Objective: This literature review aimed to identify the range of methods used in after action reviews (AARs) of public health emergencies and to develop appraisal tools to compare methodological reporting and validity standards.

Methods: A review of biomedical and gray literature identified key approaches from AAR methodological research, real-world AARs, and AAR reporting templates. We developed a 50-item tool to systematically document AAR methodological reporting and a linked 11-item summary tool to document validity. Both tools were used sequentially to appraise the literature included in this study.

Results: This review included 24 highly diverse papers, reflecting the lack of a standardized approach. We observed significant divergence between the standards described in AAR and qualitative research literature, and real-world AAR practice. The lack of reporting of basic methods to ensure validity increases doubt about the methodological basis of an individual AAR and the validity of its conclusions.

Conclusions: The main limitations in current AAR methodology and reporting standards may be addressed through our 11 validity-enhancing recommendations. A minimum reporting standard for AARs could help ensure that findings are valid and clear for others to learn from. A registry of AARs, based on a common reporting structure, may further facilitate shared learning. (Disaster Med Public Health Preparedness. 2018;page 1 of 8)

Key Words: after action reviews, emergency preparedness, epidemics, outbreaks, public health

Public health emergencies, such as infectious disease outbreaks, floods, and terrorist attacks, impact societies severely but are relatively rare for individual countries. However, this national rarity provides an impetus to systematically learn from emergencies when they do occur, so as to strengthen public health emergency preparedness and response planning.

One such learning approach is to conduct an after action review (AAR), or a lessons learned document. These documents are completed after a public health emergency has occurred and draw on quantitative and qualitative methods to identify strengths and weaknesses in the public health emergency preparedness system. By addressing any weaknesses identified, they aim to improve preparedness, response, and recovery capacities and capabilities, ultimately lessening the impact of future incidents.

Typically, documentation and other quantitative fact-finding methods help establish a skeleton timeline of events, whereas different forms of qualitative investigation, such as personal testimony, provide richer insights into how and why events unfolded. Combined, these approaches aim to establish the root causes of the event and to identify what lessons can be learned for the future.

Despite the crucial role of AARs in linking the past with the present and future, there is no widely used or standardized approach to conducting AARs of public health emergencies. Particularly, there is no indication of whether insights gained are valid or based on robust methodologies.

This literature review aimed to identify the range of methods used to produce AARs to improve emergency preparedness planning and to develop appraisal tools to compare their methodological reporting and validity standards, with a focus on qualitative methods.

METHODS

Literature Search
We searched biomedical databases (Medline, Embase, Scopus) and gray literature sources (Google Advanced, Google Scholar) for AARs that described an enacted
response to an emergency (theoretical or “table-top” exercises were excluded), were within the geographic scope of the literature review (the European Union, Australia, Canada, New Zealand, and the United States), and were published in English from January 2000 to August 2015.

Search strategies were structured around 2 major concepts: AARs and emergency preparedness. Searches combined free text and thesaurus terms (where available), including synonyms such as “post-event analysis” and “critical incident review” and techniques used within AARs such as “facilitated look back” and “root-cause analysis” (Supplemental Information [SI] 1). Additional search terms and synonyms were identified by scanning the abstracts of articles identified through a scoping search. Additional AARs were identified by searching the Endnote Library for a previous review undertaken for the European Centre for Disease Prevention and Control (ECDC), looking for evaluations of emergency response.10,11

Reviews were sifted for relevance first on title and abstract and then on full-text review (Figure 1, PRISMA diagram). Studies excluded at the full-text stage can be found in SI-2.

Development of Appraisal Tools
We developed 2 appraisal tools to systematically document the methods used in AARs, to compare methodological reporting and validity between diverse AARs, and to act as a benchmark of theoretical best practice.

We adapted the approach of Woloshynowych12 – which related to the analysis of after actions in health care – to an emergency public health context by triangulating it with 9 contemporary AAR templates.5,13–20 The templates were identified through targeted scoping searches in Google, using synonyms for AARs and templates. These templates were multi-sectorial, coming from after action reports, a significant event analysis, and peer assessments in the fields of US national defense,14 US state government,13 UK medicolegal,17 Canadian health care insurance,20 international emergency public health,5,16 a UK hospital,15 and patient safety agencies (See SI-2).18,19 Further tool modifications were made in consultation with an expert advisor to increase its relevance to emergency public health. This resulted in a 50-item appraisal tool (SI-3).

Adapting the approach of Piltch-Loeb,5 we developed an additional 11-point summary tool of factors that boosted methodological rigor in case study and qualitative data collection and analysis.

