TECHNICAL REPORT

Best practice recommendations for conducting after-action reviews to enhance public health preparedness

www.ecdc.europa.eu

ECDC TECHNICAL REPORT

Best practice recommendations for conducting after-action reviews to enhance public health preparedness

This report was commissioned by the European Centre for Disease Prevention and Control (ECDC), coordinated by Jonathan Suk, Graham Fraser, and Massimo Ciotti. The literature review described in this report was produced by Robert Davies and Elly Vaughn, Bazian Ltd., for ECDC.

This report expands upon elements of the following paper:
Davies, R., Vaughan, E., Fraser, G., Cook, R., Ciotti, M., & Suk, J. (2018). Enhancing Reporting of After-action Reviews of Public Health Emergencies to Strengthen Preparedness: A Literature Review and Methodology Appraisal. Disaster Medicine and Public Health Preparedness, 1-8. [51]

Suggested citation: European Centre for Disease Prevention and Control. Best practice recommendations for conducting after-action reviews to enhance public health preparedness. Stockholm: ECDC; 2018.

Stockholm, November 2018

ISBN 978-92-9498-293-3
doi: 10.2900/467275
Catalogue number TQ-03-18-524-EN-N

© European Centre for Disease Prevention and Control, 2018

Cover picture: © iStock

Reproduction is authorised, provided the source is acknowledged.

For any use or reproduction of photos or other material that is not under the EU copyright, permission must be sought directly from the copyright holders.

Contents

Figures

Tables

Abbreviations

AAR	after-action reviews
PHEP	public health emergency preparedness
ECDC	European Centre for Disease Prevention and Control
WHO	World Health Organization

Executive summary

Background

The threat of serious, cross-border infectious disease outbreaks in Europe is a significant challenge in terms of public health emergency preparedness (PHEP). An integral part of improving preparedness and response planning is learning from past experience. Conducting after-action reviews (AARs) on public health emergencies is one such approach, but there is no standardised approach to conducting such assessments. This report identifies the common features of AARs, develops a validity assessment tool for appraising AARs, and proposes a set of best-practice recommendations.

Methods

A focussed scoping review was conducted to identify the key common features of AARs and an 11-item tool was then developed to appraise and rank the validity of AAR approaches. The tool was applied to 24 papers describing 22 AARs. Based on this review, a best-practice framework was developed.

Results

AARs consistently seek to achieve their primary aims by asking five common questions: what happened? Why did it happen? What can be learned? What should change? Have changes taken place?

The 11-item appraisal tool for assessing the methodological vigour of AARs included the following criteria: Prolonged engagement with the subject of inquiry; use of theory; data selection; information sampling; multiple data sources; triangulation; negative case analysis; peer debriefing and support; respondent validation; clear report of methods of data collection and analysis (audit trail) and depth and insight (Table 1).

When the appraisal tool was applied systematically to 24 papers describing AARs, significant shortcomings emerged between the theory and practice of AARs. For example, current AAR practice does not often appear to be guided by available after-action investigation theory, or common good practice used in qualitative research. Reference to formal accident investigation approaches, such as facilitated look-backs, peer assessments or root cause analysis, was typically absent. Basic methodological detail was also frequently missing, calling into question the validity of the AARs. Methods to reduce bias in qualitative data collection and analysis, such as data triangulation, negative case analysis or respondent validation, were rarely reported. Based on the shortcomings identified, a best-practice framework was developed (Figure 1).

Conclusions

The best choice of AAR methodology depends on the impact and severity of the incident itself, the immediacy of the improvements required, and the resources available for the review. Therefore, this report does not recommend a single best methodology for the purpose of preparedness planning. Instead it outlines key stages, good practice and minimum standard dimensions to consider when planning and undertaking an AAR, alongside 11 validity-boosting recommendations that every AAR practitioner should consider to improve their AAR methodology.

1. Introduction

1.1 Context

At the 2014 European Centre for Disease Prevention and Control (ECDC) annual meeting of National Focal Points for Preparedness and Response, the importance of learning from experience in the areas of public health emergency preparedness and response planning was highlighted. The use of structured, qualitative methods to review the performance of responses to emergencies was presented as a key method for identifying the strengths and weaknesses of existing preparedness systems. However, within the field of public health preparedness, there are a large range of approaches used and few guidelines describing the best qualitative techniques to use. Originally, ECDC used the terminology 'critical incident reviews' to describe this field activity. However, in order to harmonise the usage of language throughout the public health world and to be aligned with World Health Organization on this topic, ECDC has now adopted the term 'after-action reviews' (AARs).

ECDC commissioned a literature review to identify best-practices in developing AARs of actual emergencies, with a focus on qualitative methods. To further investigate this topic, ECDC hosted an expert meeting on 17-18 February 2016 entitled 'State-of-the art use of critical incident reviews in public health emergency preparedness'. During this meeting, the preliminary results from the literature review were presented, and the feedback obtained from participants was incorporated into the final report. In addition, during an expert consultation on critical incident reviews, organised by ECDC in Berlin, Germany in March 2017, ECDC National Focal Points for Preparedness and Response were asked about good practices in AARs (see Section 1.3).

This report is designed to provide guidance to public health practitioners about to embark upon an AAR, whether by designing their own protocol or using existing approaches.

1.2 Aims and scope

This literature review aimed to identify the range of methods used to produce AARs for the purposes of improving emergency preparedness planning and to evaluate the most robust methodologies. This report was therefore based on the research question 'what is the best methodology and type of study design for after-action reviews for the purposes of preparedness and response planning?'

By addressing this question, the report outlines the key questions and methods typically pursued by AARs; develops an 11-point summary validity appraisal tool for evaluating AARs (based upon a more comprehensive 50-item AAR methodological checklist tool) and presents a best-practice framework.

Aside from AARs, other mechanisms exist for evaluating and improving preparedness, such as simulation exercises. Although they may also be the subject of lessons-learned reports [1], they fall outside the scope of this specific report.

1.3 Public health emergencies and after-action reviews

For the purposes of this report, a working definition of 'public health emergencies' and 'after-action reviews' has been developed. The definition is a composite of multiple definitions, designed to describe incidents of relevance to the public health emergency preparedness and response community.

'After-action review' is also not a standardised term. In public health, a good definition of an AAR is provided by World Health Organization (WHO):

- An AAR is a qualitative review of actions taken to respond to an emergency as a means of identifying best-practices, gaps and lessons learned. Following an emergency response to a public health event, an AAR seeks to identify what worked well or not and how these practices can be maintained, improved, institutionalised and shared with relevant stakeholders.[1]

These types of review have alternatively been called 'after event reviews' or 'critical incident reviews'. However, in order to encourage harmonisation within the field of public health emergency preparedness, this report will henceforth also refer to after-action reviews (AARs).

AARs should be undertaken in the wake of any public health emergency response. In order to provide some context for when emergency responses might be initiated, a working definition of a public health emergency is useful. This report defines a public health emergency as 'an unexpected event that has a noticeable impact on society (disruption, trauma, injury/loss of life, damage/impact on infrastructure) and where public health played a significant role in the response' [2,3,4].

