity was taken to include ‘Recovery Indicators’. This new typology has emerged and is used in the recovery phase of the disaster timeline. Recent literature suggests an attempt to identify factors that may lead to effective or good recovery and strategies to monitor the community’s progress towards recovery after major events (430, 433). This approach has evolved into developing and validating indicators of disaster recovery. In the USA, Dwyer and Horney have led an extensive process to validate key indicators for disaster recovery (555). In the Australian setting, a set of outcomes, indicators and standards of successful recovery are referenced from the *Australian Monitoring and Evaluation Framework for Disaster Recovery Programs, 2018* (434). These are shaping the direction of data collection and measurement in this sector. The Australian model is already into version two and the Australian Institute for Disaster Resilience (AIDR) is calling for submissions from recovery studies based on Argyrous’s framework.

Overall, three changes were made to outcomes. These are reflected in Figure 34, which shows the modified outcome typologies.

![Figure 34: CFDET 2018 Outcome Typologies](image-url)
8.5.5 Impact Evaluations

The original CFDET 2017 impact evaluations section (shown in Figure 35) consisted of the following:

- Two ‘diamonds’ labelled ‘IE + / C’ – ‘Impact Evaluations + / - Causality of Interventions’. Impact evaluations can be undertaken during any phase of the disaster timeline.

Figure 35: CFDET 2017 Impact Evaluations

Eighteen comments were received from participants (refer to Table 43); however, they reflected different interpretations and definitions related to impact evaluations. There appeared to be confusion, contention and disagreement regarding definitions and terminology, and this was supported by the responses received. A mean of 3.80 (3.42-4.18) revealed that respondents did not know how to display impact evaluations differently and the low acceptance rate of 63.3%, which was the lowest score received during the survey, reflected the international controversy or disagreement surrounding these. At an international level, there is ongoing contention about the concept of impact and impact evaluations (312, 509, 556) and this was supported by the results received.
Although there remains ongoing debate about the exact definition of impact evaluations (312, 509, 556), they are particularly well suited to answer important questions such as whether interventions do or do not work, whether interventions have a positive or negative impact, whether there are intended or unintended consequences, and how cost effective they are (305, 310). It is believed impact evaluations greatly improve the effectiveness of interventions delivered in the disaster setting by identifying what works for whom, why and at what cost (16, 312).

No modifications were made regarding impact evaluations.

8.5.6 Accountability

Originally, accountability, as shown in Figure 36, was portrayed in CFDET 2017 as an overarching theme across the framework, and consisted of the following:

- Accountability:
  - donors
  - stakeholders
  - beneficiaries.

Figure 36: CFDET 2017 Accountability
A review of the 20 comments received from participants reflected different interpretations and definitions related to accountability. As mentioned in outcomes (refer to Section 8.5.4), ‘accountability’ was removed as an overarching theme and consequently as a stand-alone diagram. In recognition, however, of its importance in the evaluation process and responsibility to each level of stakeholders, the term ‘accountability’ was included under outcome typologies. This shift also supports the comments received as part of the survey.

There appeared to be disagreement regarding definitions, terminology and whether accountability could be included as a stand-alone ‘evaluation typology’; this was supported by the responses received. This was the first time during the survey that the number of participants responding fell below the required sample size of 30. The mean of 3.93 (3.49-4.37) was the second lowest mean score of all survey questions and this highlighted the fact that respondents did not know how to display accountability differently. Further, the low acceptence rate of 65.5% was the second lowest score received during the survey and this supports the complex nature of accountability (557) and lack of international agreement on the definition of ‘accountability’ (459).

Accountability and learning are often mentioned as two main goals of evaluation (479). Accountability was included in CFDET 2017 because of its significance as a cross-cutting theme and the overarching concept of effecting better quality and performance in humanitarian action. Since the 1990s, there has been a push to increase quality and accountability in the humanitarian sector, including all aspects of work/interventions delivered to the communities and people affected (279, 558) by disasters. It is a belief of the researchers that there is a moral imperative to ensure accountability is at the forefront of all that we do in the disaster setting, whether that accountability pertains to the affected community, the donors or the stakeholders, and however so defined. This is in keeping with the evaluation standards underpinning the conceptual framework of this thesis.

In summary, one change was made to accountability. This is reflected in Figure 37, which shows the modified outcome typologies.
8.5.7 Evaluation Standards and Evidence

The original CFDET 2017 Evaluation Standards and Evidence, as shown in Figure 38, consisted of three elements situated within its own box, comprised as follows:

1. Box 1 – Evaluation standards and guidelines
   a. Responsible and ethical evaluation practice
2. Box 2 – Evidence-based reviews and registries
3. Box 3 – Knowledge management
   a. Cross-sectoral research
   b. Collaboration
   c. Dissemination.
Ten comments were received from participants. Following review, an opportunity was taken to enhance how evaluation standards and evidence were depicted in the comprehensive framework. The following modifications were made:

- The position of the three individual boxes highlighting the independent concepts of evaluation standards and guidelines, evidence-based reviews and registries and knowledge management were moved and are now included in a larger box to indicate their relationship and that these concepts, under the umbrella of evaluation standards and evidence, are applicable across the entire framework.
- The next change is within knowledge management, where ‘cross-sectoral research’ was replaced with ‘cross-sectoral collaboration’.
- The term ‘translation’ was added.

One of the aims of the comprehensive framework is to encourage collaboration on knowledge management, and importantly, to promote the translation of evaluation findings into improved policy and practice, which may lead to improved outcomes for the communities affected by disasters. Cross-sectoral collaboration, translation and dissemination of information is vital as the sharing of information can help improve
policy and practice, and greatly improve disaster management efforts (559). Using the
information gained from previous events will help generate current and future responses
and plays a crucial role in disaster risk reduction, which is aimed at reducing mortality,
morbidity and increasing costs associated with disasters (560, 561).

Three modifications were made to evaluation standards and evidence to promote clarity
and understanding. It is possible that these results reflect the current debates related to
evaluation and the way in which evaluation standards and evidence were illustrated within
CFDET 2017. The modified evaluation standards and evidence is shown in Figure 39.

![Figure 39: CFDET 2018 Evaluation Standards and Evidence](image)

Participant comments provided an opportunity to modify CFDET 2017. The majority of
comments were considered by the team of researchers to be more ‘general’ in nature and
were often a commentary on the current field of evaluation in the disaster setting. Changes
were made based on the feedback received. It was exciting to note that only one major
change was considered necessary – the removal of ‘accountability’ as a stand-alone
overarching theme to be included in outcome typologies. The remaining comments
provided an opportunity to make minor modifications by refining the text included in
each element or overarching theme.
The opportunity was also taken to make additional changes as a result of an environmental scan of the literature. These changes reflect the continued work and contemporary nature of the disaster sector. It is worth noting that the aim of this research is not to include every example of evaluations in the literature, but it is to identify all evaluation typologies used or applicable to evaluations in the disaster setting and to provide sample exemplars of all typologies.

8.5.8 Survey Responses to CFDET 2017 as a Unifying Framework

The survey concluded by seeking comments related to the framework as a whole. The first question asked about whether the relationships between the evaluation typologies and the disaster timeline were appropriately demonstrated. Nineteen comments were provided (refer to Table 47) and did not lead to any changes.

The second question asked if the comprehensive framework would potentially hold value for the respondent’s work when undertaking evaluations in the disaster setting. Nine comments were provided (refer to Table 48), but again did not lead to any changes to the comprehensive framework.

The third question asked if the comprehensive framework would be useful for supporting and promoting evaluation in the disaster setting. Nine comments were provided (refer to Table 49), but no changes were made to the comprehensive framework.

The fourth question asked if the comprehensive framework would be useful for teaching evaluation in the disaster setting. Nine comments were provided (refer to Table 50), but no changes were made to the comprehensive framework based on these comments.

The next question asked if the participants had any additional thoughts on how the comprehensive framework could be improved. Eleven comments were received (refer to Table 51); however, significantly, no comments resulted in changes to the comprehensive framework.

The remaining two questions asked participants to comment on any barriers or enablers they perceived in undertaking evaluations in the disaster setting. These questions resulted in 13 and 8 comments, respectively (refer to Tables 52 and 53, respectively). No changes were made to the comprehensive framework based on these comments.
8.5.9 Review of the Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017)

A mixed methods approach was used to validate CFDET 2017. The quantitative responses validated the comprehensive framework at an international level. The qualitative responses highlighted opportunities to improve the comprehensive framework. Comments led to modifications of CFDET 2017. The majority of comments received were inclined to be ‘general’ in nature and were often a commentary on the current field of evaluation in the disaster setting. Changes were made based on the feedback received and it was exciting to note that only one major change was considered necessary. The remaining comments provided an opportunity to make modifications by refining the text included in each element or overarching theme. Based on the feedback obtained, in conjunction with the results from an environmental scan, modifications were made to CFDET 2017; the resulting comprehensive framework is shown in Figure 40. To differentiate between the original framework and the modified framework, the latter is referred to as the Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018).

The importance of this comprehensive framework is that it provides consistency in terminology and standards for evaluation and reporting across the different phases of a disaster. This in turn provides comparability to better understand the process, outcomes and impacts of the efficacy and efficiency of interventions. These findings will help build the evidence-base of disaster medicine and public health. CFDET 2018 has been improved with input from international experts received as part of the validation process. There is an expectation that CFDET 2018 would achieve higher acceptance rates if the validation process was replicated on the modified framework.
Figure 40: Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018)
Important lessons were learned while implementing the online survey, which will be valuable for any future surveys undertaken. The limitations of the survey included attracting an acceptable number of participants, such as a required sample size of 30, when only approximately one in five completed the survey; the electronic format of the survey and the quality of the technology used by different participants; the estimated length of the survey (30-45 minutes to complete); and the fact that the accompanying paper *Disaster Metrics: A Comprehensive Framework for Disaster Evaluation Typologies* (1) describing the development of CFDET 2017 was in press. These limitations, however, do not negate the results obtained.

### 8.5.10 Submitted Publication

A follow up publication, *Peer Reviewed Validation of the Comprehensive Framework for Disaster Evaluation Typologies 2017*, has been submitted to PDM (refer to Appendix 10).

### 8.6 Conclusion

To add credibility and global relevancy to CFDET 2017, an international peer-reviewed validation process was undertaken. This process ensured CFDET 2017 was not just the perception of the researchers but an internationally collaborative view of evaluation typologies suitable for use in the disaster and humanitarian setting.

This chapter addresses the overarching Study 4 research question on the necessary steps to validate a comprehensive framework for evaluation typologies in the disaster setting. This validation process also finalises the solution for the research aim, which was to develop and validate, at an international level, a comprehensive framework that structured disaster evaluation typologies along the disaster timeline.

Specifically, this chapter presented the research design and methodology employed to conduct the validation and which included a literature review on validating processes and an online survey to gain international input and acceptance of CFDET 2017. Despite having to run the survey three times to gain an acceptable sample size, the results clearly indicate international acceptance and validation of CFDET 2017. The international peer-review validation process also sought comment and feedback from experts in the fields of disasters and emergencies; humanitarian and development; and evaluation. The feedback provided an opportunity for reflection on the way that typologies and overarching themes were depicted within the framework. On a positive note, participants
did not identify any additional evaluation typologies that should have been included in the framework. Decisions were made to incorporate some modifications that would enhance the comprehensive framework by adding strength to the elements and overarching themes. This would also result in improved usability and applicability to disaster evaluations and research while ultimately assisting with improved policy and practice in delivering interventions in the disaster setting.
Chapter 9: Discussion

9.1 Introduction

The primary research aim underpinning this thesis was to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline. This research is organised through the evaluation conceptual framework lens using the evaluation standards of utility, feasibility, propriety and accuracy to structure the discussion.

The risk of disasters resulting in mass casualties and affecting the lives of many communities continues to increase as the global population grows in number and density. The scale, frequency and impact of crises that demand international humanitarian response is also increasing. There is a moral imperative to improve the way in which disaster evaluations are undertaken and reported with the aim of reducing preventable mortality and morbidity in future events. To reduce the risk of harm and improve the health response, an evidence-based approach to disaster evaluation is urgently required. One way of reducing disaster risk to the international community is by using a framework to structure disaster evaluations.

Existing frameworks that support consistent reporting in a disaster lack validation, consistent terminology and standards for reporting across the disaster timeline. Research indicates that an opportunity exists for the development of a comprehensive framework to structure disaster evaluation typologies. To create such a framework to provide a structure to support disaster evaluation and research, the following overarching research questions were defined and answered in this thesis:

- What does the peer-reviewed and grey literature report on comprehensive frameworks and evaluation typologies in the disaster setting?
- What key elements would support a core structure to develop a comprehensive framework to represent evaluation typologies along the disaster timeline?
- What comprehensive framework design could enable mapping evaluation typologies along the disaster timeline?
- What are the necessary steps to validate the comprehensive framework for evaluation typologies in the disaster setting?
9.2 Reflection

Reflecting upon the results, this section pulls together the themes from the multiple studies.

Direct involvement in the 3ie Scoping Paper 1, as discussed in Chapter 2, highlighted a gap and an opportunity to address the lack of high-quality evidence in the disaster and humanitarian sectors, and the inconsistent use of terminology. The findings supported the need for the development of a comprehensive framework as well as the premise that there is an urgent requirement to increase the evidence-base of what works, for whom and at what cost.

This thesis makes a direct contribution to the evidence-base by developing and validating, at an international level, a unique and unifying comprehensive framework that structures disaster evaluation typologies along the disaster timeline. As with any journey, thought and structure needed to be applied to ensure the resulting framework met the evaluation standards of utility, feasibility, propriety and accuracy. This was important to increase the adoption and application of the final framework.

To understand the current state of evaluation frameworks and typologies that are used in the disaster setting, a literature review was conducted and discussed in Chapter 4. A scoping literature review searched both the peer-reviewed and grey literature. The search strategy was structured to support two distinct themes and designed to address the following research sub questions:

Theme 1. What frameworks for evaluation and research exist in the disaster setting?
Theme 2. What evaluation typologies exist in the disaster setting?

A scoping literature review allowed for the comprehensive mapping of peer-reviewed and grey literature and was supported by the additional processes of seeking expert opinion including consultation with stakeholders and snowballing (142). Peer-reviewed literature was searched using major electronic databases specific to health. Inclusion and exclusion criteria were set and key terms for both themes of the literature review were established and structured to support identifying literature that would help answer the specific research questions. The second part of the literature review included searching the grey literature to uncover additional untapped resources that may not have been published in peer-reviewed literature. Generic search engines and industry-specific
repositories were searched, and although the same keys words could not be used as in the peer-reviewed literature search, every attempt was made to use comparable terms to support the standards of propriety and accuracy.

The scoping literature review was supplemented and enhanced by a convenience sample of international colleagues who commented on evaluation frameworks and evaluation typologies in the disaster setting. Further to the structured scoping literature review and seeking expert opinion, a technique called snowballing was used. Additional references were identified by examining the bibliographies and reference lists from the most recent full-text publications and through scrutiny of the contents pages of highly relevant journals.

Key findings related to Theme 1: Evaluation frameworks in the disaster setting identified the works of several disaster and emergency experts. A review of these findings identified a pivotal guideline that was referenced and taught at an international level – *Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style, 2003* (17). This provided a starting point for the development of a comprehensive framework for disaster evaluation typologies.

Key findings related to Theme 2: Evaluation typologies in the disaster setting included the categorisation of evaluation types that could be mapped along the disaster timeline, including baselines, consequences, outcomes and impact evaluations. Due to the wide scatter of different evaluation typologies in existence, the search was challenging as there is currently no internationally agreed classification available to assist during the search process. Disaster evaluation typologies are poorly defined, they lack consistent definitions, terminology and structure, and are often found in the grey literature. These factors made searching difficult and the final analysis of the findings complicated.

The findings of the peer-reviewed literature review identified that no core, unifying framework to structure disaster evaluations and research was being used to structure disaster evaluations, nor had been validated. A significantly higher number of references were found in the grey literature. The results from the literature review strengthened the initial findings of the 3ie Scoping Paper 1 in that there was a lack of quality evidence in the disaster and humanitarian sectors to guide decision making and inform policy and practice.
To address the issues identified, in particular, the lack of quality evidence in the disaster sector to guide decision making when delivering interventions, a comprehensive framework was developed, starting with a core structure. This core structure would form the foundation for a more comprehensive framework to promote consistent terminology and standards for reporting across the disaster timeline. Further, this work promotes an environment for constructive dialogue on evaluation in the disaster setting and adds to the evidence-base of disaster health.

Chapter 5 introduces the steps undertaken to identify key elements to support the establishment of a core structure that would form the foundation of a comprehensive framework to represent evaluation typologies mapped along the disaster timeline. Specifically, this chapter presented the research structure and outcome of the two studies. Study 1 involved a more in-depth analysis of the Utstein Guidelines 2003, which were identified in the previous chapter as the pivotal framework for disaster evaluation and research.

Study 1 provided answers for the following sub questions:

- How do professionals active in disaster evaluation and research view the Utstein Guidelines?
- How do professionals active in disaster evaluation and research use the Utstein Guidelines?
- What are the barriers and facilitators to professionals active in disaster evaluation and research using the Utstein Guidelines?

To support Study 1, semi-structured interviews were undertaken. The sample included professionals active in disaster evaluation and research to understand how they viewed the Utstein Guidelines 2003. An analysis of the results assisted in identifying key elements to include in the core structure. The results revealed:

- Health professionals active in disaster evaluation and research are aware of the Utstein Guidelines 2003 and deemed them to have value, even though they were not being used to structure disaster health evaluation and research.
- The Utstein Guidelines 2003 were used in a variety of formats other than structuring evaluation studies; for example, in teaching and as a source of a conceptual framework and definitions.
• While many comments were offered on improving disaster health evaluations, there were no consistent themes identifiable within the comments.
• A number of similar guidelines were named by respondents; however, most of these had already been identified earlier in Chapter 4.

The findings from Study 1 were augmented by additional key elements found to be contemporary in the disaster sector, as identified from the literature review. The resulting 10 key elements were used to inform Study 2, a thematic analysis structured to answer the following sub question:

• In what ways are the key elements of the Utstein Guidelines 2003, as amended, reflected in selected reviews of Australian disaster events?

The sampled Australian disaster event reviews in Study 2 were examined to identify whether the key elements from the Utstein Guidelines 2003 (17), as amended, were evident. This was important to test if the identified key elements had merit or worth in their inclusion in the core structure that would form the foundation of CFDET.

The findings from Study 2 demonstrated a preliminary validation of the key elements identified from Study 1 and as amended. The final list of key elements to be included in the core structure are:

• pre-event status of society
• hazard
• risk including prevention, modification, vulnerability and exposure
• event
• impact / damage
• disruptions / disturbances / changes in function / consequences
• response
• recovery
• resilience
• post-event status of society.

These key elements are essential as they provide structure and consistency in terminology while providing the foundation for a more comprehensive framework. This work informed Chapter 6, which was structured to answer the following sub questions:
In what ways could Study 1: Utstein Guidelines 2003 and Study 2: Reviews of Australian disaster events inform establishing a core structure for a comprehensive framework?

How can these key elements be mapped along the disaster timeline to support the development of the core structure?

The core structure is the foundation of a comprehensive framework for disaster evaluation typologies. It is diagrammatically represented as a horizontal line. Diagrammatic representation of disasters is a linear process, shown as separate and sequential phases or elements along the disaster timeline, but in reality there is often an overlap and blurring of when one phase starts and another stops. While depicted as a linear process, disasters are in no way ‘straightforward’. The core structure provides the system of concepts and assumptions that underlie the disaster timeline. It diagrammatically represents the main ‘things’ to be studied; that is, the different disaster phases and the presumed relationships between them. The core structure also helps to assess and refine the development of the comprehensive framework by providing structure, consistency in terminology and understanding.

The complete core structure consists of the following three layers:

Layer 1. provides a preliminary and simplistic view of the disaster timeline
Layer 2. provides an expanded view of the disaster timeline
Layer 3. introduces the concept of strengthening resilience as an overarching theme.

Each layer builds on the previous layer, and these layers demonstrate the evolution of the way in which disasters are viewed and also acknowledges the significant international frameworks that are pushing agendas of disaster risk reduction (58), sustainable development (59), climate (112) and humanitarian challenges (406-408). The final layer of the core structure is the concept of strengthening resilience, which is an international imperative in the drive to reduce and manage disaster risk.

Building on the research undertaken in Chapters 5 and 6, which resulted in the establishment of the core structure, Chapter 7 outlines the steps taken to develop a comprehensive framework for disaster evaluation typologies. The importance of this comprehensive framework is that it provides a structured and consistent approach to undertaking and reporting disaster evaluations. One of the aims is to reduce preventable
mortality and morbidity in future events while also supporting the improvement of disaster health science. Further, to mitigate some of the previous issues identified with earlier frameworks, this new framework defines and provides consistent terminology along with standards for reporting across the different phases of a disaster. This, in turn, provides comparability to better understand the process, outcomes and impacts of the efficacy and efficiency of health interventions delivered in the disaster setting.

