ABSTRACT

Objectives To systematically document measurement approaches used in the monitoring and evaluation of gender-based violence (GBV) risk mitigation activities, categorise the types of available literature produced by sector, identify existing tools and measures and identify knowledge gaps within the humanitarian sector.

Design Systematic mapping and in-depth review.

Data sources: PubMed, Global Health, PsychInfo, ReliefWeb, OpenGrey (grey literature), Google Scholar, Web of Science (Social Science Index)

Eligibility criteria: a structured search strategy was systematically applied to 17 databases as well as registers, websites and other resources to identify materials published between 1 January 2005 and 15 May 2019.

Data extraction and synthesis: Those resources that met the inclusion criteria underwent a comprehensive full-text review. A detailed matrix was developed and key data from each resource were extracted to allow for the assessment of patterns in thematic areas.

Results: A total of 2108 documents were screened. Overall, 145 documents and 112 tools were reviewed, representing 10 different humanitarian sectors. While numerous resources exist, many lack sufficient information on how to monitor outputs or outcomes of GBV risk mitigation activities. There is also limited guidance on how to integrate the measurement of GBV risk mitigation into existing monitoring and evaluation frameworks. Those reports that aimed to measure GBV risk mitigation activities mostly employed qualitative methods and few measured the impact of a GBV risk mitigation with robust research designs.

Conclusions Recent efforts to adapt humanitarian response to COVID-19 have highlighted new and existing challenges for GBV risk mitigation. There is a significant gap in the evidence base around the effectiveness of GBV risk mitigation across all sectors. Understanding and strengthening measurement approaches in GBV risk mitigation remains a critical task for humanitarian response.

INTRODUCTION

Despite efforts to ensure that humanitarian programming is safe, effective and responsive to all affected populations, response efforts can unintentionally increase risks of gender-based violence (GBV). In response to this challenge, a number of actors in the humanitarian community have sought to build a discourse and area of practice to integrate risk mitigation related to GBV into programming—referred to as ‘risk mitigation’ or ‘GBV risk mitigation’ for the remainder of the article. This effort is intended to help anticipate, identify and address GBV risks associated with humanitarian programmatic interventions. A number of GBV risk mitigation resources and guidelines have been created to address this issue. GBV risk mitigation aims to reduce GBV-related risks in the immediate environment as well as those directly linked to humanitarian programming.
and service delivery, while simultaneously strengthening women and girls’ safety and resilience. Concretely, this means ensuring humanitarian service delivery: does not increase the likelihood of GBV occurring; seeks to identify and mitigate GBV risks and conducts ongoing monitoring of access and barriers to services, particularly those faced by women and girls.11–13 This is distinct from GBV prevention interventions, which seek to address the root causes of GBV in the longer term by empowering women and girls economically and socially, supporting legal and policy reform and transforming harmful social norms.

In 2005, the Guidelines for Integrating Gender-Based Violence Interventions in Humanitarian Action (‘IASC GBV Guidelines’) were endorsed by the Inter-Agency Standing Committee and released by a global reference group as the first, seminal guidance for humanitarian programming on how to reduce and mitigate the risk of GBV in emergency settings.11 The guidelines are divided into 13 sections, 1 for each of the following humanitarian sectors: Camp Coordination and Camp Management (CCCM); Child Protection; Education; Food Security and Agriculture; Health; Housing, Land and Property; Humanitarian Mine Action; Livelihoods; Nutrition; Protection; Shelter, Settlement and Recovery and Water, Sanitation and Hygiene (WASH). The 2015 revision of the IASC GBV Guidelines was led by UNICEF in partnership with a 15-member reference group. This updated document provides global guidance for all humanitarian sectors on how to implement, monitor and evaluate actions to mitigate GBV in emergency settings.12

With this published resource, and increased awareness and training, progress has been made to integrate GBV risk mitigation throughout the humanitarian programme cycle and across sectors.4 14 However, a significant gap remains in how to measure the effectiveness of GBV risk mitigation in terms of improving safety for women and girls and contributing to sectors’ own outcomes (eg, reducing malnutrition, increasing school attendance/retention for girls etc.).15

This question of measurement is made challenging by a number of considerations. A key challenge is that GBV risk mitigation is intended to be implemented across all aspects of humanitarian response, including by those sectors that do not have specialised knowledge related to the protection of women and girls in humanitarian crises or GBV issues. In many cases, humanitarian practitioners may not feel that they have the specific skills or capacity to collect, analyse or interpret data on such a sensitive topic. Second, measurement of outcomes related to GBV risk mitigation deals with assessing counterfactuals—specifically, would GBV risk be higher in the absence of mitigation strategies? This is a challenging methodological problem even for measurement specialists and requires technical knowledge and appropriate, valid methods. Furthermore, analysis of this data and interpretation of its relevance to programmatic outcomes requires training and capacity-building. Finally, assessing risk of GBV—even through proxy measures such as women’s perceptions of safety—can be challenging. First, there is no standard definition of what safety comprises within the context of GBV risk mitigation. In addition, asking these questions has the potential to induce disclosure of negative experiences, including GBV, and could cause respondent distress. For this reason, assessments or routine monitoring efforts with questions on safety perceptions should be conducted with support and input from trained GBV specialists. Thus, even while initial intentions may be to follow recommended guidance like the IASC GBV Guidelines, a lack of capacity or availability of GBV specialists can contribute to ineffective measurement, unintended consequences and further harm.

