RESEARCH ARTICLE

Defining, measuring and interpreting the appropriateness of humanitarian assistance

Nada Abdelmagid* , Francesco Checchi, Sylvia Garry and Abdihamid Warsame

Abstract

This paper presents findings from a literature review of methods that explicitly assess the appropriateness of a humanitarian response. We set out to highlight the key features and limitations of each method and introduce a definition and conceptual framework for the measurement and interpretation of the appropriateness of humanitarian responses. This review is part of a broader project to enhance the accountability of humanitarian responses through developing auditing approaches for real-time monitoring. We identified eight methods that explicitly analyse the appropriateness of a humanitarian response. The review revealed that existing methods vary considerably in their definitions of 'appropriateness', provide insufficient guidance on measurement, are vulnerable to interpretive bias and frequently report findings on 'appropriateness' in an ambiguous manner. These findings suggest that, as a matter of accountability, more structured and systematic approaches to measuring the appropriateness of humanitarian response are needed. We propose a definition and conceptual framework for the measurement and interpretation of the appropriateness of humanitarian response that seeks to address the limitations identified in the review. We provide a brief overview of the main components and features of a systematic approach and audit tool for assessing the 'appropriateness' of a humanitarian response. The use of this and other systematic approaches is essential for enhancing governance and accountability in humanitarian responses.

Keywords: Appropriateness, Evaluation, Assessment, Humanitarian, Emergency

Introduction

Measuring and reporting the appropriateness of humanitarian assistance is a matter of accountability, and is critical in guiding the achievement of impact and value for money. Accountability is a widely adopted concept in the humanitarian sector, and humanitarian actors increasingly regard it as essential to delivering humanitarian assistance responsibly. In the context of affected communities, accountability recognises their dignity, capacities and abilities; for donors and the wider humanitarian community, it is concerned with impactful and quality programming (United Nations High Commissioner for Refugees 2015). Appropriateness is 'the quality of being suitable or proper in the circumstances' (Oxford Dictionaries 2018). In the context of humanitarian assistance, this pertains to the suitability of several factors to the broader crisis context; including a response's objectives, choice of interventions, scale or geographical scope of a response, targeted beneficiaries or cultural acceptability of interventions.

Measuring appropriateness entails defining an acceptable standard; most humanitarian actors would agree that such a standard should be primarily based on the needs of the affected population. However, analysing the scope and scale of need is in itself a complex process. Darcy argues that the humanitarian sector has 'ambiguous and inconsistent' approaches to situation analyses in crises and that a wide range of factors, including the political interests of donors and marketing interests of humanitarian agencies, affect the analysis and presentation of need (Darcy 2003). The result is a lack of consensus and disparate views on which needs should be prioritised and, consequently, what constitutes an 'appropriate' response. This process is complicated further by the dynamic nature of risks and needs in crises.

* Correspondence: Nada.Abdelmagid@lshtm.ac.uk
Faculty of Epidemiology and Population Health, The London School of
Hygiene & Tropical Medicine, London, UK

Criteria other than need also affect the appropriateness of a response, such as the nature of the crisis and the context it occurs in (Darcy and Hofmann 2003). Even where there is a consensus on priority risks, these criteria influence the choice of available interventions and the modality of delivery; in responding to an identified need, the impact of a response on local services, or the coping mechanisms of the affected population may influence humanitarian actors' decisions. For example, an identified need for primary health services in a crisis can be addressed through running mobile clinics, setting up community-based case management, supporting existing health facilities or a combination of modalities. The appropriate choice will be affected by contextual factors, such as availability of the local health workforce, health-seeking behaviour of the population, status of existing health service infrastructure and physical accessibility to the affected community. The decision could also be influenced by the nature of the crisis, for example, whether it is a protracted armed conflict or a rapid-onset natural disaster.

While many approaches to evaluating and assessing humanitarian action exist, it is unclear if and how many methods specifically examine the 'appropriateness' of assistance. Also unclear is how appropriateness is defined and measured by different approaches, and how intended audiences of evaluations interpret findings on appropriateness. This review of the literature seeks to address these gaps in knowledge about measuring and interpreting the appropriateness of humanitarian assistance. Understanding how appropriateness is currently used and defined would be the prequel to proposing a refined definition and a systematic approach to the measurement of appropriateness: an important step to ensure quality, accountability and transparency of humanitarian action.

This review highlights the key features and limitations of each approach and introduces a conceptual framework for the definition, measurement and interpretation of appropriateness of humanitarian responses or projects.

Methodology

As part of a broader project to develop auditing approaches for real-time monitoring of humanitarian response, we reviewed published methods that explicitly analyse the appropriateness of humanitarian response.

We aimed to identify methods that (i) were designed for use in humanitarian settings, and (ii) explicitly assess the appropriateness of one or more aspects of the humanitarian response/intervention.

After exploratory searches, we selected the following title search terms and synonyms: 'appropriate/appropriateness', 'humanitarian/disaster', 'evaluation/audit/real-time review'. Keywords were combined using 'AND' or 'OR' Boolean operators. The search was not limited by language or publication date and was conducted from 10th to 14th October 2018. We did not restrict the search to specific humanitarian sectors, e.g. health, nutrition, shelter, as we sought to identify all generic and sector-specific methodologies, where lessons could be shared across sectors. The search was carried out by the first author (author 1).

Publications were included in the review if they provided information on one or more of the following: (i) the definition of 'appropriateness' according to the method; (ii) a description of how 'appropriateness' is measured; (iii) how results on 'appropriateness' are presented.

We first searched the Medline, PubMed, Web of Science and Google Scholar electronic databases, generating 1057 results. We imported search results into EndNote X9. After removal of 231 duplicates, a review of titles and abstracts for keywords removed an additional 820 results. The full text of the final six results was reviewed, and two publications were included in the review.

The scarcity of relevant results on online electronic databases indicated that the majority of information on the review subject might be available in grey literature and organisational databases. A targeted search using the keywords was thus carried out in the following websites and databases: Active Learning Network for Accountability and Performance (ALNAP), Overseas Development Institute (ODI) and ReliefWeb. The search generated 279 results (59 from ALNAP, 41 from ODI and 179 from ReliefWeb). After removal of one duplicate, a review of titles and executive summaries removed 261 results, and the final 17 publications were included in the analysis.

A further nine publications were identified from the bibliographies of the original set of publications. Where insufficient information on a specific method was available, examples of its use were used to provide additional information on definition, methodology or presentation of findings (note: the publications for these examples are not part of the original search results).

