

# ANNEX: DETAILED METHODOLOGY

## RESEARCH APPROACH

This research follows the approach set out in ALNAP's [Lessons Papers: A methods note](#), combining a systematic evidence review with a Delphi consultation process to validate the findings as follows.

### STEP 1: SCOPING AND RESEARCH DESIGN

The study began with a scoping phase to refine and focus the research questions. We:

- consulted six humanitarian practitioners working directly on drought response to refine the research questions
- defined the scope as humanitarian response to drought, excluding broader environmental management issues
- developed sub-questions covering early warning, response, recovery and resilience, and cross-cutting themes including gender, protection and localisation.

### STEP 2: EVIDENCE SEARCH AND SCREENING

We conducted systematic searches of more than 20 humanitarian, academic and institutional databases and libraries, including ALNAP, Cash Learning Partnership (CaLP), Food and Agriculture Organization of the UN (FAO), International Federation of Red Cross and Red Crescent Societies (IFRC), ODI Global, ReliefWeb, UN Evaluation Group and World Food Programme (WFP), and organisational repositories.

Search parameters included:

- **Date range:** 2012–2024
- **Document types:** Evaluations, case studies, 'lessons learned' reports, operational reviews, guidelines
- **Full-text availability.**

We identified and screened over 1,000 documents using a two-stage process:

1. **Initial screening:** We assessed titles, keywords, summaries and study dates against inclusion/exclusion criteria.
2. **Full-text screening:** We evaluated documents for relevance, presence of lessons or recommendations and alignment with thematic and temporal scope.

We retained documents indirectly related to drought if they offered relevant cross-

cutting insights (for example, on gender equality, inclusion of persons with disabilities or older people, and localisation).

We excluded documents if they:

- lacked substantive lessons or recommendations
- were based entirely on data collected before or after 2012–2024
- were duplicates or overly general (for example, CVA or GBV guidance unrelated to drought).

A full list of the inclusion and exclusion criteria can be found [below](#).

**Outcome: 224 documents selected for detailed analysis.**

### STEP 3: QUALITY ASSESSMENT

We scored each retained document (from 1–5) against three criteria:

1. Relevance to drought-response learning
2. Methodological rigour, including transparency and robustness
3. Scope and generalisability across contexts.

We calculated an average evidence score for each document to enable weighted analysis in subsequent stages. Detailed descriptions of the quality assessment criteria can be found [below](#).

### STEP 4: DATA EXTRACTION AND CODING

We followed a structured extraction and coding process to capture consistent, traceable lessons aligned with the research framework. Using a standardised template, we extracted text from each document and coded it across more than 25 thematic areas, including anticipatory action, CVA, gender, livelihoods, localisation, protection and resilience.

Extraction involved two steps:

#### 1. Extract text

- We recorded direct quotes summarising findings, lessons, challenges or recommendations.
- We linked each extract to its source for full traceability.

#### 2. Code the extract

- We assigned each extract to a thematic area.
- We tagged each extract as a lesson on what worked, what did not work, a

challenge or a recommendation.

- We assigned each extract a document ID.

**Outcome: 1,587 unique extracts generated.**

### STEP 5: ANALYSIS, SYNTHESIS AND DRAFTING OF LESSONS LEARNED

The analytical process transformed raw evidence into a coherent longlist of preliminary lessons. This involved four stages:

- 1. Sense-making:** We reviewed and clustered the 1,587 extracts to identify patterns, recurring themes and contrasts, shifting from descriptive findings to actionable learning.
- 2. Generating the longlist:** We synthesised extracts into clear, evidence-based statements, producing an initial set of 293 lessons learned.
- 3. Consolidation:** We refined the lessons by reducing duplication, sharpening vague statements and ensuring each lesson was evidence-based and traceable. This resulted in 150 draft lessons, of which 82 directly related to drought response. We largely set aside more generic lessons to maintain a focus on drought-specific learning.
- 4. To improve accessibility,** we further refined the core lessons and grouped them into 17 overarching lessons and sub-points.

### STEP 6: DELPHI VALIDATION AND PRIORITISATION OF LESSONS

We used a structured Delphi process to validate and refine the lessons emerging from the evidence review. Objectives of this step were to:

1. validate the relevance, accuracy and completeness of the draft lessons
2. refine the language and framing of lessons for clarity and practical application
3. identify gaps or missing perspectives, especially from underrepresented regions or sectors.

This process involved two rounds:

- 1. Written review:** Panel members reviewed the longlist of lessons and provided written feedback on clarity, accuracy and relevance.
- 2. Virtual discussion:** After revisions based on Round 1 feedback, the panel convened virtually to discuss remaining differences, refine wording and agree on a final prioritised list.

We invited panel members to:

- flag lessons that were unclear, less relevant or insufficiently supported

- suggest improvements or consolidations
- highlight missing lessons
- offer feedback on the structure and usability of the draft paper.

### STEP 7: FINAL REPORT

Following the Delphi validation, we incorporated all feedback into the final synthesis and report drafting. This step involved:

- reviewing and integrating all written comments and discussion points from both Delphi rounds
- revising lesson wording to improve precision and usability
- editing and structuring the full report, including narrative sections, thematic analysis and references
- conducting a final internal review for coherence, clarity and consistency with ALNAP guidelines.

## HOW TO READ THE EVIDENCE SCORES

To help interpret the lessons and their supporting evidence, the following indicators are used throughout this report:

### NUMBER OF DOCUMENTS CONTRIBUTING TO THE LESSON

This shows how many documents were used to identify or support each lesson. A higher number indicates that the lesson is supported by a broader range of sources.