Yet, understanding whether GBV risk mitigation is effective and has contributed to better outcomes for affected communities is vital for creating a more inclusive, safe and effective humanitarian response. There is limited evidence on whether effective GBV risk mitigation can improve GBV-related outcomes as well as how it may contribute to better sector-specific outcomes. For instance, when nutrition services are built to ensure the safety of the women and girls who access them, this not only helps mitigate the risk of GBV but may also improve access to and use of services and, eventually, nutrition-related outcomes as well.13 Latrines in refugee camps that are well-lit, sex-specific and have working locks not only promote safety for users but also advance better health and sanitation.13 Strengthening measurement approaches are critical to better assess effectiveness of GBV risk mitigation strategies.

As a first step in addressing this measurement gap, this paper presents a comprehensive desk review to document the current knowledge related to measurement and evaluation of GBV risk mitigation activities. The objectives of this study were to systematically identify and document the range of measurement approaches being used across the humanitarian field to monitor and evaluate GBV risk mitigation activities, categorise the types of available literature produced by sector, identify existing tools and measures and identify knowledge gaps.

METHODS

Methods

This study, undertaken between March and December 2019, used a systematic mapping and in-depth review approach to inform its findings. The methodology for systematic mapping was originally developed by the Evidence for Policy and Practice Information and Co-ordinating Centre and is increasingly used in the social sciences as a means of characterising the evidence base and identifying knowledge gaps.16–20 A systematic mapping review aims to collate and describe the available evidence to identify the type of results available within a given topic. As a result, systematic mapping is much broader in scope than a systematic review. Systematic mapping is particularly valuable when synthesising evidence for policy and practice-based questions due to the breadth of research.
Inclusion and exclusion criteria

The inclusion criteria were defined in consultation with the study’s TAG. Potentially eligible studies and documents included journal articles, published and unpublished reports, web-based guidelines and grey literature published in English since 2005. We considered all geographical regions, sectors and phases of humanitarian response to make the search comprehensive. Data were obtained through formal database searches as well as through an open call for resources from humanitarian professionals, as described in the following section.

As previously noted, systematic maps are less limited in the types of evidence that may be included than systematic reviews and, therefore, can include a wider range of primary and secondary research. The critical appraisal stage in mapping reviews is also optional for this same reason. Hence, this study did not formally appraise the quality of included studies. This is consistent with standard systematic mapping review methodology. In addition, multiple references that were based on the same sample were not excluded, as would have been done to minimise bias in a systematic review. The inclusion criteria used to screen the literature are illustrated in table 1. Although the inclusion criteria were explicitly defined for most key elements, some elements were also iteratively included during the review (eg, in the GBV Risk Mitigation column in table 1). This study chose to retain in the analysis prevention-focused articles that drew lessons applicable to risk mitigation. Similarly, only publications published after 2005 were considered since this was when the IASC GBV Guidelines were first rolled out.

Search strategy

A structured search was undertaken of electronic databases, registers and websites to identify existing peer-reviewed articles, grey literature and other sources that would be relevant to this review. As much relevant information was not expected to be available in online databases, a public call for information and papers was also widely disseminated to public health and humanitarian professional networks and communities, including key organisations and experts, to gather additional relevant tools and reports. This dissemination was done through social media channels as well as humanitarian email listservs.

An explicit search strategy defined in the protocol was systematically applied to 17 databases and other resources to identify materials published between 1 January 2005 and 15 May 2019. Since the aim was to widely capture existing approaches and tools, both published and unpublished literature as well as specialist websites and electronic databases were included in the search process.

Table 2 provides the full list of databases, registries and websites that were searched. The specific search criteria included terms related to GBV, humanitarian response, GBV risk mitigation or safety, programme utilisation,
measurement, assessment and evaluation. Table 3 provides the final search terms used. GBV terms, humanitarian response terms, risk mitigation and risk mitigation assessment terms were combined through “AND” searches. The two types of risk mitigation were combined through “OR” searched; the three types of risk mitigation assessment terms were also combined through “OR” searches. Both British and American spellings of key words were searched (eg, both ‘program’ and ‘programme’ were included in the search terms). Additional materials from the open call for resources submitted until 3 October 2019 were included in the study and combined with the electronic search results. Duplicate references were excluded. An example of the full search strategy for OVID is provided as an online supplemental table 1.