Emergency public health situations cover a range of potential events, including bioterrorist incidents (e.g. intentional anthrax release), emerging and re-emerging pathogens (e.g. Ebola, H1N1 influenza), foodborne disease outbreaks (e.g. large-scale *E. coli* outbreaks), natural disasters (e.g. hurricanes, flooding), and chemical threats (e.g. large-scale fires). Public heath emergencies are distinct from those occurring in clinical settings, such as hospital investigations. AARs in clinical settings are not included in this review, as these tend to focus on clinical errors, such as administration of the incorrect medication, performance of surgery on the incorrect part of the body, or administrative errors that lead to a poorer prognosis. However, AARs from other sectors – such as civil protection – were included because they involve large-scale, cross-disciplinary emergency responses. Therefore, such incidents may provide lessons for the public health emergency preparedness and response community, and including them in the review ensures that such lessons are shared across sectors.

[1] World Health Organization. Definition available at: https://extranet.who.int/sph/after-action-review

2. Project methodology

2.1 Stage 1: theoretical grounding

Ultimately, the aim of AARs is to improve preparedness, response and recovery capacities and capabilities through a continuous quality improvement cycle, in order to lessen the impact of future incidents [5,6]. Different AARs may focus on distinct elements within the preparedness, response and recovery chain of events (for example, reviewing vaccine strategy and response as part of a wider response to H1N1 pandemic flu) [7].

Prior to developing an appraisal tool for AARs (Section 2.1), a focussed, rapid review of published research and grey literature on AAR methodological theory was conducted using Google Advanced and Google Scholar. These databases were selected due to their extensive coverage of both academic and grey literature, enabling the search to be conducted with precision. Search terms included all major synonyms for AAR-type exercises, such as 'after-action reports' and 'lessons learned reviews'. A total of 28 articles relating to AAR theory and 10 templates for AARs were included in the theoretical grounding phase. Relevant articles were reviewed by one analyst and common themes were identified. To prevent the exclusion of potentially relevant documents that did not describe themselves as 'after-action reviews', we did not determine an a priori definition, but rather developed the definition iteratively during the theoretical grounding stage. This definition was incorporated into the sifting criteria used to select AARs for appraisal and analysis.

Whether dealing with small or large incidents, the review highlighted that AARs consistently sought to achieve their primary aims by establishing five common elements [6, 8-15].

1. What happened? The fact gathering phase. This seeks to establish the details of a given incident and to establish what happened before (preparedness) during (response) and after (recovery) the incident in detailed chronological order. The focus is on collecting as much factual information as possible in an effort to establish objective truths. A timeline of events is typically based on one or more of three common data collection approaches:

- gathering documentation (emergency plans, protocols, action plans),
- gathering personal testimony (individual or group interviews, discussion or consultations),
- site visits (data collection from the incident site itself, taking photos, measurements, developing site maps, also surveillance data collection, laboratory results, air quality measures, etc.)

2. Why did it happen? Why did the event happen? Why did the emergency response and recovery efforts happen as they did? Was there deviation from emergency protocols? This phase seeks to establish the main (immediate cause) and contextual reasons (contributory factors) as to why the emergency public health response and recovery operations happened the way they did. This is usually more qualitative in nature, relying on personal testimony. Best practice seeks to go beyond identification of immediate cause (see glossary), unsafe acts or latent failures and explores the array of contributory factors leading to system success, failures or omissions - the root causes.

3. What can be learned? What can be learned from an evaluation of the previous two elements in order to increase PHEP in the future? What was effective, what was not? This often takes the form of a 'lessons learned' section describing different successes and failures and their relative impact on the events and future PHEP. Lessons tend to be specifically related to the incident reviewed, but can include more generic emergency preparedness capacities and capabilities, such as information sharing.

4. What should change? What policies or ways of working need to change to mitigate any problems identified, reduce the impact of similar events in the future, and generally improve PHEP? For example, was surveillance capacity sufficient to identify the threat of the incident in good time, does it need increasing? Good practice dictates SMART recommendations: specific, measurable, achievable, relevant and time-bound.

5. Have changes taken place? A plan to monitor and establish whether the changes suggested by the AAR have been implemented – completing the feedback and continuous quality improvement cycle. Have the lessons learned materialised in the form of real improvements in emergency preparedness capabilities or capacities? Or were the recommendations and lessons learned read and forgotten?

These five broad categories capture the essence of many types of AAR and approaches including The after-action technique [16], hospital-based system analysis [11], and accident or incident investigation more generally [17]. In terms of the main stages of conducting an AAR, the five-step sequence still applies, but some approaches have defined additional earlier stages, such as selecting the people for the investigation team [11] and planning after action investigation and analysis [17].

When to conduct an AAR

AARs should ideally be undertaken following all emergency responses, regardless of how successful the response is perceived to have been [6] because they contribute to the cycle of continuous quality improvement in emergency preparedness planning and response.

The temptation is to only review incidents where the response was deemed inadequate, but this misses crucial opportunities to understand and repeat effective responses [16,17]. Firstly, the perception of success could be wrong, based on false assumptions, or perceived only from a single viewpoint. Secondly, there is the risk of oversimplifying a complex subset of effective and ineffective actions or inactions into a summary outcome – 'we did well', or 'we didn't do well'. This misses valuable opportunities to identify root causes and important contributory factors within the overall response. Thirdly, if the response is successful, failure to analyse may miss the true reason for success, which may be different to that assumed. In short, lessons can be learned from all responses to after actions and initial perceptions of success and failure should be challenged [16].

Methods used in AARs

There is no standard format for conducting an AAR and a wide range of methodologies are typically deployed. AARs identified in this study used a wide variety of common qualitative and quantitative research methods to achieve their aims, alone and in combination [2, 17-22]. These include:

- Questionnaires
- Interviews
- Focus groups
- Workshops
- Public discussion forums
- Formal public consultation
- Document review
- Site visits.

As such, AARs are open to the same biases and pitfalls as the qualitative methods they rely on. For example, the subjectivity and potential biases involved in interviewing groups or selected individuals, and the potential bias introduced when selecting a sample group. There are a number of ways to reduce these limitations, such as triangulation and cross-validation of data sources. For example, triangulating interview data with email records to build a timeline of events, or cross-validating multiple viewpoints of the incident to highlight areas of consensus and difference [8,16]. These methods aim to reduce the subjectivity and biases of the methods used, thereby increasing the overall validity of the AAR.

The pros and cons of the common qualitative methods listed above, as well as methods to minimise their potential biases, are well documented in the general research methods literature. We identified no evidence to suggest that these limitations would be different in the context of an AAR. At the time of the study, we also found no evidence that less common qualitative approaches, such as using ethnographic perspectives, case studies, and conversational analysis, were applied to develop AAR theory or used in practice.[2]

Different combinations of these common methods formed the basis of a range of theoretical approaches to after-action investigation, and accident investigation more generally. The most common in the literature we reviewed were:

- Root cause analysis [23,24,25]
- The after action technique [8,16] and after action analysis [3,26]
- Realistic evaluation [2,12]
- Peer assessment approach [10]
- Facilitated look-backs [27]
- Serious case reviews [28]
- Case study research [29].