Findings: description of available approaches

A total of eight unique approaches were identified, of which seven were generic approaches and one was specific to the health and nutrition sector. In what follows, we present a synthesis of what the literature reveals about methods for assessing the 'appropriateness' of humanitarian assistance. In this section, we grouped the eight methods into four categories, based on the institutions which developed original definitions for 'appropriateness' or institutions that developed their unique approaches for assessing 'appropriateness'. The four categories are:

- Approaches based on the Organisation for Economic Co-Operation and Development (OECD) standard evaluation criteria (three methods)
- The Core Humanitarian Standard on Quality and Accountability (CHS) and related self-assessment tool (one method)
- Approaches used by the Inter-Agency Standing Committee (IASC) (two methods)
- Other approaches (two methods)

In what follows, we present a description of each method, including the definition of 'appropriateness', assessed aspects of a humanitarian response/intervention, measurement of 'appropriateness' and presentation of findings on 'appropriateness'. We do not present findings on which types of humanitarian agencies use which approaches as such a mapping exercise would require a separate search strategy, which is beyond the scope of this study.

Approaches based on the OECD standard evaluation criteria

The most commonly referenced criteria for evaluation of humanitarian assistance are based on the Guidance for Evaluation of Humanitarian Assistance in Complex Emergencies (Development Assistance Committee 1999), developed in 1999 by the OECD's Development Assistance Committee (DAC). This approach aimed to reduce the 'methodological anarchy' of evaluations of humanitarian assistance funded by the OECD Member States. The approach accounts for the particular complexities of humanitarian response, and defines 'appropriateness' as the 'tailoring of humanitarian activities to local needs, increasing ownership, accountability, and cost-effectiveness accordingly'. The approach describes it as a criterion that complements 'relevance', where 'relevance' refers to the overall goal and purpose of a programme and 'appropriateness' refers to activities and inputs. Other DAC evaluation criteria include efficiency, effectiveness, impact, sustainability/connectedness, coverage and coherence. Below, we describe three inter-agency and sector-wide approaches that use the OECD-DAC's definition of 'appropriateness'.

ALNAP's EHA

Setting a framework for interpretation of OECD-DAC criteria within a humanitarian context, guidelines for evaluation of humanitarian action were published in 2006 (ALNAP 2006) and updated in 2016 (Buchanan-Smith et al. 2016).

The guidelines elaborate further on aspects of 'appropriateness' including:

i.   Needs-based response design
ii. Choice of interventions
iii. Modality of intervention delivery
iv. Participation by the affected population in response design
v.   Design catering for vulnerabilities and capacities of different groups in the affected community
vi. Cultural appropriateness of interventions
vii. Response design informed by gender analysis

In addition to retrospective evaluations, the guidelines can be applied to real-time exercises, to inform decision-making during the response and instigate adaptations to changing conditions (Cosgrave et al. 2009).

Several agency-specific evaluation guidelines and frameworks adopt the OECD-DAC's definition of 'appropriateness' and Evaluation of Humanitarian Action (EHA) guidelines. However, none of them explicitly include all aspects of EHA-defined 'appropriateness'. For example, the International Federation of Red Cross and Red Crescent Societies' (IFRC) Framework for Evaluation (IFRC Secretariat and Planning and Evaluation Department 2011) and the Austrian Development Agency's Guidelines for Project and Programme Evaluations (Evaluation Unit 2009) adopt the OECD-DAC's definition for 'appropriateness' (IFRC Secretariat and Planning and Evaluation Department 2011) but do not offer further explanations on which aspects are evaluated and how to measure them. Other guidelines elaborate on the definition but do not detail or specify components of 'appropriateness' as defined in the EHA guidelines. For example, Action Against Hunger's (ACF) Evaluation Policy and Guidelines explain 'relevance/appropriateness' as whether interventions are not only suited to the needs of the affected population but also donor policies (ACF International 2011). Other guidelines go further to propose evaluation questions or specify the components of 'appropriateness'. Médecins Sans Frontières' (MSF) Evaluation Manual recommends assessing the appropriateness of an intervention against the expressed priorities of the affected population, likelihood of the strategy to achieve the desired objectives and adaptability of the design in response to changes in the environment (Vienna Evaluation Unit 2017). The World Food Programme's (WFP) Technical Note on Evaluation Questions and Criteria adopts the same components as MSF's guidance but also encompasses the extent to which the design and implementation of the intervention is gender-sensitive (World Food Programme 2017). Both MSF and WFP guidelines propose potential key evaluation questions to assess 'appropriateness'. A real-time evaluation of a MSF response to a meningitis outbreak in Niger assessed the appropriateness of (i) MSF resources mobilised to support the response; (ii) emergency

preparedness plans; (iii) choice of vaccination strategy and plan; and (iv) advocacy objectives around the outbreak (Froud 2016). Since the focus of the evaluation was to assess the intersectional and joint coordination between three MSF operational centres simultaneously responding to the outbreak, rather than the impact of the response itself, the choice of strategy was the only aspect of 'appropriateness' assessed. In another example, an evaluation of ACF's response to food and nutrition insecurity in Yobe in north-eastern Nigeria assessed the appropriateness of interventions in relation to the nature and scale of population needs (interpreted by the evaluator as the location and expressed priorities of the targeted population), and the practices/culture of the targeted population (interpreted as beneficiary satisfaction) (Yila 2017). Although the ACF evaluation guidelines do not provide details of the components of appropriateness, the approach by the evaluator addresses the majority of the components of 'appropriateness' as described in the EHA guidelines.

None of the generic and agency-specific guidelines recommend specific data collection methods; however, a qualitative approach is typically adopted, using a combination of methods such as desk reviews, focus group discussions and key informant interviews. While the EHA guidelines provide examples of data collection instruments (Buchanan-Smith et al. 2016), and some agency-specific guidelines propose generic evaluation questions (Vienna Evaluation Unit 2017; World Food Programme 2017), evaluating a response or project usually requires development or adaptation of contextualised questions by the evaluator. For example, an EHA of IFRC's Ebola Response in Sierra Leone and Liberia focussed on the appropriateness of the response to the needs of affected communities, and the appropriateness of the response strategies to the mandate, core capacities and comparative advantage of the IFRC (Ayoo et al. 2018). The IFRC evaluation guidelines do not offer suggested evaluation questions, and this allows the evaluator(s) to interpret 'appropriateness' in line with the evaluation objective and the context of the specific response.

The EHA guidelines and some agency-specific guidelines provide general advice on presenting findings of evaluations (Buchanan-Smith et al. 2016; IFRC Secretariat and Planning and Evaluation Department 2011; Vienna Evaluation Unit 2017), while others provide specific templates for evaluation reports (ACF International 2011; Evaluation Unit 2009). The presentation of findings on 'appropriateness' is usually in a narrative format and is sometimes structured around the OECD-DAC criteria. For example, MSF proposes the presentation of findings around evaluation questions. Where a question addresses one or more of OECD-DAC criteria, findings on 'appropriateness' will be presented in more than one section of a report. The Austrian Development Agency's Guidelines for Project and Programme Evaluations report template suggest the structuring of evaluation findings around the OECD-DAC criteria, rather than the evaluation questions (Evaluation Unit 2009), which allows narrative findings on 'relevance/appropriateness' to be identified easily.

