Research Methodology Note

Assessment of Hard-to-Reach Areas (H2R)

SOM1901

Somalia

October 2022
V2

REACH Informing more effective humanitarian action

1. Executive Summary

Country of intervention	Somalia					
Type of Emergency	x	Natural disaster	x	Conflict	□	Other *(specify)*
Type of Crisis	□	Sudden onset	x	Slow onset	x	Protracted
Mandating Body/ Agency	Office for the Coordination of Humanitarian Affairs (OCHA), Inter-Cluster Coordination Group					
IMPACT Project Code	27ANW					
Overall Research Timeframe (from research design to final outputs)	15/09/2022 to 15/12/2022					

Research Timeframe *(from research design to final outputs / M&E)*	1. Pilot/ training: 11/10/2022 (Training of Trainers – ToT); 12/10/2022 and 13/10/2022 (Enumerator Training)	6. Start output drafting: 13/11/2022
	2. Start collect data: 16/10/2022	7. Outputs sent for validation: 28/11/2022
	3. Data collected: 06/11/2022	8. Outputs published: 15/12/2022
	4. Data sent for validation: 13/11/2022 (cleaned and aggregated)	9. Final presentation: December 2022
	5. Data analysis on validated data: 20/11/2022	

Humanitarian milestones *Specify **what** will the assessment inform and **when** e.g. The shelter cluster will use this data to draft its Revised Flash Appeal;*	**Milestone**		**Deadline**
	X	Somalia Humanitarian Fund (SHF) Allocations	Ongoing
	X	Humanitarian Needs Overview (HNO)	31/11/2022
	□	Cluster plan/strategy	_ _/_ _/_ _ _ _
	□	NGO platform plan/strategy	_ _/_ _/_ _ _ _
	Audience type		**Dissemination**

Audience Type & Dissemination *Specify **who** will the assessment inform and **how** you will disseminate to inform the audience*	x Strategic x Programmatic □ Operational □ [Other, Specify]	**x** General Product Mailing (e.g. mail to NGO consortium; HCT participants; Donors) □ Cluster Mailing (Education, Shelter and WASH) and presentation of findings at next cluster meeting **x** Presentation of findings (e.g. at HCT meeting; Cluster meeting) **x** Website Dissemination (Relief Web & REACH Resource Centre) □ [Other, Specify]
Detailed dissemination plan required	□ Yes	x No Results will be shared proactively with coordination bodies, at the national and state levels, for wider sharing with operational partners. REACH will aim at having the preliminary results ready to inform the 2023 Humanitarian Needs Overview (HNO) and next IPC round
General Objective	To inform humanitarian planning and operations by providing information on needs, displacement dynamics, agricultural livelihood conditions and access to basic services in hard-to-reach settlements. With this October round, REACH will aim to inform the 2023 Humanitarian Needs Overview (HNO) and next IPC round regarding food security, agricultural livelihoods, displacement, basic services, markets and WASH conditions in hard-to-reach areas of Somalia.	
Specific Objective(s)	• To determine the severity of food insecurity, basic service needs and Water, Sanitation and Hygiene (WASH) needs of hard-to-reach populations. • To determine push and pull factors that contribute to displacement from hard-to-reach areas. • To determine the vulnerabilities to and influence of climatic hazards (particularly drought and floods) on hard-to-reach populations. • To provide up-to-date information on services accessible to households in the hard-to-reach areas, the languages used in hard-to-reach areas, what information and humanitarian services are reaching hard-to-reach populations.	
Research Questions	• What are the food security needs and coping strategies of the populations in H2R areas? To which foods, sources of livelihood and markets have populations in H2R areas access to? What are the food, livelihood coping strategies and market access constraints which may increase their vulnerability to climatic shocks? How are climatic shocks affecting access to food and livelihoods? • What are the needs and coping strategies of the populations in H2R areas regarding Water, Sanitation and Hygiene? What are the constraints to access to WASH? How does the WASH situation affect the vulnerability of populations in H2R areas? What are the impacts of climatic shocks on WASH in H2R areas? • What are the needs and coping strategies of the populations in hard-to-reach areas regarding Livelihoods, Health and Protection? • Are there basic services, humanitarian aid programs and information regarding basic services/humanitarian aid in hard-to-reach areas? Which	

	groups maybe unable to access these services and humanitarian aid programs among the hard-to-reach populations? • What are the social, economic and environmental vulnerabilities to climatic shocks of the populations in hard-to-reach areas? What are the direct and indirect social, economic and environmental impacts of climatic shocks in hard-to-reach areas? • What are the factors that contribute to displacement from the hard-to-reach areas? What are the movement intentions of populations in hard-to-reach areas?
Geographic Coverage	For October round, 17 districts will be covered through the H2R approach, as these districts could not be accessible as part of the MSNA in-person data collection. These districts present "high access constraints" or are considered "inaccessible" by the OCHA Access Working Group. In addition, further triangulation with REACH field team helped to classify these district as "hard-to-reach".[1] The final list of districts can be found in section 3 - Methodology.
Secondary data sources	• PRMN (Protection and Return Monitoring Network) • Camp Coordination and Camp Management (CCM) Cluster settlement verifications • REACH DSA (Detailed Site Assessment) • IOM DTM (Displacement Tracking Matrix) • SWALIM (Somalia Water and Land Information) • Satellite imagery • Food Early Warning Systems Network (FEWSNET), Somalia, publications • Cash Based Programming in Somalia • Somalia Health Cluster, 4W Dashboard, 2022