Despite the variation in methods and nomenclature, there was relative consensus that an AAR should seek to establish more than the immediate cause of response and recovery issues. An AAR should systematically analyse the contributory factors behind the immediate causes, aiming to get to the root causes [9,17]. This consensus seems to be rooted in the 1954 paper by Flanagan on the technique, which underpins many subsequent

[2] Since this review was undertaken, ECDC has conducted qualitative case studies on a variety of preparedness topics that may be of interest to readers of this report, but these were published after this particular literature review was conducted. Examples below:
https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/Preparedness%20planning%20against%20respiratory%20viruses%20-%20final.pdf
https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/Preparedness_planning_polio_Poland_Cyprus_Dec_2017.pdf

developments [8]. This systematic approach to root cause analysis is at the heart of most approaches in some form, including peer assessment approaches [9]. Other techniques had very similar aims but were less explicit in actively pursuing the 'root causes' [2,12,29].

2.2 Stage 2: development of an appraisal tool

An 11-item appraisal tool was designed to assist the systematic documentation of methods used in AARs, compare their validity and act as a potential starting point for a standard data collection template in the future.

The development of the appraisal tool involved:

- creating an initial 50-item tool to systematically document AAR methods. This was developed from:
  - an existing appraisal form in a 2005 systematic review on best methods of after action investigation from high-risk industries and healthcare [17]
  - nine additional multi-sector after-action report templates identified through targeted searches [4,11-15, 30-32]
  - consultation with an expert in emergency public health [6]
- summarising the 50-item tool into a more succinct 11-item validity score:
  - adapted from a 10-item tool [12]
  - summary validity scores informed by methods documented in the 50-item tool
- pilot-testing both appraisal tool elements on five real-world after action reviews and optimising the number of criteria and their definitions [18-22,33].

After these development stages we used the tools to appraise an additional 17 AARs, 22 in total, after which no further changes were made to the tools. Further details on this process are outlined in Annex 1 and have also been published elsewhere [51]

The 11-item tool is presented in Table 1.

Table 1. The 11-item tool for assessing AAR methodological rigour

Validity category description	Reviewer guidance notes
1. Prolonged engagement with the subject of inquiry. Has the review included lengthy and perhaps repeated interviews with respondents, and/or days and weeks of engagement within a case study site or group?	**Fully met:** gives the sense that engagement with incident has been thorough and deep as a result of long or repeated interviews, large sample sizes, prolonged or multiple site visits or stages of engagement, etc. **Partially met:** AAR has engaged with the subject well but does not appear comprehensive. **Not met:** they had superficial engagement or it was unclear.
2. Use of theory. Has theory been used to guide sample selection, data collection and analysis, and to draw guide interpretive analysis?	Does the AAR specify theoretical models, frameworks or approached used to inform their work? Does this include any of those recorded in the theoretical grounding work.
3. Data selection. Has purposive selection been used to allow prior theory and initial assumptions to be tested or to examine 'average' or unusual experience?	**Fully met:** clear sample rationales are given, providing a clear sense they have deliberately and purposively interviewed/studied their subjects. **Partially met:** rationale for who they have interviewed is brief or superficial, lacks detail, making it unclear why they have chosen the sample they have, or why it is limited in the way it is - e.g. 'meetings with key entities at the national, regional and local level, including health trusts, county governors and municipalities.' This example tells us the organisations interviewed but not their roles. **Not met:** rationale unclear or not reported.
4. Information sampling. Has the review gathered views from a wide range of perspectives and respondents rather than letting one viewpoint (person, organisation or specialty) dominate? Does it sample from enough people, places, times, etc. to ensure the influence of these factors on the behaviour and views of those people providing information is minimised. Is sampling expanded in the light of early findings?	Who did they select? Irrespective of whether a sample rationale has been given above, does the AAR appear to have picked an appropriately diverse sample? Fully met: wide and varied sample perspectives gathered e.g. the 'who contributed to the report' list is large and diverse. Partially met: key detail of the sample is missing - for example, the number interviewed, participants' roles or affiliated organisations. Not met: who they have interviewed is unclear or not reported.
5. Multiple data sources. Does the review seek multiple information sources (documents, personal testimony, site visits) and collate multiple examples of each? For example, are duplicate formal interviews with all sampled staff undertaken? Does it use researcher observation and informal discussion; are interviews conducted with people of different roles and levels of seniority?	**Fully met:** three main methods (testimony, records/reports, and sit visit) fully met unless site visit is not applicable - e.g. looking at the role of leadership in a response, would not necessarily need a site visit. Fully met can be two methods but multiple examples of two methods - e.g. focus group, plus in-depth interviews, plus document review. **Partially met:** two methods, commonly testimony and document review. **Not met:** one method only - e.g. document/data review without personal testimony.

Validity category description	Reviewer guidance notes
6. Triangulation. Does the review look for patterns of convergence and divergence by comparing results across multiple sources of evidence (e.g. across interviewees, and between interview and other data), between researchers, and across different methodological approaches? Does it also include comparisons within data – e.g. comparing different interview accounts.	**Fully met:** used words similar to triangulation or described methods of formally comparing and contrasting insight between, and/or within, data sources - e.g. do CCTV accounts verify eye witness accounts. **Partially met:** collected multiple sources of data and do not state how they synthesised them but it is implicit that they did as they talk about one evidence base. **Not met:** not reported.
7. Negative case analysis. Does the review look for evidence that contradicts its initial findings, explanations and theory, and refine them in response to this evidence?	Looking for specific mention of 'negative case analysis' or 'deviant case analysis' or reference to very similar approach described left.
8. Peer debriefing and support. Does the AAR include a step where the findings and reports are reviewed by other researchers or investigators?	Looking for specific mention of this or reference to very similar approach. Includes public consultation.
9. Respondent validation. Review of findings and reports by respondents to check investigator interpretation of their input.	**Fully met** if respondents have validated/had the opportunity to comment on the report findings of their views. Must be respondents. Other commentators = peer debriefing and support.
10. Clear report of methods of data collection and analysis (audit trail). Has the review kept and reported a full record of activities available to others and presented a full account of how methods evolved and were applied?	**Fully met**: clear and comprehensive methodological detail giving sense their methods could be replicated independently. **Partially met**: methods are brief and somewhat superficial but they are at least documented. Similarly, if the report links to full methods elsewhere and hard to find. Fully met if the methods appear in an appendix or if there are links to another document that is easy to find.
11. Depth and insight. Has the AAR established the direct/indirect root causes and underlying contributory factors linked to errors, inaction or latent failures?	**Fully met:** the results clearly discuss root causes alongside and contributory factors throughout and in a systematic way. **Partially met:** some causal factors behind errors are discussed, but not throughout, or systematically. Patchy insight. **Not met:** recommendations/results seem superficial - e.g. largely describing what happened without insight into why or how.