The outcome from this chapter provides a partial answer to the overall research aim, which was to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline. The development of the comprehensive framework formed Study 3, and addressed the following sub questions:

- What process enables the development of a unifying, comprehensive framework for disaster evaluation typologies within the disaster setting?
- How can these evaluation typologies be mapped to support agreement on the identification, structure and relationships between various evaluation typologies within the disaster setting?

This chapter mapped the evaluation typology classifications, identified in Chapter 4, along the disaster timeline as represented in the core structure. Evaluation typologies identified during the literature review were categorised into two overarching themes and four evaluation typologies. The two overarching themes are accountability and evaluation standards and evidence. The four evaluation types include baselines, consequences, outcomes and impact evaluation. The mapping of these along the disaster timeline also highlighted their relationships and interdependencies.

The resulting unique and unifying Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017) identifies the different evaluation typologies that can be used in the disaster setting and demonstrates key relationships in a single diagram. Importantly, it also suggests interdependencies and relationships that exist between various evaluation typologies along the disaster timeline and within the disaster setting. The comprehensive framework includes feedback loops from consequences and outcomes to improve baselines, reduce disaster risk and strengthen resilience. The feedback loop helps to strengthen the concept that the disaster timeline is a continuum and should be viewed in this light. The comprehensive framework is not limited to any
one phase of the disaster timeline and can be used for responding to disasters (regardless of the cause; that is, an ‘all-hazards approach’), humanitarian crises or in the development sector. It is relevant to any sized disaster.

It is important to note that CFDET 2017 is not an evaluation typology in itself, nor is it a list of evaluation studies; rather, it is a framework that outlines different disaster evaluation typologies that can be mapped along the disaster timeline. This unique framework has relevance at an international level and is expected to benefit the disaster, humanitarian and development sectors. This work promotes an environment for constructive dialogue on evaluation in the disaster setting and adds to the evidence-base of disaster evaluation and research. To gain international acceptance and credibility, and to complete the overall research aim, a validation process was conducted.

Chapter 8 discussed the process undertaken to validate CFDET 2017 at an international level. This process ensured CFDET 2017 was not just the perception of the researchers but an internationally collaborative view of evaluation typologies suitable for use in the disaster and humanitarian setting. This process also supported the evaluation standards previously mentioned, including utility, feasibility, propriety and accuracy.

This chapter addresses the overarching Study 4 research question, regarding the necessary steps to validate a comprehensive framework for evaluation typologies in the disaster setting. This validation process also finalises the solution for the research aim, which was to develop and validate, at an international level, a comprehensive framework that structured disaster evaluation typologies along the disaster timeline.

Specifically, this chapter presented the research design and methodology employed to conduct the validation, and included a literature review on validating processes and an online survey to gain international input and acceptance of CFDET 2017. The international peer-review validation process sought comment and feedback from experts in the fields of disasters and emergencies, humanitarian and development and evaluation. Feedback provided an opportunity for reflection on the way that typologies and overarching themes were depicted within the framework.

On a positive note, participants did not identify any additional evaluation typologies that should have been included in the comprehensive framework. Comments were reviewed by four researchers and the decision was made to incorporate some modifications that would enhance CFDET 2017. This would also result in improved usability and
applicability to disaster evaluations and research while ultimately assisting with improved policy and practice in delivering interventions in the disaster setting.

To differentiate between the original 2017 comprehensive framework and the modified framework, the updated and validated comprehensive framework was labelled the Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018).

9.3 Evaluation Criteria

Evaluation is the systematic investigation of program, merit and worth and is underpinned by program evaluation standards (103). The following discussion is centred around the four Joint Committee on Standards for Educational Evaluation (JCSEE) evaluation attributes of: utility, feasibility, propriety and accuracy (103).

9.3.1 Utility

Utility standards aim to increase the value that program stakeholders find in evaluation processes and products (473). These standards have been used to guide the research presented in this thesis. The criteria that make up the Utility Standards and how it has supported the research process are outlined in Table 54.

Table 54: Utility Standards Criteria Applied to the Research

<table>
<thead>
<tr>
<th>JCSEE Utility Standards Criteria</th>
<th>Impact on This Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1 Evaluator Credibility</td>
<td>The PhD candidate has experience in disaster and emergency health. Supervisors to this research also have international acknowledgement of their expertise in the health and disaster sector while also having experience in evaluation.</td>
</tr>
<tr>
<td>U2 Attention to Stakeholders</td>
<td>Surveys conducted throughout this research sought input from a range of international experts from different fields related to disaster medicine; humanitarian; and evaluation. These people were included to give a depth to the survey responses which would translate into the finalised comprehensive framework.</td>
</tr>
<tr>
<td>U3 Negotiated Purposes</td>
<td>The gap identified as part of the 3ie Project, and supported by a literature review, noted that there was no agreed, comprehensive framework to structure disaster evaluation typologies. This gap stimulated the theme for this thesis. The research undertaken during confirmed this gap and stimulated the subsequent development and validation of a comprehensive framework that would effectively inform disaster evaluations and thereby contribute to closing this gap.</td>
</tr>
</tbody>
</table>
Given the international and evolving nature of disaster evaluations, it was imperative to gain input, comment and acceptance from key international stakeholders. Opportunities were taken to present the evolving framework and to seek comment to ensure the final comprehensive framework would meet stakeholder needs.

U4 Explicit Values

Ethics approval was obtained for all appropriate studies prior to implementation.

U5 Relevant Information

Refer to U3.

U6 Meaningful Processes and Products

Refer to U3.

U7 Timely and Appropriate Communicating and Reporting

A publication has been submitted to PDM that details the international validation process, modifications made as a result of stakeholder comments, and introduces CFDET 2018.

U8 Concern for Consequences and Influence

The comprehensive framework developed improves the validity of disaster evaluation by clarifying the various evaluation typologies with the intent to strengthen the evidence-base of health interventions delivered in response to disasters by identifying what works for whom, and why.

From a utility perspective, this research has usefulness over a broad base due to its applicability to an international audience, as well as the engagement of international participants in the research process. The resulting CFDET 2018 is valuable in meeting the disaster evaluation needs of the international community and has the capacity to inform policy and practice; that is, assist in the choice of an appropriate evaluation typology in the disaster setting.

9.3.2 Feasibility

Feasibility standards aim to increase evaluation effectiveness and efficiency (103). These standards have been used to guide the research presented in this thesis. The criteria that make up the Feasibility Standards and how it has supported the research process are outlined in Table 55.

<table>
<thead>
<tr>
<th>JCSEE Feasibility Standards Criteria</th>
<th>Impact on This Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 Project Management</td>
<td>Project management techniques including the establishment of Gantt charts (for PhD progress) and setting milestones and goals have been used.</td>
</tr>
</tbody>
</table>
Milestones were used in the development, implementation and analysis of surveys to ensure timely completions.

F2 Practical Procedures
The validation survey needed to be extended in order to obtain a suitable sample size. A suitable sample size was necessary to ensure that the comprehensive frameworks developed during this research, were not just the perception of the researchers.

F3 Contextual Viability
The language used in the surveys was carefully crafted to remove any cultural or political bias, while using internationally accepted terminology for the disaster sector. Invitations to participate in the validation survey were offered to all six WHO regions.

F4 Resource Use
A resource used during the research process included the Qualtrics Insight Platform to host the validation survey. This provided an effective and efficient tool for distributing the survey at an international level as well as collating the results for further analysis.

From the perspective of one of the thesis outputs, that is, CFDET 2018, this in itself is a resource that is intended for use in the international community. Given the rising socio-economic costs to the global community as a result of disasters, ensuring efficiency and effectiveness in the delivery of interventions is vital. This can be achieved through the application of a consistent approach to disaster evaluations and strengthening the evidence-base of health interventions. The importance of CFDET 2018 is that it provides consistency in terminology, standards for evaluation and reporting across the different phases of a disaster, and has been validated at an international level. This in turn provides comparability to better understand the process, outcomes and impact of the efficacy and efficiency of interventions. These findings will help build the evidence-base of disaster medicine and public health.

One of the major justifications for the development of CFDET 2018 was to promote consistent and replicable disaster evaluations through the use of common terminology and a structured approach. A strong evaluation framework for disaster settings, such as CFDET 2018, is extremely important given the significant impact of disasters and rising socioeconomic costs globally. Efficiency and effectiveness can be achieved through the application of a consistent approach to undertaking disaster evaluations.
9.3.3 Propriety

Propriety standards are structured to support what is proper, fair, legal, right and just in evaluations (103). These standards have been used to guide the research presented in this thesis. The criteria that make up the Propriety Standards and how it has supported the research process are outlined in Table 56.

Table 56: Propriety Standards Criteria Applied to the Research

<table>
<thead>
<tr>
<th>JCSEE Propriety Standards Criteria</th>
<th>Impact on This Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 Responsive and Inclusive Orientation</td>
<td>Disasters are a global issue and reducing disaster risk is an imperative of the Sendai Framework. This thesis aims to reduce preventable mortality and morbidity as a result of disasters by helping to build the evidence-base of disaster medicine by improving the way in which disaster evaluation typologies are structured. Refer to U2 and U3 above.</td>
</tr>
<tr>
<td>P2 Formal Agreements</td>
<td>Refer to U2, U3 and U4.</td>
</tr>
<tr>
<td>P3 Human Rights and Respect</td>
<td>Refer to U4.</td>
</tr>
<tr>
<td>P4 Clarity and Fairness</td>
<td>In addition to this, anonymity of respondents to the validation survey was ensured as no identifying factors were recorded. All surveys had ethics approval from Monash University and explanatory statements were provided up front. Pilot studies were implemented for both surveys to test structure and content before implementation to the wider sample. All participants were given the opportunity to provide their consent to enter the survey. Participants were given the option to withdraw or opt-out at any stage.</td>
</tr>
<tr>
<td>P5 Transparency and Disclosure</td>
<td>All findings from the research process, including literature reviews and surveys have been included in this thesis without change. Comments provided by respondents to the surveys were reviewed by four researchers and are included within the relevant Chapters. This ensures transparency and removes any opportunity for bias. Where relevant, limitations to the research have also been identified in relevant chapters. A publication has been submitted to PDM for peer-review that details the international validation process, modifications made as a result of stakeholder comments, and introduces CFDET 2018.</td>
</tr>
</tbody>
</table>
| P6 Conflicts of Interest                            | Given the nature of this research, no conflicts of interest were identified. From a researcher point of view, full disclosure of publications, conference presentations,
From a propriety perspective, this research has adhered to the highest principles, both within the conduct of the research and the development and validation of the finalised comprehensive framework. Ethics approval was obtained from Monash University to conduct Studies 1 and 4. Within the comprehensive framework, the inclusion of the overarching theme ‘evaluation standards and evidence’ also promotes the concept of propriety by ensuring a focus on professional ethics.

### 9.3.4 Accuracy

Accuracy standards focus on strengthening the dependability and truthfulness of evaluation representations, propositions, and findings, especially those that support interpretations and judgments about quality (103). These standards have been used to guide the research presented in this thesis. The criteria that make up the Accuracy Standards and how it has supported the research process are outlined in Table 57.

<table>
<thead>
<tr>
<th>JCSEE Propriety Standards Criteria</th>
<th>Impact on This Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 Justified Conclusions and Decisions</td>
<td>The research design and methodology have been explained including the rationale for their selection. This is particularly important given the small sample size. The criteria for establishing the validation and acceptance of CFDET 2017 were identified.</td>
</tr>
<tr>
<td>A2 Valid Information</td>
<td>The overall research aim of developing and validating, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline has been achieved.</td>
</tr>
<tr>
<td>A3 Reliable Information</td>
<td>Four studies were undertaken to achieve the overall research aim. Each study supported, substantiated and informed the subsequent study. This is evidenced by the 3ie Scoping Paper 1 findings being supported by a scoping literature review. The literature review findings informed Studies 1 and 2. The results of these studies supported the establishment of a core structure as well as the development of a comprehensive framework for disaster evaluation typologies. The inclusion of the validation process also lends weight to the reliability of the framework in that</td>
</tr>
</tbody>
</table>
no additional or new evaluation typologies were offered by participants of the validation survey. Analysis of the survey results demonstrated international acceptance.

A4 Explicit Program and Context Descriptions
All studies have been detailed to the level that they should be replicable in the future.

A5 Information Management
A consistent and logical approach was used when conducting literature reviews. Survey responses are stored in accordance with the ethical standards of MUHREC.

A6 Sound Designs and Analyses
Refer to A1.

A7 Explicit Evaluation Reasoning
The review of all survey results was undertaken by four researchers. Refer to P4.

A8 Communication and Reporting
Refer to P5.

Research design and methods used for this research have been disclosed throughout this thesis under the relevant chapters. All methods have been selected and reviewed in consultation with three senior academics. The results from studies have been used to inform the next stage of the research, which, in itself, is a form of validation.

Study 4, the validation of the comprehensive framework, in particular, lends support to the accuracy of this research, as participants did not offer any additional evaluation typologies that could be included in the framework. Quantitative results demonstrated acceptance and validation of the comprehensive framework by the international audience.

9.4 General Outcomes

This thesis provides a higher order unifying framework in which to place the various evaluation systems already in use.

The framework developed built on the strengths identified in these evaluations and focussed attention on operational utility in the disaster setting. This utility was supported in the validation and is yet to be tested in real-time. Strong comments of support by disaster, humanitarian and evaluation experts were received during the validation process. Whilst many of the contributors provided positive critiques, few actually suggested specific changes. In this sense, the framework is regarded as a living document and will continue to evolve over time as other researchers and operational leaders test and utilise the framework.
The iterative process underpinning this thesis with rich engagement with the domain, including published papers and presentations in international forums, contributed to the acceptance of the international validity of this new framework as it evolved on the international stage. The 3ie Scoping Paper 1 provided a rich underpinning for this thesis by identifying and confirming the gap in quality evidence in this sector and the need for this improved evaluation framework as developed in the thesis and lays the basis for further research.

This comprehensive framework closes the gap identified within this thesis and has relevance and use at an international level in the disaster setting.

9.5 Research Translation

CFDET 2017 is the first unifying framework that has been internationally validated. It incorporates bodies of work from experts in the field of disasters and emergencies and includes contemporary internationally significant disaster frameworks. The comprehensive framework promotes consistent terminology and standards for reporting across the disaster timeline, and provides structure for disaster evaluation types. By providing consistency in terminology and structure for reporting across the disaster timeline, evaluations undertaken in the disaster setting will have a better comparability of results and this consistency will increase the evidence-base of health interventions delivered in the disaster setting. This would also result in improved usability and applicability to disaster evaluations while ultimately assisting with improved policy and practice in delivering interventions, by consolidating past experiences and improving future responses (85).

As a consequence of the research undertaken to complete this thesis, opportunities to present this work have been embraced and evidenced, with the presentation of this work at the UNISDR Science and Technology Conference in Geneva (refer to Appendix 2), active participation in the international 3ie Scoping Project (refer to 2.4.2), and active involvement in an international workshop ‘Health and Disaster Risk Reduction Regarding the Sendai Framework’ (refer to Appendix 1). Further, as a result of the research, an invitation was received to attend the 3rd World Conference on Disaster Risk Reduction in Sendai Japan in 2015. This invitation was extended from Professor Virginia Murray (Public Health England), who was the Co-Chair of UNISDR Science and Technology.
The comprehensive framework has been used as the core framework to structure the disaster evaluation components of two coursework units offered at Monash University:

- Master of International Development Practice delivered by the Monash University Disaster Resilience Initiative on behalf of the Faculty of Arts, Graduate Unit: APG5120 Guiding Principles for Professionals Engaged in Disasters and Humanitarian Crises (2017, 2018), [http://www.monash.edu/pubs/2018handbooks/units/APG5140.html](http://www.monash.edu/pubs/2018handbooks/units/APG5140.html).

**9.6 Limitations**

The limitations for individual studies have been discussed in the relevant chapters. It is worth noting that the limitations identified earlier do not detract from the research findings.

**9.7 Future Actions**

Future actions proposed to strengthen CFET 2018 as part of the quality improvement process include:

- Validation, implementation and evaluation of CFDET 2018,
- Preparing an interpretive guide to accompany CFDET 2018,
- The implementation of toolkits to support disaster evaluation typologies, for example ‘How To Guides’,
- Collaboration with international organisations, such as World Association for Disaster and Emergency Medicine (WADEM) to schedule promotion, training and implementation of CFDET 2018,
- Facilitate, support and publish an initial suite of disaster evaluations using the Comprehensive Framework to structure the evaluation methodology,
- Evaluating the link between CFDET 2018 and the significant international disaster risk reduction frameworks identified in the Core Structure,
- Monitor and evaluate the operational utility of CFDET 2018.
9.8 Summary

The overall research aim of this thesis was to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline. Through a series of four studies, this has been achieved, with the resulting CFDET 2017 accepted and validated at an international level.

The comprehensive frameworks developed throughout this research are the first frameworks of their type and make a unique contribution to current knowledge at an international level. The comprehensive framework provides structure and consistency in terminology and standards to help increase the evidence-base of interventions and disaster science. It has relevance at an international level and, importantly, across multiple sectors, including disasters, emergencies, humanitarian and development. This work promotes an environment for constructive dialogue in the disaster setting.
Chapter 10: Conclusion

This applied thesis has ‘disaster evaluation typologies’ as its unifying theme and is applicable to multiple disciplines. The main research aim was to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline. To support a solution to this aim, four overarching research questions were identified.

In Chapter 1, the background and need for this research was identified. The significance of the study at an international level was highlighted and the structure of the thesis was outlined.

Chapter 2 provided the background and context for this research. Key definitions and concepts were defined to provide consistency in understanding the context in which they are used. This chapter introduced the 3ie Scoping Paper 1 (3), which the researcher was directly involved in. This internationally funded project provided an independent analysis of the evidence-base of impact evaluations in the humanitarian sector. The results demonstrated a paucity of evidence and a lack of high-quality impact evaluations to inform policy and practice (3, 16). This report identified gaps and inconsistencies in methodologies and terminologies in disaster evaluations, and a need to develop and validate a comprehensive framework to structure disaster evaluation typologies.

Chapter 3 identified evaluation as the conceptual framework underpinning this thesis. Research design and methodologies used to support the four studies were specified.

Chapter 4 detailed the scoping literature review of evaluation frameworks and evaluation typologies in the disaster setting undertaken to address the first overarching research question. The methodology, including key terms, databases searched and results were described in detail. An analysis of the results showed that while some evaluation frameworks had been developed to support consistent disaster evaluations, they were fragmented, uni-focused, not validated and not widely used to structure disaster evaluation and research. Consequently, there was no agreed on comprehensive framework to structure disaster evaluation typologies.

Chapter 5 outlined the methodology used to identify the key elements that should be included in the core structure, which formed the foundation for the new and unified
comprehensive framework for disaster evaluation typologies. Two studies were structured to identify the key elements that would support a core structure. This core structure would provide the foundation for a comprehensive framework to represent evaluation typologies mapped along the disaster timeline.

Chapter 6 described the establishment of the core structure, which consists of three layers. Significantly, this core structure includes globally important disaster, humanitarian and development frameworks and agendas. The core structure resulting from the work performed in Chapters 5 and 6 addresses the second overarching research question.

The continued development of the comprehensive framework was discussed in Chapter 7. The resulting Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017) identifies the different disaster evaluation typologies and demonstrates key relationships in a single diagram, providing an answer to the third overarching research question. Interdependencies and relationships between the various evaluation typologies along the disaster timeline and within the disaster setting are identified.

Chapter 8 presented the steps undertaken to validate the comprehensive framework. This addresses the fourth overarching research question as well as finalising the solution for the overall research aim, which was to develop and validate, at an international level, a comprehensive framework that structures disaster evaluation typologies along the disaster timeline. The survey results validated the comprehensive framework.

Chapter 9 presented a generalised discussion of the research journey and summarised the research undertaken and its contribution to the area of disaster science. It provided a stimulus for a more structured and rigorous approach to developing the evidence-base of disaster preparedness and management. This chapter considered the research questions and united the literature, theory and framework. It explored research translation to policy, practice and education in the disaster evaluation setting, while considering limitations of this research and opportunities for future research. This chapter also recognised the completion of the research aim and four overarching research questions.

Chapter 10 presents the conclusions of this research.