Population(s) *Select all that apply*	□	IDPs in camp	x	IDPs in informal sites
	x	IDPs in host communities	□	IDPs [Other, Specify]
	□	Refugees in camp	□	Refugees in informal sites
	□	Refugees in host communities	□	Refugees [Other, Specify]
	x	Host communities	□	[Other, Specify]
Data collection tool(s)	**x**	Structured (Quantitative)	**x**	Semi-structured (Qualitative)

	Sampling method	**Data collection method**
Structured data collection tool *Quantitative tool*	**x Purposive** □ Probability / Simple random □ Probability / Stratified simple random □ Probability / Cluster sampling □ Probability / Stratified cluster sampling x Snowballing	**x Key informant interview (Target #): 1,032 KI interviews[2]** **(minimum 2 KIs per settlement, maximum 6)** □ Group discussion (Target #):_ _ _ _ _ □ Household interview (Target #):_ _ _ _ _ □ Individual interview (Target #): _ _ _ _ _ □ Direct observations (Target #):_ _ _ _ _ □ [Other, Specify] (Target #):_ _ _ _ _

[1] The level of access may change by the time of data collection, depending on the security situation in the country. In case of a deteriorating or improving security and access context, REACH teams might reconsider the areas to assess.
[2] Four teams of enumerators (3 teams with 4 enumerators each, 1 team with 8 enumerators) will be collecting quantitative data over a period of 15 working days. Three teams of 2 enumerators each will be collecting qualitative data over a period of 2 days. In total, 26 enumerators will be involved in data collection, for a period of 15 (+ 1 if required) days.

Semi structured data collection tool *Focus Group Discussion (FGD).*	**x Purposive** **x Snowballing**	□ Key informant interview (Target #):_ _ _ _ _ □ Individual interview (Target #):_ _ _ _ _ **x Focus Group Discussion (Target #): 12 sessions (Mogadishu 4 sessions, Baidoa 4 sessions, Kismayo 4 sessions) with 6 – 8 participants per group.**
Target level of precision if probability sampling	NA	NA

Data management platform(s)	x	IMPACT	□	UNHCR		
	□	[Other, Specify]				
Expected ouput type(s)	**x**	**Situation overview #: 1**	□	Report #: _ _	□	Profile #: _ _
	□	Presentation (Preliminary findings) #: _ _	**x**	**Presentation (Final) #: 4 (TBD) – at the regional level**	□	Factsheet #: _ _
	□	Interactive dashboard #:_	□	Webmap #: _ _	**x**	**Map #: 5 (TBD)**
	□	[Other, Specify] #: _ _				
Access	x	Public (available on REACH resource center and other humanitarian platforms)				
	□	Restricted (bilateral dissemination only upon agreed dissemination list, no publication on REACH or other platforms)				
Visibility *Specify which **logos** should be on outputs*	***REACH***					
	***Donor:** USAID*					
	***Coordination Framework:** NA*					
	***Partners:** OCHA*					

2. Rationale

2.1 Background

Somalia's prolonged, complex, and multi-faceted humanitarian crisis is characterised by ongoing conflict, climate-related shocks, communicable disease outbreaks and fragile social protection mechanisms. Droughts in a fifth failed rainy season continue to exacerbate the effects of periodic natural disasters and insecurity caused by armed conflicts. Parts of Somalia have seen localized storms and flooding during April and May while also experiencing a fifth consecutive below average rainy season. With possibly continuing rising food prices, the risk of localized famine is heightened in several areas in Somalia. This continuously deteriorating humanitarian situation means that an estimated 7.7 million people need humanitarian support, with at least 7 million people directly affected by the drought conditions, of whom 918,000 have been displaced from their homes in search of water, food and pasture; including minority groups.[3]

While information about the severity of needs in accessible areas is often available, thanks to partners' assessment efforts in the context of rapid assessments and broader, periodic, country-wide assessments, information on needs in hard-to-reach (H2R) areas remains very limited. Further, while country-wide analyses on climatic shocks are available, these are rarely cross-referenced with primary data on populations' needs in the field.

[3] OCHA, Somalia: The Cost of Inaction, July 2022

The October 2022 round of H2R data collection aims to align the H2R with the 2022 Somalia MSNA while continuing to monitor the economic, social, and the impact of climatic shocks (drought and flooding, in particular) on the H2R communities assessed to inform the Humanitarian Needs Overview (HNO) 2023 and humanitarian actors operating in H2R areas. This will be done by leveraging KI knowledge of H2R areas, and cross-referencing information provided by KIs with satellite imagery, where relevant. The October round of data collection will also continue to focus on drought-related indicators. Analysis from last round indicates that water sources are drying up, water prices are rising, and there is a definitive impact on agricultural and livestock activities which is placing immense pressure on already vulnerable communities' ability to access food and water are all seen to be results of drought conditions. People in these areas are exhausting their coping mechanisms with displacement as the ultimate result, with protection and security incidents rising in prevalence in drought impacted areas.[4] To better understand the needs of minority groups (i.e. groups which may speak minority languages), this round of H2R has also introduced some questions regarding access and languages.

2.2 Intended impact

This round of the H2R assessment will happen in October 2022, during the Deyr rain season when land preparation and planning for crops occurs. The assessment will contribute to better understanding of the multi-sectoral needs of the H2R communities and be an asset for the 2023 HNO/HPC and IPC analysis; while also supporting the overall advocacy of and humanitarian response for hard-to-reach populations.