Adapted from Piltch-Loeb et al. 2014 [12] and Davies et al [51]

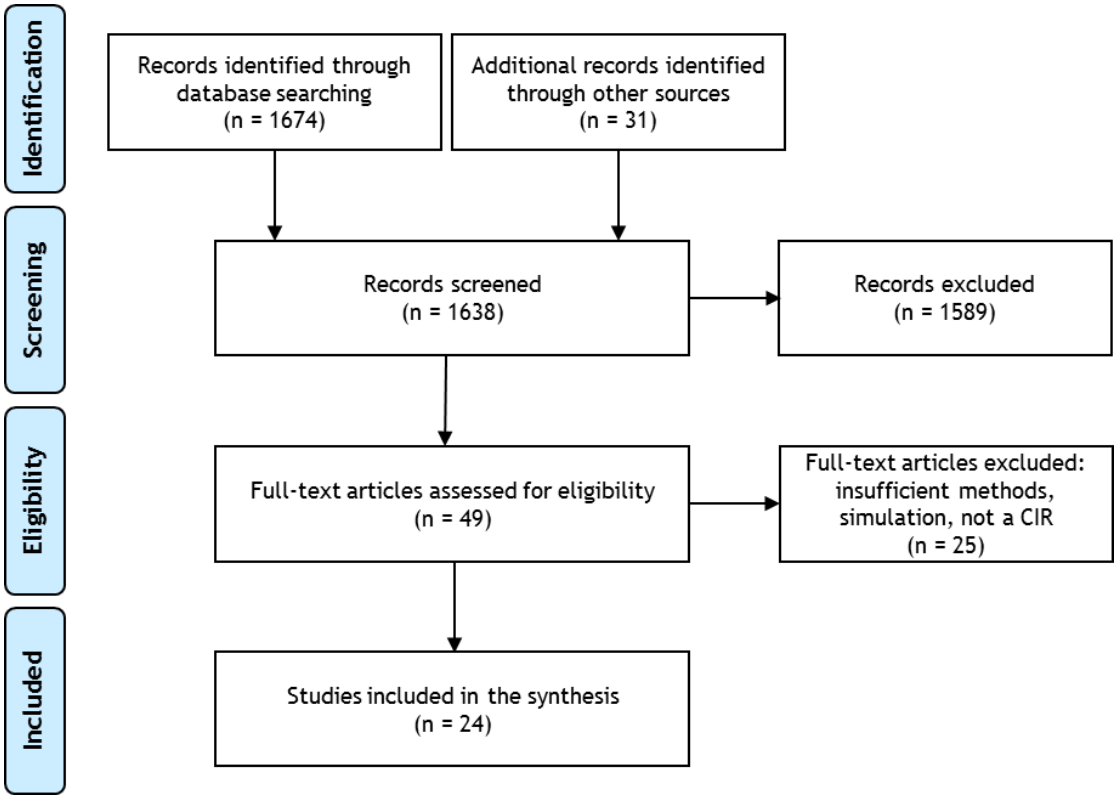
2.3 Stage 3: development of the best-practices framework

The development of a best-practices framework involved a systematic literature review of AARs which were subsequently evaluated with the tool developed in Section 2.2. A total of 24 documents were included in the appraisal analysis, including 22 AARs for appraisal and two linked annexes (Figure 2). The insights from this analysis were then used to develop a best-practices framework.

The AARs included in the analysis included:

- Biological
  - Avian influenza A virus H1N1, 2009: Canada [34], Norway [18], Sweden [35], UK [36], WHO vaccine deployment [37], EU pandemic vaccine strategies [7], EU response in general [38], and WHO Regional Office for Europe [22].
- Chemical
  - Toulouse (France) and Buncefield (UK) fuel storage depot explosion and subsequent fire [39, 19, 40,41] (four), including two annexes to AARs that reported methodological information [33, 42]
- Physical
  - Hurricanes Andrew and Katrina [43], Katrina only [44]
  - Burns disasters after Bali bombing and explosion on a boat, Australia [45]
  - Boston Marathon Bombing, 2015 [46,20] (two)
  - Heatwave France, 2003 [47]
  - London Bombings, 2005 [48]
  - Oklahoma bombing [49]
  - Water emergency, Massachusetts, USA [50]
  - UK floods [21].

Figure 1. PRISMA diagram for literature review



The 11-item validity tool (Table 1) was then used to assess the 24 AAR documents identified above. Each item on the summary validity tool was assigned a score from 2 to 0 using the following coding:

- 2 – Fully met, these criteria have been fully and often comprehensively met. We have little doubt that these criteria have been met.
- 1 – Partially met, some of the criteria have been met but there are elements missing that prevent a higher rating. The criteria have been met in some regards, but there is significant doubt about the comprehensiveness, or there are clear elements missing.
- 0 – Not met; these criteria are not met or have not been reported.

A sample of three (16%) AARs were independently rated by a second reviewer to test the reliability of the rating instrument and clarify initial rating definitions. The second rater was blind to the first rater's scores and rationales. Differences between the two raters were discussed and changes agreed by consensus. This led to revisions in the wording of some criteria and scoring guidance to improve clarity. Table 1 shows the working definitions of the 11 validity measures.

The analysis of the 24 AAR documents highlighted clear gaps between theoretical good practice and what was reported in real-world AARs. Based on this analysis, a narrative synthesis was conducted to draw attention to the ways in which AARs could be conducted to better bridge the gap between methodological validity and real-world practice.

Annex 2 presents an example of how individual studies were appraised by reviewers. The observations were collated and synthesised into the narrative synthesis.

3. Best-practice framework for conducting AARs

3.1 Key themes from the literature review

There was huge diversity in the structure, methods and reporting of AARs, something that may benefit from standardisation (see Table 2 in [51]). The non-use of theory, unclear methods and negative case analysis stood out as areas where theory and practice consistently diverged. For example, only three AARs reviewed explicitly mentioned using theory to inform their review, and next to none used negative case analysis or similar methods to increase the validity of their analysis. Meanwhile, very few AARs reported clear methods of data collection and analysis (a clear audit trail). Most described these methods very briefly, leaving many gaps and uncertainties about validity.

Sampling a range of views, using multiple methods of data collection, and triangulating data analysis were among the most consistently described elements of good practice seen in the AARs. However, they were generally reported in brief, with few AARs scoring full marks for all three basic validity dimensions.

In terms of triangulation, many AARs were not clear how they analysed their results or dealt with areas of consistency and inconsistency in the data collected, therefore scoring partially met instead of fully met. As such it was not clear to what extent certain views or data had been discounted (e.g. if they did not fit with the emerging consensus.) This could lead to biases being introduced into the analysis, or important observations being overlooked.

It was rare to find AARs reporting elements of peer debriefing and support. We identified this in only three cases, which included national public consultation as well as expert consultation on draft reports and recommendations.