**Screening process**

A title and abstract review of all resources that were identified through the above channels was carried out. Each resource was assessed against specific eligibility criteria in order to determine inclusion in the desk review. Where inclusion or exclusion could not be determined on the basis of title and abstract, the full text was screened. At the full-text screening stage, any items that a team member considered borderline or problematic were noted, and the team decided together whether the resource should...
be included. When questions remained, decisions were made in favour of an inclusive approach.

**Data extraction and analysis**

Those resources that met the inclusion criteria underwent a comprehensive full-text review. A detailed matrix was developed and key data from each resource including the tools were extracted into the matrix to allow for assessment of patterns in key thematic areas. The extraction themes were sector, objectives of the document, data type, risk mitigation activities described, approach to measurement, tools employed, measurement challenges, good practices, recommendations and additional notes. Two researchers independently coded the key variables for each document (sector, inclusion of GBV risk mitigation information, inclusion of measurement information) and these were then compared. Any disagreements or uncertainties were resolved through discussion. As the data were synthesised, the matrix enabled the research team to draw conclusions about the state of knowledge and the existence of current practices to address the research questions outlined above. A complete list of articles and tools included in the desk review is included as online supplemental table 2, with the exception that some documents in this list were obtained from third parties under the condition that they not be shared more broadly.

**Patient and public involvement**

Patients were not involved in this study. As previously mentioned, this study was reviewed by the study’s TAG. A public call for information and papers was also disseminated to public health and humanitarian professional networks and communities to assist the structured search for relevant tools and reports.

**RESULTS**

Overall, searches of databases, grey literature and the web, as well as materials shared through the open call for materials, yielded a total of 2108 documents for formal screening. These 2108 documents included 1788 peer-reviewed journal articles, 299 documents from the grey literature and 21 resources submitted through the open call for materials (see figure 1).

The majority of items reviewed were not relevant for the desk review. Of the 311 documents that met the inclusion criteria during the initial abstract and title search, 145 documents met the final inclusion criteria. Of the 145 documents, 5 were toolkits, 32 were reports, 11 were journal articles, 17 were guidance documents, 7 were case studies, 26 were research reports or evaluations, 33 were standalone tools and 14 were classified as other (see figure 2). In some cases, certain documents such as toolkits comprised multiple separate documents that were packaged together as one resource. In such cases, we counted the document as one resource. Of note is the Violence, Gender and Water, Sanitation and Hygiene (WASH) Practitioners Toolkit, which comprises of over 50 standalone toolsets, briefs and other resources that were developed by different organisations. We did assess each of these components separately within the matrix. In our document totals, we have counted each toolkit as one resource to avoid skewing the analyses. However, if we were to count each component of this toolkit separately, then the total number of documents reviewed totals 196 documents. In addition to the standalone tools, many of the reports, toolkits and other documents also contained multiple relevant tools (ranging from 1 to 18 tools across the documents). In total, there were 112 tools within the resources.

The desk review documents span 10 different humanitarian sectors. A large number of documents reviewed are cross-sector tools or reports and are, therefore, represented more than once within the figure. For this reason, the figure totals exceed 145 documents. Close to 25% of the documents reviewed are broadly applicable across the humanitarian field. The sectors that are most represented in the desk review include the Protection, WASH, Shelter and CCCM sectors (see figure 3). The 112 tools reviewed follow a similar sector breakdown as the documents, with the notable exception that the WASH sector has by far the highest number of tools (n=33) and more than two times as many as CCCM, the next highest sector (n=15). Nutrition, food security and health had the fewest number of tools relevant for GBV risk mitigation with one each (see figure 3).

**OVERVIEW OF CONTENT AND FOCUS OF RESOURCES**

**Tools**

Tools—defined by the research team as instruments designed for data collection and M&E—represented the greatest proportion of documents included in the desk review (n=112). A great number of these tools were found within larger toolkits such as the USAID Toolkit for M&E. Gender-based Violence Interventions Along the Relief to Development Continuum and the UNICEF Gender-Based...
Figure 2  Types of documents included in the desk review (n=145). The All Tools Extracted bar includes all tools identified (n=112) across all of the other document types. Standalone tools are defined as single tools or instruments, while a toolkit refers to multiple separate tools and other documents packaged together as one resource. Research/evaluations are non-peer reviewed research reports or evaluations. Journal articles include peer-reviewed literature.