The assessment aims to contribute to the following:

- Improve the understanding of the current situation in H2R areas to inform the humanitarian planning cycles (HPC) and Somalia Humanitarian Fund allocation, by providing information from inaccessible areas in Somalia.
- To assist the humanitarian community in prioritization and planning of the humanitarian response for hard-to-reach populations.

3.Methodology

3.1 Methodology overview

The H2R employs an Area of Knowledge (AOK) methodology, which offers settlement-level data on needs to shape the response in a situation where direct household surveys are currently limited or unavailable due to security constraints. The AoK approach consist of identifying key informants (KIs) who can testify about the humanitarian needs of a specific area. These KIs are identified based on their knowledge of the specific area that is being assessed, either because they have been displaced from this area recently or because they travel a lot in this area, or because they still have family or friends residing in the area that they can contact. This technique is meant to support strategic planning and contribute to a more focused and evidence-based humanitarian response in areas where humanitarian workers cannot go on a regular basis due to logistics and security constraints.

Data will be collected from 16th October 2022 to 06th November 2022, by the enumerator teams in the informal sites[5] around the 17 hard-to-reach districts identified as part of the MSNA assessment and the Access Working Group (AWG) under the Humanitarian Coordination Team (HCT), supervised by one field officer in each of the four location. IDP camps that received new arrivals from H2R settlements[6] in the previous three months, or new IDP settlements (set up in the last three months by people who arrived from hard-to-reach areas) will be identified through ACTED, local authorities, partners (CCCM cluster, IOM DTM), and REACH field networks.

[4] REACH, Hard-to-reach Assessment, Advocacy brief on Somalia drought, August 2022

[5] Site where more than 15 displaced households have settled collectively. Sites not built to accommodate people, but serving that purpose, set up on state-owned or private land/buildings – CCCM Cluster Somalia.

[6] Settlements are eligible to be covered by this assessment if they are listed under OCHA's directory of inaccessible settlements. However, additional settlements may become eligible based on REACH Field Officers' feedback from the field; this is notably the case when the shifting access situation results in new locations becoming inaccessible shortly before, or during, data collection.

The methodology is articulated into a quantitative component, whereby enumerators will conduct interviews with KIs (target of 1,032 KI interviews in total), pre-identified by field officers (FOs), through a structured KOBO questionnaire, and a qualitative component, consisting of 12 semi-structured focus group discussions that will provide a more in-depth understanding of the dynamic setting and supplement data gathered via the quantitative tool.

As part of this the October round of the H2R will include Normalized Difference Vegetation Index (NDVI) hexagon maps to highlight vulnerabilities and follow previous rounds of H2R assessments in Somalia. Further remote sensing activities will include analysis on expected precipitation during the recall period of July, 2022. These results will then be triangulated with the findings from FGDs and KI interviews.

3.2 Population of interest

Geographical area assessed: The assessment targets the inaccessible areas of Somalia that are located within the inaccessible districts identified as part of the MSNA assessment (see table 1 below).

Table 1 – List of districts presenting high access constraints or considered inaccessible and the data collection method used in 2022:

State	Region	District	OCHA Access list	Covered in DSA 2021	Covered in H2R 2021	REACH field team comments	Covered in MSNA 2022	Targeted for H2R OCT 22
Hirshabelle	Middle Shabelle	Adan Yabaal	Inaccessible	No	Yes	Inaccessible	No	Yes
Jubaland	Lower Juba	Badhaadhe	High access constraints	No	Yes	Inaccessible	No	Yes
Jubaland	Middle Juba	Bu'aale	Inaccessible	No	Yes	Inaccessible	No	Yes
Hirshabelle	Hiraan	Bulo Burto	High access constraints	No	No	Inaccessible	No	Yes
Galmudug	Galgaduud	Ceel Buur	High access constraints	No	No	Inaccessible	No	Yes
Galmudug	Galgaduud	Ceel Dheer	Inaccessible	No	No	Inaccessible	No	Yes
Jubaland	Lower Juba	Jamaame	High access constraints	No	Yes	Inaccessible	No	Yes
Jubaland	Middle Juba	Jilib	Inaccessible	No	Yes	Inaccessible	No	Yes
SWS	Lower Shabelle	Kurtunwaarey	High access constraints	No	Yes	Inaccessible	No	Yes
Puntland	Bari	Qandala	High access constraints	No	No	Inaccessible	No	Yes
SWS	Bakool	Dhuure	Inaccessible	No	Yes	Inaccessible	No	Yes
SWS	Lower Shabelle	Sablaale	Inaccessible	No	Yes	Inaccessible	No	Yes
SWS	Bakool	Tayeeglow	Inaccessible	No	Yes	Inaccessible	No	Yes
Galmudug	Mudug	Xarardheere	Inaccessible	No	No	Inaccessible	No	Yes
Puntland	Bari	Caluula	Moderate access constraints	No	No	Inaccessible	No	Yes
Jubaland	Middle Juba	Saakow	Inaccessible	Yes	No	Inaccessible	No	Yes
SWS	Lower Shabelle	Qoryooley	Moderate access constraints	Yes	Yes	Inaccessible	No	Yes

Population assessed: The target populations are IDPs and host communities residing in H2R settlements.

Unit of measurement: For the quantitative tool, the primary focus of the assessment is the settlement level. Given that some indicators cannot be collected at the settlement level some individual-level questions will be also asked.