### AVERAGE EVIDENCE SCORE

This represents the average strength or quality of the evidence across all documents contributing to the lesson. It gives an overall sense of how strong the supporting evidence is.

### MEDIAN EVIDENCE STRENGTH

This is the middle evidence score when all contributing documents are ranked from weakest to strongest. It reflects the **typical** strength of the evidence, and is less influenced by very high or very low scores.

Together, these indicators provide a balanced picture of both the breadth and strength of the evidence underpinning each lesson.

## EVIDENCE SCORES AND STRENGTH

- **Evidence score:** The evidence score of each document is calculated as the average of three criteria: evidence relevance, methodological rigour and evidence quality, and evaluation scope. Each criterion is scored out of 5, with 5 representing the highest possible score.
- **Evidence strength:** Evidence strength is calculated as the total score across the three criteria (out of 15), which is then converted into a percentage. This provides a standardised measure of the overall strength of each document, allowing for easier comparison across documents.

**Note:** We scored the evidence based on the reviewers' assessment against pre-agreed criteria developed with ALNAP. Detailed descriptions of the quality assessment criteria are presented in the '[Quality assessment criteria](#)' section below.

### NOTE ON THE DISTRIBUTION OF MEDIAN EVIDENCE-STRENGTH SCORES

A number of lessons have a median evidence strength of 80%. This reflects the structure of the evidence scoring system and the distribution of scores across the evidence base.

Evidence strength is calculated from three criteria (scored from 1 to 5), resulting in discrete percentage values, with 80% (12 out of 15) representing a common outcome for documents demonstrating consistently good performance across relevance, methodological rigour and evaluation scope.

Where a large number of documents support a lesson, the median therefore often falls at this value, reflecting the typical strength of the contributing evidence rather than an absence of variation.

## INCLUSION AND EXCLUSION CRITERIA

When screening the initial database of resources, we applied the following inclusion and exclusion criteria to ensure the selected documents were relevant, high-quality and aligned with our research goals.

### INCLUSION CRITERIA

- **Relevance to drought response:** Documents must focus on drought interventions, evaluations or lessons learned.
- **Evaluation-based:** Documents must include an element of evaluation or impact assessment, or documented lessons from drought-response programmes or policies.
- **Date of evaluation:** The evaluation period must be within the study timeframe (2012–2024).
- **Complete access:** Only full-text documents will be included.

- **Peer-reviewed and grey literature:** Both peer-reviewed articles and credible grey literature (for example, NGO reports and working papers) will be included to maintain a balance between academic rigour and practical field-based evaluations.

### EXCLUSION CRITERIA

- **Irrelevant topics:** Documents that discuss general environmental issues, water resource management or disaster response but do not specifically address drought and humanitarian response.
- **Scientific papers:** Documents that focus on evaluating mathematical models for anticipating droughts.
- **Outdated information:** Documents that were published within the study period but that provide a historical perspective or evaluation period before 2012.
- **Lack of evaluation:** Documents that do not offer evaluations, impact assessments or lessons learned (for example, theoretical papers or news articles).
- **Incomplete or abstract-only:** Documents where full text is unavailable or that only provide abstracts without sufficient detail.
- **Conference papers, theses, presentations, speeches and news articles.**

### QUALITY ASSESSMENT CRITERIA

We used the following quality assessment criteria.

#### EVIDENCE RELEVANCE

This criterion assessed the extent to which each document relates to the various stages of the drought response and our research questions:

- **High score (5):** Documents that provide direct and in-depth analysis or lessons learned on multiple stages of the drought response and explicitly address the research questions.

*Example: A document that thoroughly evaluates a drought-response intervention – covering preparedness, emergency aid, recovery and long-term impact – and directly answers key research questions.*

- **Low score (1):** Documents that only tangentially mention drought response or focus on a narrow aspect unrelated to the main research objectives.

*Example: A document that focuses primarily on water resource management and only briefly mentions drought response in a broader context of environmental planning.*

## METHODOLOGICAL RIGOUR AND EVIDENCE QUALITY

This criterion assessed the robustness of the research design, data collection and analysis methods, along with the overall quality and richness of evidence presented.

- **High score (5):** Documents with strong methodologies (for example, mixed methods or robust qualitative or quantitative approaches), well-triangulated data from multiple credible sources and transparent analysis.

*Example: A report that synthesises data from interviews, surveys and secondary data sources and demonstrates consistency across different data sets.*

- **Low score (1):** Documents with weak designs, limited data or poor evidence quality that rely on single, unverified sources or anecdotal evidence without thorough analysis.

*Example: A document that bases its conclusions on a small, non-representative sample, with no verification from other data sources or cross-checking of information.*

## DOCUMENT SCOPE

This criterion assessed the potential for generalising lessons learned. We weighted documents more heavily that have greater potential for generalisation.

- **High score (5):** Documents with broad scope, covering diverse geographical areas or populations or multiple aspects of drought response, with potential for generalising lessons learned to other contexts.

*Example: A document that evaluates drought response across multiple countries in the Horn of Africa, covering various population groups and stages of the response, providing insights that could be applied globally.*

- **Low score (1):** Documents with narrow scope, limited to a single region or population or focusing on a very specific aspect of drought response, with little potential for wider applicability.

*Example: A document that focuses exclusively on a local community's water intervention without assessing the broader implications for other regions or aspects of drought response.*