The original Piltch-Loeb 10-point tool remained intact with minor revisions in definitions to better reflect the context of AARs in emergency public health. We added an 11th factor to capture whether the AAR had ultimately achieved its aim.
of uncovering the root causes of preparedness, response and recovery activities, rather than more superficial causes. Definitions of the 11 points are included in SI-4.

**Appraising the After Action Reviews**

The 50-item appraisal tool (SI-3) and 11-item summary measure (SI-4) were applied sequentially to each AAR. First, the 50-item tool was used to systematically document the methods undertaken by each AAR, before being summarized in the 11-item measure, allowing for a simpler comparison of methodology and validity across diverse reviews.

AARs were reviewed against each item on the summary validity tool and assigned one of 3 codes. Fully met (+ +): These criteria have been fully and often comprehensively met, and we have little doubt that these criteria have been met. Partially met (+): The criteria have been met in some regards, but there is significant doubt about the comprehensiveness or there are clear elements missing, preventing a higher rating. Not met (-): These criteria are not met or have not been reported.

A sample of 3 AARs was independently coded by a second reviewer to test the reliability of the coding instrument and to clarify initial rating definitions. The second rater was blind to the first rater’s scores and rationales. Given the size of the sample, inter-coder agreement was not calculated. Differences between the 2 raters were discussed and changes agreed by consensus. This led to revisions in the wording of some criteria and scoring guidance to improve clarity and therefore scoring consistency. Definitions of the criteria and additional notes used to guide rating decisions are described in SI-3.

**RESULTS**

**Overview**

Our search identified 24 published AAR documents, relating to 22 distinct AARs (Table 1).

The reviews covered national and international responses to the 2009 A(H1N1) influenza pandemic (n = 8),21–28 terrorist bombing incidents (n = 5),29–33 industrial explosions (n = 6),34–39 hurricanes (n = 2),40,41 chemical contamination of drinking water (n = 1),42 a heat wave (n = 1),43 and large-scale flooding (n = 1) (see Table 1).44

**Appraisal of After Action Reviews**

There was great diversity in the structure, scope, and level of methodological reporting in the 24 reviews identified, potentially reflecting a lack of a standardized approach (Table 2).21–44 The majority drew heavily on qualitative research best practice.

Validity boosting measures most frequently reported in the 24 reviews included spending adequate time to observe the setting, people, and incident documentation; sampling a diverse range of views; using multiple sources of data collection; and utilizing multiple perspectives during the analysis.21–44 However, these techniques were generally reported in brief, with few reviews fully meeting all 4 basic validity dimensions.

The criteria that were most commonly unmet in these reports were acknowledging a theoretical basis for the review methodology; describing how the reviewers handled discordant evidence; having an external peer-review process; and ensuring respondents to the reviews had an opportunity to validate that their views had been reflected accurately in the final analysis and report (see Table 2).

The majority of AARs showing depth and insight (9 fully met this validity measure) also clearly reported using multiple data sources (7 of 9) and sustained engagement (5 of 9). Other AARs demonstrated depth and insight without reporting clear methods (see Table 2).29,34,35,44

**Suggestions**

Based on the systematic assessment of methods and validity measures in 24 AARs, we suggest 11 measures to improve the reporting and validity of reviews more widely (Table 3).

**DISCUSSION**

To our knowledge, this is the first review to systematically document methods used in public health emergency preparedness AARs across a range of hazards and to formulate suggestions to improve future practice based on principles of qualitative research best practice.

The strengths of this review include our inclusive definition of an AAR, our inclusion of non-health-care specific after actions and reporting templates, and the development of tools rooted in after action methodological research. These tools were applied to a variety of real-world AARs in the field of emergency preparedness spanning multiple hazard types.

The most common data collection methods used by the 24 AARs were document review (typically preparedness plans and protocols compared to execution), focus groups, formal public consultations, in-depth interviews, public discussion forums, questionnaires, site visits, and workshops.

