HOST COMMUNITY MULTI-SECTOR NEEDS ASSESSMENT
TERMS OF REFERENCE

SUMMARY

Mandating Body/ Agency	Inter-Sector Coordination Group secretariat	
Overall Research Timeframe (*from research design to final outputs / M&E*)	30/09/2018 to 28/02/2019	
Research Timeframe	1. Start collect data: 12/11/2018	4. Preliminary analysis presentation released: 05/12/2018
	2. Data collected: 02/12/2018	5. Products released: Jan-Feb 2018
	3. Data analysed: 04/12/2018	
Humanitarian milestones to inform	**Milestone**	**Deadline**
	Inter-cluster plan/strategy *2018 JRP reporting, 2019 JRP drafting*	01/12/2018
	Sector plan/strategy *2019 sector strategic plans*	01/12/2018
Detailed dissemination plan required	Yes (to be developed following presentation of preliminary results)	
General Objective	To inform evidence-based strategic planning of humanitarian response activities by the ISCG, sectors, and sector partners, through the provision of up-to-date, relevant and comparable information on the multi-sectoral needs of host community populations across areas hosting refugees in Cox's Bazar District, Bangladesh.	
Specific Objective(s)	1. Provide a comprehensive evidence base of multi-sectoral needs among host community populations, including: a. Comparison between affected Unions and Upazilas	

	b. Comparison at Union level between host community populations directly and indirectly affected by the refugee influx[1] 2. Provide a common dataset to inform reporting on 2018 JRP, development of 2019 JRP, and development of 2019 sector strategies.	
Research Questions	1. What are the needs, vulnerabilities and service gaps within host communities that can be addressed by humanitarian interventions across food security/livelihoods, education, protection, nutrition, WASH, shelter/NFI, health, and communication with commuities? 2. What is the severity of need within and across sectors? 3. How do these needs differ according to geographic area and proximity to refugee population?	
Geographic Coverage	All Unions in Ukhia and Teknaf Upazilas[2]	
Secondary data sources available	• 2011 Population and Housing Census • Bangladesh Multiple Indicator Cluster Survey 2012-2013 (UNICEF, 2014) • Livelihoods in the Teknaf – Ukhia Peninsula (FEG/WFP, 2017) • 2017 Refugee Influx Emergency Vulnerability Assessment (WFP, 2017) • Information Needs Assessment (Internews, 2017) • Impact of the Rohingya Crisis on Host Communities in Cox's Bazar and Bandarban Districts (UNDP/PRI, Forthcoming) • Integrated SMART nutrition survey 2018 February March 2018 (ACF, 2018) • Rohingya Crisis Host Communities Overview (ACAPS/NPM, 2018) • Rapid Needs Assessment in Cox's Bazar Phase I: March 2018 (FXB/BRAC, 2018) • Support to Bangladesh Host Communities and Institutions in the Joint Response Plan for the Rohingya Humanitarian Crisis (ISCG, 2018) • Crisis within the Crisis: A study of impact of Rohingya influx on the host community (COAST, 2018) • Ground Truth Solutions assessment of social cohesion in host communities in close proximity to camps (forthcoming) • NPM assessment of refugees living in dispersed sites within host community (forthcoming)	
Population(s)	All Bangladeshi populations living in Ukhia and Teknaf Upazilas	
Stratification	Geographical #:12	Gender: #: 2

---

[1] Directly affected communities are those experiencing direct contact with and living alongside Rohingya refugees. Indirectly affected populations are defined as those not experiencing direct contact with Rohingya refugees.
[2] St. Martin Dwip Union in Teknaf will not be included in the sample due to resource and time constraints, coupled with the low likelihood of humanitarian or early recovery partners engaging in this location.