Figure 3  Number of documents and tools included in the desk review categorised by the humanitarian sector of focus. The chosen sectors reflect those included in the 2015 IASC GBV Guidelines. Note that many documents and tools were produced by or for more than one sector and are thus counted more than once in the figure. Some documents and tools were designed to be broadly applicable across the humanitarian field. These documents are included in the All Sectors category. CCCM, Camp Coordination and Camp Management; IASC GBV Guidelines; Integrating Gender-Based Violence Interventions in Humanitarian Action; WASH, Water, Sanitation and Hygiene.
Violence in Emergencies Programme Resource Pack On Assessment. While there were a relatively large number of these documents, they varied substantially in approach, level of detail and robustness. The majority of the tools do not provide specific guidance on when to use the tool, who should use it, during which phase of an emergency response it should be used or how to analyse and interpret data collected using the tool. Very few of the tools are specifically designed to measure change in outcomes over time. With respect to perceptions of safety, most do not define safety, and there is wide variation in the way questions about safety are being asked, both in terms of the phrasing of questions and conceptualisation of safety, and whether questions ask the respondent directly about their own safety (ie, using a direct approach) or ask them about other people who are similar to them (ie, using an indirect approach). In addition, most of the tools focus on general safety and do not attempt to differentiate between GBV-related safety perceptions and other safety concerns. For example, a latrine may be considered unsafe because of the physical integrity of the building or because of risk of insect or snake bites, concerns that are not related to GBV risk. Questions that ask about safety more generally may not be able to distinguish these concerns. There are variations in the recall period employed (with most lacking a specified recall period) and few of the tools aimed to dissect safety perceptions by time of day or by other factors such as location (ie, at a facility vs on the way to a facility).

Reports, case studies and assessments/evaluations
Several different types of reports (n=32) were reviewed for this study ranging from field reports containing programmatic highlights to technical reports that provide recommendations and lessons learnt across specific case studies. Key reports include UNICEF’s report on protection against sexual exploitation and abuse, which makes recommendations for measurement of risk mitigation activities and a report on WASH sector activities in the Rohingya response, which reflects on inclusion of safety-related questions in routine monitoring of facilities. Overall, the reports spanned a variety of sectors, but the majority (n=21) target humanitarian action more broadly and are applicable to all sectors. The next most common sectors represented in the reports are Protection followed by Livelihoods and Shelter. A large number of reports pointed to good practices and recommendations for GBV risk mitigation activities as well as identified existing gaps in targeted efforts responding to gender-sensitive needs. For example, a Refugee International report from Burundian camp settings found that in many instances, humanitarian responses had fallen short of minimum standards and did not adhere to guidelines to reduce the risk of GBV. Multiple reports have echoed this concern together with a lack of knowledge—on the part of those leading humanitarian responses—in how to translate the IASC GBV Guidelines into practice.

Many reports focused their recommendations on how to mitigate key GBV risks rather than ways of measuring the impacts of risk mitigation activities. As noted in one report reviewing CARE’s work on identifying successful approaches to reducing GBV from 2011 to 2013, some of the evaluations reviewed described project activities but did not include any reference to information on outcomes, results or impacts achieved. Participatory methods are often recommended across reports for programme design and evaluation, yet these methods are rarely described in detail within the reports themselves. The quality of participatory ‘community consultations’ is also called into question in some report evaluations.

Of the non-peer reviewed literature reviewed, only a handful provided examples of rigorous approaches used to assess changes resulting from GBV risk mitigation activities. Two Oxfam reports examined the effects of installing improved lighting around sanitation facilities. The project measured changes in outcomes over time through a combination of Key Informant Interviews (KII), Focus Group Discussions (FGD) and baseline and endline surveys. Another report described how integrating safety considerations into the monitoring of sanitation facilities could support measurement of GBV risk mitigation. In general, almost every report reviewed reflected on the need to strengthen M&E frameworks to monitor the effectiveness of established mitigation measures and assess ongoing GBV risks.

Peer-reviewed journal articles
A relatively limited number of articles in the desk review were from the peer-reviewed literature (n=11). Four of these articles focused on calling for greater attention to be paid to the importance of GBV risk mitigation and a need to align measurement approaches and indicators, including more effective M&E. Two papers represented systematic reviews of existing interventions. Robbers and Morgan conducted a systematic review of evaluated sexual violence prevention and response interventions published in the peer literature between 2000 and 2016. The article found 29 studies that met the inclusion criteria, of which 7 studies addressed prevention, 14 studies response and 8 addressed both. The authors did find that, generally, the interventions screened lacked adequate M&E and were often short term, making robust impact assessment of GBV interventions quite challenging. Noble et al conducted a systematic review of academic and grey literature in September 2015 to examine the evidence base for programming that seeks to reduce violence against adolescent girls in humanitarian contexts. At the time of their search, they found only three adolescent girl programme evaluations from humanitarian settings. Each of these were pre/post-test evaluations that examined changes in measures such as social assets, self-esteem, decision-making, livelihood skills and financial assets, gender norms and feelings of safety. The authors concluded that there is a significant gap in rigorous research on the topic and called for
improved evidence to inform programming decisions aimed at reducing GBV among this population.

One article assessed changes associated with distributing handheld solar lights to women in an internally displaced persons camp in Haiti (n=754). The study found that those who received lights reported going outside at night more frequently at the end of the study compared with the beginning. However, neither safety nor GBV-related outcomes were specifically assessed in the study.33

Two qualitative studies explored dynamics related to GBV risk in low resource or humanitarian settings. A qualitative study undertaken with health professionals in Uganda (n=36 participants) underlined the GBV risks associated with poor water and hygiene infrastructure and explored the extent to which healthcare professionals expressed an understanding of these risks.34 A participatory study was conducted across three countries with individuals living in refugee settings who presented with physical, intellectual, sensory and mental disabilities and identified barriers to accessing services among this often under-represented population.35 The study was able to effectively identify specific risks, needs and barriers for those accessing sexual and reproductive health services using participatory methods in order to find practical ways to address and mitigate these challenges.