Apart from ACF's Evaluation Policy and Guidelines, none utilise a quantitative measure of 'appropriateness'. In addition to a narrative presentation of findings on 'appropriateness', ACF requires evaluators to provide a rating of 1 (low) to 5 (high) on a Likert scale for each DAC criterion, including 'relevance/appropriateness', with a brief rationale of the rating selected (ACF International 2011). For example, Yila allocated a score of 5 to the relevance/appropriateness of ACF's food and nutrition security ACF response in Yobe, justified by the 'highly-relevant intervention modalities that were used and adequately adapted to the local context' (Yila 2017).

Interagency Health and Nutrition Evaluations in Humanitarian Crises

The guidelines for Interagency Health and Nutrition Evaluations (IHE) in Humanitarian Crises, published in 2007, adapt OECD-DAC criteria to assess the performance of sector-wide humanitarian health and nutrition responses (Interagency Health and Nutrition Evaluations in Humanitarian Crises Initiative 2007). IHE guidelines define 'appropriateness and relevance' as pertains to the choices of, and balance among, various health and nutrition services (i.e. whether the response offered the right package of services offered). Specific areas that are assessed to evaluate the 'appropriateness and relevance' of the response include:

- Whether the top causes of morbidity and mortality formed the basis of interventions
- Whether the response has shown timely adaptability to contextual changes
- Whether the response meets the expressed needs of the population (including cultural appropriateness)
- The extent of participation of the affected community in the design and delivery of the response

The IHE guidelines require all OECD-DAC criteria to be assessed and suggest core questions for each criterion. The questions require further adaptation for each IHE and in-line with the context of the crisis. The guidelines recommend a mixed methods analysis approach,

collecting data through document reviews, interviews with key stakeholders, informal interviews with regional and local actors, attendance at interagency planning meetings, analysis of epidemiological trends using secondary data or spot checks of health facilities. A review of IHE reports (see below) shows a combination of desk reviews with qualitative data collection, but none mentioned conducting epidemiological analysis to identify the top causes of morbidity and mortality.

A review of three IHE reports shows variation in how 'appropriateness' was interpreted and assessed by evaluators. For example, while an IHE in Liberia assessed and reported on the appropriateness of specific interventions (e.g. care for sexual and gender-based violence survivors and access to free health services) (Msuya and Sondorp 2005), a IHE in Chad did not explicitly analyse 'appropriateness' but commented on the 'relevance' of intervention choices and modality of intervention (Michael et al. 2006). Another IHE in Burundi did not refer to the OECD-DAC criteria or indeed to 'appropriateness' (Deboutte et al. 2005).

According to the guidelines, the evaluation framework's components, rather than the OECD-DAC criteria, form the basis for structuring the findings; these are (i) health and nutrition outcomes; (ii) provision of health and nutrition services; (iii) risks to health and nutrition; (iv) health and nutrition sector policy and strategic planning; and (v) humanitarian context. Therefore, findings on 'appropriateness' are present in one or more of these sub-sections. IHE presents findings in a narrative format, and the guidelines do not propose composing quantitative measures of 'appropriateness'. Interestingly, none of the three IHE reports reviewed followed this proposed structuring of findings verbatim. For example, the IHE in Liberia, authors presented findings of 'appropriateness' of specific interventions, each under a separate chapter dedicated to that intervention (Msuya and Sondorp 2005). The IHE in Chad report structured its findings around sections dedicated in humanitarian sectors (e.g. health, nutrition, water and sanitation) and cross-cutting issues, but presented a conclusion section structured around OECD-criteria, with brief statements for each criterion, including relevance (Michael et al. 2006).

Evaluating Humanitarian Innovation
The authors of the working paper on Evaluating Humanitarian Innovation (EHI) argue that not all OECD-DAC criteria are relevant for the evaluation of humanitarian innovation—for example, including criteria such as sustainability when the innovation aims to address a timebound crisis-specific challenge can unnecessarily

complicate the evaluation process (Obrecht et al. 2017). 'Appropriateness', however, is considered a relevant criterion and has a specific application in evaluating an innovation process. In this context, the paper defines 'appropriateness' as the extent to which the innovation responds to a recognised problem. The authors add that 'relevance' complements 'appropriateness' as a measure of the extent to which the innovation not only responds to a clear need but one that is recognised and prioritised by its beneficiaries.

EHI considers the following aspects of an innovation when assessing 'appropriateness': (i) The method by which innovators identified the problem as a need; (ii) the extent to which beneficiaries were able to influence the design of the innovation; and (iii) whether beneficiaries accept the innovation as meeting one of their priority needs. Point (iii) may sometimes be included in relevance rather than appropriateness, and Obrecht et al. acknowledge that there is an overlap between these two concepts in EHI. A case study of a project to improve wheelchairs in emergencies (Thomas and Obrecht 2015) concluded that the wheelchair design was 'appropriate' because innovators recognised the problem through their experience; feedback from partners and wheelchair users guided prototype improvement; and, following Typhoon Haiyan, 86% of users of the new wheelchair users reported that it met their needs (Xavier 2014).

The working paper proposes an overarching framework to help evaluators plan their evaluation, but recommends that evaluators define project-specific methodologies. Similarly, it does not provide suggestions on how findings on 'appropriateness' should be presented, as that would depend on the methodologies used. As seen in the case study above, the findings were reported primarily in narrative format, but also provided a quantitative measure of user satisfaction.

The Core Humanitarian Standard on Quality and Accountability and related self-assessment tool
The Core Humanitarian Standard on Quality and Accountability (CHS) (CHS Alliance 2014) sets out nine commitments (also known as standards) that organisations and individuals involved in humanitarian response should use to improve the quality and effectiveness of their assistance. The first commitment is that 'communities and people affected by crisis receive assistance that is appropriate and relevant to their needs'. According to the CHS, 'appropriate assistance' is defined as one that is based on an impartial assessment of needs and risks and an understanding of the vulnerabilities and capacities of different population groups (e.g. women, men, girls,

boys, youth, older persons, persons with disabilities and specific minority or ethnic groups). The CHS also alludes to the concept of cultural appropriateness (CHS Alliance 2015).

The CHS's self-assessment tool evaluates overall organisational performance in terms of accountability and quality in humanitarian response (CHS Alliance 2016); it is not intended for a response, programme or project evaluation. The tool is structured as follows:

- For each of nine CHS commitments/standards, a set of pre-defined indicators are used to assess the status of each CHS commitment. For CHS commitment 1 on 'relevance/appropriateness', there are eight pre-defined standard indicators.
- The tool solicits opinions of humanitarian staff, affected communities and partners, and, based on feedback, each standard indicator is allocated a score between zero and five in a scoring matrix.
- These indicators collectively reflect the status of three key areas (for each CHS commitment/standard):

i. Organisational policies: existence of policies, guidelines and procedures, and extent of staff awareness of them.
ii. Key actions: translation of principles into practice.
iii. Feedback from communities: perception of the organisation and its interventions by the affected communities.