Most reviews (17 of 24) did not report a theoretical framework to guide investigation; of those that did, all reported a comparative or case study methodology. This represents a small fraction of the diverse range of approaches available to after action investigators, including the after action technique,4,8 after action analysis,7,45 root-cause analysis,46–48 facilitated look-backs,49 peer assessment approach,6 realist evaluation,5,9 bow-tie analysis,39 and serious case reviews.50
### TABLE 1

**Summary of 22 Included AARs**

<table>
<thead>
<tr>
<th>Authors (year)</th>
<th>Hazard</th>
<th>Geography</th>
<th>Description</th>
<th>Method(s)</th>
<th>Theoretical Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socialstyrelsen (Swedish National Board of Health and Welfare (2011)</td>
<td>Biological</td>
<td>Sweden</td>
<td>H1N1 flu pandemic: Swedish response</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>WHO (2010)</td>
<td>Biological</td>
<td>Global</td>
<td>H1N1 flu pandemic: global survey on national vaccine deployment and vaccination plans</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>European Commission (2010)</td>
<td>Biological</td>
<td>Europe</td>
<td>H1N1 flu pandemic: EU member states pandemic vaccine strategies and implementation</td>
<td>Questionnaires</td>
<td>Comparative methodology</td>
</tr>
<tr>
<td>Health Protection Agency (2010)</td>
<td>Biological</td>
<td>Europe</td>
<td>H1N1 flu pandemic: individual member state and EU-wide responses</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>WHO (2010)</td>
<td>Biological</td>
<td>Europe</td>
<td>H1N1 flu pandemic: WHO Europe evaluation of the response to pandemic</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>Health and Safety Executive (2008)</td>
<td>Chemical</td>
<td>England</td>
<td>Industrial explosion/fire at Buncefield fuel depot</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>Tapster (2007)</td>
<td>Chemical</td>
<td>England</td>
<td>Industrial explosion/fire at Buncefield fuel depot</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>Paltrinieri N, Dechy N, Salzano E, et al. (2012)</td>
<td>Chemical</td>
<td>UK and France</td>
<td>Industrial explosion/fire at Buncefield (UK) fuel depot and Toulouse (France) fertilizer factory</td>
<td>Document review</td>
<td>Bow-tie analysis*</td>
</tr>
<tr>
<td>Terenzini C. (2007)</td>
<td>Chemical</td>
<td>USA</td>
<td>Chemical contamination of a public water supply (Massachusetts, USA)</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>Massachusetts Emergency Management Agency (2014)</td>
<td>Physical</td>
<td>USA</td>
<td>Two bombs at the 2013 Boston marathon</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>Goralnick E, Halpern P, Loo S, et al. (2015)</td>
<td>Physical</td>
<td>USA</td>
<td>Two bombs at the 2013 Boston marathon (focus on leadership)</td>
<td>Interviews</td>
<td>Not reported</td>
</tr>
<tr>
<td>Little M, Cooper J, Gope M, et al. (2012)</td>
<td>Physical</td>
<td>Bali and Australia</td>
<td>Royal Perth Hospital Emergency Department response to Bali bombing and refugee ship fire</td>
<td>Document review</td>
<td>Comparative case study</td>
</tr>
<tr>
<td>Knox CC. (2013)</td>
<td>Physical</td>
<td>USA</td>
<td>Hurricane Andrew (1992) and Katrina (2005); comparison of AARs and recommendations</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>Brevard SB, Weintraub SL, Aiken JB, et al. (2008)</td>
<td>Physical</td>
<td>USA</td>
<td>Hurricane Katrina flood and full evacuation of a Level 1 Trauma Center</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
<tr>
<td>Oklahoma Department of Civil Emergency Management (1995)</td>
<td>Physical</td>
<td>USA</td>
<td>Truck bomb at Oklahoma City Federal Building</td>
<td>Document review</td>
<td>Not reported</td>
</tr>
</tbody>
</table>

*A risk evaluation method that can be used to analyze and demonstrate causal relationships in high risk scenarios.*
### TABLE 2

Summary Validity Measures Reporting for 22 AARs (Including 2 Annexes Appraised Alongside the AAR)\textsuperscript{21–44}

<table>
<thead>
<tr>
<th>After Action Review</th>
<th>Sustained Engagement</th>
<th>Validation of Conclusions</th>
<th>Selection of Respondents</th>
<th>Multiple Data Sources</th>
<th>Multiple Observers</th>
<th>Case Selection</th>
<th>Theoretical Construct</th>
<th>Validation by Respondents</th>
<th>External Peer Review</th>
<th>Discordant Evidence</th>
<th>Depth and Insight</th>
<th>Overall Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flood, England\textsuperscript{44}</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>17</td>
</tr>
<tr>
<td>H1N1, Canada\textsuperscript{21}</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>16</td>
</tr>
<tr>
<td>Heat wave, France\textsuperscript{43}</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>15</td>
</tr>
<tr>
<td>Boston bomb, leadership, US\textsuperscript{30}</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>H1N1, vaccines, global\textsuperscript{25}</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Water contamination, US\textsuperscript{42}</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>H1N1, Europe\textsuperscript{28}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Buncefield explosion, England\textsuperscript{34, 35}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Boston marathon bomb, US\textsuperscript{29}</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Hurricane Katrina hospital evacuation, US\textsuperscript{24}</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>H1N1, vaccine strategy, EU\textsuperscript{26}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Chemical blasts, France and England\textsuperscript{39}</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1N1, Sweden\textsuperscript{23}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>H1N1, Norway\textsuperscript{22}</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Burn disasters, Australia\textsuperscript{31}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Hurricanes Andrew and Katrina, US\textsuperscript{40}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>H1N1, UK\textsuperscript{24}</td>
<td>+</td>
<td>-</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Buncefield, social impact assessment, England\textsuperscript{28, 37}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>H1N1, EU-wide\textsuperscript{27}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>London bombing, England\textsuperscript{32}</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Oklahoma bombing, US\textsuperscript{33}</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Buncefield multiagency review, England\textsuperscript{38}</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