	Unions	Male respondents Female respondents
Data collection tool(s)	Structured (Quantitative)	Semi-structured (Qualitative)
	Sampling method	**Data collection method**
Structured data collection tool # 1	Probability / Stratified simple random sample	Household interview (Target #): 3,143 (incorporation 10% non-response / ineligibility rate)
Target level of precision if probability sampling	95% level of confidence	Union level +/- 6 % margin of error (overall) +/- 10% margin of error (gender of respondent) Upazila level +/-5 % margin of error (overall) +/- 5% margin of error (gender of respondent) Whole of target population +/- 2% margin of error (overall) +/- 3% margin of error (gender of respondent)
Semi-structured data collection tool # 1	Purposive	Focus group discussion (Target #): 16 (8 per Upazila)
Expected output type(s)	Report #: 1	Factsheet #: 12	
	Presentation (Preliminary findings) #: 1	Presentation (Final) #: 1	
	Cleaned raw dataset #: 1	Analysed dataset #: 1	
Access	Public (available on REACH resource center and other humanitarian platforms)	
Visibility *Specify which **logos** should be on outputs*	ECHO, ISCG, REACH, IOM NPM/ACAPS Analysis Hub, Translators Without Borders	

RATIONALE

Since August 2017, an estimated 728,306 Rohingya refugees have arrived in Bangladesh's Cox's Bazar District from Myanmar, bringing the total number of Rohingya refugees residing in Bangladesh to approximately 895,631.[3] The

---

[3] UNHCR. Population data and key demographic indicators (as of 30 September 2018).

rapid and massive increase of the refugee population, concentrated in the south of the district in Ukhia and Teknaf, has had an enormous impact on Bangladeshi host communities' food security, economic vulnerability, market access, labour opportunities and environment. The dramatic increase in population has strained resources, infrastructure, public services and the local economy. The most affected areas have been the Unions in Ukhia and Teknaf Upazilas, but impacts are being felt throughout the district. Increased pressures include rising food, firewood and transport prices, pressure on water, basic services and the environment and competition for jobs. The education system has been impacted due to the hiring of both teachers and students to work on the refugee response. Increased traffic congestion on the roads has led to access and safety concerns. At the same time, the crisis has generated some positive impacts in terms of new labour and livelihood opportunities provided by the presence of a rapidly expanding humanitarian sector, and increased demand for goods and services by Rohingya themselves.

The majority of available data on the host community populations of Ukhia and Teknaf pre-dates the current influx of refugees. While a number of post-influx assessments have included host community populations, the information they provide has significant limitations in terms of being able to serve as the basis for strategic prioritisation and decision-making. First, assessments have often used differing definitions of "host communities" and hence differing sampling approaches, meaning findings are not comparable across different assessments. Second, coverage across sectors has been uneven. Third, assessment design and indicator selection have generally focused on the information needs of specific actors, offering little opportunity for the ISCG and its sectors to identify and define the information needs of the humanitarian response in a strategic and systematic manner. Overall, a rapid secondary data review conducted by the extended Information Management Working Group for coordinated assessments in support of the 2018 Joint Response Plan mid-term review has found that data on host communities for key indicators identified by ISCG sectors is slim to non-existent. Based on this data gap, a coordinated inter-sector Multi-Sector Needs Assessment (MSNA) focusing on host communities has been mandated by ISCG in order to better understand the needs of this population, and bring a system-wide and collective approach to data collection and analysis.

The assessment will aim to identify the severity and geographical spread of acute needs within the host community population. Its primary audience of this assessment is the ISCG, sectors, sector partners, and humanitarian donors. In this respect, the long-term development needs of the population that are more likely to be met by government or development agency intervention are not the core focus of this assessment. Similarly, this assessment is not designed to assess the impact of the August 2016 refugee influx on the host community (though may serve as a point of triangulation for actors wishing to do so).

METHODOLOGY

Methodology overview

The MSNA will implement a mixed methods approach in order to collect data on key indicators, involving a household survey and semi-structured focus group discussions (FGDs). The initial phase of the study will involve a household survey, which will aim to provide quantitative data on needs at the household level, generaliseable to the populations of the geographical and social strata identified for inclusion in the assessment. A second phase of FGDs/IDIs will be implemented subsequent to the household survey. Informed by joint analysis of assessment findings, FGDs will aim to provide data on more sensitive or complex topics (such as how/why questions) that cannot adequately be captured through survey approaches, as well as follow up on questions raised through analysis of survey data.