- Once the scores are entered, the tool automatically computes three percentage scores (each with a maximum of 100%), one for each of organisational polices, key actions and feedback from communities. The tool displays the three percentage scores in a bar chart, but it does not calculate a composite score. The graph allows the audience to discern strengths and weaknesses in 'relevance/appropriateness' in each of the three areas. If repeated periodically, it can be used to track progress over time.

A CHS self-assessment exercise carried out by Christian Blind Mission (CBM) provided a score and justifying narrative for each of six indicators for CHS commitment 1 (two indicators capturing feedback from communities were excluded for unknown reasons) (RED: Agency for Resilience Empowerment and Development 2018). The narrative comments and scoring matrix were included in the report but the bar chart was not. There was no overall narrative for CHS commitment 1. The CHS self-assessment tool and the adapted community scorecard methodology (see below) are the only reviewed approaches that quantify 'appropriateness' and present it in a manner that is easy to interpret by decision-makers, as it can be compared across time and organisations.

Approaches used by the Inter-Agency Standing Committee

The Inter-Agency Standing Committee (IASC) uses complementary instruments to evaluate different aspects and phases of the humanitarian response: The operational peer review (OPR) and the Inter-Agency Humanitarian Evaluations of Large-Scale System-Wide Emergencies (IAHE). Both are triggered at specific points of the humanitarian programme cycle and use qualitative approaches, including desk reviews, consultations with intended beneficiaries and interviews with humanitarian staff. Neither instrument provides quantitative measures of 'appropriateness'.

Operational peer review

The OPR seeks to identify areas for immediate corrective action early in a response (United Nations Office for the Coordination of Humanitarian Affairs 2013). It replaces the IASC's earlier Inter-Agency Real-Time Evaluations (IA-RTE) as a lighter and quicker option for early course-correction. The OPR reviews the processes of response management at the interagency level, including leadership arrangements, coordination mechanisms and the humanitarian programme cycle. The OPR analyses the appropriateness of response coordination mechanisms in the context of the crisis (referred to as focus area 3), but does not critique the appropriateness of the response interventions themselves. Four key evaluation questions are proposed in the guidance for consideration, including whether coordination structures are appropriate to the country context and operational situation.

As with most reviewed approaches, it presents findings in a narrative format, and provides a recommended report template, which includes a dedicated section for focus area 3. In the report of an OPR in Central African Republic in 2014, the reviewers did not comment on the overall appropriateness of coordination structures to the context, but made specific comments and recommendations on different aspects of coordination, including strengthening the inter-cluster coordination group, promoting cross-sectoral collaboration between clusters and increasing investment and attention to coordination hubs outside the capital Bangui (Operational Peer Review Mission Team 2014).

Inter-Agency Humanitarian Evaluations of Large-Scale System-Wide Emergencies

IAHE guidelines assess the appropriateness of the response in relation to the wishes of the affected population(Inter-Agency Standing Committee 2014). The IAHE guidelines do not propose a generic definition of 'appropriateness'—instead, a tailored definition and interpretation for each of the OECD-DAC criteria, including appropriateness, is defined by the evaluators at the start of each IAHE (United Nations Office for the Coordination of Humanitarian Affairs 2018). The method analyses two aspects: the appropriateness of the strategic response plan's objectives in relation to expressed needs, and the appropriateness of the actual services offered. An IAHE of the response in the Central African Republic in 2016 (Lawday et al. 2016) found that 'appropriateness' was one of the weaker aspects of the response, largely because of insufficient engagement of communities in prioritisation, design and delivery of assistance. Strategic response plan objectives were found to insufficiently consider the wishes of displaced populations to return to their areas of origin and progress towards development. The evaluation report noted that measuring 'appropriateness' was complex for various reasons. For example, they reported uncertainty about whether response objectives can be deemed 'appropriate' (or not), as an assessment of the objectives was not sufficiently informative without examining the appropriateness of the overall response strategy and actual services delivered. The evaluators also questioned how far a humanitarian response should go to match the perceived priorities of the affected population, as it was not seeking to replace a health service but rather mitigate the effects of the emergency.

Other approaches
The Independent Commission for Aid Impact
Independent Commission for Aid Impact (ICAI) examines the United Kingdom's (UK) aid spending through independent evaluations. ICAI carries out rapid, performance, impact and learning reviews of UK-funded responses. Rapid reviews aim to provide timely feedback on the appropriateness and effectiveness of a response. We could not identify a specific definition of 'appropriateness' for these evaluations; despite this, rapid review reports do assess 'appropriateness' with regards to the needs and priorities of those worst affected (Independent Commission for Aid Impact 2014). Despite explicitly mentioning 'appropriateness' in the objectives, a rapid review of the UK Government's response to Typhoon Haiyan in the Philippines did not describe the method of assessing 'appropriateness' and did not refer to 'appropriateness'

when reporting the findings (Independent Commission for Aid Impact 2014).

ICAI performance reviews do not explicitly refer to 'appropriateness' as a criterion but do consider appropriate resourcing (funding and staffing) and appropriate targeting of beneficiaries (prioritising populations with the most severe needs) (Independent Commission for Aid Impact 2018). ICAI's approach to assessing effectiveness and value for money also refers to the appropriateness of objectives and plans to achieve the intended impact (Independent Commission for Aid Impact 2011).

The methodologies used in different ICAI reviews are typically qualitative, including desk reviews, consultations with beneficiaries, and interviews with key stakeholders and humanitarian staff. While ICAI uses a traffic light system to rate performance, effectiveness or other aspects of a response (depending on the type of review), it does not explicitly provide a rating or quantitative measure for 'appropriateness'.

Modified community scorecard methodology
In another unique approach, following perennial flooding in Northern Ghana, Apanga et al. used a community scorecard (CSC) methodology to assess the performance of the main responder agency (Apanga et al. 2017). The authors used a modified form of the CSC methodology (Singh and Shah n.d.) with focus group discussions. Performance looked at eight different aspects of the response, one of which was appropriateness of relief items. 'Appropriateness' was defined by affected community members as 'whether relief given is what the victim needed/lost due to disaster'. The researchers held focus group discussions with the communities, and groups allocated a collective score between 1 and 100 for the appropriateness of assistance received. The study found that appropriateness of relief items received was considered inadequate by all communities in the study, because the assistance received was different from what they had requested.

Discussion
Comparison of existing approaches
The above review reveals differences in how 'appropriateness' is defined, measured and interpreted. The methods are not mutually exclusive—in fact, they are sufficiently different so that multiple approaches may be used in the same response or intervention. Of the eight approaches, only one was sector-specific (IHE). Despite these variations, a few features and themes recur (Tables 1 and 2).