In conclusion, the resulting comprehensive frameworks have relevance at an international level and will benefit the disaster, humanitarian and development sectors. This work
promotes an environment for constructive dialogue on evaluations in the disaster setting to strengthen the evidence-base for interventions across the disaster spectrum.
Appendices

Appendix 1: Health and Disaster Risk Reduction Regarding the Sendai Framework

Health and disaster risk reduction regarding the Sendai Framework

Lentar Reife, Paul Arbore, Anthony Capon, John Hardme, Alistair Humphrey, Virginia Murray, Caroline Spencer, and Diana F. Wong

An expert workshop was held at the University of Melbourne in July 2017 to consider disaster risk reduction for the health sector under the Sendai Framework. Outcomes were recommendations for alliances and partnerships to link researchers and government across disaster risk reduction and health to inform policy and practice.

Introduction

Health is a pivotal dimension to be addressed within all-hazards approaches to disaster risk reduction. It is also a key point of convergence across global and national policy frameworks. The recent synchronous adoption of the five landmark UN agreements, the Sendai Framework for Disaster Risk Reduction, Sustainable Development Goals, COP21’s Paris Climate Conference, World Humanitarian Summit, and Habitat III, has created a rare and significant opportunity to build coherence across different but overlapping policy areas. Extreme weather events are projected to increase in frequency, intensity and duration over the coming decades. It is apparent that these events could potentially increase the vulnerability of individuals, communities, and regions and lead to longer recovery times. Taken together these UN agreements make a more complete resilience agenda as building resilience requires action spanning development, humanitarian, climate and disaster risk reduction areas and for multi-hazard assessments. These develop a dynamic, local, preventive and adaptive urban governance system of the global, national, and local levels.

The Sendai Framework for Disaster Risk Reduction 2015-2030 is the principal global treaty to guide disaster risk reduction efforts. The Sendai Framework reflects an important shift away from managing disasters and towards reducing disaster risk. Health resilience is strongly promoted throughout.

The Sendai Framework calls for broad disaster risk reduction (DRR) activities that reduce the effects of disasters with respect to loss of life, injury and health impacts as well as on the socioeconomic determinants that affect population health. These include property damage, loss of livelihoods and services, social and economic disruption and environmental damage. The use of scientific evidence to inform policy and formulate effective initiatives and interventions is crucial to DRR within health. The importance of health as a core dimension in DRR was emphasized within the Doha Declaration Principles following the UNISDR International Conference on the implementation of the Health Aspects of the Sendai Framework. These principles are further developed in the UNISDR fact sheet: “Health in the Context of the Sendai Framework for Disaster Risk Reduction and in the WHO Technical Guidance Series on Health Emergency and Disaster Risk Management.”

Effective DRR hinges on concerted national implementation and it is critical to examine the implications of the DRR paradigm across societal sectors and health domains. The 2031 targets of the Sendai Framework call for substantial global reductions in disaster-related mortality, number of affected people, direct economic loss and damage to critical infrastructure. The UN General Assembly endorsed indicators to measure progress against the Sendai Framework’s seven global targets. Using these indicators, Australia has already prepared an initial report on its Sendai Framework data readiness. The benefits of this approach to the Australian emergency management sector are clear: improved preparedness, more effective response, rehabilitation and reconstruction and more effective post-disaster recovery and reconstruction to build back better. However, it is considered a significant challenge for Australia to fully engage with the international monitoring and reporting process. Nonetheless, at the UNISDR Global Platform for Disaster Risk Reduction in Cancun, Mexico in May 2017, Senator Concetta Fierravanti-Wells, Minister for International Development and the Pacific, in delivering Australia’s official statement, reaffirmed that the Australian Government is firmly committed to implementing the Sendai Framework.

Following the Global Platform meeting, an expert workshop “Health and Disaster Risk Reduction: State of the Art and Implications for Australia” was held at the University of Melbourne in July 2017. The workshop was jointly hosted by the Centre for Mental Health, Melbourne School of Population and Global Health and the European Union Centre on Shared Complex Challenges. The workshop was conducted in collaboration with partners of Flinders University, RMIT, University of Sydney and Public Health England. The expertise of national and international
experts and practitioners was sought from the health and emergency management sectors. The intent was to explore the critical intersections of the fields of health and DRR and implications of the Sendai Framework for Australia. A number of participants who attended the Global Platform meeting, and two research papers led by the WHO Thematic Platform for Health Emergency and Disaster Risk Management Research Group informed the structure and process of the expert workshop.

What was discussed

A review of the Sendai Framework pointed towards health, science and technology to engage with transdisciplinary and interdisciplinary partners to provide evidence to inform policy and practice. The implementation of the Sendai Framework requires national reporting on indicators every two years. A summary of Australian-based resources and disaster databases was included. The need for partnerships within localities and across decision-making areas within government at all levels and with all health care, academic and private organisations was key within Australia. Mental health effects arising from all hazards have been identified as a major area of concern as all disasters impact on the health of the population, bringing about substantial losses and disruptions to health systems. The example of the impact of a recent incident ("thunderstorm asthma" in November 2016 in Victoria) on the population was used to demonstrate the complex nature of such events. The preparedness for health care response in the US was shared and the role of primary care in disasters was discussed. The Australian Red Cross reported on its work to encourage people-centred action in their FedIPlan.

Workshop discussions focused on identifying principal risks and hazards across health domains and fields of practice and key strategies to mitigate these risks. Following lively discussions between the four working groups, outputs recommended that it was important to know the hazards and risks that exist but plan and train for an all-hazard approach recognising that interagency communication for preparing, warning and informing Australian communities and the wider public requires trust. It was essential to listen to and understand local community issues and to have a dialogue with mutual trust and respect. A call for the recognition of the central place of health across all national and global policy frameworks was made.

The following recommendations were made:

- Consider producing an interpretive statement of the Sendai Framework to assist all levels of government to understand its implications for Australia and its relevance to global, national and local initiatives.
- Consider developing local hazards risk assessments to develop an Australian National Risk Register, possibly using the UK National Risk Register as a model.
- Consider creating an Australian DRR research network/alliance that maintains a research registry that could reflect the UK Alliance for Disaster Risk Reduction. Suggestions for how such an alliance could be facilitated include linking to support decision makers at all levels of government and building partnerships between academics, their discipline and their universities or other relevant organisations and to celebrate the rich and diverse Australian disaster research community.
- Consider creating a partnership to enhance foresight and early warning, possibly using as a model the UK Natural Hazards Partnership, which was established in 2011. This provides a network of government and academic partners to support early warning and other activities requested by the UK Cabinet Office for communications and services for civil contingencies, governments and the responder community. This is important because no such partnership exists in Australia and it would appear from the UK experience that such a collaboration between similar organisations strengthens consistent DRR standards and guidelines and improves outcomes.

Fulfilling the Sendai Framework objectives requires concerted action from key stakeholders across government, academic, sectoral and community levels to address existing research gaps to reflect the all-hazard approach. The WHO Official Statement at the Global Platform Cancun May 2017 states it values collaboration and partnerships and that the ‘recent development of WHO Thematic Platform for Health Emergency and Disaster Risk Management Research Group brings together representatives of Member States and academia who are committed to strengthening the evidence base for health policy and practice’ is important. It would be beneficial if Australian academic health professionals were encouraged to engage in this activity.

This abridged article will be published in full in Monograph 2, February 2018 on the Australian Disaster Resilience Knowledge Hub.

Author affiliations

Lunnart Raible – Melbourne School of Population and Global Health, University of Melbourne.
Fouad Al-Ghouti – Torrens Resilience Institute, Flinders University.
Anthony Capan – Planetary Health Initiative, University of Sydney.
John Handmer – Risk and Community Safety, RMIT University.
Alistair Humphrey – Canterbury District Health Board, New Zealand.
Virginia Murray – Public Health England, United Kingdom.
Caroline Spencer – Monash University Disaster Resilience Initiative, Monash University.
Diana F Wong – Monash University Disaster Resilience Initiative, Monash University.

24 Australian Institute for Disaster Resilience
Appendix 2: UNISDR Submission

DISASTER EVALUATION TYPOLOGIES

Authors: Diana Wong (1), Caroline Spencer (1), Leanne Boyd (1), Dudley McArdis (1), Frederick Burke Jr (1), and Frank Archer (1).
(1) Monash University, Melbourne, Australia

INTRODUCTION:
Evaluations in the disaster setting are diverse, lack definitional clarity and methodological agreement, and, require a stronger evidence-base. The aim of this paper is to outline an evolving framework of typologies to underpin the diversity of disaster evaluations.

METHODS:
A literature review of contemporary disaster evaluation standards and guidelines was undertaken and used to identify evaluation typologies and inform an initial comprehensive framework.

RESULTS/OUTCOMES:
The core structure of the evolving framework was adapted from the TFCDM ‘Guidelines for Research and Evaluation in Health Disaster Management’ (2003). Four evaluation domains and two overarching concepts were identified. Shown below first as thumbnails and then as a single comprehensive framework.

<table>
<thead>
<tr>
<th>Core Structure of the evolving framework</th>
<th>Baseline Evaluations providing essential pre-event data</th>
<th>Consequence Evaluations outlining the structural and functional damage as a result of a disaster</th>
<th>Outcome Evaluations often in the context of formal inquiries</th>
<th>Impact Evaluations measuring causality of specific interventions at any stage in the disaster</th>
<th>Accountability: donors, stakeholders, beneficiaries</th>
<th>Evaluation Standards and Evidence</th>
</tr>
</thead>
</table>

**EVALUATION STANDARDS AND GUIDELINES**
- Responsible and ethical evaluation practice
- Cross-sectional research, collaboration and dissemination

**EVIDENCE-BASED REVIEWS AND REGISTRIES**
- Accountability: donors, stakeholders, beneficiaries

**DISASTER EVALUATION TYPOLOGIES: Comprehensive Framework**
- Rapid Needs Assessment (Damage), Day 1
- Detailed Needs Assessment (Functional), Days 2-3
- Post Disaster Needs Assessment (PDNA)
- Independent Real Time and Collaborative Joint Evaluations
- Operational/Strategic/Institutional Reviews (Internal): Disaster, After Action Reviews, Lessons Learnt
- Operational/Strategic/Institutional Reviews (External): Audit, Key Performance Indicators, Quality Improvement
- Government Inquiries
- Multi-Disciplinary Event Reports/Reviews
- Process and Outcome Evaluations
- Health Impact Assessment (HIA) Evaluations

**STRENGTHENING RESILIENCE**
- “Culture of Safety”: risk reduction, prevention and mitigation
- Reduce exposures and vulnerabilities; build antipathetic, absorbing and adaptive capacities; and promote community development (Bahadur, 2015)

**PRE-EVENT STATUS**
- HAZARD
- RISK REDUCTION
- EVENT
- DAMAGE
- CORRUPTION CHARGE IN FUNCTIONS CONSEQUENCES
- RESPONSE: (Prepared and alerted, ‘Never Again!’)
- POST-EVENT STATUS

**PRE-EVENT**
- PRE-C
- BASELINES

**EIA-C**
- EVENT
- POST-EVENT

**CONSEQUENCES**

**OUTCOMES**

**DISCUSSION**
As a ‘work-in-progress’, this framework of typologies has been successfully used to guide the development of graduate education programs in disaster evaluations and to structure associated research projects and evaluations. The framework facilitates common communication and structuring the science and evidence-base of disaster health research and practice. A future international Disaster project will be undertaken to validate the framework. Subsequently, for each evaluation type identified and included in the final framework, a Gold Standard ‘how to guide’ will be identified or will need to be developed.

**CONCLUSION**
The framework demonstrates the different typologies of disaster evaluations that are available and brings them together into one framework. It reveals interdependencies and relationships that exist between various evaluation methodologies in the disaster setting. This evolving framework provides an opportunity to better structure the science of disaster health.

**Acknowledgements:** With thanks to our Monash colleagues, my family for their home feedback on earlier versions of the comprehensive framework, and Lauren Vassallo for her amazing graphic design work.

For more information please contact Diana Wong
dianawong@peakmail.com.au
©2015 Diana Wong. All Rights Reserved

327
International Savannah Fire Management Initiative
Sam Johnston, The United Nations University Institute for the Advanced Study of Sustainability (UNU-IAS)

Innovative tools and effective scheme to support establishing effective disaster damage and loss database -Strategy of Global Centre for Disaster Statistics-
Yuri Utsu, International Research Institute of Disaster Science, Tohoku University

Forecasts for triggering humanitarian actions: science to support the Forecaster based Financing pilot projects
Liz Stephens, University of Reading

The use of Multistor and Flood Warnings UAVs to Assess Damage Generated by Tropical Cyclone Pam (Category 5) in Vanuatu 2015.
Bernd Dammann, Pacific Community - Geoscience Division

Strategic Collaboration between Scientists and Civil Protection in DRR: The Example of the VeTOOLS Project
Joan Marti, Horta, Group of Seismology, Institute of Earth Sciences Jaume Almera, CSIC, Barcelona, Spain

Building Resilience through climate information in Burkina Faso and Ethiopia
Sophie Rigg, King’s College London

Disaster Risk Governance in Bangladesh to Manage River Flood risk
Md. Salimu Ahm Shathi, TU Dortmund, Germany

Detection of water based Bifhenol A using real-time microwave sensing
Mohammed Russej, Deakin University of Technology

The Socioeconomic Component of OpenQuake: Assessing Earthquake Risk using the Integrated Risk Modelling Toolkit
Christopher Burton, Global Earthquake Model (GEM) Foundation

Disaster Mental Health Risk Reduction: An Upstream Paradigm Shift for Disaster Mental Health
Lehant Kesler, University of Melbourne / Free University of Berlin

Work stream 4: Leveraging science through capacity development and research

Building a Culture of Resilience through Cyberinfrastructure Technologies
Ahmed A. E. Ahmed, President, International Association for Wind Engineering

Making advances in science and technology available for Disaster Risk Reduction planning in middle and low-income countries
Christopher Wardle, DFID Foundation

System Enabled Real-Time Coordinated Decision Making in Crisis and Emergency Management Using the Sendai Framework
Aleaks Amaya, University College Cork, Centre for Security and Emergency Management

Projections, Prospects, and Challenges of the Community data
Shahzana Khan, International Social Science Council

Trends in science and technology for Disaster Risk Reduction and the implementation of the Sendai Framework 2015-2030: a case study-based analysis
Ann-Mari Aal, Public Health England

Measuring Disaster Resilient Communities: A Case Study of 63 Coastal Communities in Arusha and on Nias Island, Indonesia
Mejia Kafu Disaster and Climate Change Study Centre, Nepal (ICCCSC Nepal)

Strengthening adaptation and resilience to climate variability and change in Kenya
Nyree Ninger, Met Office

Sea Seeds of Change: Local Best Practices in Disaster Risk Reduction Efforts
Evangelos E. Lamberts, De La Salle University-Manila

WHO/WHRP 10DWeather project to harness global science for better local warning systems
Brian Goulding, Met Office

In the Heat of the Moment: Scientists, Scientific Risk and Expertise during Disasters and Hazard Events.
Deborah Branson, Brainstorm Center

Governance & advanced regionalism: keys to integrated risk management in Morocco
Abdeslam Badre Mohammed, V University of Rabat

Reducing Flood Disaster Risk using Participatory Mapping as Capacity Building
Haris Rahadando, Electronics Engineering Polytechnic Institute of Surabaya

The Influence of Civil Society in Negotiation Processes at the United Nations: the Case of the Women’s Major Group on the Sendai Framework for Action
Lian Kimball, UNICEF

Advancing the understanding of Creeping Disasters for Resilience Building in Africa: The Role of Partnerships and Capacity Building through BRACED in Uganda
Shaba Oloya, Makerere University

Leveraging Science and Policy on Disaster Risk Reduction through Regional Networks – A Case Study
Snezana Krsic
Appendix 3: Ethics Approval 1 for the Utstein Study

Project: Disaster Health Evaluation: A Review of the Utstein Style Guidelines/Template, MUHREC Approval CF14/158 - 2014000041

Monash University Human Research Ethics Committee (MUHREC)
Research Office

Human Ethics Certificate of Approval

This is to certify that the project below was considered by the Monash University Human Research Ethics Committee. The Committee was satisfied that the proposal meets the requirements of the National Statement on Ethical Conduct in Human Research and has granted approval.

Project Number: CF14/158 - 2014000041
Project Title: Disaster Health Evaluation - Review of Utstein Style Guidelines/Template
Chief Investigator: Emeritus Prof Frank Archer
Approved: From: 17 January 2014 To: 17 January 2019

Terms of approval - Failure to comply with the terms below is in breach of your approval and the Australian Code for the Responsible Conduct of Research.

1. The Chief Investigator is responsible for ensuring that permission letters are obtained, if relevant, before any data collection can occur at the specified organisation.
2. Approval is only valid whilst you hold a position at Monash University.
3. It is the responsibility of the Chief Investigator to ensure that all investigators are aware of the terms of approval and to ensure the project is conducted as approved by MUHREC.
4. You should notify MUHREC immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.
5. The Explanatory Statement must be on Monash University letterhead and the Monash University complaints clause must include your project number.
6. Amendments to the approved project (including changes in personnel): Require the submission of a Request for Amendment form to MUHREC and must not begin without written approval from MUHREC. Substantial variations may require a new application.
7. Future correspondence: Please quote the project number and project title above in any further correspondence.
8. Annual report: Continued approval of this project is dependent on the submission of an Annual Report. This is determined by the date of your letter of approval.
9. Final report: A Final Report should be provided at the conclusion of the project. MUHREC should be notified if the project is discontinued before the expected date of completion.
10. Monitoring: Projects may be subject to an audit or any other form of monitoring by MUHREC at any time.
11. Retention and storage of data: The Chief Investigator is responsible for the storage and retention of original data pertaining to a project for a minimum period of five years.

Professor Nip Thomson
Chair, MUHREC

cc: Dr Caroline Spencer, Assoc Prof Leanne Boyd, Mrs Diana Wong
Appendix 4: Ethics Approval 2 for the Validation of The Comprehensive Framework for Disaster Evaluation Typologies 2017

Project: Validation of ‘A Comprehensive Framework for Disaster Evaluation Typologies’, Project Number 7737

Monash University Human Research Ethics Committee

Approval Certificate

This is to certify that the project below was considered by the Monash University Human Research Ethics Committee. The Committee was satisfied that the proposal meets the requirements of the Minimal Statement on Ethical Conduct in Human Research and has granted approval.

Project Number: 7737
Project Title: Validation of ‘A Comprehensive Framework for Disaster Evaluation Typologies’
Chief Investigator: Emeritus Prof Francis Acker
Enquiry Date: 07/03/2012

Terms of approval: Failure to comply with the terms below is in breach of your approval and the ‘Instructions for the Responsible Conduct of Research’.

1. The Chief Investigator is responsible for ensuring that permission letters are obtained, if relevant, before any data collection can occur at the specified organisation.
2. Approval is only valid whilst you hold a position at Monash University.
3. It is the responsibility of the Chief Investigator to ensure that all investigators are aware of the terms of approval and to ensure the project is conducted as approved by MUREC.
4. You should notify MUREC immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.
5. The Exploratory Statement must be on Monash letterhead and the Monash University complaints dance must include your project number.
6. Amendments to approved projects including changes to personnel must not commence without written approval from MUREC.
7. Annual Report - continued approval of this project is dependent on the submission of an Annual Report.
8. Final Report - should be provided at the conclusion of the project. MUREC should be notified if the project is discontinued before the expected completion date.
9. Monitoring - project may be subject to an audit or any other form of monitoring by MUREC at any time.
10. Retention and storage of data - The Chief Investigator is responsible for the storage and retention of the original data pertaining to the project for a minimum period of five years.

Thank you for your assistance.

Professor My Thomson
Chair, MUREC

O/C Ms Diana Wong, Dr Assoc Prof Leonie Boyd, Dr Caroline Spencer, Prof Frederick Biddle

List of approved documents:

<table>
<thead>
<tr>
<th>Document Type</th>
<th>File Name</th>
<th>Date</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploratory Statement</td>
<td>DansiWong_22334318_Exploratory_Report_V1_2017_02_13</td>
<td>13/02/2017</td>
<td>V1</td>
</tr>
<tr>
<td>Consent Form</td>
<td>DansiWong_22334318_Quarantine_V1_2017_02_13</td>
<td>13/02/2017</td>
<td>V1</td>
</tr>
<tr>
<td>Questionnaire / Surveys</td>
<td>DansiWong_22334318_Quarantine_V1_2017_02_13</td>
<td>13/02/2017</td>
<td>V1</td>
</tr>
</tbody>
</table>
Appendix 5: Explanatory Statement for the Utstein Study

EXPLANATORY STATEMENT

Disaster Health Evaluation - Review of the Utstein Style Guidelines/Template

Project: Disaster Health Evaluation: A Review of the Utstein Style Guidelines/Template

Chief Investigator: Emeritus Professor Frank Archer
Student: Diana Wong
Department: Monash University Disaster Resilience Initiative

You are invited to take part in this study, which is titled: ‘Disaster Health Evaluation: A Review of the Utstein Style Guidelines/Template’. Please read this Explanatory Statement in full before deciding whether or not to participate in this research. If you would like further information regarding any aspect of this project, you are encouraged to contact the researchers via the phone numbers or email addresses listed above.