As part of the FGD component, multiple KIs from different settlements are gathered to report on the needs of their community living in a H2R area. As the FGD is a new component of the H2R assessment, the unit of measurement remains the hard-to-reach areas as a whole. However, REACH will consider refining this aspect for the upcoming rounds; REACH will aim at conducting the FGDs with multiple KIs from the same settlements, gathered to report on the needs of their community. Using KIs from the same settlement will aim to create a discussion environment where one view might either be supported or challenged, giving possibilities for important insight. Doing so will also allow REACH to elaborate a narrative around needs situations in specific locations and thus advocate for more targeted aid interventions in areas of concern.

Unit of analysis:

- For most indicators collected through the quantitative and qualitative methodology, the aim for the October round will be to produce one general situational overview which will include key findings for the specific regions assessed.

- Results for a selection of WASH, mortality and Food Security indicators, as detailed in the data analysis plan in section, will be disaggregated according to the month in which the KI left or last visited the settlement of origin or last contacted friends/family in the settlement of origin; this is to enable a trends analysis on key indicators that can provide an indication of changing nutrition situation.

- Maps produced for the planned outputs will present results aggregated at the hexagon level (12km-sided) to avoid showing the exact locations of assessed settlements, due to protection considerations.

3.3. Secondary data review

In addition to the secondary data outlined in the original ToR of the assessment[7], most of which will be used for sampling purposes, additional sources will be mobilized:

- Food Security and Nutrition Analysis Unit (FSNAU) publications;
- Somalia Water and Land Information Management (SWALIM) Combined Drought Index;
- Integrated food security Phase Classification (IPC) reports
- WASH Severity Classification (to be published);
- OCHA, Drought Situation reports;
- United Nations University - Institute for Environment and Human Security, Understanding and reducing agricultural drought risk: Examples from South Africa and Ukraine, 2018;
- United Nations Office for Disaster Risk Reduction, Global Assessment Report on Disaster Risk Reduction, Special Report on Drought 2021
- FEWSNET, Somalia publications
- World Food Program, Annual Country Report 2021, Somalia
- Somalia Health Cluster, 4W Dashboard, 2022
- UN OCHA, Somalia drought response plan, 2022
- World Bank, HDX Data
- WASH Cluster, Somalia WASH Cluster Drought Response, 2022
- Drought Displacement Monitoring Dashboard (July 2022) - Somalia | ReliefWeb

Broadly speaking, as much as possible secondary sources will be used to contextualise findings, such as the Integrated food security Phase Classification (IPC) reports providing information on food security and nutrition needs as well as projections, per region. The World Food Programme (WFP) Annual country report adds more context to the food security situation in further context and complement the IPC. Somalia Water and Land Information Management (SWALIM) can provide climatic data such as the combined drought index capturing information on crops, pasture, fire danger, water shortages, livestock migration. FEWSNETs seasonal monitoring adds important value to the monitoring of seasons, especially during drought conditions, contributing with

[7] REACH, Research Terms of Reference, Assessment of H2R areas, October 2021.

information on rainfall during Somali seasons. In particular, this and the SWALIM data will be triangulated with reported impact of drought and floods in hard-to-reach areas. The WASH and CCCM clusters' products will be used to provide key definitions. For the Health sector, the Somalia health cluster Dashboard provide an interactive alternative to better understand the distribution of health care services and the reach of the health-related programming in Somalia. Finally, stock satellite imagery will be used to triangulate findings (e.g NDVI, rainfalls).

A selection of secondary data resources, including existing REACH tools, informed indicator and questionnaire design.

3.4 Primary Data Collection

Sampling

Given that physical access to the target locations is limited and there is no possibility of drawing a representative sample, purposive sampling will be adopted for all data collection methods, namely the Key Informant (KI) interviews and the Focus group discussions (FGDs).

KIs and FGD participants will be selected based on their knowledge of their settlement of origin that must be in hard-to-reach areas mentioned previously in this document, as well as based on the period they have displaced.

Notably, the following eligibility criteria will apply (and will be integrated in data collection tools):

1) Being newly displaced from the hard-to-reach settlement (=< 3 months before the start of data collection) **or** having visited the hard-to-reach settlement in the 3 months prior to the start of data collection **or** having been in contact with friends/family living in the settlement of origin in the 3 months prior to the start of data collection.

2) **And** come from a hard-to-reach settlement where at least one household still remain. Since the assessment of hard- to-reach areas aims to fill in gaps in understanding the humanitarian context, targeting settlements that are no longer inhabited would not contribute to this.

Additionally, for REACH to assess the robustness of the information given by the interviewed KI, the questionnaire will ask the duration of the stay of the KI before leaving the H2R settlement. This won't be considered as criteria of eligibility but rather as an indication of the level of knowledge of the interviewed KI.

Finally, the questionnaire includes a section to allow a snowballing approach. The interviewees will be asked if they can refer REACH team to another KI that matched the eligibility criteria, from their settlement of origin or any other settlement in a hard-to-reach area. The contact details of these additional KIs will be collected for sampling purposes. FOs will make sure to keep an up-to-date contact list of potential KIs and local guides in order to build a strong network at the field level. This information will not be shared externally and will be stored only on REACH assets, protected by a password.