Apart from the modified CSC methodology, all approaches apply multiple components to the definition of 'appropriateness'. The modified CSC methodology

is the only approach whereby affected communities define 'appropriateness'. All other approaches rely on the perspective of humanitarian actors, whether implementers, evaluators or donors, though the CHS and IHE do require beneficiary feedback on how far the response meets their needs and how culturally appropriate it is.

As mentioned above, one of the complexities of measuring 'appropriateness' is the difficulty in setting a standard or benchmark. All approaches reviewed assess 'appropriateness' with regards to the needs of the affected population or the broader context of the crisis. However, most of the reviewed approaches do not provide sufficient guidance on assessing or interpreting 'appropriateness'.

The review shows that approaches assess the 'appropriateness' of different aspects of humanitarian assistance. Some focus on response objectives (EHA, IAHE,

ICAI performance review), while others evaluate humanitarian organisations (CHS) or the actual assistance delivered (modified CSC methodology, IAHE). The EHA, IHE, CHS and EHI also evaluate response design (for example, the extent to which affected communities influence the design or the extent to which the design is needs-based).

Some approaches assess the interventional aspects of the response, such as the appropriateness of services received (CHS and modified CSC methodology), the choice of interventions in relation to needs (IHE and EHA), the modality of interventions (ICAI and EHA) and the extent to which the response caters to the needs of specific groups (EHA, CHS and ICAI). Other approaches solely consider innovations (EHI) or response coordination mechanisms (OPR).

Table 1 Main features of approaches measuring the appropriateness of humanitarian assistance

Name of approach	Definition of 'appropriateness'	Methodology	Data collection tools	Format for reporting of findings 'appropriateness'
Using OECD-DAC Criteria: ALNAP's Evaluation of Humanitarian Action and Real-Time Evaluations	The tailoring of humanitarian activities to local needs, increasing ownership, accountability and cost-effectiveness accordingly	Qualitative	No specific tool(s) Requires development or adaptation of contextualised questions by evaluators	Primarily in narrative format (with the exception of ACF who uses a 1–5 Likert scale for rating) Some agency-specific reporting templates Sometimes structured around OECD-DAC criteria
Using OECD-DAC Criteria: Interagency Health and Nutrition Evaluations in Humanitarian Crises	The choices of, and the balance between, various health and nutrition services (i.e. whether the right things were done).	Mixed methods: qualitative and quantitative data	No specific tool(s)	Narrative format No specific template Structured around the evaluation framework's components
Using OECD-DAC Criteria: Evaluating Humanitarian Innovation	The extent to which the innovation responds to a recognised problem	Recommends defining methodologies for each evaluation	No specific tool(s) (depends on methodologies used)	No specific template
Core Humanitarian Standard on Quality and Accountably	A response based on an impartial assessment of needs and risks, and an understanding of the vulnerabilities and capacities of different groups	Qualitative	Tools used should collect information to support scores allocated to pre-defined indicators	Scoring summary for 9 indicators and graphical display (bar chart) of 3 percentage scores for commitment 1 (relevance/appropriateness)
IASC: Operational Peer Review (OPR)	No specific definition for 'appropriateness'	Qualitative	No specific tool(s)	Narrative format Recommended reporting template ('Appropriateness' findings under section for Focus Area 3)
IASC: Inter-Agency Humanitarian Evaluations of Large-Scale System-Wide Emergencies	A context-specific definition of 'appropriateness' is developed by evaluators at the start of each evaluation	Qualitative	No specific tool(s)	Narrative format Recommended reporting template (no dedicated section for 'appropriateness' findings)
UK's Independent Commission for Aid Impact Rapid and Performance Reviews	No specific definition for 'appropriateness'	Qualitative	No specific tool(s)	Narrative format No specific template Uses a traffic light system to rate aspects of a response, but not 'appropriateness'
Adapted community scorecard methodology	Whether relief provided is what the victim needed/ lost due to disaster	Qualitative	Modified community scorecard	Quantitative: score between 1 and 100 for 'appropriateness'

Some approaches use 'appropriateness' and 'relevance' synonymously while others draw a clear distinction between them and consider them complementary to each other

Table 2 Aspects of 'appropriateness' considered by different evaluation approaches

Approach	Response is based on an impartial and comprehensive needs assessment	Response caters for the vulnerabilities and capacities of different groups	Response engages and ensures participation of affected communities	Response meets the expressed needs of affected communities	Response is culturally acceptable to the affected community	Response uses the appropriate modality of intervention	Other
Using OECD-DAC Criteria: ALNAP's Evaluation of Humanitarian Action (EHA) and Real-Time Evaluations (RTE)	✓	✓	✓		✓	✓	Response has appropriate choice and balance of interventions Response is based on gender analysis
Using OECD-DAC Criteria: Interagency Health and Nutrition Evaluations in Humanitarian Crises (IHE)	✓		✓	✓	✓		Response shows timely adaptability to a changing context
Using OECD-DAC Criteria: Evaluating Humanitarian Innovation (EHI)	✓		✓	✓			
Core Humanitarian Standard on Quality and Accountably (CHS)	✓	✓	✓	✓	✓		Response builds on local capacities in affected communities
IASC: Operational Peer Review (OPR)							Response coordination mechanisms are appropriate, given the context
IASC: Inter-Agency Humanitarian Evaluations of Large-Scale System-Wide Emergencies (IAHE)				✓			Services offered by the response are appropriate
UK's Independent Commission for Aid Impact (ICAI) Rapid and Performance Reviews		✓				✓	Response objectives are appropriate to the context Response is appropriately-resourced
Modified community scorecard (CSC) methodology				✓			

The majority of methods reviewed adopt a qualitative approach to collecting data and information, except for IHE which recommends using quantitative and qualitative data. The CHS and modified CSC approaches are the only approaches that generate quantitative or numerical scores on 'appropriateness'.

Limitations of current approaches
An important limitation is the absence of or broad definitions of 'appropriateness'. There is considerable variation in terms of the aspects considered, and evaluation reports show diverse interpretations of 'appropriateness'. Even the most widely utilised definition of 'appropriateness' (that of the OECD-DAC evaluation criteria) had not been interpreted in a consistent manner by evaluators. In several of the reviewed approaches, guidelines tend to provide generic guidance that will apply to all contexts and responses and therefore avoid being overly rigid or prescriptive. While this has its advantages, it does not sufficiently address the persistent knowledge gap in assessing the 'appropriateness' of a given humanitarian response. The inconsistent interpretation and presentation of findings on 'appropriateness' prevent both assessment of changes in 'appropriateness' of a given response over time, as well as comparison between humanitarian responses.

Another fundamental limitation of current approaches to measuring 'appropriateness' is that they primarily rely on qualitative information and the

evaluators' judgment. As a result, they are vulnerable to bias, depending on information availability, knowledge and expertise of evaluators and the perspective of both the evaluators and those who contribute to the evaluation.