Some AARs scored well for depth and insight but only obtained an average score for other validity measures because their methods were not clear. This suggests many AARs may not have clearly reported the methodology and various data sources, which may or may not be linked to underlying methodological limitations. The big issue is that validity cannot be judged without clearly reported methodologies – and these were rare to find. Validity underscores the reliability of the findings and potential effectiveness of implementing the recommendations in future public health preparedness planning. This strengthens the case for a minimum reporting standard in relation to methodological detail for AARs in order to address this frequent limitation.

Each of the 11 validity dimensions are discussed in Table 2 below by drawing on illustrative examples from across the 22 AARs. Insights from the literature review are thematically grouped across these 11 dimensions, leading to the development of 11 validity-boosting considerations that practitioners may wish to consider when conducting after-action reviews.

Table 2. 11 validity-boosting considerations for improving after-action review methods and reporting

	Consideration	Recommendation
1	Prolonged engagement	AARs should spend adequate time observing the setting and incident documentation and speaking with a range of people to build a deep understanding of the after-action and its context. Prolonged and repeated engagement with the people and processes involved in the after-action over time offers a greater chance of uncovering deeper and more valid insight than brief and sporadic engagement with the study subject.
2	Use of theory	After-action reviews may benefit from being more closely aligned with after action review theoretical frameworks – such as the after-action technique – to ensure PHEP improvement plans from AARs address root causes. Furthermore, AARs should consider applying basic qualitative methods validity checks as a standard – such as those in the 11-item validity tool – to boost the validity of the insights gained. This - and additional concepts such as data saturation - could also help improve efficiency where an 'ask everyone, gather everything' approach is not pragmatic.
3	Data selection	Study sample rationales should be clearly described in all AARs to allow readers to easily understand how data informing the review, for example, participants in interviews, were targeted and selected. This is important to enable assessment of potential selection bias.
4	Information sampling	Study samples should be clearly described in all AARs to allow readers to understand which individuals, groups or data were used to inform the review. For example, documenting the number of people interviewed, their job titles and their role in the after action. Without this, readers do not know whether important views, reports or data were excluded, so are less able to evaluate the review for selection bias.
5	Multiple data sources	AARs should use multiple approaches for data collection to ensure a variety of information is considered, reducing the risk that one potentially biased data source dominates, and increasing the likelihood that root causes and relevant contributory factors will be appropriately uncovered. It is common for the most comprehensive AARs to include a combination of personal testimony (through different types of interviews, questionnaires etc.), document review (PHEP protocols, guidelines, relevant reports on the incident, safety reports before the incident etc.), and where relevant, one or more site visits.
6	Triangulation	Triangulation can help uncover perception bias and ensure insights are more roundly developed. It is recommended that multiple analysts, observers or reviewers be used to check interpretation of data; specifically looking at consistencies and divergence among different, but also within similar, data sources.
7	Negative case analysis	AARs should clearly report how discordant evidence (from personal testimony, reports, site visits or in forming improvement plans) has been reconciled. AARs should discuss any evidence that contradicts initial findings, explanations and developing theories alongside the consensual views. This encourages open and critical assessment of emergent themes and conclusions when forming the AAR, and may discourage seeking only harmonious views as part of the AAR sample.
8	Peer debriefing and support	Sharing preliminary or draft findings of the AAR with PHEP experts outside of the event for critical comment may increase the validity by introducing a fresh and independent perspective on the findings of an AAR, as well as pointing out any gaps in the review or analysis. This may also serve to facilitate learning across different sectors and geographies, increase awareness and build and expand professional networks.
9	Respondent validation	Initial insight and findings should be checked by those who contributed to the review (respondent validation) to ensure the accuracy and relevance of the AAR findings. This technique increases the likelihood that the AAR accurately represents the views of those contributing to it.
10	Audit trail	As a minimum standard, after-action reviews should report the methods they have used to gather information, analyse it, and clearly report how these led to the recommendations made. To aid readability, these can be in an appendix, but should be easily available for those wanting to evaluate the validity of the AAR. The development of evidence-based minimum reporting standard for after-action reviews, similar to the CONSORT statement for randomised control trials[3], may facilitate this process and comparisons between AARs.
11	Depth and insight	AARs should seek to uncover and report active and latent failures, contributory factors and root causes of the after-actions and make specific recommendations to improve PHEP as a result of significant depth and insight into the issues at hand. They should be explicit in their methods of doing so in terms of data collection and analysis. AARs should be explicit in stating how they interpreted data to gain the insights they have into improvement processes, including any attempts to increase the validity of their insights – e.g. independent interpretive checks through peer review.

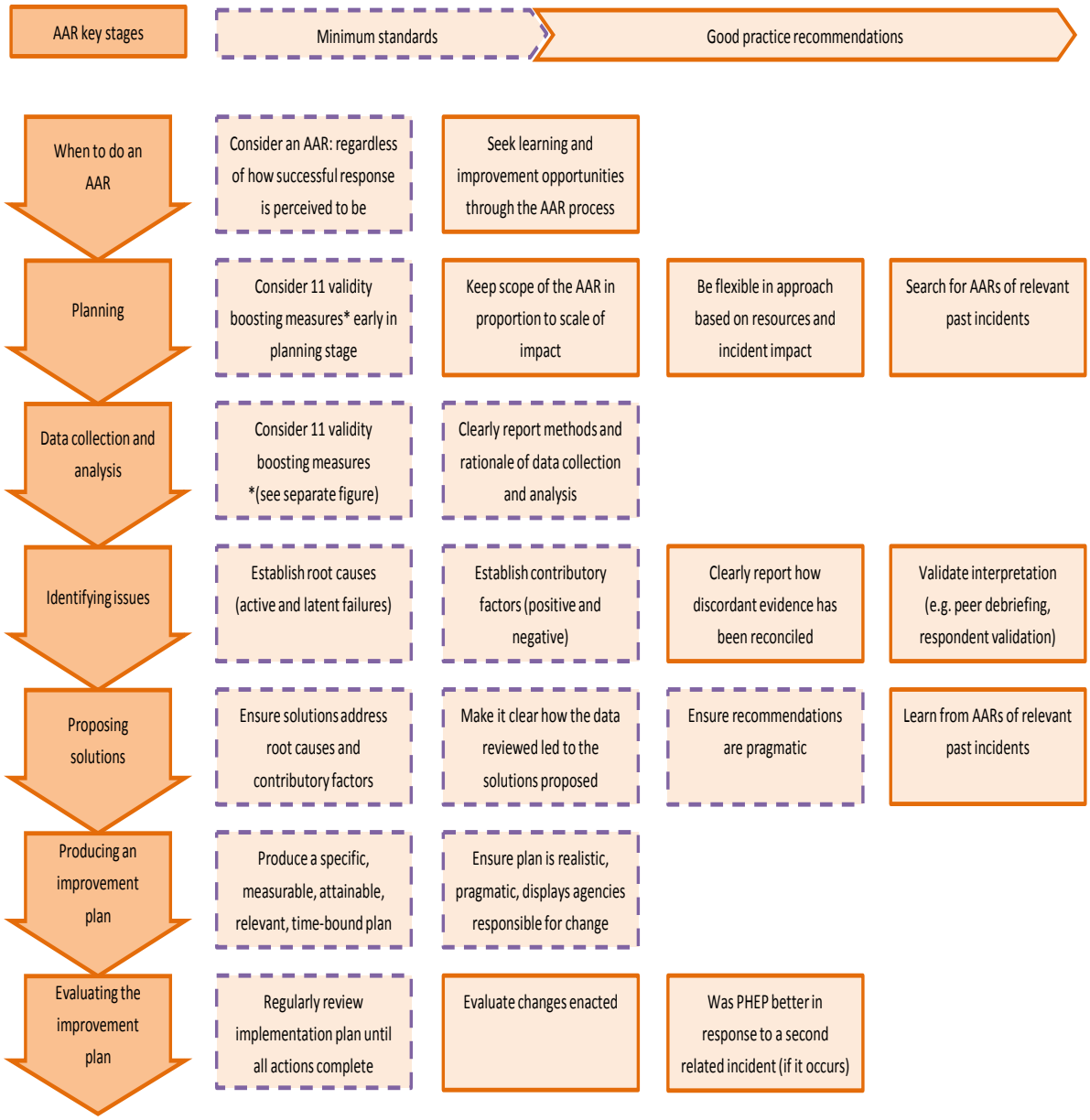
Adapted from Table 3 in [51]