The assessment will be implemented as a collective, inter-agency initiative in order to ensure buy-in and use of findings across the response. It will be coordinated by the Inter-Sector Coordination Group secretariat and facilitated by REACH, with additional support to data collection and analysis respectively provided by IOM NPM and ACAPS. The assessment will be designed according to the inputs of Sector Coordinators, information management focal points, and sector partners, with technical support from REACH and ISCG. Data collection will be implemented jointly by REACH and IOM NPM teams. Preliminary analysis will aim to inform the development of the 2019 Joint Response Plan. Subsequent detailed Joint Analysis and interpretation of findings will be a coordinated process involving ISCG,

sectors/partners, REACH, and the IOM NPM/ACAPS Analysis Hub. The Joint Analysis process will inform the development by REACH of final products aimed at informing detailed sector and partner strategy and programming through 2019.

Population of interest

The assessment will aim to provide data on the Bangladeshi population of Teknaf and Ukhia Upazilas as the areas most directly affected by the refugee influx and thus most relevant to the operations of humanitarian and early recovery interventions. Within these areas, it will aim to collect data stratified the Union level in order to understand the differences in needs across different geographical areas. It will also aim to compare data based on host communities' proximity to refugee camps, thereby providing comparison of needs between populations more and less directly affected by the refugee influx.

Secondary data review

As discussed above in the Rationale section, several previous studies have focused on the host community population, with varying methodologies and uneven sector coverage. As part of the assessment design process, ACAPS will conduct a rapid secondary data review in order to inform finalisation of indicators selected by sectors. This will ensure that selected indicators are contextually relevant, and avoid duplication in cases where recent data are available.

Phase I: Survey Component

Indicator selection and tool design

Involvement of sector leadership and partners will be critical during the study design stage in order to ensure that data collected is relevant to their information needs, with REACH serving in a facilitating role. Sectors will initially be asked to provide lists of key indicators for inclusion in the assessment. Indicators should serve to identify key needs, vulnerabilities and barriers to service access within the host community that can be addressed by sectors' humanitarian programming within the timeframe of the 2019 Joint Response Plan and beyond. They should therefore be relevant to informing sectors' 2019 strategic planning and programme design, or monitoring outcomes that are expected to change as a result of sector interventions during 2019. Sectors will be encouraged to categorise indicators according to the sector-specifc outcomes of the 2018 JRP in order to provide a clear link between information to be collected and its intended use. REACH technical staff will meet with sector coordinators following initial development of indicator lists in order to refine and finalise them. Following the endorsement of all indicator lists by sector coordinators, ACAPS will conduct a rapid review for relevance and appropriateness, based on the results of the secondary data review described above. Based on this, REACH will submit a finalised list of indicators (including documentation of modifications from sectors' original requests, together with justifications for doing so) to ISCG for final endorsement, together with a finalised assessment ToR.

Following endorsement, REACH will develop question framings for the survey tool. These will then be fed back to sectors for input and endorsement. Prior to the start of data collection, finalised tools will be translated into Chittagonian; REACH will work with Translators Without Borders—supported through ECHO's Common Service for Community Engagement and Accountability—to ensure accurate translation and training of teams. REACH will code the final household tool into Kobo for use with smartphones.

If necessary, ISCG will facilitate presentation of the assessment approach and tools at relevant government forums (DC's office, RRRC) and solicit input from government actors.

Primary data collection

Data collection will be a joint initiative involving REACH and IOM NPM field teams. This support will serve the dual function of i) ensuring that data can be collected quickly according to the tight timeframe necessary to inform 2019 JRP development; and ii) encouraging collaborative and inclusive approaches to inter-agency assessment as a norm moving forward. The assessment team will aim for a 50:50 gender balance of male/female members. Research

teams will aim to mix REACH staff and IOM NPM, under overall coordination of senior REACH team members. Overall team strength will be 64 enumerators (32 REACH, 32 IOM NPM) split into eight teams of eight. Each team will be coordinated by a team leader. Team leaders will be overseen by the REACH Senior Field Coordinator, who will be supported in the overall management of field data collection by two Field Assistants. Additional specifics on protocols regarding joint data collection will be outlined in an additional SoP document to be jointly prepared by REACH and IOM NPM, if necessary.