The review shows that, with few exceptions, approaches primarily present findings on 'appropriateness' in a narrative format. Furthermore, with few exceptions, findings related to 'appropriateness' are not explicitly reported, and must be inferred from evaluation reports. This problem can partly be attributed to the number and breadth of key evaluation questions on 'appropriateness', driven by different interpretations by both evaluators and respondents, which then affect analysis and presentation of findings. The method of presenting or displaying information to humanitarian decision-makers can be crucial to its uptake and use. Humanitarian practitioners prefer information to be presented concisely and in a format that is easy to understand, especially by non-technical decision-makers (Darcy et al. 2013). For example, CHS and modified CSC approaches provide graphic displays or quantitative measures which are easier to decipher than narrative findings.

One way of preserving objectivity is favouring an external approach to project evaluation. However, there also remains a clear gap in the form of self-assessment tools to compliment external evaluations with which response staff can critically self-reflect on their work. While most external approaches attempt to overcome this by requiring the participation of humanitarian staff, not all humanitarian teams will have the capacity to engage fully with evaluations or to make use of the findings (Hallam and Bonino 2013). Approaches such as the CHS self-assessment tool however, encourage inward reflection, which adds a complementary dimension to external evaluations and may enable more systematic 'appropriateness' evaluations at the project level.

Another limitation, particularly relevant in large responses, is that response-wide inter-agency evaluations rarely consider different geographical areas separately (e.g. districts or camps) or delve into detail on the appropriateness of specific interventions, modalities of delivery or choice of targeted beneficiaries. Reports of such evaluations rarely provide information and recommendations on specific projects or sectors (e.g. health, education). For example, while the IHE is a sector-specific response evaluation, its reports may not provide actionable findings for each of the actors involved in the response.

A proposed definition and conceptual framework

The first step in this process is to refine the definition of 'appropriateness', and then outline a more systematic measurement process. Building on the strengths of existing approaches, and addressing the challenges identified in this review, we propose a definition and conceptual framework (Fig. 1) for assessing the 'appropriateness' of humanitarian assistance.

The framework is based on the premise that the appropriateness of a response or intervention is determined by the extent to which it is designed to save
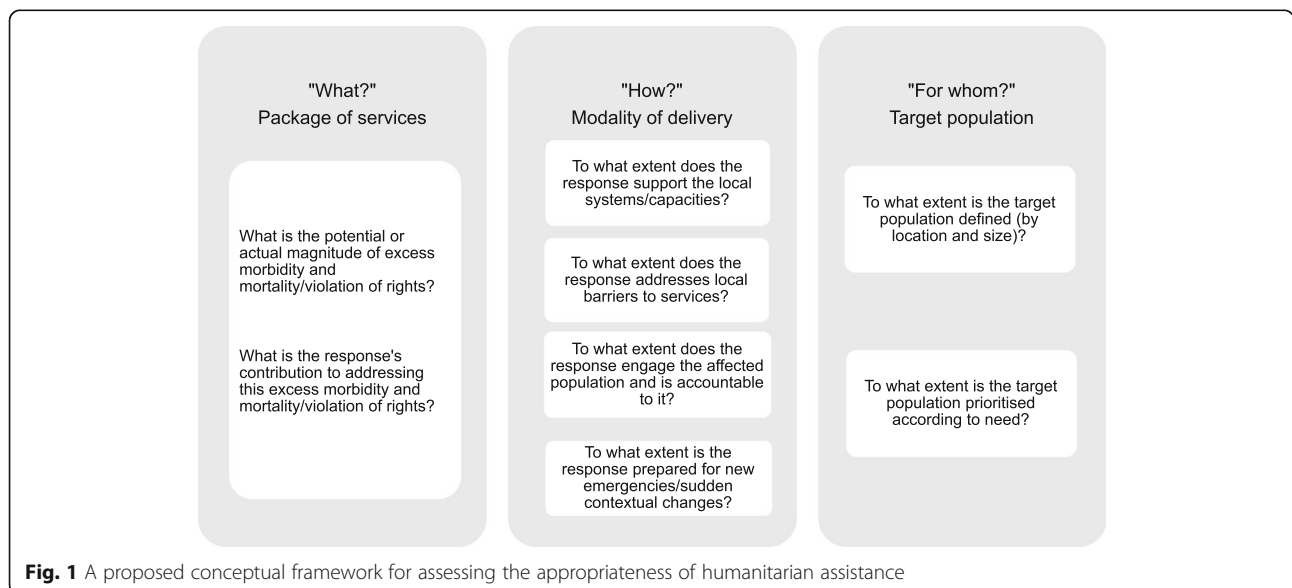

Fig. 1 A proposed conceptual framework for assessing the appropriateness of humanitarian assistance

lives, alleviate suffering and maintain human dignity. Similarly, a protection intervention's appropriateness is determined by the extent to which it is designed to prevent and address violations of individuals' rights and ensure and promote respect for laws related to human rights, international humanitarian assistance and refugees.

We define 'appropriate humanitarian assistance' as a combination of (i) an intervention/package of services that addresses objective needs and threats to the health or welfare of crisis-affected populations; (ii) a modality of delivery that reflects the context, enhances user acceptability and promotes sustainability where possible; and (iii) a target beneficiary population that is clearly defined, sufficient in size and prioritised according to need.

In the framework, we propose a specific set of questions relating to the 'what/how/to whom' domains of a humanitarian project or response. The domains reflect the three components of the proposed definition: appropriateness of the interventions, the modality of the delivery and the targeted population. Although we acknowledge that additional or alternative questions could be considered, we hypothesise that the proposed ones are the most critical for the three domains. The main output of the methodology is a semi-quantitative scorecard that provides a score for each of the questions/domains, accompanied by a brief narrative contextualisation of the findings, e.g. constraints posed by sectoral capacity, resource availability and the external environment (e.g. security)

In theory, this design caters for any crisis setting and any potential humanitarian response and would be applicable at the level of an individual project, and single organisation or an inter-agency coordinated response. The framework does not make assumptions about the scale or scope of the response and so data at project, agency or response level can be used to answer the questions posed by the methodology.

Figure 2 shows how the conceptual framework can be adapted to the humanitarian health and nutrition sector. Based on this adaptation, we are currently developing a data collection tool and operational guidance to test the method in a number of ongoing responses. Additional adaptation of the framework would be required for other humanitarian sectors (e.g. protection, food security, livelihoods, shelter, education).

The approach is designed for self-assessment by response teams for early course-correction and seeks to simplify the measurement and interpretation of 'appropriateness' by decision-makers. We propose that humanitarian actors conduct regular real-time 'appropriateness audits' to enable comparability over time and to inform adaptations to contextual changes. The method is designed to be a lightweight exercise, with a target implementation period of less than a month, from start to finish, including reporting.

Our ultimate aim is to use this approach to enhance governance and accountability in humanitarian response, and an important feature of this audit tool is the promotion of transparency. The authors encourage users of this approach to make results publicly available and to embed the method into organisational governance and accountability processes—particularly those aimed towards increasing impact and value for money.