[3] http://www.consort-statement.org/

3.2 Best-practices framework

Based on the literature review described above, practical and actionable measures that could be undertaken to enhance the methodological validity of AARs were identified. Taking into consideration the information in Table 2, and based on the key stages involved in undertaking an AAR, a best-practices framework is proposed (Figure 2).

Figure 2. Best-practices framework for undertaking an AAR



4. Conclusion

This report is the outcome of a systematic literature review which addresses gaps identified in AAR practice through expert consultations organised by ECDC.

After-action reviews are an important element of the public health emergency preparedness cycle, and there is currently a global impetus, led by WHO, to promote their wider usage[4]. A systematic review of the mobilisation of public health emergency response system represents an important opportunity for improvement within a public health emergency response system. The best choice of AAR methodology depends on the impact and severity of the incident itself, the immediacy of the improvements required, and the resources available for the review. Therefore, this report does not recommend a single best methodology for the purposes of preparedness planning. Instead it outlines key stages, good practice and minimum standard dimensions to consider when planning and undertaking an AAR (Figure 1), alongside 11 validity-boosting recommendations that every AAR practitioner should consider to improve their AAR methodology.

In the coming years, ECDC will build upon the work presented here to promote the wider usage of AARs among public health emergency preparedness and response professionals.

[4] https://extranet.who.int/sph/after-action-review

References

1. Biddinger PD, Savoia E, Massin-Short SB, Preston J, Stoto MA. Public health emergency preparedness exercises: lessons learned. Public Health Reports 2010, 125
2. Stoto MA. Getting from what to why: using qualitative methods in public health systems research. 2012. Available at: https://www.academyhealth.org/publications/2013-11/getting-what-why-using-qualitative-methods-public-health-systems-research
3. Schwester RW. Handbook of after-action analysis. New York (NY): Routledge (Taylor & Francis Group); 2014
4. National Patient Safety Agency. Significant event audit: guidance for primary care teams. 2008. Available at: http://www.nrls.npsa.nhs.uk/EasySiteWeb/getresource.axd?AssetID=61501
5. Stoto MA. Measuring and assessing public health emergency preparedness. Public Health Managment and Practice 2013, 19:S16-S21
6. National Health Service Education for Scotland. Significant Event Analysis. Guidance for Primary Care Teams. 2011. Available at: https://www.nes.scot.nhs.uk/media/346578/sea_-_full_guide_-_2011.pdf
7. European Commission. Assessment report on EU-wide pandemic vaccine strategies. 2010. Available at: http://ec.europa.eu/health/communicable_diseases/docs/assessment_vaccine_en.pdf
8. Flanaghan JC. The after action technique. Psychological Bulletin 1954, 51
9. Piltch-Loeb RN, Nelson CD, Kraemer JD, Savoia E, Stoto MA. A peer assessment approach for learning from public health emergencies. Public health reports 2014, 129:28-34
10. Piltch-Loeb RN, Nelson CD, Kraemer JD, Savoia E, Stoto MA. A peer assessment approach for learning from public health emergencies. Improving Systems and Practices 2014, 129
11. Taylor-Adams S, Vincent C. Systems analysis of clinical incidents: the London protocol. Available at: https://www1.imperial.ac.uk/resources/C85B6574-7E28-4BE6-BE61-E94C3F6243CE/londonprotocol_e.pdf.
12. Piltch-Loeb RN, Nelson C, Kraemer J, Savoia E, Stoto MA. Peer assessment of public health emergency response toolkit. 2014. Available at: https://blogs.commons.georgetown.edu/lamps/files/2013/10/Peer-Assessment-toolkit-Updated-1-17-14.pdf
13. International SOS. Significant event analysis reporting form. 2015. Available at: Not publicly available.
14. Medical Defence Union. Medico-legal guide to serious event analysis. 2014. Available at: http://www.themdu.com/~/media/files/mdu/publications/guides/significant%20event%20analysis/medico_-_legal_guide_to_significant_event_analysis.pdf
15. National Patients Safety Agency (UK). A quick guide to conducting a significant event audit. 2008. Available at: http://www.nrls.npsa.nhs.uk/EasySiteWeb/getresource.axd?AssetID=61502&type=full&servicetype=Attachment
16. Serrat O. The after action technique. Knowledge Solutions 2010, 86
17. Woloshynowych N, Rogers S, Taylor-Adams S, Vincent C. The investigation and analysis of after actions and adverse events in healthcare. Health Technology Assessment 2005, 9
18. Directorate for Civil Protection and Emergency Planning (Norway). New influenza A virus (H1N1): a summary of a study on the national response in Norway. 2011 Available at: https://www.dsb.no/globalassets/dokumenter/rapporter/new-influenza-a-virus-h1n1.pdf
19. Health and Safety Executive (UK). The Buncefield incident 11 December 2005: the final report of the major incident investigation board volume 2. 2008. Available at: http://www.hse.gov.uk/comah/buncefield/miib-final-volume2a.pdf
20. Massachusetts Emergency Management Agency. After-action report for the response to the 2013 Boston marathon bombings. 2014. Available at: http://www.mass.gov/eopss/docs/mema/after-action-report-for-the-response-to-the-2013-boston-marathon-bombings.pdf
21. Pitt M. Learning lessons from the 2007 floods: full report 2008. Available at: http://webarchive.nationalarchives.gov.uk/20100807034701/http:/archive.cabinetoffice.gov.uk/pittreview/_/media/assets/www.cabinetoffice.gov.uk/flooding_review/pitt_review_full%20pdf.pdf
22. World Health Organization. Recommendations for good practice in pandemic preparedness: identified through evaluation of the response to pandemic (H1N1) 2009. 2010. Available at: http://www.euro.who.int/__data/assets/pdf_file/0017/128060/e94534.pdf
23. Berry K, Krizek B. Root cause analysis in response to a "near miss". Journal for Healthcare Quality: official publication of the National Association for Healthcare Quality 2000, 22:16-18
24. Iedema R, Jorm C, Braithwaite J. Managing the scope and impact of root cause analysis recommendations. Journal of Health, Organisation and Management 2008, 22:569-585
25. Singleton CM, Debastiani S, Rose D, Kahn EB. An analysis of root cause identification and continuous quality improvement in public health H1N1 after-action reports. Journal of Public Health Management and Practice: JPHMP 2014, 20:197-204
26. Scottish Institute for Excellence in Social Work Education. Evaluation of an innovative method of assessment: After-action analysis. 2005. Available at: https://pure.strath.ac.uk/portal/files/11177868/sieswe_nam_evaluation_critical_incident_analysis_2005_02.pdf