Extensive training to all participating staff will be provided by REACH. This will include objectives and methodology of the assessment, field data collection protocols, clarification of tools/agreement on standards for recording responses, and multiple rounds of practice with tools. As part of this process, Translators without Borders will be asked to provide inputs on relevant socio-linguistic nuances of tools. The Gender-Based Violence (GBV) sub-sector will also provide a training on referral pathways in instances where GBV survivors self-report to the assessment team. Where needed, REACH will also seek input from ISCG sectors on the need for additional training (e.g. in prevention of sexual exploitation and abuse) in order to uphold the principle of do no harm during data collection. Tools and data collection protocols will then be piloted to identify and rectify problems before full roll-out of data collection. Data collection will take place throughout the month of November. Prior to data collection, REACH teams will liaise with Upazila Nirbahi Officers and Union Parishad chairs to inform them of the purpose and modality of the assessment and facilitate access to sampling locations.

For the household survey component, REACH will employ a stratified, simple random sampling approach to provide data generaliseable at the Union level at 95% confidence level and 6% margin of error. These levels have been determined, based on the resources available, to increase the likelihood of obtaining a large enough sample for statistically significant correlations based on populations' proximity to camps to be made at the Upazila level.[4] When weighted and aggregated, these data will also allow for aggregation at the Upazila and whole-of-population level at 95% confidence level and 5% margin of error. Sample size for each stratum will be calculated using 2011 census data, and is expected to yield a final sample of approximately 3,143 interviews based on a 10% non-response rate (see Table 1).

Table 1: Sample size by Union at 95% confidence level, 6% margin of error and 10% non-response rate

Upazila	Union	Population (HHs)[5]	Total sample	Male HH members	Female HH members
Teknaf	Teknaf Paurashava	4,752	278	139	139
Teknaf	Baharchara	4,822	276	138	138
Teknaf	Nhilla	8,235	290	145	145
Teknaf	Sabrang	9,960	290	145	145
Teknaf	Teknaf	8,431	290	145	145
Teknaf	Whykong	8,867	292	146	146
Ukhia	Haldia Palong	9,006	290	145	145
Ukhia	Jalia Palong	8,511	284	142	142
Ukhia	Raja Palong	10,596	286	143	143
Ukhia	Ratna Palong	4,238	280	140	140
Ukhia	Palong Khali	5,589	292	146	146
TOTAL		**83,007**	**3,148**	**1,574**	**1,574**

In the absence of easily obtainable household lists for each union, REACH will construct a sample frame that is based on a shapefile of the known settled area for each Union. Within the settled area, REACH will randomly

---

[4] Due to the lack of available population data, it has not been possible to stratify the sample according to distance from camps.

[5] Bangladesh Population and Housing Census 2011, Community Report, Zila: Cox's Bazar. Bangladeshi Bureau of Statistics and Information Division, November 2014.

generate a set of GPS sample points, with point distribution weighted according to estimated population density. A variety of sources including Open Street Map shelter footprints, and GHS and WorldPop raster population datasets, will be triangulated against each other to provide best guess estimates of settled area and population density. A back-up set of points will be generated for use in the event that the initial set of points does not yield the required sample size. Refugee camps and spontaneous sites will be excluded from the settled area to be sampled. Given the fact that several camps include parts of host community settlements, REACH will modify existing camp boundary shapefiles to ensure that areas within camp boundaries populated by host communities are included in the assessment.[6] Full details of the GIS-based sampling methodology are outlined in Annex II.

Enumerator teams will be provided with GPS bookmarks of sample points to be assessed, uploaded onto Maps.Me navigation software on smartphones. Enumerators will be instructed to navigate to the sample point, and select the nearest inhabited shelter for interview. Rationale and approach to selection of individuals to interview within each household is outlined in Box 1 below. Interviews will be administered using a structured questionnaire uploaded onto smartphones. Informed consent will be sought prior to each interview. In sampling areas that fall within the boundaries of camps, enumerators will be instructed to verbally check whether hosueholds self-identify as Bangladeshi or Rohingya before starting interviews, with self-identifying Rohingya not eligible for interview. Due to the sensitivities of collecting such data, household self-identification status will not be documented.