What does the research involve?
The aim of the study is to investigate the utilisation of and the barriers or facilitators to the Utstein Style Guidelines/Templates for the reporting of evaluation and research in health disasters. Participants will be contacted via Skype at a time that is convenient for you and will participate in a one on one semi structured interview which will be audio recorded.

Why were you chosen for this research?
You were chosen for this research based on your experience in the fields of Disaster Medicine, Disaster Management, Emergency Management or the humanitarian arena based on your expertise in or undertaking evaluation studies.

Source of funding
There is no funding for this project.

Consenting to participate in the project and withdrawing from the research
Please read and sign the consent form. The form can be returned either by fax or scanned and emailed back. You have the right to withdraw from further participation at any stage. It may not be possible to withdraw the data once the responses have been submitted.

Possible benefits and risks to participants
There is no expectation of any physical/psychological stress, inconvenience or discomfort with the interview process.
Confidentiality
Your responses will remain anonymous and only the researchers will have access to the original data.

Storage of data
Data will be retained for five years in a private office with security and in a locked filing cabinet that only the researcher will have access to.

Use of data for other purposes
The data provided may be used by the researchers in future projects including, but not limited to: PhD thesis, journal article/book/chapter, conference presentation, online web based and oral presentation. Only aggregate de-identified data will be used for other projects where ethics approval has been granted.

Results
Participants can contact the researchers and request a summary of the findings.

Complaints
Should you have any concerns or complaints about the conduct of the project, you are welcome to contact the Executive Officer, Monash University Human Research Ethics (MUHREC):

Executive Officer
Monash University Human Research Ethics Committee (MUHREC)

Thank you,

Chief Investigator’s signature:

Chief Investigator’s name: Emeritus Professor Frank Archer
Appendix 6: Consent Form for the Utstein Study

CONSENT FORM

Review of Utstein Style Guidelines/Template
Project: Disaster Health Evaluation: Review of Utstein Style Guidelines/Template

Chief Investigator: Emeritus Professor Frank Archer, Dr Caroline Spencer, Associate Professor Leanne Boyd and Diana Wong (PhD Candidate)

I have been asked to take part in the Monash University research project specified above. I have read and understood the Explanatory Statement and I hereby consent to participate in this project.

<table>
<thead>
<tr>
<th>I consent to the following:</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio and/or video recording during the interview / focus group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taking part in a focus group of up to 12 - 15 people</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The data that I provide during this research may be used by Emeritus Professor Frank Archer, Dr Caroline Spencer, Associate Professor Leanne Boyd and Diana Wong in future research projects.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Name of Participant

Participant Signature ____________________________ Date ___________
Appendix 7: Theme List of Initial Questions for the Utstein Study

Theme List of Initial Questions

Project: ‘Disaster Health Evaluation: Review of Utstein Style Guidelines/Template’

Chief Investigator: Emeritus Professor Frank Archer, Dr Caroline Spencer, Associate Professor Leanne Boyd and Diana Wong (PhD Candidate)

Welcome

Introduction: My name is Diana Wong and I am a PhD student at Monash Injury Research Institute (MIRI), Monash University Disaster Resilience Initiative (MUDRI) at Monash University looking at ‘Disaster Health Evaluation’.

Can I confirm with you that you have read the explanatory statement, signed a consent form to participate in this research and returned this signed consent form and that you are willing to have this interview recorded?

If at any stage you wish to terminate the interview, please let me know. Would it be ok to commence the interview now?

Definition of ‘Disaster’:

- For the purpose of this interview, the word ‘disaster’ will include – disasters (natural and man-made), catastrophes, complex emergencies, humanitarian crises (conflict, famine, epidemics) and public health emergencies

Initial Questions

Demographics of Participant:

- Prime discipline; general appointments and general experience in disaster / emergency settings and evaluation/s

Question One:

- Are you familiar with the 2003 version of ‘Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style? (Now referred to as the ‘Utstein Guidelines’). The ‘Utstein Guidelines’ were published as the ‘red book’ supplement to PDM in 2003 and can be downloaded from the World Association for Emergency and Disaster Medicine (WADEM) website at: http://www.wadem.org/guidelines.html
- If yes, how did you gain this information/awareness?
- If yes, do you use them?
• If yes, are they valuable?
• How have you found them to be valuable?
• Are you aware of other colleagues using them?
• If no, move onto Q2

**Question Two:**

• From the literature review we can see that the ‘Utstein Guidelines’ are being referenced, but not used. Do you have any suggestions to offer on why they are not being used?

**Question Three:**

• Are you aware of any other similar guidelines that are being used?
• Are you aware of any recent updates on the ‘Utstein Guidelines’?
• 2 recent updates on the ‘Utstein Guidelines’ include:
  
  [http://journals.cambridge.org/download.php?file=%2FPDM%2FS1049023X14000211a.pdf&code=2364a433c0032b299ea5e022a6df7e32](http://journals.cambridge.org/download.php?file=%2FPDM%2FS1049023X14000211a.pdf&code=2364a433c0032b299ea5e022a6df7e32)
  
  
  [http://sjp.sagepub.com/content/42/14_suppl.toc](http://sjp.sagepub.com/content/42/14_suppl.toc)

**Question Four:**

• Do you have any suggestions on how we can improve the evaluation of disaster health?

Thank you for your time today.

Do you have any questions?

Are you happy for me to come back to you if I need further clarification?

Disaster Metrics: A Comprehensive Framework for Disaster Evaluation Typologies

Diana F. Wong, MCP Ngs;1,2 Caroline Spencer, PhD;1 Lee Boyd, PhD;1,4 Frederick M. Bunker, Jr. MD, MPH;1,2,4 Frank Archer, MBBS, MPH1a

Abstract

Introduction: The frequency of disasters is increasing around the world with more people being at risk. There is a moral imperative to improve the way in which disaster evaluations are undertaken and reported with the aim of reducing preventable morbidity and mortality in future events. Disasters are complex events and undertaking disaster evaluations is a specialized area of study at an international level.

Hypothesis/Problem: While some frameworks have been developed to support consistent disaster research and evaluation, they lack validity, consistent terminology, and standards for reporting across the different phases of a disaster. There is yet to be an agreed, comprehensive framework to structure disaster evaluation typologies.

The aim of this paper is to outline an evolving comprehensive framework for disaster evaluation typologies. It is anticipated that this new framework will facilitate an agreement on identifying, structuring, and relating the various evaluations found in the disaster setting with a view to better understand the process, outcomes, and impacts of the effectiveness and efficiency of interventions.

Methods: Research was undertaken in two phases: (1) a scoping literature review (peer-reviewed and “grey literature”) was undertaken to identify current evaluation frameworks and typologies used in the disaster setting; and (2) a structure was developed that included the range of typologies identified in Phase One and suggests possible relationships in the disaster setting.

Results: No consensus framework to structure disaster evaluation and research was identified in the literature. The authors propose a “Comprehensive Framework for Disaster Evaluation Typologies” that identifies, structures, and suggests relationships for the various typologies detected.

Conclusion: The proposed Comprehensive Framework for Disaster Evaluation Typologies outlines the different typologies of disaster evaluations that were identified in this study and brings them together into a single framework. This unique, unifying framework has relevance at an international level and is expected to benefit the disaster, humanitarian, and development sectors. The next step is to undertake a validation process that will include international leaders with experience in evaluation, in general, and disasters specifically. This work promotes an environment for constructive dialogue on evaluations in the disaster setting to strengthen the evidence base for interventions across the disaster spectrum. It remains a work in progress.


Keywords: disaster; evaluation; framework; typology

Abbreviations:
BII: International Initiative for Impact Evaluation
ALNAP: Active Learning Network for Accountability and Performance in Humanitarian Action
HRA: health risk assessment
IFRC: International Federation of the Red Cross and Red Crescent Societies
NGO: nongovernment organization
PDNA: post-disaster needs assessment
SFDRR: Sendai Framework for Disaster Risk Reduction 2015 – 2030
UN: United Nations

UNEG: United Nations Evaluation Group

Received: July 2, 2016
Retrieved: October 17, 2016
Accepted: November 12, 2016
Online publication: May 8, 2017
doi:10.1017/S1049023X17000471

References
improve the approach to understanding and reporting disaster evaluations, with the aim of reducing preventable mortality and morbidity in future events. Improvement of the quality of disaster evaluations and strengthening accountability is urgently required. While some frameworks have been developed to support consistent disaster research and evaluation, they are fragmented and uni-focused. There is yet to be an agreed, comprehensive framework to structure disaster evaluation typologies. Such a framework could provide consistency in terminology and standards for reporting across the different phases of a disaster, with a view to providing comparability to better understand the process, outcomes, and impacts of the efficacy and efficiency of interventions. Sharing methodological experiences would contribute to the further development of these standards and guidelines to systematically build disaster science.

Undertaking disaster evaluations is a specialized area of study at an international level. Different approaches to evaluation over the years have led to a variety of definitions being offered when describing the term "evaluation." An earlier definition of evaluation put forward by the Joint Committee on Standards for Educational Evaluation (JCSEE) in 1994 states that "evaluation is the systematic assessment of the worth or merit of an object." The "object" in this case is the program, project, or intervention under review. Other recent definitions focus more on active purposes such as accountability assessment, decision making, program improvement, or organizational learning. Regardless of the definitions used, evaluations are largely conducted to find areas for improvement and to generate an assessment of overall quality and value, usually for reporting or decision making purposes.

The aim of this paper is to outline an evolving Comprehensive Framework for Disaster Evaluation Typologies. It is anticipated that this new framework will facilitate an agreement on organizing and describing the various evaluations found in the disaster setting. While continuing to be a work in progress, it is intended that this work will add structure to the current understanding and help to underpin the diversity of disaster evaluation typologies that currently exist.

When considering the title of the framework and how best to describe this body of work, the authors considered two words: methodology and typology. The word "methodology" or "methodology" is defined as "a particular procedure for accomplishing or approaching something." The preferred term for the framework was "typology," which refers to "a structure of different types," and is a closer match to describing the classification of the variety of disaster evaluation styles that are currently available.

Methodology

This research was undertaken in two phases. Phase One was designed to identify current evaluation frameworks and typologies in the disaster setting. A scoping literature review was undertaken in two parts. Firstly, the peer-reviewed literature was searched using major electronic databases, including PubMed/Medline (US National Library of Medicine, National Institutes of Health; Bethesda, Maryland USA); CINAHL (EBSCO Information Services, Ipswich, Massachusetts USA); EMBASE (Elsevier; Amsterdam, Netherlands); PsycINFO (American Psychiatric Association USA); Science Web (Thomson Reuters; New York, New York USA); Scopus (Elsevier; Amsterdam, Netherlands); and Web of Knowledge (Thomson Reuters; Philadelphia, Pennsylvania USA). These databases were searched to identify contributions to the history and development of disaster/disaster health evaluation frameworks/models/dispositional. The key search words used included "disaster OR emergency," AND "health," AND "guidelines OR framework OR models OR repositories OR evaluation OR evaluation typology." Including criteria consisted of articles in English, published after 2003, and included frameworks, models, or methodologies rather than exemplars of specific evaluations. Additional references were identified through examination of bibliographies from the most recent publications (snowballing) and through scrutiny of the contents pages of highly relevant journals. This scoping review was supplemented by a convenience sample of international colleagues who commented on the evolving framework to identify additional relevant typologies.

Secondly, a review of the "grey literature" also was undertaken, including similar key words, using Google and Google Scholar (Google Inc; Mountain View, California USA) and supplemented by "ReliefWeb," a resource maintained by the United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA; New York, USA and Geneva, Switzerland) and the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP; London, United Kingdom). The ALNAP is an organization that is dedicated to improving humanitarian performance through learning and accountability. The scoping review provided examples of a wide range of typologies used under the general label of "evaluation." The following Comprehensive Framework lists at least one example of each evaluation type identified as an exemplar; however, it does not intend to list all evaluations identified. One hundred twenty-two papers were used in developing this Comprehensive Framework.

In Phase Two, all co-authors contributed to developing a structure that included the range of typologies identified in Phase One, and suggested potential relationships in the disaster setting. The resulting "Comprehensive Framework for Disaster Evaluation Typologies" not only identifies and structures different disaster evaluation typologies, but it also suggests relationships between these typologies and all phases of the disaster cycle. Various disaster evaluation typologies are mapped across the disaster timeline, demonstrating their inter-relationships. It is not the intent of this paper to describe perceived strengths or weaknesses of any particular evaluation typology. It is important to note that Baselines, Consequences, and Outcomes evaluation typologies are related but not hierarchical; that is, one is not more important that the other and are to be interpreted within the context of a specific disaster.

Results

The literature review revealed that more information can be found in the "grey literature" and humanitarian arena than in peer-reviewed literature. There were very few evaluations of health interventions during disaster reported in the literature. Most evaluation reports were descriptive, process-focused, and lacked a core conceptual framework. Recent research undertaken by Stratton in 2014 identified that the majority of papers submitted to Prehospital and Disaster Medicine (PDM) were surveys or descriptive in nature. The published reports did not demonstrate a consistent and structured approach to evaluations of interventions, and the impact of interventions on the affected population was rarely measured. Many nongovernment organizations (NGOs), such as the International Federation of Red Cross and Red Crescent Societies (IFRC; Geneva, Switzerland) and various United Nations (UN) agencies, have their own internal standards for evaluations. Attempts are being made to consolidate standards and guidelines across the sector, as evidenced by the work of ALNAP and the Inter-Agency Standing Committee (IASC). National government disaster organisations, however, were noticeably absent in this activity.
One influential guideline identified during the literature review was the "Health Disaster Management Guidelines for Evaluation and Research" (hereafter referred to as "The Guidelines"). The Guidelines were co-authored by the Task Force on Quality Control and Disaster Management (TFFQCDM), the World Association of Disaster and Emergency Medicine (WADEM), Madison, Wisconsin USA, and the Nordic Society for Disaster Medicine. It provided a conceptual framework for undertaking research and evaluation in the disaster setting. The core of this conceptual framework was frequently referenced in peer-reviewed papers and scholarly journals, and was used extensively in the field of disaster research and in higher degree research theses. But it was rarely used as the methodological framework for undertaking disaster evaluations and research.

The literature review revealed an article and a book that utilized the "Conceptual Framework" and terminology used in "The Guidelines". The three articles/books were based on the Sumatra-Andaman Earthquake and subsequent Asian Tsunami that occurred in December 2004.

Consideration of The Guidelines to undertake the "Comprehensive Framework" included a validation stage. In-depth interviews of 18 experts in the field of disaster and emergency medicine undertaken by the lead author in 2014 and 2015 revealed that the core framework of The Guidelines was deemed to be valuable and was being referenced. It was not, however, being used to structure research and evaluation in the disaster setting.

In an attempt to test the validity of the core framework, the authors undertook a thematic analysis of seven Australian disaster reports/queries from 2006-2014 to see if the core elements of the conceptual framework in "The Guidelines" were present in all reports/queries. The disasters occurred in four different Australian States, covered four different types of events, included four different types of reports, and were chaired by six different Chairpersons. Results from the thematic analysis were reviewed by two researchers and identified that all elements of the "Conceptual Framework" were present in each of the seven Australian disaster reports/queries.

Given this support for "The Guidelines" from both the internationals experts and the thematic review of Australian reports, it was decided to use its core structure, with some modifications, to undertake this Framework for Disaster Evaluation and Research (DFER).

Other frameworks or guidelines that were identified included work by Stephenson, Powers and Daily; Kullig et al.; Debauche; Fatima; Sundnes; and Birnbaum et al.

While Fatima identified more than 10 frameworks, she also discovered that some had been validated and they were not commonly used to structure evaluations and research in the disaster setting.

It became evident that a core unifying framework did not exist to structure disaster evaluation and research. In an attempt to create a tool to consolidate the diverse non-standardized frameworks together, the authors utilized key components from "The Guidelines" and "The Impacts Framework" (comprising of event, event characteristics, object, harm, and impacts) from Stephenson into their framework. Disaster Evaluation Typologies was created by linking and integrating various typologies into a unifying core structure that aimed to inform and support a Comprehensive Framework for Disaster Evaluation Typologies.

The Comprehensive Framework for Disaster Evaluation Typologies is presented using the following headings:

1. Figure One: Core Structure;
2. Figure Two: Baselines;
3. Figure Three: Consequences;
4. Figure Four: Outcomes;
5. Figure Five: Impact Evaluation;
6. Figure Six: Accountability;
7. Figure Seven: Evaluation Standards and Evidence;
8. Figure Eight: Disaster Evaluation Typologies: Comprehensive Framework.

Core Structure

The Core Structure outlines the fundamental framework of Disaster Evaluation Typologies to which all other entries will be related (Figure 1: Core Structure).

The Core Structure consists of three important layers. The first layer is found at the bottom of the diagram and provides a preliminary and simplistic view of the disaster continuum or timeline. At its most basic level, this layer has three core elements or phases that are represented by a pre-event phase, an event phase, and the post-event phase of an emergency or disaster. While each phase can be identified individually, their timing is not necessarily sequential and the phases can overlap. Post-event phases will influence the pre-event phase of subsequent events.

The second layer is represented by an expansion of the earlier pre-event, event, and post-event phases and is based on a modified representation of "The Guidelines". Additional detail is evident and the relationships of key disaster phases are demonstrated, including:

- Pre-event status (of the community);
- Hazard;
- Risk reduction;
- Event;
- Damage;
- Disruption, changes in functions, consequences of damage;
- Response (respond and adapt), rescue, surge, and relief;
- Recovery (recover, grow/thrive/transform), reconstruction, development, renewal, and regeneration (early recovery and persistent recovery); and
- Post-event status (of the community).

The third layer introduces the concept of "Strengthening Resilience" in an overarching theme. It enhances and enriches the Core Structure of Disaster Evaluation Typologies and is an emerging, international imperative embedded within the Sendai Framework for Disaster Risk Reduction 2015-2030 (SFDRR). Key elements of "Strengthening Resilience" include:

- System structure, governance, coordination, and leadership;
- "Culture of Safety" with consideration to: risk reduction, prevention, and mitigation;
- Reduce exposure and vulnerabilities; build anticipative, absorbing, and adaptive capacities; and promote community development;
- Elements are identified in five contemporary and influential international frameworks:
  1. Sendai Framework for Disaster Risk Reduction 2015-2030 (SFDRR);
  2. Sustainable Development Goals (SDGs);
  3. Climate Change Conference (COP21).
Figure 1. Core Structure.

4. Global Facility for Disaster Reduction and Recovery (GFDRR) Recovery Framework, also known as Post-Disaster Needs Assessment (PDNA) Recovery Framework, and
5. Rockefeller Foundation 100 Resilient Cities (RC) Framework.

The Core Structure for the Comprehensive Framework is illustrated in Figure 1.

Baselines

Baselines are a series of evaluations or assessments that occur during the pre-event phase of disasters and provide information about the current state of the community (Figure 2: Baselines).

Baseline evaluations include any information or data that have been collected prior to an event or disaster occurring. It covers both the pre-event status of a community and the actual hazard itself. Obtaining baseline information in the pre-event phase is critical in understanding the state of the community, how it has been affected by a disaster, and the subsequent damage that has occurred. This information assists in identifying community strengths, weaknesses, and vulnerabilities to disasters. Additionally, this information will assist in developing appropriate disaster management and disaster risk-reduction strategies.

Innovations in science and technology have made it easier in recent years to collect information that helps reduce disaster risk, and therefore, plan for the future. The SDRRR, ratified in Sendai, Japan in March 2015 by 187 UN Member States, acknowledges that there is a growing demand for science and technology to play a more prominent and effective role in providing evidence for policy and decision making. Knowledge is essential to the process. A strengthened evidence base to support the implementation of disaster risk-reduction strategies also is required.

Furthermore, Priority 4, paragraph 34(b) of the SDRR supports the "further development and dissemination of instruments, such as standards, codes, operational guides, and other guidance instruments to support coordinated action in disaster preparedness and response to facilitate information sharing on lessons learned and best practices for policy practice and post-disaster reconstruction programmes." Examples of Baseline evaluations include, but are not limited to:

- Baseline evaluations and evaluability,
- Demographics and infrastructure (such as Geographic Information Systems/GIS),
- Epidemiology and emerging threats,
- Define minimum standards/criteria: baselines, targets, and indicators,
- Preparedness, resilience capacity with a special focus on the SDRR,
- Risk management, surveillance, and early warning, and
- Health Impact Assessment (HIA) Predictive.
Baseline information and evaluations are illustrated in Figure 2.