One article used mathematical modelling techniques to link the risk of sexual assault to the number of sanitation facilities and the amount of time women had to spend walking to sanitation facilities. The study emphasised that improving access to sanitation facilities may be more cost-effective overall and could reduce the risk of GBV by cutting down on the time travelled.36

Guidance documents
Several of the guidance documents reviewed contain information on M&E for GBV risk mitigation. Some do provide clear guidance and recommendations for sectoral programme staff on how to consider the M&E of GBV risk mitigation at various stages of the project cycle, and some also provide references to other relevant reports and materials. However, many of the tools have been reviewed as part of this desk review and suffer from the limitations described above. The Women in Displacement Toolkit developed by the International Organization for Migration and the Women’s Refugee Commission (WRC), for example, is very comprehensive, including detailed guidelines for different strategies to measure GBV risk mitigation interventions.37 Other guidance documents are broader, providing only general guiding principles, or are extremely focused on one specific measure or intervention.25 26 Naturally, the desk review represents the material available at a particular moment in time, and the guidance on GBV risk mitigation measurement has continued to evolve since then.

Measurement approaches and information sources
While the rigour and ambition of the documents reviewed vary widely, the most common measurement approaches that are described or were used were safety audits. Safety audits can be observational (ie, involve observing but not engaging directly with service providers and affected populations) or participatory (ie, involve engaging with service providers and affected populations through focus groups, interviews and other participatory methods such a mapping or issue ranking exercises). In addition to safety audits, quantitative surveys were also used to collect data among affected communities, staff or other actors. Other methods that were described include FGD, KII, service and stakeholder mapping and participatory methods such as participatory ranking, pile sorting and pocket voting. Qualitative methods dominated among the tools were reviewed (n=112). Approximately 62% of the tools included qualitative approaches, 33% included quantitative methods and 21% included observational components.

Many of the reports from the qualitative assessments employed small sample sizes of respondents and did not describe whether sample size calculations were undertaken or the techniques employed for sampling the respondents that were surveyed. Others did not describe well the other methodological and data collection procedures. These types of shortcomings limit the ability to accurately measure the necessary indicators for M&E GBV risk mitigation strategies and restrict the generalisability of any data collected.

The Gender-Based Violence Information Management System (GBVIMS) was described in several documents as an important source of existing data that could be analysed in terms of trends. The GBVIMS is an information management system that enables GBV service providers to safely and ethically collect, store, analyse and share data related to reported incidents of GBV. Though it is a crucial tool for informing and adapting GBV response services, data generated by the GBVIMS are generally not as useful for understanding GBV-related risks, identifying priority risk mitigation activities or determining the effectiveness of risk mitigation interventions. This data source

FINDINGS RELATED TO MEASUREMENT OF GBV RISK MITIGATION
Measurement of the effectiveness of GBV risk mitigation strategies, both in terms of GBV-specific outcomes as well as sector-specific outcomes, generally, was not a primary focus of the documents reviewed. In fact, while most documents describe the importance of or need for GBV risk mitigation strategies, few suggest specific measurement tools for monitoring and assessing such activities for the purpose of adjusting interventions, informing decision-making and evaluating success or failure. Few documents had the objective of assessing outcome changes due to GBV risk mitigation activities and/or employed sufficiently rigorous methodologies to be able to demonstrate the effectiveness of GBV risk mitigation strategies. Among the few that included rigorous approaches include two Oxfam reports from 2018 examining the effects of improved lighting of GBV risk reduction.27 28
was typically described as most relevant for GBV practitioners rather than other sectors.

In the programme design phase, most guidance documents and toolkits recommend a baseline analysis of the situation on the ground prior to implementation of response activities. Constructing an accurate picture of GBV risks and safety issues helps design an effective intervention and also provides preimplementation data that could later be compared with data collected during and after the rollout of the programme. In order to be most effective and tailored to the community’s needs, many tools recommend involving affected communities at the design stage of the programme.

During the implementation phase, regular monitoring was described as important in order to correct course immediately if need be, to document the implementation of the programme and measure output indicators. While evaluations are needed to assess outcome-level changes of an intervention, the majority of the tools in this desk review do not include specific guidelines on how to incorporate evaluation, particularly at the early stages of project design. Those that do suggest a number of ways that measurement of GBV risk prevention can be incorporated into M&E activities throughout the programme duration. Approaches for both staff-focused and affected population-focused inputs at the design phase (when programmes are being conceived and planned) and implementation phases (when programmes are being carried out) are outlined in Table 4.