27. Aledort JE, Lurie N, Ricci KA, Dausey DJ, Howard S. Facilitated look-backs: a new quality improvement tool for management of routine annual and pandemic influenza. 2006. Available at: http://www.rand.org/pubs/technical_reports/TR320.html
28. Office for Standards in Education, Children's Services and Skills (UK). Learning lessons from serious case reviews 2009–2010: Ofsted's evaluation of serious case reviews from 1 April 2009 to 31 March 2010. 2010. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/381110/Learning_20lessons_20from_20serious_20case_20reviews_202009-2010.pdf
29. Yin RK. The Case Study Anthology. London: SAGE Publications 2004.
30. California Governor's Office of Emergency Services. Standardized emergency management system: after-action report. 2013. Available at: http://www.caloes.ca.gov/PlanningPreparednessSite/Documents/Local%20Government%20After%20Action%20Report%20Template.doc
31. Homeland Security Exercise and Evaluation Program (USA). After-action report/improvement plan. 2013. Available at: www.in.gov/dhs/files/AAR-IP_Template_Apr-13-2_Clean.docx
32. Healthcare Insurance Reciprocal of Canada. After-actions and multi-patient events: risk resource guide. 2015. Available at: https://www.hiroc.com/getmedia/c110b394-c5a6-4e93-80b3-73da14436dbd/HIROC-Management-of-Critical-Incidents-April-2015.pdf.aspx?ext=.pdf
33. Health and Safety Executive (UK). The Buncefield incident 11 December 2005: the final report of the major incident investigation board volume 1. 2008. Available at: http://www.hse.gov.uk/comah/buncefield/miib-final-volume1.pdf
34. Masotti P, Green ME, Birtwhistle R, Gemmill I, Moore K, O'Connor K, et al. PH1N1: a comparative analysis of public health responses in Ontario to the influenza outbreak, public health and primary care: Lessons learned and policy suggestions. BMC Public Health 2013, 13
35. Socialstyrelsen. A(H1N1) 2009: an evaluation of Sweden's preparations for and management of the pandemic. 2011.
36. Hine D. The 2009 influenza pandemic: an independent review of the UK response to the 2009 influenza pandemic. 2010. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/61252/the2009influenzapandemic-review.pdf
37. World Health Organization. Global survey on national vaccine deployment and vaccination plans for pandemic A(H1N1) 2009 vaccine - 2010: report of findings. 2013. Available at: http://www.who.int/influenza_vaccines_plan/resources/2010_H1N1_NVDP_WHO_Survey.pdf
38. Health Protection Agency. Assessment report on the EU-wide response to pandemic (H1N1) 2009. 2010. Available at: http://ec.europa.eu/health/communicable_diseases/docs/assessment_response_en.pdf
39. Paltrinieri N, Dechy N, Salzano E, Wardman M, Cozzani V. Lessons learned from Toulouse and Buncefield disasters: from risk analysis failures to the identification of atypical scenarios through a better knowledge management. *Risk Analysis* 2012, 32:1404-1419
40. SQW. Buncefield social impact assessment: final report January 2007. 2007. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/78983/Buncefield-Social-Impact-Assessment-Final_20Report_0.pdf
41. Tapster C. Buncefield: multi-agency debrief report and recommendations. 2007. Available at: http://anaesthesiaconference.kiev.ua/downloads/bunrepdebrief_2007.pdf
42. SQW. Buncefield social impact assessment: annex A key information sources. 2007. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/78984/buncefield-social-impact-assessment-final-report-annexes.pdf
43. Knox CC. Analyzing after-action reports from Hurricanes Andrew and Katrina: repeated, modified, and newly created recommendations. Journal of Emergency Management 2013, 11:160-168
44. Brevard SB, Weintraub SL, Aiken JB, Halton EB, Duchesne JC, McSwain Jr NE, et al. Analysis of disaster response plans and the aftermath of Hurricane Katrina: lessons learned from a level i trauma center. Journal of Trauma - Injury, Infection and Critical Care 2008, 65:1126-1132
45. Little M, Cooper J, Gope M, Hahn KA, Kibar C, McCoubrie D, et al. 'Lessons learned': a comparative case study analysis of an emergency department response to two burns disasters. EMA - Emergency Medicine Australasia 2012, 24:420-429
46. Goralnick E, Halpern P, Loo S, Gates J, Biddinger P, Fisher J, et al. Leadership during the Boston marathon bombings: a qualitative after-action review. Disaster Medicine and Public Health Preparedness 2015, 9:6
47. Adrot A. Crisis response, organizational improvisation and the dispassionate communicative genre during the 2003 French heat wave. 2011. Available at: https://basepub.dauphine.fr/bitstream/handle/123456789/7969/Adrot_aim2011.PDF?sequence=1
48. Aylwin CJ, König TC, Brennan NW, Shirley PJ, Davies G, Walst al. Reduction in critical mortality in urban mass casualty incidents: analysis of triage, surge, and resource use after the London bombings on July 7, 2005. Lancet 2006, 368:2219-2225

49. Oklahoma Depatment of Civil Emergency Management. After-action report: Alfred P. Murrah Federal Building bombing 19 April 1995 in Oklahoma City, Oklahoma. 1995. Available at: https://www.ok.gov/OEM/documents/Bombing%20After%20Action%20Report.pdf.
50. Terenzini C. The report of the Blue Ribbon Committee on the water emergency of April 25-27, 2007 in Spencer, Massachusetts. 2007. Available at: http://www.spencerma.gov/Pages/SpencerMA_Fire/WaterEmergency.pdf
51. Davies, R., Vaughan, E., Fraser, G., Cook, R., Ciotti, M., & Suk, J. (2018). Enhancing Reporting of After-action Reviews of Public Health Emergencies to Strengthen Preparedness: A Literature Review and Methodology Appraisal. Disaster Medicine and Public Health Preparedness, 1-8.

Annex 1. Development of the appraisal tool

During the theoretical grounding literature search a key review was identified and formed the basis of the 50-item appraisal tool development [17].