Targets per hard-to-reach area are as follows:

H2R DISTRICT	NUMBER OF TOTAL SETTLEMENTS IN H2R DISTRICT	SETTLEMENT MINIMUM (10% of total settlements) IN H2R DISTRICT	KII TARGET FROM EACH H2R DISTRICT
Adan Yabaal	36	4	24
Badhaadhe	66	7	42
Bu'aale	86	9	54

Bulo Burto	177	18	108
Caluula	69	7	42
Ceel Buur	56	6	36
Ceel Dheer	69	7	42
Jamaame	125	13	78
Jilib	126	13	78
Kurtunwaarey	80	8	48
Qandala	72	7	42
Qoryooley	331	33	198
Rab Dhuure	88	9	54
Saakow	55	6	36
Sablaale	89	9	54
Tayeeglow	122	12	72
Xarardheere	41	4	24
Total	**1,688**	**172**	**1,032**

Structured KI interviews

Based on information from UNHCR PRMN (Protection and Return Monitoring Network) and CCCM cluster settlement verifications, as well as based on information obtained in the field, Field Officers (FOs) will identify eligible KIs among new arrivals in the sites around Baidoa, Garowe, Kismayo and Mogadishu.

- The target for each of the four bases differs based on preliminary discussions with ACTED, CCCM and IOM DTM colleagues. Targets per base are as follows:

Location	**Target number of KIIs**
Baidoa	414
Garowe	42
Kismayo	210
Mogadishu	216
Total	**1,032**

The quantitative tool will include questions relating to the following sectors: Accountability for Affected Persons, Displacement, Food Security and Livelihoods, Health, Protection and WASH. An indication about the nutrition situation in assessed settlements will be provided by running a trends analysis on a selection of WASH, Food Security and Health indicators. The tool has been adapted to specifically capture vulnerabilities of target populations to climatic shocks, as well as the impact of these on the humanitarian situation in H2R areas; relevant stakeholders were consulted to provide feedback on the modifications.

Most indicators will be collected at the settlement level, except for a selection of individual-level indicators regarding the KI's profile and displacement. KIs which report displacement from the hard to reach settlement in September or October will also be asked a set of indicators regarding household food security and water security indicators (Food Consumption Score (FCS), Household Hunger Scale (HHS), Water Insecurity Experiences Scales (WISE)).

For most indicators, KIs will be reporting on the period between July-September and the moment they last left the settlement of origin, while the nutrition-related indicators KIs will be reporting on the one month preceding their displacement. The latter indicators, which will be subject to a trends analysis, include:

- % of settlements by categories of items reportedly available in the market where the majority of the population were going in the recall period
- % of settlements where food item prices reportedly increased in the recall period
- % of settlements where water prices increased in the recall period
- % of settlements where NFIs prices reportedly increased, in the recall period
- % of settlements where MOST people were reportedly skipping 2 or more meals a day in the recall period
- % of settlements with reported excess mortality in recall period
- % of settlements by type of primary source of drinking water used by most households
- % of settlements where ANY people were using surface water as their main source of drinking water in the recall period
- % of settlements reporting coping due to a lack of food
- % of households having enough water for drinking, cooking, bathing and washing
- % of HHs with access to functioning handwashing facilities with water available

Data collection will be organized as follows:

- **Field Officers (FOs) and enumerator training:** 2 full days for the training for the upcoming round in October, including 1 day of pilot.
- **Data collection:** 3 weeks of daily face-to-face data collection in informal IDP sites around Baidoa, Garowe, Kismayo and Mogadishu.
- **Data cleaning:** Daily data checking and cleaning will be conducted by the field and assessment team during data collection.

Semi-structured focus group discussions

The quantitative survey tool will be complemented by semi-structured Focus Groups Discussions (FGDs) aiming to aid interpretation and provide more information on the context of hard-to-reach areas. The FGD topics will be based on the information from the quantitative data to provide a wider understanding of the community's needs in the settlement of origin. The overall objective of FGD is to deepen REACH team's understanding of the context and provide a narrative text to quantitative indicators.

A total of 12 FGDs will be conducted, disaggregated by gender, 4 in each base, and each consisting of a minimum of 7 and maximum of 9 participants; the FGD participants will be purposefully selected using a hotspot-approach based on the H2R May round, so as to include as much as possible diverse experiences from the same settlement (FOs will endeavour to include individuals from remote H2R areas and individuals from urban H2R areas, and individuals with different socio-economic backgrounds). A team of 2 enumerators, 1 facilitator and 1 note taker, will be responsible for organizing and facilitating the discussions; note taking is chosen over recording, due to protection considerations.

Location	Number of FGDs	Number of FGDs with male participants:	Number of FGDs with female participants:
Baidoa	4	2	2
Kismayo	4	2	2
Mogadishu	4	2	2
Total	**12**	**6**	**6**

Field Officers (FOs) and enumerator training: 1 day for the training for the upcoming round in October, including half a day of pilot.
Data collection: 4 days of daily face-to-face data collection in informal IDP sites around Baidoa, Kismayo, and Mogadishu;
Data cleaning: Daily debriefings between the enumerator team and the assessment team will be conducted during data collection. Simultaneously, the assessment team will conduct the qualitative analysis.