Consequences

Consequences are a series of evaluations and/or assessments that occur after the event or disaster has occurred and include assessment of damage and changes in function (Figure 3: Consequences).

Consequence evaluations include any information or assessments that have been collected after an event or disaster has occurred. It covers both the event and post-event phase of the disaster timeline. Systematic data collection and assessment is required in order to inform disaster needs analysis after an event. It is used in monitoring the effectiveness of response and recovery interventions and to aid decision making.

Examples of Consequence evaluations include, but are not limited to:

- Rapid needs assessments (damage), usually occur on Day 1 after the event; 77, 79
- Detailed needs assessments (functional), usually occur on Days 2-3 and may include PDNA; 80
- Continual assessments that include monitoring and surveillance, usually occur on multiple occasions after the event; 77 and
- Independent real-time evaluations and collaborative joint evaluations are contemporary evaluation types. 82

The information received from these evaluations will ideally be compared with previous Baseline studies and incorporated into helping to plan response and recovery for the current event, provide feedback into planning and preparing for subsequent events, and assist in disaster risk reduction. 83 Currently, damage and loss trends are difficult to monitor over time, partly due to inconsistent methodologies and the fact that very few countries keep national disaster databases. Even then, only one in five countries will have consistently recorded economic losses using validated tools and data collection methods. 84 The PDNA’s aim to provide a common approach to post-crisis needs assessments and recovery planning. 85 The Centre for Research on the Epidemiology of Disasters (CRED, Brussels, Belgium) promotes research, training, and information dissemination on disasters. 86 In the Australian context, the Australian Business Roundtable for Disaster Resilience and Community Safety provides a first-time overview of disaster data with the aim of making Australian communities safer and more resilient to natural disasters. 87

In an attempt to reduce disaster risk and strengthen resilience, a feedback loop is present in Figure 8 from Consequences to Baselines. Consequence evaluations are illustrated in Figure 3.

Outcomes

Outcomes are a series of evaluations and/or assessments that occur towards the end of the post-event phase of a disaster (Figure 4: Outcomes).
Figure 3. Consequences.

Outcome evaluations reflect information or data that have been collected after an event or disaster has occurred. These evaluations include summative reviews of processes used in managing the event and outcomes related to the post-event status of the community. This information will ideally be incorporated into planning and preparing for the next event or disaster.

Examples of Outcome evaluations include, but are not limited to:

- Operational/Strategic/Institutional Reviews that are Internal Reviews and include: Debriefs, After Action Reviews,22,29 and Lessons Learned;50
- Quality Reviews that are External Reviews and include: Audit, Key Performance Indicators, and Quality Improvement;26
- Government Inquiries;22,93
- Multi-disciplinary event reports/reviews (for example, Kelling et al. and Fattah);40,41
- Process and outcome evaluations that include:94,95 logic maps/models, theories of change, and causal links/attributions used to guide process evaluations. Each intervention and evaluation requires its own logic map that describes the sequence of actions to be undertaken and that communicates what the program is and will do;96 and
- HIA Evaluations.77

In an attempt to reduce disaster risk and strengthen resilience, a feedback loop is present in Figure 3 from Outcomes to Baselines. Outcome evaluations are illustrated in Figure 4.

Impact Evaluations

Impact Evaluations of programs, projects, and interventions are evaluations that include a measure of causality or attribution and can occur during any phase of the disaster timeline (Figure 5: Impact Evaluations).

In the disaster setting, Impact Evaluations have gained popularity for identifying causal links between specific interventions and outcomes. This is a result of the internationa community demanding accountability and improved evidence-based interventions.12,72 Although there remains ongoing debate about the exact definition of Impact Evaluations,12 they are particularly well-suited to answer important questions, such as: whether interventions do or do not work; whether interventions make a positive or negative impact; whether there are intended or unintended consequences; and how cost effective they are.12,96 It is believed they will greatly improve the effectiveness of interventions delivered in the disaster setting by identifying what works for whom, and why.12

Examples of Impact Evaluations include, but are not limited to:

- Battenstein Impact Evaluation in the Post-Disaster Setting: A Case Study of the 2005 Pakistan Earthquake.201
Figure 4. Outcomes.

- Rogers Introduction to Impact Evaluation, 112
- Pint et al What Methods May Be Used in Impact Evaluations of Humanitarian Assistance 113 and
- Chambers et al Designing Impact Evaluations: Different Perspectives. 113

Impact Evaluations are illustrated in Figure 5.

Accountability

Accountability to donors, stakeholders, and beneficiaries is a cross-cutting theme across all phases of the disaster timeline and is applicable to every evaluation undertaken in the disaster setting 124 (Figure 6: Accountability).

Over the last 20 years, there has been a call for greater Accountability in disaster and humanitarian settings. More recently, at an international level, there has been mounting pressure to strengthen quality, accountability, and learning practices, while also ensuring transparency. 105-108 The lack of an accepted definition for “accountability” in the humanitarian context remains a challenge. The term “accountability” seems to represent a whole range of concepts and principles. 109 The ALNAP is an example of an international organization dedicated to improving humanitarian performance through accountability and increased learning. 110

For the purpose of this Typology, the term “accountability” will be defined as the means in which power is used responsibly. This includes consideration of the views of all interested parties (including donors, stakeholders, and beneficiaries). 115

Examples of Accountability evaluations include, but are not limited to:

- 2013 Humanitarian Accountability Report, 116 and

Accountability is illustrated in Figure 6.

Evaluation Standards and Evidence

Evaluation Standards and Guidelines, Evidence-Based Reviews and Registers, and Knowledge Management are important cross-cutting themes that are relevant throughout the entire disaster timeline (Figure 7: Evaluation Standards and Evidence). Evaluation Standards and Guidelines include generic Evaluation Standards, such as:

- General evaluation theorists, including but not limited to Patton, 118 Scriven, 119 and Stufflebeam. 120
- United Nations Evaluation Group (UNEG) Standards for Evaluation in the UN System 114
- American Evaluation Association (AEA) Program Evaluation Standards, 123 which are based on earlier works by Yarbrough et al. 17

October 2017

Prehospital and Disaster Medicine
Evaluation Standards and Guidelines also include guidelines for responsible and ethical conduct in undertaking evaluations that include, but are not limited to:

- United Nations Evaluation Group (UNEG) Ethical Guidelines for Evaluation;¹⁵⁶
- Australian Council for International Development (ACFID) Guidelines for Ethical Research and Evaluation in Development;¹²⁷ and
- Australian Evaluation Society (AES) Guidelines for the Ethical Conduct of Evaluations.¹²⁸

Evidence-Based Reviews and Registers include meta-evaluations, systematic reviews, other types of literature review methodologies, and registries of evaluation reports.

The level and quality of evidence in this setting has been reviewed by Clarida and Darcy in Insufficient Evidence? The Quality and Use of Evidence in Humanitarian Action – A UNDP Study.¹²⁹ Despite improvements over the last 20 years, they identified that there remains room for further development in the quality and use of evidence in the humanitarian setting. The authors also suggest that "evidence matters: the use of good quality evidence improves the effectiveness and accountability of humanitarian action, and is in accordance with humanitarian ethics and principles."²⁰
Systematic Reviews are structured, comprehensive literature reviews that utilize a rigorous and published search strategy, with the aim of minimizing selection bias. Examples of Systematic Reviews in this discipline include, but are not limited to:

- Blanchet et al., *An Evidence Review of Research on Health Interventions in Humanitarian Crisis.*
- Gallardo et al., *Core Competencies in Disaster Management and Humanitarian Assistance: A Systematic Review.*
- Modela et al., *Characteristics of an Effective International Humanitarian Assistance: A Systematic Review.*

Other literature review typologies include: scoping reviews, gap analyses, and priority settings.

**Meta-Evaluations**

Meta-Evaluations are systematic and formal evaluations of evaluations and are a high-level of evidence; however, they are uncommon in the disaster setting.

Examples of Meta-Evaluations include, but are not limited to:


**Registries**

For the purpose of this paper, Registries (sometimes called repositories) are defined as publicly available, free-access collections of evaluation studies that have been undertaken in the disaster setting. Registries aim to help build capacity and strengthen disaster risk reduction and resilience. A separate review of such Registries undertaken by the lead author suggests that these are not well-known in the disaster sector but contain a large number of evaluation reports that might be of use to aid decision making and improve practice.

Examples of disaster evaluation Registries include, but are not limited to:

- Independent Evaluation Group (IEG) hosted by the World Bank.
- Humanitarian Evaluation and Learning Portal (HELP) hosted by ALNAP.
- IFRC\textsuperscript{143}
- Evaluation and Research Database (ERD) hosted by United Nations Children’s Emergency Fund (UNICEF)\textsuperscript{144} and
- Impact Evaluations hosted by 3ie\textsuperscript{145}

Knowledge management includes cross-sectoral research, collaboration, and dissemination of information to improve the evidence base of disaster science and to improve practice. Sharing knowledge enables informed decision making regarding disaster risk reduction and management.\textsuperscript{146}

Evaluation Standards and Guidelines, Evidence-Based Reviews and Registries, and Knowledge Management are relevant in advancing the science of disaster evaluations by providing scientific rigor, common terminology, and the ability to replicate various methodologies.

Evaluation Standards and Guidelines are illustrated in Figure 7.

**Disaster Evaluation Typologies: Comprehensive Framework**

Disaster Evaluation Typologies: Comprehensive Framework identifies the different typologies of disaster evaluations and demonstrates key relationships in a single diagram. It suggests the interdependencies and relationships between various evaluation typologies along the disaster timeline and within the disaster setting. This consolidates the previous Figures 1–7 (Figure 8: Disaster Evaluation Typologies: Comprehensive Framework).

A strong evaluation framework for disaster settings is extremely important given the increasing frequency and scale of disasters. It would need to utilize agreed definitions and be able to measure the impact and effectiveness of interventions. It is anticipated that Disaster Evaluation Typologies: Comprehensive Framework will create a useful and usable framework and promote an environment for constructive dialogue on an international level.\textsuperscript{147}

Figure 8 also includes feedback loops from Consequences and Outcomes to improve Baselines, reduce disaster risk, and strengthen resilience. The framework is not limited to any one phase of the disaster timeline and can be used for responding to disasters, humanitarian crises, or in the development sector.

**Discussion**

Natural disasters themselves are complex events\textsuperscript{148–150} and undertaking structured evaluations in this setting is also a complex activity. The authors intended to create a classification of disaster evaluation typologies that would provide structure, encourage common terminology, and advance the evidence base of disaster science.
Figure 8. Disaster Evaluation Typologies: Comprehensive Framework.

The role of the framework is to support the ability to measure and evaluate the effectiveness of interventions provided in the disaster setting and thereby reduce the increasing human and economic costs associated with disasters.

The Comprehensive Framework outlined in this paper is the first framework of this type and thus makes a unique contribution to current knowledge. No previous reference has been located that identifies such a wide range of evaluation typologies used in the disaster setting and further provides conceptual relationships in a single comprehensive framework. The aim of the authors is consistent with that of James J. James in his recent Editorial where he concludes “A common Disaster Medicine and Global Health taxonomy will form the foundation of a safer, more resilient world, through more effective preparedness and response, but we must first come together for the public good.” [13]

The Comprehensive Framework will undergo further research to validate the typologies and their relationships through structured interviews with targeted international experts in both general evaluation and disaster evaluations. Subsequently, additional work is needed to identify and develop toolkits of standards and guidelines for each of the evaluation typologies identified, as well as any methods that come to light as a result of the validation process. A recent example includes the framework for Disaster Research and Evaluation published by Birnbaum, Daży, O’Rourke, Loretz, and Kushneri. [41-42]

Limitations
Limitations of the evolving Comprehensive Framework include difficulty in searching the “gray literature” and identifying all evaluation typologies used in this sector. There may be evaluation typologies that have not been identified. Secondly, the suggested relationships are framed through the eyes of the authors and there may be alternative perspectives to frame these relationships. Thirdly, there is a lack of a unifying theory for disaster evaluation. Finally, the authors have not considered specific research methods that might be utilized in the disaster setting. These can be found in any standard textbook on research methods in epidemiology, social sciences, or limited disciplines. It is hoped that the validation process will address these limitations.

Conclusion
Disaster Evaluation Typologies: Comprehensive Framework identifies the different typologies of disaster evaluations that were identified in this study and brings them together in a single framework. It suggests interdependencies and relationships that exist between various evaluation typologies within the disaster setting. This unique unifying framework has relevance at an international level and is expected to benefit the disaster, humanitarian, and...
development sectors. This work promotes an environment for constructive dialogue on evaluation in the disaster setting and adds to the evidence base of disaster evaluation and research.

Acknowledgements
The authors would like to acknowledge and thank: Lauren Vassallo for her valuable graphic design work on the diagrams; and Jackie van Dam and Dr. Sarah Wong for their assistance with editing the manuscript.

Supplementary Material
To view supplementary material for this article, please visit https://doi.org/10.1111/s1049023X71006471

References


57. Bakkala A, Loend F, Willissom E, Tamei T. Resilience in the ADG Developing an indicator for Target 5.3.1 that is Fit for Purpose, United Kingdom: Overseas Development Institute (ODI); 2015.


Appendix 9: Online Validation Survey Including Explanatory Statement and Consent

Consent

MONASH University

Invitation to participate in research

Validation of a 'Comprehensive Framework for Disaster Evaluation Typologies'

Researchers: Emeritus Professor Francis Archer, Dr Caroline Spencer, Professor Leanne Boyd, Professor Frederick M. Burkle Jnr and Ms Diana Wong (PhD student)

Dear Colleague,

Now that the World Congress on Disaster and Emergency Medicine (WCDEM) has concluded and people have returned home, we are hoping you will participate in our research.

I am a PhD student at Monash University Accident Research Centre. On behalf of the research team, I invite you to participate in an international peer-review validation process of a 'Comprehensive Framework for Disaster Evaluation Typologies'. This study seeks your comments on an evolving and contemporary framework that summarises disaster evaluation typologies mapped over the disaster timeline. Your comments will contribute to informing this validation process. This research has been approved by the Monash University Human Research Ethics Committee.

Your participation in this peer-review is greatly appreciated. Instructions are provided to assist you in completing the review.
Thank you for considering this request.

**Diana Wong** on behalf of
**Emeritus Professor Francis Archer**
**Dr Caroline Spencer**
**Professor Leanne Boyd**
**Professor Frederick M. Burkle Jnr**

15th May, 2017

**Explanatory Statement and Consent**

**Research Project:** Validation of a ‘Comprehensive Framework for Disaster Evaluation Typologies’

**Chief Investigator**
Emeritus Professor Francis Archer  Diana Wong
Monash University Accident Research Centre  PhD student

You are invited to take part in this study. Please read this Explanatory Statement in full before deciding whether or not to participate in this research. If you would like further information regarding any aspect of this project, you are encouraged to contact the researchers via the phone numbers or email addresses listed above.

**What does this research involve?**
This study aims to gain a consensus on disaster evaluation typologies that evaluators currently use, or could use, in the disaster setting.

The survey questions comprise two parts, should you agree to participate:

1. Demographics
   - Provide us with some basic information about yourself, including your age, gender, country of residence and level of experience.
2. Questionnaire
   - Read the information sheet and complete the questionnaire that is attached to this document, following the instructions provided.

The survey questions should take approximately 30 - 40 minutes to complete and will be available until **Monday, 5th June, 2017**.

**Why were you chosen for this research?**
You were chosen to participate in this research based on your experience in the fields of Disaster Medicine, Disaster Management, Emergency Management; Humanitarian and Development Sector; and/or Evaluation. Your contact details were obtained through public domain resources or by reference from our international network.

**Consenting to participate in the project and withdrawing from the research**
Your participation in this research is voluntary. You may choose not to participate and you may withdraw your consent to participate at any time. It may not be possible to withdraw the data once the responses have been submitted.

**Possible benefits and risks to participants**
This unique, unifying framework has relevance at an international level and is expected to benefit the disaster, humanitarian and development sectors. There is no expectation of any physical or psychological stress, inconvenience or discomfort with completing the questionnaire.

**Confidentiality**
Your responses will remain anonymous and only the researchers will have access to the original data. Your name nor any another identifying information about you will be recorded. This survey is designed for you to complete it once, therefore, the Qualtrics platform will use your computer’s cookies and user identifiers to verify that each survey response is from a unique user. None of this technical information is collected or retained and at no time is it available to researchers or Monash University.

**Storage of data**
All the data collected will be treated in a confidential manner. Your responses will be entered into a database for analysis and will be stored on a password protected computer.

**Use of data for other purposes**
This study is being conducted in partial fulfilment of a PhD degree under the supervision of Emeritus Professor Francis Archer, Dr Caroline Spencer and Professor Leanne Boyd. The only people with access to the data are those on the research team. Results of this study will be published in professional journals, presented at conferences and will form an integral part of Diana Wong’s PhD thesis.

**Results**
Participants can contact the researchers and request a summary of the findings.

**Complaints**
If you have any concerns about the conduct of this research, you are welcome to contact Monash University Human Research Ethics Committee (MUHREC) via phone on [Insert Phone Number] or via email at: [Insert Email Address]. If you have any questions about the research please contact the Chief Investigator, Emeritus Professor Francis Archer via phone on: [Insert Phone Number] or via email at: [Insert Email Address].

I have read and understood the Explanatory Statement and I hereby consent to participate in this research. Please tick to confirm your participation in this research.

Yes
No

**Copyright**
This work is based on a recent publication Wong DF, Spencer C, Boyd L. Burkle Jr. F, Archer F. Disaster metrics: a comprehensive framework for disaster evaluation typologies. *Prehosp Disaster Med.* 2017:32(5):1-14. All rights reserved. Prehospital and Disaster Medicine copyright material or any portion thereof may not be reproduced or used in any manner whatsoever without the express written
permission of Prehospital and Disaster Medicine except for the use of brief quotations with appropriate credit in a literary review, journal article, newsletter, or book.

I have read and understood the Copyright regulations above and agree to abide by them.

Yes
No

You have chosen ‘No’ to understanding and agreeing to the Copyright regulations, in order to proceed with the survey you must comply with these regulations.

Thank you for considering this request,

Emeritus Professor Francis Archer
Dr Caroline Spencer
Professor Leanne Boyd
Professor Frederick ‘Skip’ Burkle Jnr
Ms Diana Wong

15th May, 2017

Part A: Demographics

Validation of a ‘Comprehensive Framework for Disaster Evaluation Typologies’

This study aims to gain a consensus on disaster evaluation typologies that evaluators currently use, or could use, in the disaster setting.

Part A: Demographics

This section helps us gain an understanding of the people participating in the survey and will help us assess the survey coverage and interpret the results.
What is your age (in years)?

- < 30 years
- 31 - 40
- 41 - 50
- 51 - 60
- > 60 years

What is your gender?

- Male
- Female

What is your primary country of residence?

Select the category that best describes your current field of experience (i.e., the job or role where you spend most time)

- Disaster Medicine, Disaster Management, or Emergency Management
- Humanitarian or Development Sector
- Evaluation

How many years, full time equivalent, have you had in this current field (as indicated above) ?

- < 5 years
- 6 - 10 years
- 11 - 15 years
- 16 - 20 years
- 21 - 25 years
- > 25 years
Do you undertake or supervise disaster/emergency or humanitarian/development evaluations as part of your normal work?

Yes
No

**Background**

The following questions relate to the 'Comprehensive Framework for Disaster Evaluation Typologies'. Please read the following information for insight about the development of the framework. The information is intended to help you with your decision making.

Disasters are complex events and undertaking disaster evaluations is a specialised area of study at an international level. While some frameworks have been developed over the years to support consistent disaster research and evaluation, they lack validation, consistent terminology and standards for reporting across different phases of a disaster. Currently, no agreed, comprehensive framework to structure disaster evaluation typologies exists.

This framework outlines an evolving comprehensive framework for disaster evaluation. An anticipated outcome of this new framework would be to facilitate an agreement for identifying, structuring and relating the various evaluations found in the disaster setting. Furthermore, it would help in understanding the processes, outcomes and impacts of the effectiveness and efficiency of interventions across the disaster spectrum.

During the literature review, it was identified that no core, unifying framework to structure disaster evaluation and research existed. The authors propose a 'Comprehensive Framework for Disaster Evaluation Typologies' that identifies, structures and suggests relationships for the various typologies found. The
different types of disaster evaluations that were identified have been assembled into a single framework. This unique, unifying framework has relevance at an international level and is expected to benefit the disaster, humanitarian and development sectors. As internationally experienced leaders in these areas, the authors seek your assistance in validating this framework.

Importantly, this Comprehensive Framework is not an evaluation typology in itself, nor is it a list of evaluation studies. Rather, it is framework that outlines different Disaster Evaluation Typologies.