Box 1: Selection of respondents according to gender

According to the 2017 Inter-Agency Standing Committee Gender Handbook for Humanitarian Action, in-depth joint assessments should "be aware of possible biases in information collection and analysis. For instance, if women [are] not consulted, the identified priorities do not reflect the needs and priorities of the whole community." Traditionally, household survey methodology dictates that the—predominantly male—head of household is interviewed on behalf of the entire household. However, preliminary information needs for this assessment identified by ISCG sectors involve a mix of objectively verifiable household-level indicators (such as coping strategies or water points, etc.), and perception-based indicators that are likely to reflect the specific opinions of the individual respondent (such as sense of safety, perceptions of relationship with Rohingya refugees, etc.). In these circumstances, constraining respondents to male household heads is likely to significantly bias the sample in favour of men's views.

This assessment will therefore adopt a "good enough" approach that aims to balance the need to obtain a sample that is randomised in order to be generalisable to the entire population and accurately reflects the situation of each household, with the need to obtain a gender-balanced overview of priorities and needs. Under this approach, each enumerator in a 50/50 gender-balanced team will ask to interview the single household member who is most knowledgeable about the affairs of the household, who is of the same gender as the enumerator. All interviews will be drawn from the same set of sample points (i.e. there will not be a separate set of "male" and "female" sample points). Under this approach, the sampling design is expected to yield data that can be compared according to gender of respondent at 95% margin of error and 10% confidence at the union level, as well as data that are generalisable to households across the Union at 95% confidence level and 6% margin of error.[1] A small number of indicators related to GBV will only be asked of female respondents, meaning that data for these indicators will only be generalisable to 95% confidence level and 10% margin of error at the union level. The compromises inherent in this approach with respect to randomisation will be clearly outlined in the limitations section of the study's final report.

---

[1] An alternative approach of splitting household interviews by gender, with men and women answering different questions, was rejected as impracticable in terms of both the increased number of enumerators required, and the possibility of rejection of this approach by the target population.

---

[6] In camp 23, where Rohingyas are fully dispersed throughout the host community, all areas of the camp will be included in the sample frame.

Data processing and analysis

Cleaning and checking of household survey data will be conducted on a daily basis by REACH teams according to a set of pre-established Standard Operating Procedures (SoP). Data cleaning will include outlier checks, analysis of 'other' responses, identification and removal or replacement of incomplete or inaccurate records, and recoding and standardizing entries. All changes will be recorded in a data cleaning log. Data checking will involve review of GPS points of survey instances collected against originally assigned GPS points, length of time taken per questionnaire, and number of interviews per enumerator. A daily report of identified issues will be produced by REACH's data team and provided to field teams for inclusion in daily briefings. During data collection, assessment team leaders will monitor enumerator interview practices using a quality checklist and provided feedback on an ad-hoc basis and during daily briefings. Daily briefings will take place at the start of each day of data collection.

Following the finalisation of tools, an initial data analysis plan aimed at the provision of initial results will be developed. This will ensure linkages between questionnaire questions/ responses, reporting on indicators, and stratification of the sample. The plan will be reviewed by ACAPS prior to finalisation. Based on the finalised plan, REACH will develop an initial analysis syntax using R software to run data analysis automatically, ensuring that preliminary findings from the survey component can be computed as soon as a final clean dataset is available.

Following the presentation of preliminary findings, ISCG, REACH, and IOM NPM/ACAPs will aim to facilitate a Joint Analysis Workshop of sector leads and other key stakeholders will aim to examine the implications of data in greater depth, providing space for partners to comment on the implications of finding, and proposing additional levels of analysis. As part of this process, REACH will aim to work with ACAPS and ISCG on the development of needs indexes for each sector. These will aim to synthesise data from different indicators into snapshots of the overall breadth and depth of need for each sector. A single, multi-sectoral needs index will also aim to identify overall breadth and depth of needs across all sectors. Both sectoral and multi-sectoral indices will also be assessed according to household vulnerability. Input from the joint analysis workshop will also feed into the research questions for the study's proposed qualitative component (see below), along with an initial set of publications.