3.5 Data Processing & Analysis

IMPACT data cleaning minimum standards checklist will be followed.[8] Detailed data cleaning procedures will be outlined in the data cleaning Standard Operating procedures in **Annex 1: Data Cleaning SOPs.**

Quantitative data: Every day, the surveys are uploaded on the REACH/IMPACT Kobo-server and downloaded by the Database Officer (DO) at the end of data collection. The DO anonymises and subsequently checks the dataset before it goes through to Field and Assessment Officers who will be conducting data checking and cleaning with log changes and deletions. The Assessment Officer will oversee and do the data cleaning templates for the Field Officers, who are in turn responsible for data checking and the supervision of field teams. The following protocols will be in place to ensure the quality of data collected:

- Daily data cleaning by Field Officers, who identify outliers, anomalies, and logical inconsistencies, and give regular feedback to enumerators through daily briefings and ad-hoc training. Data points that cannot be resolved through follow-ups with the enumerators or respondents will be deleted. If survey records have more than three outliers that cannot be checked, the entire record is deleted from the dataset. Also, if the duration of the survey taken is very long or short and the enumerators couldn't provide concise and clear justification, the entire survey will be deleted.
- Weekly data cleaning will be conducted by the Assessment Officer, who reviews data cleaning conducted by Field Officers and provides additional feedback to the data collection teams in regular communication, briefings, and training.
- The GIS and Database Officers do data aggregation and spatial verification, who provide feedback to ascertain settlement coverage.

Data collected with structured tools will be aggregated at the settlement level and disaggregated by KI gender, KI month of displacement, and settlements which reported drought vs settlements which did not report drought at the time of data collection.

[8] IMPACT Memo Data-Cleaning-Min-Standards-Checklist 28012020-1.pdf (reachresourcecenter.info)

Given that more than one quantitative survey will be collected for a given settlement, data from key informants reporting on the same settlement is aggregated to the settlement level using an R script which employs the following logic to calculate settlement-level responses:

- Single response questions: Most survey questions only allow a KI to select a single response. For this type of question, mode aggregation is used, whereby "I don't know" responses are dropped and then the most reported response is taken for each settlement. Should several KIs from the same settlement provide different responses to the same question, the result is reported as "No consensus".
- Multiple response questions: Mode aggregation is used, whereby "I don't know" responses are dropped and then all other responses reported by the KIs are presented.

Qualitative data: notes of the FGD discussions will be translated into English from Somali by Field Officers in case the enumerator cannot take notes in English directly. At the end of each focus groups discussion, the FO will ensure to have a debrief conversation with the enumerators who took the notes during the session. This debrief will allow the FO to read and cross-check the notes, and ask enumerators for clarification and/or complete their notes, if necessary. Qualitative data processing and analysis will be in line with this IMPACT guidance on qualitative data analysis.[9]

The notes of the FGD will be anonymised, and a data saturation grid will be developed to analyse the information gathered. Cross-case analysis will then be applied for the topics under consideration.

2. Key ethical considerations and related risks

The proposed research design meets / does not meet the following criteria:

The proposed research design…	*Yes/ No*	*Details if no (including mitigation)*
… Has been coordinated with relevant stakeholders to **avoid unnecessary duplication** of data collection efforts?	Yes	
… **Respects respondents, their rights and dignity** (*specifically by: seeking informed consent, designing length of survey/ discussion while being considerate of participants' time, ensuring accurate reporting of information provided*)?	Yes	
… Does not **expose data collectors to any risks as a direct result** of participation in data collection?	Yes	
… Does not **expose respondents / their communities to any risks as a direct result** of participation in data collection?	Yes	
… Does not involve **collecting information on specific topics which may be stressful and/ or re-traumatising** for research participants (both respondents and data collectors)?	Yes	
… Does not involve **data collection with minors** i.e. anyone less than 18 years old?	Yes	

[9] https://www.impact-repository.org/wp-content/uploads/2020/10/impact_guidance_qualitative-data-analysis-checklist_october2020_FINAL.pdf

… Does not involve **data collection with other vulnerable groups** e.g. persons with disabilities, victims/ survivors of protection incidents, etc.?	Yes	
… Follows IMPACT SOPs for management of **personally identifiable information**?	Yes	

3. Roles and responsibilities

Table 2: Description of roles and responsibilities

Task Description	*Responsible*	*Accountable*	*Consulted*	*Informed*
Research design	*REACH Assessment Officer (AO)*	*REACH Assessment Officer (SAO), REACH Research Manager (RM)*	*REACH Deput Country Coordinator (DCC), IMPACT HQ Research Desig and Data Unit (RDDU)*	*E OCHA Somalia Information Management and Assessment Working Group (IMAWG), Inter-Cluster Coordination Group (ICCG), Drought Operations Coordination Center (DOCC)*
Supervising data collection	REACH Field Officers (FOs)	REACH Senior Field Officer REACH AO REACH SAO	REACH RM REACH DCC	OCHA Somalia, ICCG, IMAWG, DOCC
Data processing (checking, cleaning)	REACH AO, FOs, GIS Officer (GISO) SDO	REACH SAO	REACH RM RDDU	OCHA Somalia, ICCG, IMAWG, DOCC
Data analysis	REACH AO, GISO and SDO	REACH SAO	REACH DCC REACH RM RDDU	OCHA Somalia, ICCG, IMAWG, DOCC
Output production	REACH AO	REACH SAO	OCHA Somalia REACH DCC REACH RM	ICCG, DOCC

			Research and Reporting Unit (RRU)	
Dissemination	REACH AO	REACH SAO	REACH DCC REACH RM	OCHA Somalia, ICCG, IMAWG, DOCC
Monitoring & Evaluation	REACH AO, REACH SAO	REACH RM	REACH DCC Humanitarian partners, OCHA Somalia HQ Research Department	OCHA Somalia, ICCG, IMAWG, DOCC
Lessons learned	All team members involved in the assessment (field team, data team, assessment team), partners if possible	REACH AO, REACH SAO	REACH DCC REACH RM HQ Research Department Humanitarian partners	OCHA Somalia, ICCG, IMAWG, DOCC

Responsible: *the person(s) who executes the task*

Accountable: *the person who validates the completion of the task and is accountable of the final output or milestone*

Consulted: *the person(s) who must be consulted when the task is implemented*

Informed: *the person(s) who need to be informed when the task is completed*

6. Data Analysis Plan

For the Data Analysis Plan for H2R October 2022, please refer to this link.