The ‘Comprehensive Framework for Disaster Evaluation Typologies’ is presented using the following headings:

1. Figure One: Core Structure
2. Figure Two: Baselines
3. Figure Three: Consequences
4. Figure Four: Outcomes
5. Figure Five: Impact Evaluations
6. Figure Six: Accountability
7. Figure Seven: Evaluation Standards and Evidence
8. Figure Eight: Disaster Evaluation Typologies: Comprehensive Framework

**Core Structure**

**1. Core Structure**

The Core Structure outlines the fundamental framework of the Disaster timelines to which all typologies will be related.

The Core Structure consists of three important layers. The first layer, at the bottom of the diagram, provides a preliminary and simplistic view of the disaster continuum, represented by a pre-event phase, an event phase and the post-event phase of an emergency or disaster.
The second layer expands the earlier, more basic layer and is based on a modified representation of the ‘Health Disaster Management Guidelines for Evaluation and Research in the Utstein Style’ (Task Force on Quality Control of Disaster Management, 2003).

The third layer introduces the contemporary concept of ‘Strengthening Resilience’ as an important and overarching theme. It underpins and enriches the Core Structure and is an emerging, international imperative embraced within the Sendai Framework for Disaster Risk Reduction 2015 – 2030, and enhanced from other sources.

Figure One: Core Structure

The elements illustrated in each layer of the Core Structure reflect contemporary thinking on disaster timeline elements.

Strongly agree
Somewhat agree

360
Neither agree nor disagree
Somewhat disagree
Strongly disagree

Are there any disaster timeline elements not covered in the Core Structure?

Yes
No
Unsure

Please add your comments

Are there any contemporary disaster timeline elements which could be substituted in the Core Structure?

Yes
No
Unsure

Please add your comments

Baselines
2. Baselines

*Baselines* comprise a series of evaluations or assessments that occur during the pre-event phase of disasters and provide information about the current state of the community.

Baseline evaluations include any information or data collected prior to an event or disaster occurring. It covers both the pre-event status of a community and the actual hazard itself. This information assists in identifying community strengths, weaknesses and vulnerabilities to disasters, while also supporting the development of appropriate disaster management and disaster risk reduction strategies. Additionally, it will also provide background information for comparison with data from consequences or outcomes following specific events.

Figure Two: Baselines
The evaluation typologies illustrated in Baselines reflect current baseline evaluation methods.

Strongly agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

Are there any baseline evaluation typologies not covered in the Baselines?

Yes
No
Unsure

Please add your comments

Consequences

3. Consequences

Consequences comprise a series of evaluations and/or assessments that occur after a disaster event and include assessment of damage and change in function.

Consequence evaluations cover both the event and post-event phase of the disaster timeline. Systematic data collection and assessment informs disaster needs analysis after an event. It is used in monitoring the effectiveness of response and recovery interventions and to aid decision making.
The information received from these evaluations would ideally be compared with previous baseline studies and incorporated into helping to plan response and recovery for the current event, provide feedback into planning and preparing for subsequent events and assist in disaster risk reduction.

In an attempt to reduce disaster risk and strengthen resilience, a feedback loop is present in Figure Eight: Disaster Evaluation Typologies: Comprehensive Framework from *Consequences to Baselines*.

Figure Three: Consequences

The evaluation typologies illustrated in *Consequences* reflect current consequence evaluation methods.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
Somewhat disagree
Strongly disagree

Are there any consequence evaluation typologies not covered in the Consequences?

Yes
No
Unsure

Please add your comments

Outcomes

4. Outcomes

Outcomes comprise a series of evaluations and/or assessments that occur towards the end of the post-event phase of a disaster.

Outcome evaluations reflect information or data collected after an event or disaster. These evaluations include summative reviews of processes used in managing the event and outcomes related to the post-event status of the community. This information would ideally be incorporated into planning and preparing for the next event or disaster.

In an attempt to reduce disaster risk and strengthen resilience, a feedback loop is present in Figure Eight from Outcomes to Baselines.
The evaluation typologies illustrated in *Outcomes* reflect current outcome evaluation methods.

Strongly agree  
Somewhat agree  
Neither agree nor disagree  
Somewhat disagree  
Strongly disagree

Are there any outcome evaluation typologies not covered in the *Outcomes*?

Yes  
No  
Unsure
Impact Evaluations

5. Impact Evaluations

Impact Evaluations of programs, projects and interventions comprise evaluations that include a measure of causality or attribution and can occur during any phase of the disaster timeline.

In the disaster setting, Impact Evaluations have gained popularity for identifying causal links between specific interventions and outcomes resulting from the international community demanding accountability and improved evidence-based interventions. The two diamonds on the diagram reflect the inclusion of impact evaluations which can occur at any time along the disaster timeline, and not just at the two points displayed.

Figure Five: Impact Evaluations
The evaluation typology of Impact Evaluations illustrated above reflects current impact evaluation methods.

Strongly agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

Could Impact Evaluation typologies be portrayed in a different manner?

Yes
No
Unsure

Please add your comments
Accountability

6. Accountability

Accountability to donors, stakeholders and beneficiaries comprise a cross-cutting theme across all phases of the disaster timeline and applies to every evaluation undertaken in the disaster setting.

Over the last 20 years there has been a call for greater accountability in disaster and humanitarian settings and more recently, at an international level. Accordingly, there has been mounting pressure to strengthen quality, accountability and learning practices, while also ensuring transparency.

The lack of an accepted definition for accountability in the humanitarian context remains a challenge. The term ‘accountability’ represents a wide range of concepts and principles. Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) is an example of an international organisation dedicated to improving humanitarian performance through accountability and increased learning.

For the purpose of this typology, the term ‘accountability’ refers to the means in which power is used responsibly. This includes consideration of the views of all interested parties including donors, stakeholders and beneficiaries.

Figure Six: Accountability
The evaluation typologies of Accountability illustrated above reflect current accountability methods.

Strongly agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

Could Accountability typologies be portrayed in a different manner?

Yes
No
Unsure

Please add your comments
Evaluation Standards and Evidence

7. Evaluation Standards and Evidence

Evaluation Standards and Guidelines; Evidence-Based Reviews and Registries; and Knowledge Management reflect important cross-cutting themes relevant throughout the entire disaster timeline.

Evaluation Standards and Guidelines include generic evaluation standards, disaster specific evaluation standards and guidelines and ethical guidelines in undertaking evaluations. They also include reference to the use of Logic Models and Theory of Change. Evidence-Based Reviews and Registries include meta-evaluations, systematic reviews, other types of literature review typologies and registries of evaluation reports. Knowledge Management includes cross-sectoral research, collaboration and dissemination of information to improve the evidence-base of disaster science and to improve practice. Sharing knowledge enables informed decision making regarding disaster risk reduction and management.

Evaluation Standards and Guidelines; Evidence-Based Reviews and Registries; and Knowledge Management are relevant in advancing the science of disaster evaluations by providing scientific rigour, common terminology and the ability to replicate various methodologies.

Figure Seven: Evaluation Standards and Evidence
The evaluation typologies illustrated in *Evaluation Standards and Evidence* reflect current evaluation standards and evidence methods.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Are there any evaluation standard and evidence typologies not covered in *Evaluation Standards and Evidence*?

- Yes
- No
- Unsure

Please add your comments
Comprehensive Framework

8. Comprehensive Framework

*Disaster Evaluation Typologies: Comprehensive Framework* identifies the different typologies of disaster evaluations and demonstrates key relationships in a single diagram. It suggests the interdependencies and relationships between various evaluation typologies along the disaster timeline and within the disaster setting and consolidates the previous Figures One – Seven.

A strong evaluation framework for disaster settings demands extreme rigor, given the increasing frequency and scale of disasters. The framework would need to utilise agreed definitions and have the capacity to measure the impact and effectiveness of interventions. An anticipated outcome of the *Disaster Evaluation Typologies: Comprehensive Framework* would be the creation of a useful and usable framework to promote an environment for constructive dialogue on an international level, while also structuring the science of disaster evaluations.

Figure Eight includes feedback loops from *Consequences* and *Outcomes* to improve *Baselines*, to inform reducing disaster risk and strengthening resilience. The framework is not limited to any one phase of the disaster timeline and can be used for responding to disasters, humanitarian crises or in the development sector.

You are reminded that this Comprehensive Framework is not an evaluation typology in itself, nor is it a list of evaluation studies. Rather, it is framework that outlines different Disaster Evaluation Typologies. Following this validation process, where guidelines and standards do not exist, our next step is to promote the development of guidelines for the remaining typologies.
Taken from an article in print - Wong DF, Spencer C, Boyd L, Burke Jr. F, Archer F. Disaster metrics: a comprehensive framework for disaster evaluation typologies. Prehosp Disaster Med. In Press.

Other references supporting the above information are available upon request from Diana Wong at: Diana.F.Wong@monash.edu

Figure Eight: Disaster Evaluation Typologies: Comprehensive Framework

The relationships between the evaluation typologies and the disaster timeline are appropriately demonstrated in the Comprehensive Framework.
The Comprehensive Framework potentially **holds value** for my work when undertaking evaluations in this setting.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Please explain your reason(s) for this rating.

The Comprehensive Framework would be useful for **supporting and promoting evaluation** in this setting.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Please explain your reason(s) for this rating.
Please explain your reason(s) for this rating.

The *Comprehensive Framework* would be useful for **teaching evaluation** in this setting.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Please explain your reason(s) for this rating.

Do you have any additional thoughts on how we can **improve** the *Comprehensive Framework*?

Please comment on any **barriers** you perceive in undertaking evaluations in this
setting.

Please comment on any enablers you perceive in undertaking evaluations in this setting.
Appendix 10: Peer Reviewed Validation of a Comprehensive Framework for Disaster Evaluation Typologies (submitted to PDM)

Original Research

Peer Reviewed Validation of a Comprehensive Framework for Disaster Evaluation Typologies

Authors: Diana F. Wong, MCP Nsg;¹,² Caroline Spencer, PhD;¹ Leanne Boyd, PhD;³,⁴ Frederick M. Burkle Jr., MD, MPH;¹,⁵,⁶ Frank Archer, MBBS, MPH¹
1. Monash University Disaster Resilience Initiative (MUDRI), Monash University, Melbourne, Australia
2. New South Wales (NSW) Health, Sydney, New South Wales, Australia
3. School of Nursing and Midwifery, Monash University, Melbourne, Australia
4. Cabrini Health, Malvern, Victoria, Australia
5. Harvard Humanitarian Initiative, Harvard University, Cambridge, Massachusetts USA
6. Woodrow Wilson International Center for Scholars, Washington, DC USA
Corresponding author: Diana Wong at Diana.Wong2@health.nsw.gov.au

Abstract

Introduction

The Comprehensive Framework for Disaster Evaluation Typologies, developed in 2017 (CFDET 2017), aims to unify and facilitate agreement regarding the identification, structure and relationships between various evaluation typologies found in the disaster setting. A peer reviewed validation process sought input from international experts in the fields of disaster medicine, disaster/emergency management, humanitarian/development, and evaluation. This paper discusses the validation process, its results and outcomes.

Research Problem

Previous frameworks, identified in the literature, lack validation and consistent terminology. To gain credibility and utility, this unique Framework needed to be validated by international experts in the disaster setting.
Methods

A mixed methods approach was designed to validate the Framework. An initial iterative process informed an online survey which used a combination of a 5-point Likert scale and open-ended questions. Pre-determined consensus thresholds, informed by a targeted literature review, provided the validation criteria.

Results

A sample of 33 experts from eleven countries responded to the validation process. Quantitative measures largely supported the elements and relationships of the Framework, and strongly supported its value and usefulness for supporting, promoting and undertaking evaluations, and, its usefulness for teaching evaluation in the disaster setting. Qualitative input suggested opportunities to strengthen and enhance the Framework. There were limited responses to better understand the barriers and enablers of undertaking disaster evaluations. Feedback suggested changes to modify and improve the Framework. A potential for self-selection bias of respondents may be a limitation of this study. The attainment of high consensus thresholds, however, provides confidence in the validity of the results.

Conclusion

For the first time, a Framework of this nature has undergone a rigorous validation process by experts in three related disciplines at an international level. The modified Framework, CFDET 2018, provides a unifying framework within which existing evaluation typologies can be structured. It gives evaluators confidence to choose an appropriate strategy for their particular evaluation in the disaster setting and facilitates consistency in reporting across the different phases of a disaster to better understand the process, outcomes, and impacts of the efficacy and efficiency of interventions. Future research could create a series of toolkits to
support improved disaster evaluation processes and to evaluate the utility of the Framework in the real-world setting.

Keywords

disasters; evaluation; framework; validation
Peer Reviewed Validation of a Comprehensive Framework for Disaster Evaluation Typologies

Introduction
The Comprehensive Framework for Disaster Evaluation Typologies 2017, developed in 2017 (CFDET 2017), unifies and facilitates agreement on the identification, structure and relationships between the various evaluation typologies found in the disaster setting. Despite the evolution of disaster evaluation and research, and concerted efforts since 2003 to develop consensus on disaster research and terminology, little agreement from experts exists for a comprehensive and validated framework that structures disaster evaluation typologies. Previous frameworks, while recognized in the literature, lacked validation and were rarely used as a methodological framework to structure disaster evaluation and research. To add credibility and global relevancy to CFDET 2017, an international peer-reviewed validation process was conducted. The primary purpose of this paper discusses the validation process, results and outcomes.

Background
CFDET 2017 was developed around two key elements: the different phases of disasters (also referred to as the disaster timeline), and, evaluation typologies that have been or could be used in the disaster setting. The extensive, underpinning literature review of evaluation typologies previously reported were categorized into the following elements:

- Core Structure;
- Baseline Evaluation Typologies;
- Consequence Evaluation Typologies;
- Outcome Evaluation Typologies;
- Impact Evaluations;
• Accountability; and,
• Evaluation Standards, Evidence and Knowledge Management.

While ‘Accountability’ and ‘Evaluation Standards, Evidence and Knowledge Management’ are not true evaluation types, as recognized by the evaluation sector, their importance to the process was deemed significant enough for inclusion in CFDET 2017.

Each element was brought together into a single, unifying framework, CFDET 2017. The full development and rationale behind the structure of CFDET 2017, including the scoping literature review, was published in *Prehospital and Disaster Medicine* (PDM) in 2017 and is shown in Figure 1.

---

**Figure 1: Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017)**
CFDET 2017 promotes an environment for constructive dialogue on evaluation typologies currently in use, or that could be used, in the disaster setting, and used as a tool to support decision-making when undertaking evaluations before, during or after a disaster. Furthermore, it provides structure to improve the selection and reporting of valid disaster evaluation and research typologies thereby strengthening the evidence base for interventions delivered across the disaster spectrum.

This study addresses the research problem that previous frameworks, identified in the literature, lacked validation and consistent terminology and to gain credibility and utility, this unique Framework, CFDET 2017, needed to be validated by international experts in the disaster setting.

**Validation Process**

A focussed literature review, including peer-reviewed and grey literature, identified articles related to the validation of ‘comprehensive frameworks’, similar to CFDET 2017. Such articles were limited in the disaster or evaluation settings. The literature highlighted multiple perspectives and confusing terms and typologies for ‘validation’ and how to undertake a ‘validation process’.

This paper defines ‘validity’ as ‘the ability of the instrument to measure the attributes of the construct under study’. In this paper, the instrument under study is CFDET 2017, including its ‘attributes’ that are the concepts and elements (disaster phases and evaluation typologies found in the disaster setting) and their relationships.

Validity can be further divided into ‘external validity’ which measures the generalisability of the findings, and ‘internal validity’ that refers to the confidence placed in the cause and effect relationship. To make a framework useful, it needs both internal and external validity.

The literature review identified no ‘gold standard’ guidelines or consensus for undertaking a validation of CFDET 2017. The literature identified a wide variety of validation practices
ranging from complex three step Delphi processes\textsuperscript{35-38} to opinion reviews by, sometimes as few as two experts.\textsuperscript{34}

Davis-Stober et al\textsuperscript{35} and Jorm\textsuperscript{36} independently reviewed the ‘wisdom-of-the-crowd’ effect. Jorm argues that although evidence-based medicine has gained strength over the past two decades, especially for informing clinical decisions, and despite ‘expert consensus’ rating lowly on the Levels of Evidence Scale, ‘expert evidence’ should not automatically be classed as an ‘inferior method’.\textsuperscript{36} To support his argument, Jorm summarized the literature on Surowiecki’s concept of the ‘Wisdom of Crowds’, where the term ‘crowd’ referred ‘to any collection of individuals with some expertise’. Jorm quotes Surowiecki who proposed that certain conditions must be met for a ‘crowd to be wise’ and for expert consensus to produce good answers; namely: 1. diversity of expertise; 2. independence: the expert panel members make their decisions independently, uninfluenced by others; 3. decentralization of panel members; and 4. a mechanism of co-ordinating and aggregating the crowd’s expertise.\textsuperscript{36-37} Davis-Stober et al also modelled ‘wisdom-of-the-crowd’ effect with a mathematical approach and concluded that ‘while the members are individually biased and the crowd not particularly accurate, the crowd is still wise relative to the individual’. They also supported the importance of two of Surowiecki’s conditions, namely: diversity of the crowd; and aggregation of the crowd’s expertise.\textsuperscript{35}

The CFDET 2017 validation process primarily aimed to capture the ‘wisdom-of-the-crowd’ represented by a wide group of international experts. Surowiecki’s conditions guided the CDFET 2017 validation methodology.

The first challenge was to pre-determine the size and nature of the ‘Crowd’, or expert panel. Jorm in reviewing the ‘Wisdom of Crowds’ literature, specifically focussing on determining
expert panel size in Delphi studies, concluded that there was little firm guidance and that the stability of responses was demonstrated with panels of around 20 members. The CFDET 2017 validation aimed for 30 responses. Jorm also advised that panel attrition in Delphi studies may be larger if the questionnaire is long or requires a substantial time commitment. An observation that proved prophetic in this validation study.

The second challenge was to pre-determine the consensus thresholds. Jorm also reviewed the quantitative definition of ‘consensus’ used in Delphi or consensus studies, concluding that no single definition existed for ‘consensus’ and researchers would need to define consensus and provide a rationale in the context of their own studies. The definition for ‘consensus’ varied between 70% and 90% to determine endorsement as ‘essential or important’. For this CFDET 2017 validation, four definitions of consensus thresholds were used and are defined in the methods section of this paper.

**Methods**

A mixed methods approach validated the Comprehensive Framework. A preliminary iterative process of presenting CFDET 2017 at international conferences, national seminars and other academic meetings provided opportunities to discuss and receive formative feedback on the evolving CFDET 2017 prior to undertaking the summative validation process. This approach embraced the underlying principle of Action Research as an iterative process seeking input from end-users. Evolving versions were presented and discussed at three international conferences, including the World Congress on Disaster and Emergency Medicine (WCDEM), Cape Town, South Africa, April 2015; the United Nations Office for Disaster Risk Reduction (UNISDR) Science and Technology Conference, Geneva, Switzerland, January 2016; and the World Library and Information Congress, Kuala Lumpur, Malaysia, August 2018; and at three further professional and research forums at Monash University.
Feedback received contributed to progressive modifications to the evolving CFDET 2017, for example the addition of arrows to indicate feedback processes and ‘complete the circle’ of information. Other comments highlighted an opportunity to further clarify the theory underpinning the structure. These iterative events and the feedback received strengthened the development of CFDET 2017 and were a key component of the progressive validation process.

The formal validation process used an online survey, hosted on the Qualtrics Insight Platform,\textsuperscript{40} to collect quantitative (5-point Likert Scale) and qualitative data. An online survey, rather than interviews or focus groups, was selected due to the intended sampling size and the international nature of the participants.\textsuperscript{36} Given that participants were provided with an overview of CFDET 2017 as the instrument under study, which included the key concepts and elements, the survey was intended as a one-step survey. There was the potential to use a second-round survey had the first-round survey failed to reach consensus, however, this was not necessary as the first survey attained consensus.

A questionnaire was developed and pilot tested (n = 9). Results indicated no major structural changes or modifications were required, however, some minor edits were made to the survey instrument. Human Research Ethics approval was obtained from Monash University.

One hundred and forty experts with backgrounds in the areas of: disaster medicine and disaster/emergency management; humanitarian/development; or evaluation generalists were approached to participate in the validation via email. Both non-probability and purposive sampling were used to recruit the target population.\textsuperscript{41,42} The stratification of participants was based on the following characteristics of importance: primary country of residence with participants allocated into one of the six World Health Organisation (WHO) Regions of the
World; current field of experience (participants self-selected one of the following three subgroups: Disaster Medicine, Disaster Management, Emergency Management sectors; Humanitarian and Development sectors; or Evaluation sector).