### Participatory methods

Most of the documents reviewed described or included at least one form of input from the local community through focus groups or surveys. However, it is important to acknowledge that this is not necessarily an indication of meaningful engagement with the community, or with women and girls. FGD and individual interviews were labelled as participatory in nature even when they were not fully inclusive, did not include specific strategies to foster collaboration between affected populations and humanitarian actors or when they were conducted once only (such as during risk assessment) and were not part of an ongoing process of collaboration. Simply questioning affected populations is not a truly participatory approach, and will not help to create a sense of ownership of a programme or evaluation. For example, the Evaluation of Implementation of 2005 IASC Guidelines for Gender-Based Violence Interventions in Humanitarian Settings in the Syria Crisis Response (2015), which examined the integration of the IASC GBV Guidelines across three sectors in Lebanon, Jordan and Kurdistan Iraq—pointed out that affected populations in the projects concerned were often frustrated by the lack of a two-way flow of information. They were required to respond in FGDs and surveys but did not receive sufficient information in return and did not feel like they had any influence on the programme design or outcome. In some cases, community consultations may replicate existing gender inequality or fail to be truly inclusive with respect to gender, age and disability.

Several of the documents did highlight successful implementation of participatory approaches to design and measurement. For example, the International Federation of Red Cross and Red Crescent Societies case studies on disaster response described how communities were involved in data collection and in designing their own approaches to resilience and GBV risk mitigation. Other examples of successful implementation of participatory approaches are the two Oxfam studies on lighting from 2018. In both cases, community members were invited to identify priority areas and were consulted in the M&E processes as well. However, studies where participatory approaches are employed from the design through to the M&E phase were rare.

Approximately, 41% of the tools included participatory approaches (see Figure 4). The tools developed by or for the WASH and CCCM sectors were more likely to include participatory methods. Some of the more innovative participatory methods (24-hour clock, pocket voting, participation ladder) were found among the WASH tools. Surprisingly, only 15% of the 13 Protection tools included participatory elements. None of the tools reviewed from several of the sectors including Health, Education, Food Security and Nutrition had participatory methods.

### Table 4 Summary of data collection methods by activity phase and target audience (source of input)

<table>
<thead>
<tr>
<th>Design phase: baseline research</th>
<th>Staff/project focused input</th>
</tr>
</thead>
<tbody>
<tr>
<td>► Review of existing data sources</td>
<td>► Desk review</td>
</tr>
<tr>
<td>► Baseline quantitative surveys</td>
<td>► Service mapping</td>
</tr>
<tr>
<td>► Focus group discussions</td>
<td>► Safety mapping</td>
</tr>
<tr>
<td>► Participatory safety mapping/audits</td>
<td>► Safety audits</td>
</tr>
<tr>
<td>► Other participatory methods: community story-telling, role-playing, pocket voting</td>
<td>► Observational approaches</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Implementation and evaluation phase</th>
<th>► Midline and endline quantitative surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>► Participatory safety mapping/audits</td>
<td>► Repeated safety audits</td>
</tr>
<tr>
<td>► Post-distribution monitoring</td>
<td>► Key informant interviews</td>
</tr>
<tr>
<td>► Focus group discussions</td>
<td>► Staff quantitative surveys</td>
</tr>
<tr>
<td>► Key informant interviews</td>
<td>► Observational approaches</td>
</tr>
<tr>
<td>► Consultations with local organizations such as women’s groups</td>
<td></td>
</tr>
</tbody>
</table>


BMJ Open: first published as 10.1136/bmjopen-2021-050887 on 8 September 2021. Downloaded from http://bmjopen.bmj.com/ on October 11, 2021 by guest. Protected by copyright.
However, it should be noted that these sectors had few tools that were reviewed overall (see Figure 3).

**Age, gender and disability considerations**

Few of the studies or reports directly emphasise the disaggregation of data according to diversity criteria beyond sex or gender such as age and disability status. Two documents, both from the protection sector, explicitly refer to the age, gender and diversity (AGD) approach.42 43 Three different studies refer to reporting sex and age-disaggregated data (SADD) in order to better account for vulnerable groups.44–46 However, all of these documents report challenges with analysing existing data because SADD elements are not reported. Only two other reports mention the disaggregation of data according to diversity criteria in their study design.47 48 The Oxfam Lighting the Way study (2018) does not explicitly describe their approach to SADD, but nevertheless presents SADD data within the reported findings.27

This lack of emphasis on ensuring diversity and consideration of subpopulations is also reflected in the toolkits and guides reviewed. Few of the tools specifically describe the need for collecting disaggregated data. One tool that explicitly included guidance on diversity is a toolkit developed by the WRC, International Rescue Committee and Mercy Corps to optimise cash-based interventions for displaced persons while also strengthening the prevention and response of GBV. It aims to assist practitioners in collecting protection information according to the AGD approach. USAID’s Beyond Access Toolkit for Integrating Gender-Based Violence Prevention and Response into Education Projects (2014) also describes the need for collection of disaggregated data. A limited number of additional tools explicitly mention diversity concerns,8 10 49 one of these includes specific guidance and specific stories and questions to guide discussions with different populations including men and women with disabilities, adolescent girls and boys and Lesbian, Bisexual, Gay, Transgender, Queer, Intersex, and Asexual (LGBTQIA) individuals.13

**Implementation of GBV risk mitigation measurement**

While the toolkits and guidance documents reviewed provide recommended approaches to measuring GBV risk mitigation activities, they do not provide information on the extent to which the recommendations are actually being implemented in the field. Project reports, articles, evaluations and case studies provide valuable insight as to whether and how recommendations are being applied on the ground.