Annex 1: Data Cleaning SOPs

Data Cleaning Procedures for Key Informant (KI) Surveys

Below are the data cleaning standard operating procedures (SoPs) for KI surveys. SoPs are subject to change depending on feedback from Assessment Officers, the Database Manager or Field Coordinator during data collection.

OVERVIEW OF DAILY RESPONSIBILITIES

Field Coordinator (FC)

- Responsible for communicating security concerns from Senior Field Officers (SFOs) and implementing partners to larger Senior Management Team (SMT), who can then decide on whether or not to change the sampling framework and communicate that to Assessment Officers (AOs) and Database Manager (DM)
- Knowledgeable of AO to SFO communications regarding data cleaning issues via teams group and e-mails

Senior Field Officers (SFOs)

- Responsibility for Somalia regions divided between SFOs, each assigned to contact enumerator team leaders, individual enumerators and/or implementing partners if any issues with the data
- In constant communication between the database manager (DM) and field SFOs / individual enumerators regarding issues with data collection and data quality issues
- Ultimately responsible for progress tracking in google sheets
- Responsible for following up on feedback given to enumerators via google sheets

Database Manager (DM)

- Responsible for downloading, deleting and anonymizing raw data and passing to S/DBOs for cleaning
- Runs daily R data checking script with clean data, and raw data from most recent day to identify errors for SFOs to follow up on
- Reviews daily cleaned data outputs
- Updates Daily Progress Tracker (HTML)
- Makes final call on survey deletions. Updates deletion columns in online tracking sheets
- Responsible for overseeing changes to sampling framework and adjusting sampling targets accordingly

Lead Database Officer (Lead DBO)

- Responsible for supervising the visual check, split data for visual checks for J/DBOs and consolidate all the visual check results in single file.
- Runs data cleaning and checking scripts on daily basis
- Sends cleaning result to AOs for providing feedback for and follow up on it
- Consolidate all cleaning logs which feedback provided for them

Assessment Officer (AO)

- Responsible for reviewing cleaning result on daily basis and providing feedback to either DM, SFOs, GISO
- Responsible for compiling all subsequent feedback from DM, SFOs, and GISO and sending to DM
- In communication regarding security or logistical concerns that change sampling framework – communication lead by FC, DM, and SMT

Database Officers (J/DBOs)

- Visually check for patterns in the data in questions to see if there is any evidence of enumerators developing habits of always entering the same values.
- DBO's responsible for ensuring daily backups of their cleaned data occur and that proper file naming protocol is followed for cleaned data and cleaning log
- DBOs communicate all data issues to Lead DBO, who communicates feedback and issues to responsible DM to contact responsible SFO, who contacts field teams and individual enumerators to clarify any issues with the data
- Cross-checks household locations with village selected by enumerator. If location issues or spatial duplicates found, they are to be brought up with TAM, who will contact responsible SFO.

Individual Enumerators

- Ensure phones are fully charged prior to next day of data collection
- Ensure phones are set to the correct time and date prior to data collection. Achieved with steps below: SettingsàGeneral managementàDate and timeàAutomatic date and time AND Use 24-hour format ON
- Collect coordinates and finalize survey after asking final questions
- Upload forms to the kobo server daily

Additional information for DBOs:

- We've been automating most of the data processing stages and encourage our DBOs to implement their new ideas for developing this process. In addition, DBOs should familiarize themselves with all data cleaning scripts, in particular, translation, replacing cleaning log and pattern check logger, time check and etc.
- DBOs responsible for familiarizing themselves with the constraints of the kobo data collection tool
- DBOs to have the latest version of the kobo tool and this SOP open to understand each question code
- The column structure must be followed in order
- Any consistent pattern issues related to specific enumerators should be noted by DBOs
- Translations from Somali are harmonized to match existing categories - if possible. If answers do not fall into already existing categories but occur more than once, answers are harmonized as new categories

DATA PROTECTION

To maintain the safety and security of both respondents and enumerators, the following procedures will be followed:

- GIS points, settlement location, enumerator information and interview times are deleted from final dataset

- Settlement location will be removed from processed raw dataset that is shared among AOs and FC
- Province codes, district codes, and enumerators will be used for raw dataset that is shared among AOs and FC
- Database Manager is the only individual with full access to raw dataset

DELETION OF DATA

- Considering the deletion criteria, the deletion command will be coded in R script to apply on daily data processing
- Database Manager is responsible for deletion of surveys
- Database Manager receives cleaning logs from lead DBO and feedback from GISO, and compiles uuid of surveys to delete from dataset.

Criteria for deletion:

- All surveys completed in under 20 minutes or more than 80 minutes
- All surveys with a pattern match of 90% or higher to another survey
- All surveys with a respondent 17 years of age or younger
- All surveys without consent
- All surveys with 6 or more flags, without reasonable explanation
- Surveys located 5km or more from selected village, without reasonable explanation

DATA CHECKING SUMMARY

A – DUPLICATE CHECKS: PARENT WORKSHEET

Duplicate uuid's should not be present

#	STEP
1.	**Duplicate surveys are flagged for deletion**. We have already coded this in our data processing script. If there was any duplicate, it will be flagged for deletion.