The recruitment list was sourced from: the literature; the public domain; and, known professionals active in the fields of disasters/emergencies, humanitarian/development and/or evaluation. An explanatory statement was sent via email and consent to participate in the research was requested. Participants had three-weeks to complete the survey with a reminder email sent automatically at two weeks. Recruitment used a snowballing strategy until the consensus threshold of 30 responses was reached. The survey ran for a six-month period in 2017.

The online survey (refer to PDM link) consisted of two elements: an overview and explanation of CFDET 2017 which was consistent with the 2017 publication Disaster Metrics: A Comprehensive Framework for Disaster Evaluation Typologies which although was ‘in print’ and had not yet been published. Participants were asked to:

a. confirm their consent to participate and respect the intellectual property contained in CFDET 2017;

b. complete a short section on their demographics;

c. use a 5-point Likert scale (5 = strongly agree, 1 = strongly disagree) to identify their agreement or disagreement for a series of statements related to the elements and relationships found within CFDET 2017;

d. provide responses to open ended questions with opportunity to comment on the disaster timeline elements and the different evaluation typologies presented;
e. comment on the relationships and linkages between the disaster timeline and evaluation typologies as demonstrated in CFDET 2017; whether CFDET 2017 had value when undertaking evaluations; would be useful for supporting and promoting evaluations; and whether it would be useful for teaching evaluation in the disaster setting; and

f. suggest other evaluation typologies that were or could be used in the disaster setting but were not included in CFDET 2017; identify other improvements to CFDET 2017; and, identify potential barriers and enablers to undertaking evaluations in this setting.

Analysis of the quantitative responses used definitions and descriptive statistics facilitated by a Statistics Canada Platform and the Qualtrics Insight Platform. Qualitative responses utilized a thematic analysis facilitated by an Excel spreadsheet. As no single definition of consensus exists, a pre-survey determination was made that the following criteria would be used to indicate acceptance thresholds for the validation of CFDET 2017:

1. \( n = 30 \) responses;
2. Mean of 4.00, +/- 95% Confidence Intervals, on a 5-point Likert scale for each item;
3. Median of 4.00, + interquartile range, on a 5-point Likert scale for each item; and
4. 80% of participants who ‘strongly agree’ or ‘somewhat agree’ on a 5-point Likert Scale for each item.

Results

Demographics of Responders

Of the 140 experts approached, 39 (27.9%) completed the demographic section, however, only 33 (23.6%) completed all, or most, of the questionnaire related to CFDET 2017.
Demographic information showed that the number of female to male participation was almost equal, with 51.3% and 48.7% respectively. The majority of participants were aged 41 years and above (92.3%), with 33.3% greater than 60 years. While participants were invited from all six WHO Regions, responses were only received from four regions: Western Pacific (n=22), European (n=11), Americas (n=5) and Eastern Mediterranean (n=1). However, 11 countries were represented. Nearly two-thirds of participants (64.1%) identified their area of expertise as Disaster Medicine, Disaster Management or Emergency Management; 28.2% identified Evaluation; with the remaining participants (7.7%) categorizing themselves as belonging to the Humanitarian or Development sectors. Two-thirds of participants (66.67%) had greater than 11 years or more full-time experience in their self-selected area of expertise and of those, 46.1% had 16 (or more) years’ experience. The majority of participants (64.1%) indicated that they undertook evaluations or supervised evaluations as part of their normal work.

Elements and Relationships Represented within CFDET 2017

Table 1 provides the results for the responses on the elements and relationships represented with the Framework.
### Table 1: Elements and Relationships Represented Within CFDET 2017

<table>
<thead>
<tr>
<th>ELEMENTS</th>
<th>N</th>
<th>LIKERT SCORE</th>
<th>MEAN +/- 95% CI</th>
<th>MEDIAN (interquartile range)</th>
<th>ACCEPTANCE THRESHOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Structure</td>
<td>32</td>
<td>0 0 1 19 12</td>
<td>4.34 (4.15 – 4.54)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>96.88</td>
</tr>
<tr>
<td>Baselines</td>
<td>33</td>
<td>1 3 4 9 16</td>
<td>4.09 (3.69 – 5.00)</td>
<td>4.00 (3.50 – 5.00)</td>
<td>75.76</td>
</tr>
<tr>
<td>Consequences</td>
<td>32</td>
<td>0 1 3 12 16</td>
<td>4.34 (4.06 – 4.63)</td>
<td>4.50 (4.00 – 5.00)</td>
<td>87.50</td>
</tr>
<tr>
<td>Outcomes</td>
<td>32</td>
<td>0 2 3 16 11</td>
<td>4.13 (3.82 – 4.43)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>84.38</td>
</tr>
<tr>
<td>Impact</td>
<td>30</td>
<td>0 4 7 10 9</td>
<td>3.80 (3.42 – 4.18)</td>
<td>4.00 (3.00 – 5.00)</td>
<td>63.33</td>
</tr>
<tr>
<td>Accountability</td>
<td>29</td>
<td>0 5 5 6 13</td>
<td>3.93 (3.49 – 4.37)</td>
<td>4.00 (3.00 – 5.00)</td>
<td>65.52</td>
</tr>
<tr>
<td>Standards and Evidence</td>
<td>29</td>
<td>2 2 1 14 10</td>
<td>3.97 (3.57 – 4.40)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>82.76</td>
</tr>
</tbody>
</table>

Of the seven CFDET 2017 elements, the results show:

- all had between 29 and 33 responses;
- four elements had a mean of > 4.00. The remaining three elements had means close to 4.00, and all had 95% Confidence Interval (CI) which included 4.00;
- all had medians = or > 4.00, however, the interquartile range indicated the responses were skewed to the right, as may have been expected; and
- four elements received > 80% ‘strongly agree’ or ‘somewhat agree’, while three elements did not meet this consensus threshold.

Explanations for these observations are potentially apparent in the qualitative responses. In the qualitative responses, a total of 114 comments were specifically related to the elements within the Framework. There were no strongly negative comments and suggestions for major
change except related to ‘Accountability’. Some clarification suggestions were offered for the other elements and are considered in the discussion section of this paper.

**Holistic Representation as a Unifying Framework**

Table 2 provides the results for the responses on the holistic representation as a unifying Framework.

<table>
<thead>
<tr>
<th>N</th>
<th>LIKERT SCORE</th>
<th>MEAN +/- 95% CI</th>
<th>MEDIAN (interquartile range)</th>
<th>ACCEPTANCE THRESHOLD 80.00%</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0 2 2 17 9</td>
<td>4.10 (3.80 – 4.40)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>86.67</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>4.28 (3.97 – 4.58)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>86.21</td>
</tr>
<tr>
<td>29</td>
<td>0 1 3 12 13</td>
<td>4.28 (3.96 – 4.60)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>82.76</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>4.33 (4.03 – 4.63)</td>
<td>4.50 (4.00 – 5.00)</td>
<td>86.67</td>
</tr>
</tbody>
</table>

Of the four questions relating to the holistic representation as a unifying Framework, the results show:

- all had at least 29 responses;
- all exceeded a mean of > 4.00;
- all had a median of = or > 4.00, however, the interquartile range indicated the responses were skewed to the right, as may have been expected; and
- all comfortably received > 80% ‘strongly agree’ or ‘somewhat agree’.

Respondents were asked to rate and comment on whether the relationships between the disaster timeline and evaluation typologies were appropriately demonstrated within CFDET
2017. Quantitative results were met all four consensus thresholds (refer to Table 2). Qualitative comments endorsed CFDET 2017 with the following being typical comments: ‘makes sense for the way it is illustrated’; ‘great work putting it altogether in one diagram, it’s intuitive, well done’; ‘relationships are appropriately demonstrated’; and ‘I think this final framework is excellent’.

Respondents were asked to rate and comment on whether CFDET 2017 potentially held value when undertaking evaluations in the disaster or humanitarian setting. While the number of responses was slightly lower than preferred (n = 29), quantitative results met the other three consensus thresholds (refer to Table 2). Qualitative comments endorsed CFDET 2017 with the following being typical comments: ‘I think this can inform different understandings of the range of not just evaluation, but sources of data to analyse and inform better service delivery in humanitarian contexts’; ‘there is a need to show evaluation as a continuum – which your framework shows’; ‘it helps to understand the wider context and provides insight into how current and future studies relate to the overall picture’; ‘the various elements are clearly linked to the various phases of disaster’; ‘clear, thoughtful overview, higher level approach provides a useful framework’ and ‘the framework provides clear visual representation’.

Respondents were asked to rate and comment on whether CFDET 2017 would be useful for supporting and promoting evaluations in the disaster or humanitarian setting. While the number of responses was slightly lower than preferred (n = 29), quantitative results met the other three consensus thresholds (refer to Table 2). Qualitative comments endorsed CFDET 2017 with the following being a typical comment: ‘as clearly explained in the introduction of this survey, there is not a clear standard and framework for disaster evaluation. Such as well-developed systematic framework would definitely help in supporting and promoting sound disaster evaluation’. Further comments included: ‘it synthesizes the disaster evaluation
process’, ‘the framework is a useful visual representation that can be used to support and promote evaluations’; ‘identifies the facts to evaluate/measure’; ‘evaluators, particularly workers who are researching as a result of an event and are not academics would benefit from a clear framework to base their evaluation on’; and ‘by providing an internationally agreed upon framework it has the benefit of providing guidance on the research and clearer understanding of responsibilities in this field’.

Respondents were asked to rate and comment on whether CFDET 2017, as a unifying framework, would be useful for teaching evaluation in the disaster or humanitarian setting. Quantitative results met all four consensus thresholds (refer to Table 2). Qualitative comments endorsed CFDET 2017 with the following being typical comments: ‘each element is clearly linked to the various phases of disaster and provided a very useful visual representation, that is not only useful for garnering support but also for pulling apart in teaching environment’; ‘the framework graphically aids students learning the complexity and importance of relations among all elements’; ‘breaks down the system to simpler parts’; ‘clean and organised’; ‘helps to understand the process’; ‘provides a clear representation’; and a ‘very logical step by step framework’.

Additional Questions

Respondents were asked to identify any typologies for each of the elements that may have not been included in CFDET 2017. A number of suggestions were offered, however, most of these had been included in the published version of CFDET 2017 (1) which was not available to respondents at the time of the survey.

Respondents were asked to offer their thoughts on improving CFDET 2017. Only 11 additional comments were offered. No comments related to the structure of CFDET 2017.
Most comments related to operationalizing the Framework and included: ‘the best way for improvement is by testing the framework on the ground’; ‘my suggestion is not to suggest it replaced all other forms of organising ideas in this area of work, just suggest it adds to and pulls together many other elements’; ‘please be bold enough to revise once this work is out and has been utilised for a certain period of time’; and ‘you could make many subnotes for each section of the diagram to clarify what each section means, the diagrams should be easily understood with no knowledge of the field’.

Respondents were invited to comment on any barriers and enablers they perceived in undertaking evaluations in this setting. Only 13 responders commented on barriers and eight commented on enablers. Summative comments on barriers included:

- Complexity of evaluation guidelines (3) and their difficulty in being operationalised (2);
- Lack of consistency of definitions, criteria, standards and ‘on-the-ground’ tools (3);
- Poor understanding of evaluations, lack of experience/training (2);
- Disaster setting is complex and sophisticated and is the first barrier to undertaking evaluations;
- Accessibility and consistency of data;
- Ethics;
- Short term funding;
- Expectations of donors; and
- Evaluations are not seen as priorities among emergency management/humanitarian workers.
Summated comments on enablers included:

- **Clear guidance for evaluators; if we use the same framework, we can compare outcomes; recognise the generic utility and potential for comparisons that are apparent with a process such as this (3);**
- **Enthusiasm; capacity building, increasing awareness and sharing knowledge are key elements for better application of these evaluations (2);**
- **Early publications of evaluations linked to the framework (2);**
- **Be pro-active, maintain good relationships and build up confidence with the right people at ministries and sponsor organisations; and**
- **The use of a national framework will help promote the various elements and enhance understanding of the need to undertake evaluations in the various settings suggested by the framework.**

The stated aim of the Comprehensive Framework addresses many of these barriers and uses many of these enablers to move forward.

**Discussion**

Of the international experts who responded to the peer reviewed validation survey, there was overwhelming support and endorsement for the concept, structure and usefulness of CFDET 2017. The survey provided a valuable opportunity and independent input to improve the Comprehensive Framework.

**Core Structure**

The Core Structure of CFDET 2017 consisted of three layers: first, a simplified disaster timeline including stages of pre-event, event and post-event; second, a more detailed
sequence of the disaster timeline identifying: pre-event status; hazard; risk reduction; event; damage; disruption, change in functions, and consequences; response; recovery; and post event status; and a third layer, strengthening resilience.¹

Quantitative results on Core Structure met all four consensus thresholds (refer to Table 1). Qualitative responses provided two respondents’ comments influencing modifications to the Core Structure. The first was that ‘rehabilitation’ be included within the recovery section and was accepted. The Sendai Framework for Disaster Risk Reduction 2015 – 2030 (SFDRR) mentions rehabilitation as a major priority (see Priority 4: Enhancing disaster preparedness for effective response).⁴⁵ The SFDRR defines rehabilitation as ‘a set of measures aimed at restoring normal living conditions through the repair and re-establishment of vital services interrupted or degraded by a disaster or emergency’.⁴⁶ Importantly, recovery actions can occur concurrently and are not mutually exclusive.⁴⁷

The second comment recommended the addition of the International Health Regulations (IHR) (2005)⁴⁸ into the Strengthening Resilience section. This was a valid suggestion as one of the underlying themes of CFDET 2017 includes health. The IHR is a legally binding international instrument for the 196 signatory countries, including Member States of the WHO. IHR aims to help international communities prevent and respond to acute public health risks that have the potential to cross borders and threaten the health of people worldwide.⁴⁸,⁴⁹

Two further international consensus frameworks were deemed globally significant and were included: The World Humanitarian Summit (2016) (WHS)⁵⁰ and Habitat III (2016).⁵¹
All guidelines and frameworks included in the Core Structure reflect international significance and their inclusion in the updated Core Structure demonstrates the contemporary nature and living status of the modified Framework. Although not all are legally binding, these international consensus frameworks have shaped policy and practice globally and add rigor and strength to CFDET 2018.

**Baselines**

While the number of responses was slightly lower than preferred (n = 29), quantitative results met the other three consensus thresholds (refer to Table 1). Qualitative comments provided insight that the inclusion of pre-event hazard and risk reduction strategies as baseline evaluations may have caused some confusion. Recent examples of baseline evaluations, including the new Australian Vulnerability Profile, confirm the decision to retain this element in the Comprehensive Framework. Baselines evaluations and baseline assessments remain an integral part of disaster risk reduction as they provide a critical reference point for assessing changes and impact. Furthermore, they establish a basis for comparing the community situation before and after a disaster. The inclusion of the word ‘Assessments’ is reflective of a survey response and intended to provide further clarity within the Baselines section. The term ‘Evaluability’ in Baselines appears to have been misunderstood. Evaluability is a pre-evaluation assessment that the program or intervention intended for evaluation ‘is able to be evaluated’, e.g., is there adequate data and are stakeholders available and willing to participate in the evaluation? However, it is not an evaluation in itself and was removed from CFDET 2017 as a stand-alone entity.

The importance of Common Statistical and Operational Datasets was also mentioned in a responder’s comment. Reflecting this response, changes to Demographics and Infrastructure were made to include ‘Existing Common Statistical and Operational Databases’. The term
databases were preferred and used instead of the term ‘datasets’ to ensure the recognition of broader data sources. Common Operational Datasets (COD) and Fundamental Operational Datasets (FOD) provide important databases in humanitarian emergencies. They can be used to improve the effectiveness of humanitarian assistance by supporting technical standards, improving the quality of the data and strengthening interoperability and harmonization as developed and endorsed by the Inter-Agency Standing Committee (IASC).55-57

A recent environmental scan of the literature identified the expansion of the use of Targets and Indicators in this domain. CFDET 2017 Baselines already included ‘Define Minimum Standards Criteria: Baselines, Targets and Indicators’, however, to reinforce the importance of this typology and in recognition of the evolution of the Targets and Indicators developed and endorsed by the United Nations after the Sendai Conference an opportunity was taken to make an additional change to the Comprehensive Framework. The emergence of Targets and Indicators reflects the maturing of the disaster evaluation approach. Significantly, Targets and Indicators will be used every two years as the framework for all countries to report on their progress towards disaster risk reduction goals within the Sendai Framework. The updated Baseline Typologies now includes explicit Baselines Targets and Indicators, and, the Disaster Risk Reduction Targets and Indicators of the Sendai Framework.

Subsequent to and within the principles of the Sendai Framework, the WHO led a global approach to Health Emergency and Disaster Risk Management Framework (H-EDRM) which complements the Sendai Framework. This evolving WHO Framework includes a Thematic Platform, research agenda and technical guidelines and has been included in the Baseline Typologies.
The final modification made to the Baseline section included the title ‘Baseline Typologies’ as a descriptor, which provides clarification and direction to the items included within.

**Consequences**

Quantitative results on Consequences met all four consensus thresholds (refer to Table 1). Qualitative responses suggested two modifications. The timeframes mentioned in the original Rapid Needs Assessment and Detailed Needs Assessment appeared to be taken literally and generated discussion. The timeframes were intended to be indicative only, and in acknowledgement of the feedback, the term ‘for example’ (e.g.) was added to the timeframes mentioned under Rapid Needs Assessment (Damage), e.g. Day 1 and Detailed Needs Assessment (Functional), e.g. Days 2 – 3. These timeframes are also as per the IASC’s guiding document on Initial Rapid Assessments.

The second modification made to the Consequences section included the title ‘Consequences Typologies’ as a descriptor, which provides clarification and direction to the items included within.

**Outcomes**

Quantitative results on Outcomes met all four consensus thresholds (refer to Table 1). Qualitative responses provided suggested two minor modifications. Firstly, the descriptor ‘Outcome Typologies’ was added to provide further clarification and direction to the items included within ‘Outcomes’. Secondly, ‘Accountability’, which was shown as a stand-alone element in CFDET 2017, was suggested as being better located in the Outcomes section and was thus relocated. The rationale is outlined in detail in the discussion section on Accountability.
Reflecting the contemporary nature of the Comprehensive Framework, the opportunity was taken to include ‘Recovery Indicators’ in the Outcome Typologies section. This new typology has recently emerged and is used in the recovery phase of the disaster timeline. Recent literature suggests an attempt to identify factors that may lead to effective or good recovery and strategies to monitor the community’s progress towards recovery after major events.\textsuperscript{61,62} Further consideration is included later in this discussion section.

\textbf{Impact Evaluations}

Quantitative results on Impact Evaluations only met two of the four consensus thresholds (refer to Table 1). There was an expectation that the inclusion of Impact Evaluations in CFDET 2017 would generate discussion and this proved to be true. Qualitative respondents offered different interpretations and definitions related to Impact Evaluations. Contention and disagreement regarding definitions and terminology was supported by the responses received: ‘Most of the available impact evaluations of humanitarian action are deeply flawed: first, there is the problem of attribution, and second is the problem of mechanism’.

Qualitative responses suggested a number of respondents were unaware of the contemporary nature of Impact Evaluations as evidenced by one exemplar comment: ‘I think causal link assessment is closely linked with inquiry style assessments, in the pre-event stages, this would probably be referred to as a risk assessment.’

Despite ongoing definitional debates,\textsuperscript{63} Impact Evaluations are particularly well suited to answer important questions, such as: whether interventions do or do not work; whether interventions make a positive or negative impact; whether there are intended or unintended consequences; and how cost effective they are.\textsuperscript{64,65} Correctly structured and implemented
Impact Evaluations are expected to greatly improve the effectiveness of interventions delivered in the disaster setting by identifying what works for whom and why.\textsuperscript{66,67}

The trend to undertaking true Impact Evaluations is sound and utilizing the subsequent results will lead to greater accountability. The concept, however, continues to emerge and requires further education of the sector and implementation of ‘true’ impact evaluations, as defined by International Initiative for Impact Evaluation (3ie)\textsuperscript{68} and for these reasons Impact Evaluation remains in the Comprehensive Framework unchanged. The ‘diamonds’, as illustrated in Figures 1 and 2 in the Impact Evaluation section, are interpreted as indicative of Impact Evaluations, which could occur during any phase of the disaster timeline.

\textit{Accountability}

Quantitative results on Accountability met only one of the four consensus thresholds (refer to Table 1). Qualitative responses included 17 comments that provided constructive critique which led to modifications in the Comprehensive Framework. ‘Accountability’ was the most controversial element within CFDET 2017. Respondents’ qualitative comments demonstrated varied and different points of view which reflects current debates surrounding accountability.\textsuperscript{69} Disagreement regarding definitions, terminology and whether Accountability could be included as stand-alone ‘evaluation typology’ was supported by the responses received. Examples of these comments included: ‘there is no portrayal of accountability typologies at all, unless you are referring to the ‘up’, ‘down’, ‘across’, ‘down’ typology of donors, stakeholders and beneficiaries’; and ‘it would be worth linking accountability to both monitoring and evaluation typologies, i.e., you do them for both learning as well as accountability’. These responses had the biggest impact on CFDET 2017.