\*Overall validity score based on the following scoring: (++) = 2; (+) = 1; (-) = 0.
Underlying methodologies were frequently unreported, so the report validity remained ambiguous. Although a lack of reporting of basic methods to safeguard validity does not necessarily imply that they were not considered or followed, it does significantly increase doubt surrounding the methodological basis of the review and the validity of its conclusions.

**Limitations**
Our review searched for reports from a diverse range of after actions, but the analyzed sample was small (n = 24) and subject to reporting and selection bias, and may not represent the full spectrum of incident reports available. For example, we excluded 16 studies with insufficient methods for analysis (see SI-2: Excluded Studies) and all reviews not published in English.

Three of the 24 included reviews were used to test and develop early versions of both appraisal tools before their final application to the remaining 21 reports, further reducing the number of independent reviews appraised.

Most AAR reports were not clear on how their data analysis led to generalizable insights by reviewers or how discordant
information was handled.\(^{22,28,29}\) As such, it was not clear to what extent certain views or data had been explored or discounted, for example, if they did not fit with the emerging researcher consensus. This risked introducing perception bias into the analysis and conclusions drawn.

CONCLUSIONS

We suggest that the lack of methodological reporting provides a strong case for the development of evidence-based minimum reporting standard for AARs, akin to the CONSORT statement for randomized controlled trials. These standards could benefit after action reports in 2 ways. First, they may ensure that a wider range of robust methods is considered before and during the review, and, second, that methods are more clearly reported in the end report itself, allowing an external assessment of validity. The 11-point summary tool presented here allows a simple validity comparison to be made across a range of diverse AARs, which could be further developed and refined in the future.

It is noteworthy that critical incident registries have been adopted in transport, health care, and workplace safety industries, but not in emergency preparedness.\(^ {5}\) We thus advocate an AAR registry (similar in nature to the US government’s Lessons Learned Information Sharing program) in Europe, to facilitate cross-border learning that will further strengthen emergency preparedness.\(^ {51}\) The 11-point summary validity tool presented here could contribute to such an initiative by promoting an AAR design that is as robust and credible as possible.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit [https://10.1017/dmp.2018.82](https://10.1017/dmp.2018.82)

About the Authors

Public Health Directorate, St Helens and Knowsley Teaching Hospitals NHS Trust, West Midlands, England (Mr Davies), Bazian Ltd, An Economist Intelligence Unit Business, London, England (Mrs Vaughan, Dr Cook), and Country Preparedness Section, European Centre for Disease Prevention and Control, Stockholm, Sweden (Dr Fraser, Ciotti, Suk).

Correspondence and reprint requests to Dr Jonathan E. Suk, European Centre for Disease Prevention and Control (ECDC), Gustav III:s Boulevard 40, 16973 Solna, Sweden (e-mail: Jonathan.Suk@ecdc.europa.eu).

Acknowledgment and Author Contributions

This publication is based upon a report produced by Bazian Ltd and commissioned by the ECDC under Direct Service Contract ECD.5860. Robert Davies provided input into project design, performed data extraction, performed data synthesis, and coauthored this manuscript; Elly Vaughan managed the project at Bazian, provided input into project design, designed and ran literature searches, performed data extraction, contributed to the synthesis, and coauthored this manuscript; Dr Robert Cook provided input into project design, reviewed draft reports, and provided project oversight; Dr Graham Fraser, Dr Massimo Ciotti, and Dr Jonathan Suk initiated the study and commissioned the work, provided technical guidance throughout the study, and coauthored this manuscript; Dr Katie Geary provided expert advice throughout the project design and execution, including refining the appraisal tools for a public health emergency context.

REFERENCES

After Action Reviews of Public Health Emergencies