B – TIME CHECKS: PARENT WORKSHEET

Survey should take between 20 and 175 minutes

#	STEP
1.	This survey should take between **20 and 70 minutes** to complete. **Under 20 minutes: Surveys will be deleted** as invalid as they are of suspected poor quality/to be fake. **80 minutes or more: Surveys will be deleted** as invalid as they are of suspected poor quality/to be fake. Surveys taking **longer than 70 minutes will be flagged for Area FC/FOs review.** If there is no valid explanation provided, then the interview will be deleted. Operations will be notified of this deletion.
2.	We're using audit files to check the duration of each survey and using R script for doing this operation.

C - TRANSLATIONS FROM Somali

Organization and Village names may be entered in Somali

#	QUESTION	ACTION
1.	**village_other, organisation_other** *Parent Worksheet*	**'Other' answers, entered initially in Somali.** Translations from Somali done within same cell in the *Data Checking* Tab. If possible, translations are first harmonized into groups that match existing categories. If answer falls into category that already exists, the 'other' category answer is corrected in the same cell.

D – SPECIFIC QUESTION CHECKS

Parent Worksheet

#	QUESTION	ACTION
1.	'Other' columns	Check that data entered into 'Other' column is translated, logical, and consistent with the context.

		Action: If data entered into 'Other' column matches any of the potential survey responses, re-classify that entry and log the change. If the entry cannot be reclassified, just translate.
2.	hc_push_main climate_shock	If push_factor ='drought' and 'drought' is not selected for shocks_aoo If push_factor ='flood' and 'flood' is not selected for shocks_aoo **1.** Flag the entry. **2.** Follow up with the enumerator on whether there is an explanation to this contradiction. Change accordingly if an explanation is provided. **3.** If an enumerator follow-up is not possible, follow up with the respondent and change. **4.** If neither of the follow-ups is possible, delete entries from the respective columns **5.** Communicate to the enumerators to make sure the question is understood and asked properly.
3.	crop_loss_reasons climate_shock	If crop_loss_reasons = drought and 'drought' not selected under climate_shock If crop_loss_reasons = flooding and 'flooding' not selected under climate_shock **1.** Flag the entry. **2.** Follow up with the enumerator on whether there is an explanation to this contradiction. Change accordingly if an explanation is provided. **3.** If an enumerator follow-up is not possible, follow up with the respondent and change. **4.** If neither of the follow-ups is possible, delete entries from the respective columns Communicate to the enumerators to make sure the question is understood and asked properly.
4.	reason_livestock_decrease climate_shock	If reason_livestock_decrease = drought and 'drought' not selected under climate_shock If reason_livestock_decrease = flooding and 'flooding' not selected under climate_shock **1.** Flag the entry.

		2. Follow up with the enumerator on whether there is an explanation to this contradiction. Change accordingly if an explanation is provided. **3.** If an enumerator follow-up is not possible, follow up with the respondent and change. **4.** If neither of the follow-ups is possible, delete entries from the respective columns Communicate to the enumerators to make sure the question is understood and asked properly
5.	Food_nut Aid_received	If food_nut = general food distribution, 'food_distro' and 'food' not selected under Aid_received If reason_livestock_decrease = flooding and 'flooding' not selected under climate_shock **1.** Flag the entry. **2.** Follow up with the enumerator on whether there is an explanation to this contradiction. Change accordingly if an explanation is provided. **3.** If an enumerator follow-up is not possible, follow up with the respondent and change. **4.** If neither of the follow-ups is possible, delete entries from the respective columns Communicate to the enumerators to make sure the question is understood and asked properly

E – ENUMERATOR PATTERN CHECKS

#	QUESTION	ACTION
1.	Select_one and select_multiple questions	Visual check that individual enumerators are not developing patterns for answers to these questions or answering the same answers for every survey

H – DAILY R SCRIPT FOR RAW DATA CHECK

Producing reports broken by region, the purpose of the daily R script check is to provide guidance to FOs and DBOs for enumerator follow up and provide an additional check against enumerator error.

R Script incorporates cleaned data as well as raw data from the most recent day of data collection

#	QUESTION	ACTION
1.	**progress check**	At overall and district level to determine day-to-day progress and ensure a timely data collection.
2.	**time checks**	Flags surveys with illogical time stamps including an end time that is before the start time, interviews submitted in the future, interviews that are either under 20 minutes or more than 70 minutes. Interviews under 15 minutes or more than 79 minutes will be automatically deleted.
3.	**survey frequency**	Flags enumerator productivity in most recent day of data collection. Enumerators with more than 5 surveys per day maybe notified to slow down and interviews may be deleted if number exceeds 5.

SUMMARY OF SAMPLING FRAMEWORK CHANGES

- **Step 1:**
  Individual enumerator or team leader raises security or logistical concern (including active conflict, natural disaster, change in mobile connectivity in area, hostility from local actors etc.) to FO. FC then alerts DM/AO and SMT.
- **Step 2:**
  FOs escalates security or logistical concern to Operations/SLO.
- **Step 3:**
  SMT determines sampling framework change and communicates sampling change to AO and DM. Alternatively, AO communicates sampling change to FC and FOs.
- **Step 4:**
  DM adjusts sampling framework to accommodate changes.