Accountability and learning are often mentioned as two main goals of evaluation.\textsuperscript{70} Accountability was included in CFDET 2017 because of its significance as a cross-cutting
theme and the overarching concept of effecting better quality and performance in humanitarian action. Since the 1990s, a push to increase quality and accountability exists in the humanitarian sector including all aspects of work/interventions delivered to the communities and people affected \(^{71-73}\) by disasters. There is a moral imperative to ensure accountability is at the forefront of all that is delivered in the disaster setting, whether that be to the affected community, donors and stakeholders, however so defined.

Accountability as a stand-alone element was removed from CFDET 2017, however, the concept was included as an ‘Outcomes’ typology in the updated Comprehensive Framework.

**Evaluation Standards, Evidence and Knowledge Management**

Quantitative results on Evaluation Standards, Evidence and Knowledge Management met two of the four consensus thresholds (refer to Table 1). Qualitative responses were varied and included: ‘I think this is the ideal and what many organisations are striving towards, but I don’t believe it’s been widely adapted’; and ‘ethical consideration should be placed in this domain’.

These results may reflect the current debates related to evaluation and to the way Evaluation Standards, Evidence and Knowledge Management were illustrated within CFDET 2017. Importantly, CFDET 2017 illustrates contemporary disaster evaluation typologies rather than research methods, which classifies its own typologies. Ethical evaluation practice has been retained as a highlighted component of this element.

Minor modifications were made to this element to promote clarity and understanding. The first modification related to the position of the three individual boxes highlighting the independent concepts of Evaluation Standards and Guidelines; Evidence-Based Reviews and
Registries; and Knowledge Management. These three boxes have been moved and included within a single, larger box to indicate their relationships to one another and that these concepts (under the umbrella of Evaluation Standards, Evidence and, Knowledge Management) are applicable across the entire Framework.

The second and third modifications are related to comments regarding ‘Knowledge Management’. ‘Cross-sectoral research’ has been replaced with ‘cross-sectoral collaboration’; and the term ‘translation’ added. A major aim of the updated Comprehensive Framework encourages cross-sectoral collaboration on Knowledge Management and importantly promotes the translation of evaluation findings into improved policy and practice, which may lead to improved outcomes for disaster impacted communities.\textsuperscript{74,75}

Within this element, four new evaluation standards/guidelines were identified additional to those previously published in \textit{Prehospital and Disaster Medicine};\textsuperscript{1} namely:

- Moore et al., \textit{Process Evaluation of Complex Interventions: Medical Research Council Guidance},\textsuperscript{76}
- Christoplos et al., \textit{Strengthening the Quality of Evidence in Humanitarian Evaluations. ALNAP Method Note},\textsuperscript{77}
- Christoplos, I, Dillon, N and Bonino F, \textit{Evaluation of Protection in Humanitarian Action. ALNAP Method Note},\textsuperscript{78} and
- Blanchet et al, \textit{Using Research Evidence in the Humanitarian Sector: A practice guide}.'\textsuperscript{79}
**Targets and Indicators as a Cross-Cutting Theme**

As previously mentioned, targets and indicators, and their measurement, represent an emerging maturity in the sector. A Baseline Typology, ‘preparedness surveys’, included in CFDET 2017 now demonstrates more sophisticated, update exemplars: nations surveys in Canada, USA and a state survey in Queensland, Australia. This typology was included in CFDET 2017 and no changes were made, other than to identify these recent exemplars in the text of the discussion and add them to the updated bibliography.

‘Strengthening Resilience’ in the CFDET 2017 Core Structure included the ‘Rockefeller 100RC Framework’. Evolving from this international framework has been the ‘City Resilience Index’ as a companion typology designed to measure resilience within the Rockefeller City Resilience setting and was added to the Framework. Although influential in guiding the Rockefeller 100 Resilient Cities project, it may be too early to determine its use or its influence on resilience in those cities. In Australia, an Australian Natural Disaster Resilience Index is also evolving, but again it is too early to determine its use or its influence on resilience in Australia.

The ‘Recovery Indicators’ typology emerged recently for use in the recovery phase of the disaster timelines and has been added to the ‘Outcome Typologies’. The recent literature suggests an attempt to identify factors that may lead to effective or good recovery and strategies to monitor the communities progress towards recovery after major events. This approach evolved into developing a validating indicators of
disaster recovery. In the USA, Horney led an extensive process to validate key indicators for disaster recovery. In the Australian setting, a set of outcomes, indicators and standards of successful recovery are referenced from the Australian Monitoring and Evaluation Framework for Disaster Recovery Programs. These shape the future direction of data collection and measurement in this sector.

**Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018)**

The outcomes of the validation process of CFDET 2017 provided opportunities to improve the Framework through its readability, understanding and usability. Modifications were made based on feedback obtained and the resulting Comprehensive Framework (2018) is shown in Figure 2.
Figure 2: Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018)

Strong acceptance and support for CFDET 2017 as a unifying framework demonstrated it cohesively represented the relationships between the different evaluation typologies.

Limitations

Important lessons were learnt while implementing the online survey which will be valuable for future surveys. A potential limitation of this validation survey included: the difficulty in attracting an acceptable number of participants which became very time consuming and prolonged the duration of the survey; the electronic format of the survey and the quality of the technology that was used by different participants which restricted their responses; the estimated time to complete the survey (30 – 45 minutes); and the fact that the accompanying paper Disaster Metrics: A Comprehensive Framework for Disaster Evaluation Typologies
describing the development of CFDET 2017\textsuperscript{1} was in press and not actually published at the time of the survey. The potential for self-selection bias exists given the high non-response rate. Despite this, saturation of data and meeting the consensus thresholds of most elements in the Framework were evident in the contributor responses which is testament to the validity of the results. The number of responders is consistent with acceptable levels for similar research studies.

**Future Actions**

Given that the Comprehensive Framework has now been validated and improved, the direction for future actions include:

1. Preparing an interpretive guide to accompany CFDET 2018;
2. The implementation of toolkits to support disaster evaluation typologies, for example ‘How to Guides’, of which the *World Association for Disaster and Emergency Medicine (WADEM) Research and Evaluation Frameworks*\textsuperscript{10-18} and Evidence Aid’s *Use of evidence in the humanitarian sector: A practice guide*\textsuperscript{79} are leading examples;
3. Collaboration with international organisations, such as WADEM and Evidence Aid to schedule promotion, training and implementation of CFDET 2018;
4. Facilitate, support and publish an initial suite of disaster evaluations using the Comprehensive Framework to structure the evaluation methodology;
5. Evaluating the link between CFDET 2018 and the significant international disaster risk reduction frameworks identified in the Core Structure; and
6. Monitoring and evaluating the use and usefulness of CFDET 2018.

**Conclusion**

The objective of the international peer-reviewed validation process of CFDET 2017 was to seek acceptance and feedback from experts in the fields of disasters and emergencies;
humanitarian and development; and, evaluation. This objective has been achieved. For the first time, a Framework of this nature in the disaster setting has undergone a rigorous validation process at the international level and is supported by experts in three related disciplines. Respondent feedback obtained during the validation process improved the Comprehensive Framework. Modifications have enhanced the Framework and added strength to the Framework’s elements, understanding and usability. The updated CFDET 2018 provides a unique, unifying framework within which existing typologies can be structured. It gives evaluators confidence to choose an appropriate strategy for their particular evaluation in the disaster setting. Additionally, it facilitates consistency in reporting across the different phases of a disaster to better understand the process, outcomes and impacts of the efficacy and efficiency of interventions. It is expected that CFDET 2018 will improve policy and practice associated with the delivery of health interventions in the disaster setting by strengthening the evidence-base for interventions delivered across the disaster spectrum. Future research could create a series of toolkits to support improved disaster evaluation processes and to evaluate the utility of the Framework in the real-world setting.

Acknowledgements

The validation process would not have been possible without the input from all those who responded to the survey. Your contribution is acknowledged and appreciated and will help advance disaster evaluations. A special thanks to Jackie Van Dam and Dr Sarah Wong for their assistance.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3ie</td>
<td>International Initiative for Impact Evaluation</td>
</tr>
<tr>
<td>CFDET 2017</td>
<td>Comprehensive Framework for Disaster Emergency Typologies 2017</td>
</tr>
<tr>
<td>CFDET 2018</td>
<td>Comprehensive Framework for Disaster Emergency Typologies 2018</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>COD</td>
<td>Common Operational Datasets</td>
</tr>
<tr>
<td>FOD</td>
<td>Fundamental Operational Datasets</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information Systems</td>
</tr>
<tr>
<td>GFDRR</td>
<td>Global Facility for Disaster Reduction and Recovery</td>
</tr>
<tr>
<td>Habitat III</td>
<td>Urban Conference on Housing and Sustainable Urban Development</td>
</tr>
<tr>
<td>H-EDRM</td>
<td>Health Emergency and Disaster Risk Management Framework</td>
</tr>
<tr>
<td>HIA</td>
<td>Health Impact Assessment</td>
</tr>
<tr>
<td>IASC</td>
<td>Inter-Agency Standing Committee</td>
</tr>
<tr>
<td>IHR</td>
<td>International Health Regulations</td>
</tr>
<tr>
<td>PDNA</td>
<td>Post Disaster Needs Assessment</td>
</tr>
<tr>
<td>PDM</td>
<td>Prehospital and Disaster Medicine</td>
</tr>
<tr>
<td>SFDRR</td>
<td>Sendai Framework for Disaster Risk Reduction 2015 – 2030</td>
</tr>
<tr>
<td>UNISDR</td>
<td>United Nations Office for Disaster Risk Reduction</td>
</tr>
<tr>
<td>WADEM</td>
<td>World Association for Disaster and Emergency Medicine</td>
</tr>
<tr>
<td>WCDEM</td>
<td>World Congress on Disaster and Emergency Medicine</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>WHS</td>
<td>World Humanitarian Summit</td>
</tr>
</tbody>
</table>
## Tables

### Table 1: Elements and Relationships Represented Within CFDET 2017

<table>
<thead>
<tr>
<th>ELEMENTS</th>
<th>N</th>
<th>LIKERT SCORE</th>
<th>MEAN +/- 95% CI</th>
<th>MEDIAN (interquartile range)</th>
<th>ACCEPTANCE THRESHOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Structure</td>
<td>32</td>
<td>0 0 1 19 12</td>
<td>4.34 (4.15 – 4.54)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>96.88</td>
</tr>
<tr>
<td>Baselines</td>
<td>33</td>
<td>1 3 4 9 16</td>
<td>4.09 (3.69 – 5.00)</td>
<td>4.00 (3.50 – 5.00)</td>
<td>75.76</td>
</tr>
<tr>
<td>Consequences</td>
<td>32</td>
<td>0 1 3 12 16</td>
<td>4.34 (4.06 – 4.63)</td>
<td>4.50 (4.00 – 5.00)</td>
<td>87.50</td>
</tr>
<tr>
<td>Outcomes</td>
<td>32</td>
<td>0 2 3 16 11</td>
<td>4.13 (3.82 – 4.43)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>84.38</td>
</tr>
<tr>
<td>Impact</td>
<td>30</td>
<td>0 4 7 10 9</td>
<td>3.80 (3.42 – 4.18)</td>
<td>4.00 (3.00 – 5.00)</td>
<td>63.33</td>
</tr>
<tr>
<td>Accountability</td>
<td>29</td>
<td>0 5 5 6 13</td>
<td>3.93 (3.49 – 4.37)</td>
<td>4.00 (3.00 – 5.00)</td>
<td>65.52</td>
</tr>
<tr>
<td>Standards and Evidence</td>
<td>29</td>
<td>2 2 1 14 10</td>
<td>3.97 (3.57 – 4.40)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>82.76</td>
</tr>
</tbody>
</table>

### Table 2: Holistic Representation as a Unifying Framework

<table>
<thead>
<tr>
<th>CFDET 2017 AS A UNIFYING FRAMEWORK</th>
<th>N</th>
<th>LIKERT SCORE</th>
<th>MEAN +/- 95% CI</th>
<th>MEDIAN (interquartile range)</th>
<th>ACCEPTANCE THRESHOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationships</td>
<td>30</td>
<td>0 2 2 17 9</td>
<td>4.10 (3.80 – 4.40)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>86.67</td>
</tr>
<tr>
<td>Value</td>
<td>29</td>
<td>0 1 3 12 13</td>
<td>4.28 (3.97 – 4.58)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>86.21</td>
</tr>
<tr>
<td>Support Evaluations</td>
<td>29</td>
<td>0 1 4 10 14</td>
<td>4.28 (3.96 – 4.60)</td>
<td>4.00 (4.00 – 5.00)</td>
<td>82.76</td>
</tr>
<tr>
<td>Teaching</td>
<td>30</td>
<td>0 1 3 11 15</td>
<td>4.33 (4.03 – 4.63)</td>
<td>4.50 (4.00 – 5.00)</td>
<td>86.67</td>
</tr>
</tbody>
</table>
Figures

Figure 1: Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017)
Figure 2: Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018)

Colour (online version)

Figure 1: Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017)

Figure 2: Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018)

Black and white (printed version)

Figure 1: Comprehensive Framework for Disaster Evaluation Typologies 2017 (CFDET 2017)

Figure 2: Comprehensive Framework for Disaster Evaluation Typologies 2018 (CFDET 2018)
Reference List – Endnote Off


55. (IASC) IASC. IASC Guidelines: Common Operational Datasets (CODs) in Disaster Preparedness and Response. Inter Agency Standing Committee (IASC); 2010.

56. Handbook UE. Common operational datasets (CODs) and fundamental operational datasets (FODs) New York: The UN Refugee Agency (UNHCR); 2018 [Available from: https://emergency.unhcr.org/entry/50307.


71. Core Humanitarian Standard on Quality and Accountability. HAP International People in Aid Sphere Project; 2014.


References


88. Wells T. There is no such thing as a natural disaster anymore: Why wilful ignorance is not innocence. ABC Religion and Ethics, 3 October 2017: ABC; 2017.


191. DiMaggio C. A brief introduction to disaster epidemiology2011 1 March, 2018:[6 p.].


208. Cuny FC, Disaster Management Center at the University of Wisconsin-Madison. Fred C. Cuny Memorial Continuing Education Series. Principles of disaster management.


224. Debacker M. Data reporting in disasters. 2nd Evidence Aid Conference, Prioritising Evidence in Disaster Aid; 29-30 October 2012; Brussels: Evidence Aid; 2012.


229. Debacker M, editor Utstein-style template for uniform reporting of medical response in disasters and health crises. 17th World Congress on Disaster and Emergency Medicine; 2011; Beijing, China: Prehosp Disaster Med.


239. Sundnes KO. Guidelines for evaluation and research of health disaster management in the Utstein style - Background and principles. 13th World Congress for


United Nations International Drug Control Programme (UNDCP),
European Monitoring Center on Drugs and Drug Addiction (EMCDDA); 2000.
evaluation within the framework of a monitoring and evaluation system. Washington, D.
306. Buttenheim A. Impact evaluation in the post-disaster setting: A conceptual
discussion in the context of the 2005 Pakistan earthquake. New Delhi: International
Initiative for Impact Evaluation (3ie); 2009.
307. Buttenheim A. Impact evaluation in the post-disaster setting: A case study of the
308. Chambers R, Karlan D, Ravallion M, Rogers PJ. Designing impact evaluations:
Different perspectives. New Delhi: International Initiative for Impact Evaluation (3ie);
2011.
Delhi: International Initiative for Impact Evaluation (3ie); 2012. Contract No.: Version
No. 7.
310. Puri J, Aladysheva A, Iversen V, Ghorpade Y, Bruck T. What methods may be
used in impact evaluations of humanitarian assistance?2015. Available from:
311. Rogers PJ. Introduction to impact evaluation. Melbourne: InterAction,
312. White H. Some reflections on current debates in impact evaluation. New Delhi:
International Initiative for Impact Evaluation (3ie); 2009.
313. Adibhatla S, Dudek O, Ramsel-Miller J, Birnbaum M. Classification of disaster
health publications. In: Presented at 19th World Congress on Disaster and Emergency
editor. 19th World Congress on Disaster and Emergency Medicine (WCDEM); April
2015; Cape Town: Prehospital and Disaster Medicine 2015. p. s111.
314. Birnbaum ML, Adibhatla S, Dudek O, Ramsel-Miller J. Categorization and
analysis of disaster health publications: An inventory. Prehosp Disaster Med.
315. International Federation of Red Cross and Red Crescent Societies. Evaluations
Geneva: International Federation of Red Cross and Red Crescent Societies (IFRC); 2018
316. United Nations International Children's Emergency Fund. Evaluation Database:
United Nations International Children's Emergency Fund (UNICEF); 2018 [Available
from: https://www.unicef.org/evaldatabase/.
317. Active Learning Network for Accountability and Performance in Humanitarian
Action. ALNAP London: Overseas Development Institute (ODI); 2016 [Available from:
http://www.alnap.org/who-we-are/our-role.
318. Inter-Agency Standing Committee. Inter-Agency Standing Committee (IASC)
Geneva: Inter-Agency Standing Committee (IASC); 2016 [Available from:
https://interagencystandingcommittee.org/.

436


349. Worthington AC. Natural disasters and national natural disaster insurance: An Australian perspective. Melbourne: Australia Centre for Financial Services (ACFS); 2015.


Overseas Development Institute (ODI); 2015.


414. Smith N. There’s no such thing as a natural disaster New York: Social Science Research Council (SSRC); 2006 [Available from: http://understandingkatrina.ssrc.org/Smith/.


419. Basabe P. Comparing and contrasting approaches to risk governance. Lausanne: International Risk Governance Council (IRGC); 2016.


424. Association of Southeast Asian Nations. ASEAN disaster recovery reference guide. Jakarta: Association of Southeast Asian Nations (ASEAN); 2016.


449. ARUP. City Resilience Index. London: The Rockefeller Foundation
450. ARUP. City Resilience Framework. London: The Rockefeller Foundation
ARUP; 2015.


458. Brown LD, Moore MH. Accountability, strategy, and international non-governmental organizations. Cambridge, MA: The Hauser Center for Nonprofit Organizations,
The Kennedy School of Government,


489. Joanne Briggs Institute, University of Adelaide. The JBI database of systematic reviews and implementation reports Adelaide: Joanne Briggs Institute University of Adelaide; 2016 [Available from: https://journals.lww.com/jbisrir/Pages/default.aspx.


Harvard School of Public Health,

Overseas Development Institute (ODI)


545. Metcalfe J, Poon L. Everything you need to nnow about the momentous Habitat
urban-agenda/503927/.

down assessment of disaster resilience: A conceptual framework using coping and

547. Aitsi-Selmi A, Blanchard K, Murray V. London: Evidence Aid. 2015 13 October
development-goals-and-the-sendai-framework-for-disaster-risk-reduction-a-year-for-
policy-coherence/.

development knowledge platform New York: United Nations; 2018 [Available from:

549. UN Women. What is baseline assessment? New York: United Nations Entity for
Gender Equality and the Empowerment of Women; 2012 [Available from:
http://www.endvawnow.org/en/articles/1323-what-is-a-baseline-
assessment.html?next=1324

550. Inter Agency Standing Committee. IASC guidelines: Common Operational
Datasets (CODs) in disaster preparedness and response. Geneva: Inter Agency Standing
Committee (IASC); 2010.

2018 [Available from: https://www.humanitarianresponse.info/en/applications/tools/category/operational-
datasets.

552. United Nations Office for the Coordination of Humanitarian Affairs. OCHA home
page: United Nations Office for the Coordination of Humanitarian Affairs (OCHA); 2018

553. Australia: Sendai Framework data readiness review - report - Australia. Geneva:
United Nations Office for Disaster Risk Reduction (UNISDR); 2018. Contract No.: 53161.

554. Inter-Agency Standing Committee. Initial Rapid Assessment (IRA): Field
assessment form Geneva: World Health Organization (WHO); 2010 [Available from:

555. Dwyer C, Horney J. Validating indicators of disaster recovery with qualitative

556. Network of Networks for Impact Evaluation. Impact evaluations and
development. NONIE guidance on impact evaluation. Washington: The Network of
Networks on Impact Evaluation (NONIE); 2009.

557. Hilhorst D. Being good at doing good? Quality and accountability of humanitarian

558. Beck T. Evaluating humanitarian action using the OECD-DAC criteria. An
ALNAP guide for humanitarian agencies. London: Active Learning Network for
Accountability and Performance in Humanitarian Action (ALNAP); 2006.