Fig. 2 The conceptual framework adapted for the health and nutrition sector
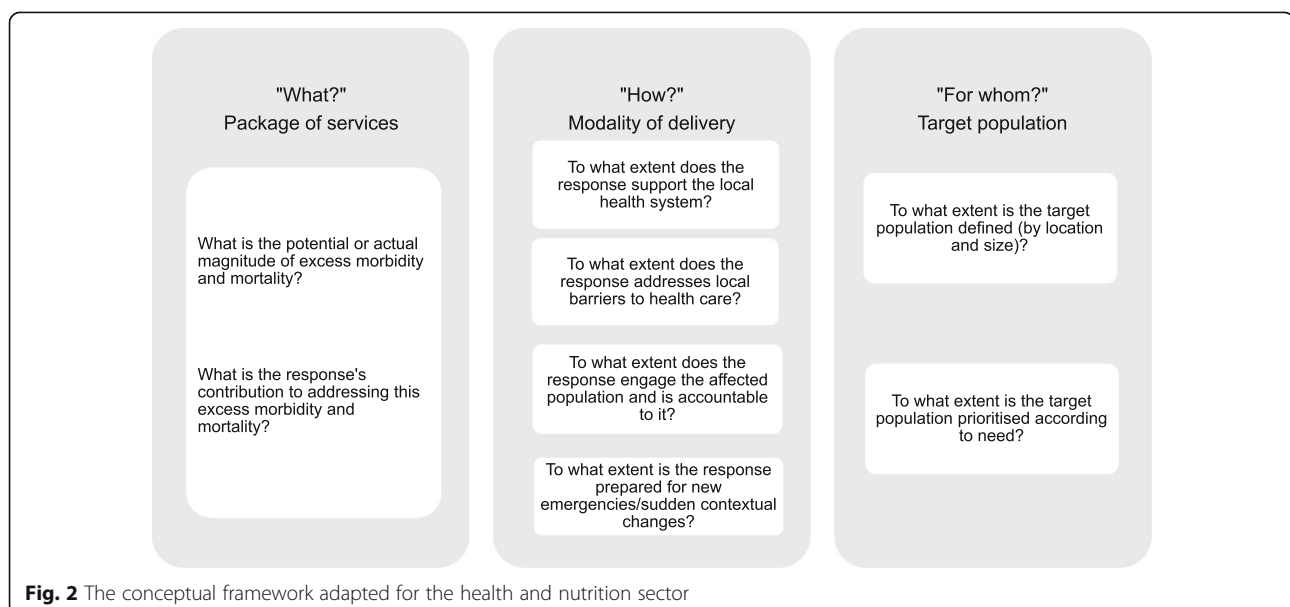
Abbreviations

ACF: Action Against Hunger/Action contre La Faim; ALNAP: Active Learning Network for Accountability and Performance; CBM: Christian Blind Mission; CHS: Core Humanitarian Standard on Quality and Accountability; CSC: Community scorecard; DAC: Development Assistance Committee; EHA: Evaluation of Humanitarian Action; EHI: Evaluating Humanitarian Innovation; IAHE: Inter-Agency Humanitarian Evaluations of Large-Scale System-Wide Emergencies; IA-RTE: Inter-Agency Real-Time Evaluations; IASC: Inter-Agency Standing Committee; ICAI: Independent Commission for Aid Impact; IFRC: International Federation of Red Cross and Red Crescent Societies'; IHE: Interagency Health and Nutrition Evaluations; MSF: Médecins Sans Frontières; ODI: Overseas Development Institute; OECD: Organisation for Economic Co-Operation and Development; OPR: Operational peer review; WFP: World Food Programme

References

ACF International (2011) Evaluation Policy & Guideline. ACF International https://www.alnap.org/help-library/acf-evaluations-policy-and-guidelines . Accessed 7 Dec 2018

ALNAP (2006) Evaluating humanitarian action using the OECD-DAC criteria. Overseas Development Institute https://www.alnap.org/system/files/content/resource/files/main/eha-2006.pdf. Accessed 7 Dec 2018

Apanga S, Titi Addebah G, Chirawurah D (2017) Listening to the voices of the people: community's assessment of disaster responder agency performance during disaster situations in rural Northern Ghana PLoS. Curr 9. https://doi.org/10.1371/currents.dis.4226abe816b2746df13d16ea307b5846

Ayoo S, Njigua L, Roiron V (2018) Evaluation of the IFRC West Africa ebola viral disease appeal response: Sierra Leone And Liberia. International Federation of Red Cross and Red Crescent https://www.alnap.org/system/files/content/resource/files/main/IFRC%20EVD%20final%20evaluation%20report.pdf. Accessed 7 Dec 2018

Buchanan-Smith M, Cosgrave J, Warner W (2016) Evaluation of humanitarian action guide. ALNAP https://www.alnap.org/system/files/content/resource/files/main/alnap-evaluation-humanitarian-action-2016.pdf. Accessed 6 Dec 2018

CHS Alliance (2014) Core humanitarian standard on quality and accountability. https://corehumanitarianstandard.org/files/files/Core%20Humanitarian%20Standard%20-%20English.pdf. Accessed 6 Dec 2018

CHS Alliance (2015) CHS Guidance Notes and Indicators. https://corehumanitarianstandard.org/files/files/CHS-Guidance-Notes-and-Indicators.pdf. Accessed 6 Dec 2018

CHS Alliance (2016) Self-assessment tool: core humanitarian standard on quality & accountability. CHS Alliance https://reliefweb.int/sites/reliefweb.int/files/resources/CHS_Alliance-Self-Assessment_Tool_v1_3-06July2016.pdf. Accessed 6 Dec 2018

Cosgrave J, Ramalingam B, Beck T (2009) Real-time evaluations of humanitarian action: An ALNAP Guide (Pilot Version). Overseas Development Institute https://www.alnap.org/system/files/content/resource/files/main/rteguide.pdf. Accessed 6 Dec 2018

Darcy J (2003) Measuring humanitarian need: a critical review of needs assessment practice and its influence on resource allocation. Overseas Development Institute - Humanitarian Policy Group https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/4899.pdf. Accessed 6 Dec 2018

Darcy J, Hofmann C (2003) According to need? Needs assessment and decision-making in the humanitarian sector. Overseas Development Institute - Humanitarian Policy Group https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/285.pdf. Accessed 6 Dec 2018

Darcy J, Stobaugh H, Walker P, Maxwell D (2013) The use of evidence in humanitarian decision making: ACAPS operational learning paper. Feinstein International Center https://www.alnap.org/system/files/content/resource/files/main/tufts-1306-acaps-3-online.pdf. Accessed 7 Dec 2018