This review stood out because it systematically drew on a large body of work reviewing best methods in after action methodologies from both healthcare and non-healthcare industries including aviation, nuclear power, chemical and fuel industries. This was directly relevant to the scope of the work which sought to review best methods among a diverse range of real-world after action reviews from different hazard types. Other potentially useful reviews were more healthcare focussed and seemed less relevant to the wide range of incident threats. The wider scope of this review can be seen as a distinct advantage as many high-risk industries have a long history of using AARs to improve safety and plan for response, whereas use of AARs in healthcare is less mature, so can be more fragmented [17].

The large and systematic 2005 review appraised published and unpublished work detailing 12 techniques of accident analysis in high-risk industries, and six approaches to incident investigation and analysis in healthcare [17]. It also developed and piloted an appraisal form and accompanying guideline for the analysis of after actions in healthcare, stating direct relevance for the hospital sector, mental health and primary care. This appraisal form had its origins in a diverse range of high-risk industries and had been adapted and piloted in a healthcare environment, thus it was chosen as the foundation for this stage of the project with the idea of further developing it specifically for after actions.

Developing the 11-item appraisal tool

The existing appraisal form was modified by triangulating it with nine additional after action report templates identified through targeted searches [4,11-15,30-32]. These additional templates were multi-sectorial, coming from after-action reports, significant event analysis, and peer assessments in the fields of US National Defence [31], US State Government [30], UK medico-legal [14], Canadian Healthcare Insurance [32], international emergency public health [12,13] and UK hospital [11] and patient safety agencies [4,15].

Key elements of these tools were listed and combined thematically into a 50-item revised appraisal form. The tool aimed to be generic enough to be applicable across a range of AARs, but specific enough to compare their relative comprehensiveness and identify good practice.

The wider literature suggested that AARs as a whole are open to the same limitations and biases as the qualitative methods they rely on [12]. For example, the subjectivity and potential biases involved in interviewing groups or selected individuals, and the biases involved in choosing the people to interview in the first place. We found no evidence that their use in AARs was any different. As such we did not attempt to replicate the wide body of research discussing the pros and cons of individual qualitative methods. We did however, attempt to summarise these into a pragmatic rating tool.

The 50-item tool was a rigorous method of systematically capturing different AAR methodologies providing us with rich data, but was too detailed to allow simple comparisons between AARs. Thus, the need to develop a summary measure that represented these rich data more concisely and allowed easier comparison.

A 2014 peer assessment of public health emergency response toolkit was identified, which included a 10-point summary of factors boosting rigour in case study and qualitative data collection and analysis [12]. This had the advantage of summarising key elements of study validity in the context of public health emergency response. The study suggested that 'although qualitative methods are often criticized insufficiently rigorous and transparent, there is a well-established body of social science methods that can help to ensure rigor in qualitative research' [12]. The 10-item suggestions for how to do this were consistent with our wider knowledge of qualitative methods and validity and other sources reviewed in the theoretical grounding stage, and utilised many of the 50-item tool dimensions. Therefore they were used to form the foundation of our summary rating system to apply to real-world AARs. While not comprehensive, this offered a pragmatic method for comparing validity in a diverse range of AARs from a recent and relevant source directly relevant to emergency public health.

In the 11-item appraisal form, the original 10-points remained intact with minor revisions in definitions to better reflect AARs in the field of emergency health. Moreover, an eleventh factor was added to capture whether the AAR had ultimately achieved its aim of uncovering the root causes of preparedness, response and recovery activities, rather than only immediate or obvious causes [51].

Annex 2. Example of validity summary appraisal

The table below presents an example of how the 11-item appraisal tool described in Section 2.1 was applied in practice.

Fully met; 2, partially met; 1, not met 0	Score	Comments
1. Prolonged engagement with the subject of inquiry. Has the review included lengthy and perhaps repeated interviews with respondents, and/or days and weeks of engagement within a case study site or group?	1	Not clear enough to give a fully met as appears mainly based on past AAR reports.
2. Use of theory. Has theory been used to guide sample selection, data collection and analysis, and to draw guide interpretive analysis?	2	References theory throughout.
3. Data selection. Has purposive selection been used to allow prior theory and initial assumptions to be tested or to examine 'average' or unusual experience?	0	Not clear, potentially not applicable as it is a AAR of AARs. Hence there was not direct sampling of new people, only past records. They do specify the addition of 'expert knowledge' but this is not further described.
4. Information sampling. Has the review gathered views from a wide range of perspectives and respondents rather than letting one viewpoint (person, organisation or specialty), dominate? Does it sample from enough people, places, times, etc. to ensure the influence of these factors on the behaviour and views of those people providing information is minimised? Is sampling expanded in the light of early findings?	0	No mention of additional sampling outside of the original accident investigation reports.
5. Multiple methods. Does the review seek multiple information sources (documents, personal testimony, site visits) and collate multiple examples of each. For example, are duplicate formal interviews with all sampled staff undertaken? Does it use researcher observation and informal discussion; are interviews conducted with people of different roles and seniorities?	1	Accident investigation reports and expert knowledge cited although the expert knowledge may be the author of the review. Unclear so do not have confidence to award a fully met.
6. Triangulation. Does the review look for patterns of convergence and divergence by comparing results across multiple sources of evidence (e.g. across interviewees, and between interview and other data), between researchers, and across different methodological approaches? Does it also include comparisons within data - e.g. comparing different interview accounts.	1	They have triangulated data across multiple reports applying multiple methods, although not formally used the term.
7. Negative case analysis. Does the review look for evidence that contradicts its initial findings explanations and theory, and refine them in response to this evidence?	0	Not reported
8. Peer debriefing and support. Does the AAR include a step where the findings and reports are reviewed by other researchers or investigators?	0	Not reported
9. Respondent validation. Review of findings and reports by respondents to check investigator interpretation of their input.	0	Not reported
10. Clear report of methods of data collection and analysis (audit trail). Has the review kept and reported a full record of activities available to others and presenting a full account of how methods evolved and were applied?	1	We know what they did, although the sources of information are not as clear as they could be.
11. Depth and insight. Has the AAR established the direct/indirect root causes *and* underlying contributory factors linked to errors, inaction or latent failures?	2	The analysis of two AARs has added an extra level of insight over and above the individual, despite the lack of clarity on some of the methodology.
Total (max 22)	8	

European Centre for Disease Prevention and Control (ECDC)

Address:
Gustav III:s boulevard 40, SE-169 73 Solna, Sweden

Tel. +46 858601000
Fax +46 858601001
www.ecdc.europa.eu

An agency of the European Union
www.europa.eu

Subscribe to our publications
www.ecdc.europa.eu/en/publications

Contact us
publications@ecdc.europa.eu

Follow us on Twitter
@ECDC_EU

Like our Facebook page
www.facebook.com/ECDC.EU

ECDC is committed to ensuring the transparency and independence of its work

In accordance with the Staff Regulations for Officials and Conditions of Employment of Other Servants of the European Union and the ECDC Independence Policy, ECDC staff members shall not, in the performance of their duties, deal with a matter in which, directly or indirectly, they have any personal interest such as to impair their independence. Declarations of interest must be received from any prospective contractor(s) before any contract can be awarded.
www.ecdc.europa.eu/en/aboutus/transparency

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

Publications Office