A great variety of these documents with different levels of ambition and academic rigour were assessed as part of the review. Most are reports (n=32), but there are several non-peer-reviewed research and evaluations (n=26) and a limited number of journal articles (n=11) and case studies (n=7). Nearly all incorporated some form of qualitative data collection or research such as FGDs and/or KIIs. However, the number of KIIs and FGDs conducted as part of specific assessments varies widely. In some cases, the sample sizes are insufficient to make strong conclusions or recommendations, but these are included in the reports. Some reports complemented qualitative research with quantitative data obtained through surveys with affected populations, including household surveys, Knowledge, Attitudes and Practices surveys or specific surveys among women and girls.

A number of studies and reports such as a recent ACTED Report (2018) also include observational methods such as safety audits in order to identify specific GBV risks. Few reports, however, use or suggest observational methods to

---

Figure 4 Percentage of tools reviewed that suggest inclusion of participatory methods (per sector). CCCM, Camp Coordination and Camp Management; WASH, Water, Sanitation and Hygiene.
measure the impact of GBV risk mitigation interventions. Those that do, recommend conducting repeated safety audits to assess changes in safety risk before, during and after implementation of GBV risk mitigation activities. Oxfam’s ‘Shining a light’ study on the impact of lighting installation in and around sanitation facilities is one of the few to include mixed methods approaches and also incorporated a pre-test post-test to assess changes in outcomes over time. This study methodology combined quantitative surveys, FGDs, KIIs and observational methods at baseline and at endline in order to measure changes in key outcomes related to the GBV risk mitigation activity (lighting installation) over time.

**DISCUSSION**

There is increasing recognition of the importance of integrating GBV risk mitigation activities into humanitarian response. Yet, there are still significant gaps in how to best measure and evaluate the effectiveness of these actions. Of the 145 documents that met the final inclusion criteria, 10 different humanitarian sectors were represented, and a large number of documents were designed to be applicable for all areas of humanitarian action. The number and sectoral span of the documents speak to the increased focus on this issue. However, measurement of the effectiveness of GBV risk mitigation strategies, both in terms of GBV-specific outcomes as well as sector-specific outcomes was often not a primary focus of these documents.

While most materials emphasised the importance of GBV risk mitigation, very few suggested specific measurement tools for monitoring and assessing such activities for the purpose of adjusting interventions, informing decision-making and evaluating success or failure. The tools reviewed lacked clear guidance on key points: including how and when to use the tool (including at which phase of the project cycle and during which types of humanitarian response), who should use the tool, or how to analyse and interpret data collected.

There is limited evidence demonstrating change in outcomes (GBV-related and sector-specific outcomes) as a result of GBV risk mitigation interventions among the documents reviewed. Only a handful of reports had the objective of assessing outcome changes due to GBV risk mitigation activities and employed sufficiently rigorous methodologies to be able to demonstrate their effectiveness. Therefore, there is a significant gap in the evidence base around the effectiveness of GBV risk mitigation across all sectors. Overall, the peer-reviewed literature reflected broader trends from the desk review: no studies represented impact evaluations of GBV risk mitigation activities. A number of the studies were commentaries highlighting the gap in measurement related to risk mitigation and calling for more standardised, field-friendly guidance.\textsuperscript{13,15,29,30}

The majority of documents reviewed recommend or use qualitative forms of data collection such as FGDs, KIIs or observational approaches. These are valuable approaches that tend to be recommended for initial risk assessments to inform programme design. However, there is a gap in collecting high-quality qualitative data in routine M&E efforts. In addition, there is limited guidance on how to successfully employ quantitative measures to monitor or evaluate GBV risk mitigation activities.

While the vast majority of GBV risk mitigation literature reviewed for this article focuses on qualitative approaches, both qualitative and quantitative methods are important for achieving accurate, holistic pictures of just how effective specific risk mitigation strategies are in operation. It is likely this deference to qualitative methods is linked to both a need for further quantitative guidance on GBV risk mitigation measurement and the potentially increased costs associated with procuring the specialised knowledge necessary to design a quantitative study.