Deboutte D, Mach O, Rossi L (2005) Interagency health evaluation burundi—health and nutrition in the humanitarian context. United Nations High Commissioner For Refugees (UNHCR) https://www.alnap.org/help-library/interagency-health-evaluation-burundi-health-and-nutrition-in-the-humanitarian-context. Accessed 7 Dec 2018

Development Assistance Committee (1999) Guidance for evaluating humanitarian assistance in complex emergencies. Organisation for Economic Co-operation and Development https://www.oecd.org/dac/evaluation/2667294.pdf. Accessed 7 Dec 2018

Evaluation Unit (2009) Guidelines for project and programme evaluations. Austrian development agency https://www.entwicklung.at/fileadmin/user_upload/Dokumente/Projektabwicklung/Englisch/Guidelines_for_Project_and_Progamme_Evaluations.PDF. Accessed 7 Dec 2018

Froud A (2016) "Strengths in unity": a real time evaluation of an MSF collaborative approach to a meningitis outbreak in Niger. Médecins Sans Frontières: Stockholm Evaluation Unit. https://www.alnap.org/help-library/%E2%80%9Cstrengths-in-unity%E2%80%9D-a-real-time-evaluation-of-an-msf-collaborative-approach-to-a. Accessed 7 Dec 2018

Hallam A, Bonino F (2013) Using evaluation for a change: insights from humanitarian practitioners. ALNAP/ODI https://www.alnap.org/system/files/content/resource/files/main/alnap-study-using-evaluation-for-a-change.pdf. Accessed 7 December 2018

IFRC Secretariat, Planning and Evaluation Department (2011) IFRC Framework for Evaluation https://www.ifrc.org/Global/Publications/monitoring/IFRC-Framework-for-Evaluation.pdf. Accessed 6 Dec 2018

Independent commission for aid impact (2011) ICAI's approach to effectiveness and value for money https://icai.independent.gov.uk/wp-content/uploads/ICAIs-Approach-to-Effectiveness-and-VFM2.pdf. Accessed 7 Dec 2018

Independent Commission for Aid Impact (2014) Rapid review of DFID's humanitarian response to typhoon Haiyan in the Philippines. https://icai.independent.gov.uk/wp-content/uploads/ICAI-Philippines-report-FINAL.pdf. Accessed 7 Dec 2018

Independent Commission for Aid Impact (2018) The UK's humanitarian support to Syria: a performance review. https://icai.independent.gov.uk/wp-content/uploads/ICAI-Syria-Report-final.pdf. Accessed 6 Dec 2018

Interagency health and nutrition evaluations in humanitarian crises initiative (2007) Guidelines for implementing IHE in humanitarian crises. https://www.unhcr.org/47692e3b4.pdf. Accessed 6 Dec 2018

Inter-Agency Standing Committee (2014) Inter-agency humanitarian evaluations of large scale system-wide emergencies (IAHE). https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/IAHE%20Guidelines.pdf.pdf. Accessed 6 Dec 2018.

Lawday A, Adjibade K, Dahrendorf N, Kubwayezu F, Morinière LC (2016) Inter-agency humanitarian evaluation of the response to the crisis in the Central African Republic. Steering Group for Inter-Agency Humanitarian Evaluations. https://interagencystandingcommittee.org/evaluations/content/inter-agency-humanitarian-evaluation-response-crisis-central-african-republic. Accessed 7 Dec 2018

Michael M, Pearson N, Daliam A (2006) Interagency health evaluation—Chad. United Nations High Commissioner For Refugees (UNHCR). https://www.alnap.org/system/files/content/resource/files/main/erd-3393-full.pdf. Accessed 6 Dec 2018

Msuya C, Sondorp E (2005) Interagency health evaluation—Liberia. United Nations High Commissioner For Refugees (UNHCR). https://www.alnap.org/system/files/content/resource/files/main/erd-3392-full.pdf. Accessed 6 Dec 2018

Obrecht A, Warner A, Dillon N (2017) Working paper: evaluating humanitarian innovation—a HIF/ALNAP working paper. ALNAP/ODI https://www.alnap.org/system/files/content/resource/files/main/alnap-hif-evaluating-humanitarian-innovation-2017.pdf. Accessed 6 Dec 2018

Oxford Dictionaries (2018) Definition of *appropriateness* in English. https://en.oxforddictionaries.com/definition/appropriateness. Accessed 7 Dec 2018.

RED: Agency for resilience empowerment and development (2018) CBM: core humanitarian standards self-assessment final report. https://www.cbm.org/fileadmin/user_upload/CHS_self_Assessment_report_Executive_Summary_and_scoring_190618.pdf. Accessed 12 July 2019

Singh J, Shah P (n.d.) Community score card process—a short note on the general methodology for implementation. World Bank, Social Development Department. Retrieved from http://siteresources.worldbank.org/INTPCENG/1143333-1116505690049/20509286/comscorecardsnote.pdf. Accessed 12 Dec 2018

Thomas M, Obrecht A (2015) Supporting disabled people in emergencies: motivation's appropriate and affordable wheelchairs. HIF/ALNAP Case Study. ODI/ALNAP. https://www.elrha.org/wp-content/uploads/2017/05/hif-alnap-motivation-wheelchairs-case-study-2015.pdf. Accessed 7 December 2018

United Nations High Commissioner for Refugees (2015) Accountability to affected populations (AAP). https://emergency.unhcr.org/entry/42554/accountability-to-affected-populations-aap. Accessed 12 July 2019

United Nations Office for the Coordination of Humanitarian Affairs (2013) Operational Peer Review Guidance. The Interagency Standing Committee (IASC). https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/Working%20Draft_Guidance_Operational%20Peer%20Review%2017%20December%202013%20v2.pdf. Accessed 6 Dec 2018

United Nations Office for the Coordination of Humanitarian Affairs (2018) Inter-agency humanitarian evaluations: process guidelines. https://interagencystandingcommittee.org/system/files/iahe_guidelines_2018.pdf. Accessed 12 July 2019

Vienna Evaluation Unit (2017) Evaluation Manual: A handbook for initiating, managing and conducting evaluations in MSF. Médecins Sans Frontières (MSF). https://evaluation.msf.org/sites/evaluation/files/evaluation_manual_online_version_2017.pdf. Accessed 6 Dec 2018

World Food Programme (2017) Evaluation questions and criteria. https://docs.wfp.org/api/documents/258036f37ecb4a17af7ea8afd212b0f1/download/. Accessed 7 Dec 2018

Xavier CA (2014) Impact assessment: emergency wheelchair response in the Philippines. Elrha https://www.elrha.org/wp-content/uploads/2015/01/Impact-AssessmentReport_TX-140730-1.pdf. Accessed 7 Dec 2018

Yila O (2017) Protecting and promoting the food and nutrition security in the Yobe State, Phase II, North Eastern Nigeria. ACF International https://reliefweb.int/sites/reliefweb.int/files/resources/201708-nigeriab3c-final-evaluation-report_0.pdf. Accessed 7 Dec 2018

Publisher's Note