Impact evaluations—research to assess causal changes that are attributable to a particular intervention—would provide the most precise understanding of the effectiveness of a particular GBV risk mitigation strategy and help to fill clear evidence gaps. However, impact evaluations may need to be conducted over longer periods of time, which may not align with traditional programme and funding cycles and demand considerable expense and expertise, which may not be available in humanitarian emergencies. The time, financial and capacity challenges of humanitarian response mean that not all humanitarian practitioners recognise GBV risk mitigation as a priority. In particular, the measurement of GBV risk mitigation may then be especially challenging, since it requires not only a commitment to GBV risk mitigation but also to collecting, analysing and interpreting data related to GBV risk mitigation.

For these reasons, impact evaluations are not always the right methodological choice in humanitarian emergencies, where project implementation is typically shorter or subjected to a high number of contextual and environmental changes that could undermine a researcher’s ability to properly execute the study design. However, some forms of quantitative measurement, as part of standard agency risk assessments or M&E practices, for example, are more broadly suitable for humanitarian contexts and should be assessed alongside qualitative results for a fuller picture of GBV risk mitigation.

A lack of contextualised guidance about how and when to implement potential methods of measurement and which methods would be most effective for collecting particular information was evident in many of the reviewed resources. Guidance on how to adapt tools for a particular context and the associated ethical considerations required for implementation in varied humanitarian settings was critically lacking. Given this, there seems to be a current need for tools that are designed with these considerations in mind and accompanied by more detailed, practitioner-focused guidance on these topics. A recent article on good practices related to the M&E of GBV risk mitigation in humanitarian contexts seeks to fill this gap. This article outlines lessons learnt and effective approaches by combining learning and...
recommendations from published and grey literature as well as key-informant interviews with practitioners.50

Recent challenges related to adapting humanitarian response to COVID-19 have highlighted new and existing challenges for GBV risk mitigation.51–53 As practitioners work to understand how to protect women and girls in highly dynamic and constrained environments, understanding the effectiveness of GBV risk mitigation activities is all the more important. Measurement approaches that allow for adaptive programming that centres the voices and needs of women and girls will be vital.

LIMITATIONS
This systematic mapping study has several limitations. First, the findings reflect information and practices that have been published in a searchable online format or shared through the open call for submissions. We anticipate that some field knowledge and experiences are not captured in the documents, reports and articles we reviewed, and this review, thus, may not fully capture field lessons and insights into the measurement of GBV risk mitigation. In addition, some relevant literature may have been missed through the search methodology. For example, any relevant literature in languages other than English may have been missed, although our team did review several French-language tools. Moreover, the search term list is not exhaustive and may have missed terms with different synonyms or British variants of spelling. It would have been infeasible to search all of the potential sources of grey literature, and, thus, our team prioritised those deemed to be the most relevant for this review. The literature review extraction process was time bound and only continued up until December 2019; hence, findings should be interpreted within this time frame. Finally, there may have been some relevant lessons to draw from in literature and reports on risk mitigation and measurement of safety perceptions from outside the humanitarian sector. While that was outside of the scope of this review, future research may benefit from examining non-humanitarian literature.

CONCLUSION
The systematic mapping review findings demonstrate that integration of GBV risk mitigation activities into non-GBV sectors is advancing despite challenges in measurement and a lack of detailed guidance. The emphasis on measurement of GBV risk mitigation activities is stronger among several key sectors including WASH, Shelter and CCCM. Overall, 145 documents and a total of 112 tools were reviewed, which well exceeded the number anticipated at the outset. While numerous resources across 10 sectors exist, many lack sufficient information on how to monitor outputs or outcomes of GBV risk mitigation activities or on how to measure their effectiveness in terms of GBV and non-GBV (sector specific) outcomes. There is also limited guidance on how to integrate measurement of GBV risk mitigation into existing M&E frameworks.

This dearth of measurement-related guidance is subsequently reflected in the reports and articles assessed as part of this desk review. Most employed qualitative methods and very few attempted to measure the impact of a GBV risk mitigation activity or employ sufficiently robust designs and methodology to be able to do so. Overall, there is a significant gap in the evidence base around the effectiveness of GBV risk mitigation across all sectors. Understanding and strengthening measurement approaches in GBV risk mitigation remain a critical task for projects that aim to address GBV.

REFERENCES

1 Potts A, Zuco V. If GBV programming is essential in emergencies, how do we do it? Developing a model to operationalise existing guidance. Humanitarian Practice Network 2014.
2 Women’s Refugee Commission. Call to action on protection from gender-based violence in emergencies 2015.
affects the risk of gender-based Violence in humanitarian crises. 

Overseas development


House S. Str AND Abuse 2018.

UNICEF.


Winter S. Enhancing UNHCR’s capacity to monitor the protection, rights and well-being of refugees 2004. 


USAID. Toolkit for monitoring and evaluating gender-based violence interventions along the relief to development continuum 2014.

UNICEF. Gender-Based violence in emergencies programme resource pack on assessment 2019.

UNICEF. Tips for strengthening protection against sexual exploitation and abuse 2018.


Oxfam. Shining a light: how lighting in or around sanitation facilities afffects the risk of gender-based violence in camps 2018. 