Throughout the assessment process, REACH's technical team in Geneva will conduct internal review and validation of tools and products in order to ensure they meet REACH's organisational quality standards. This will take place prior to, and is subordinate to, final validation by the ISCG secretariat.

Phase II: Qualitative component

Based on the results and analysis of the study's survey component, a supplementary qualitative component involving FGDs and/or in-depth interviews will be implemented to explore additional follow-up questions, explore causal relationships, and contextualise survey findings. This ToR will be updated following joint analysis to outline specific details of the study's qualitative phase.

Data protection and data sharing

All data collected during the assessment will be uploaded and stored on a dedicated ISCG server. Since GPS points constitute identifying data and hence a potential protection risk, both the server and all downloaded databases (e.g. for data cleaning, initial analysis) containing GPS points will be password-protected. GPS points will be removed prior to public release of all data. All identifying data will be stripped from databases/transcripts prior to public release. In the event that partners wish to access data complete with GPS points, they will be asked to submit a data release request to ISCG outlining how it will be used, and what steps will be taken to protect the data. ISCG will respond to these requests on a case-by-case basis according to their merits.

Products and visibility

Raw data and descriptive analysis tables will be released as soon as they are available, and will be usable under Creative Commons Attribution. Following Joint Analysis, initial data will be followed by an initial set of quantitative

factsheets at the Union level and for the assessment overall. Following the study's qualitative second phase, a final narrative report will be produced. All products will be publicly available on commonly-used web platforms including Humanitarian Data Exchange (HDX), HumanitarianResponse.info, the REACH resource centre and the IOM Bangladesh NPM portal as they are produced. Products will be drafted by REACH and shared with ISCG and NPM/ACAPS for review prior to release. All products will feature visibility of ISCG, REACH, IOM NPM/ACAPS Analysis Hub, Translators Without Borders, and ECHO as the primary donor for the assessment.

Collection of counter-trafficking data

As a condition of contributing resources to the assessment, and in order to minimuse duplication of data collection processes, IOM NPM has requested the inclusion of a short module on counter-traffickin in the main assessment questionnaire. Data from this module will not be made public and will not be included in any reports or analysis related to the main assessment. Instead, it will be separated prior to the finalisation of the study's raw dataset, and handed over to IOM NPM. In order to facilitate this process, REACH enumerators will receive a dedicated training on sensitivities of data collection around counter-trafficking issues prior to the launch of the assessment process.

TIMELINE

Given the urgent need for the availability of a summary dataset to inform the 2019 JRP drafting process, the assessment's initial timeline will be focused on producing this dataset and is available as Annex I of this document.

This timeline will be updated following publication of the initial dataset to include joint analysis and initial product drafting (January), supplementary qualitative research (February), and drafting/finalisation of the narrative report (March/April).

ROLES AND RESPONSIBILITIES

A summary of roles and responsibilities for key processes is presented in the table below:

Task Description	Responsible	Accountable	Consulted	Informed
Research design	REACH	ISCG	Sectors, IOM NPM/ACAPS, REACH HQ	Refugee Relief and Repatriation Commission (RRRC), District Commissioner's Office
Data collection	REACH, IOM NPM	ISCG	Sectors, ACAPS	Upazila Nirbahi Officers (UNOs), Union Parishads
Data processing (checking, cleaning)	REACH	REACH		
Data analysis	REACH, IOM NPM/ACAPS	ISCG	Sectors, REACH HQ	

Output production	REACH	ISCG	IOM NPM/ACAPS, Sectors, REACH HQ	
Dissemination	REACH	REACH	ISCG, IOM NPM/ACAPS	Sectors
Monitoring & Evaluation	REACH	REACH	REACH HQ	
Lessons learned	REACH, IOM NPM/ACAPS	ISCG	Sectors, REACH HQ	

ANNEX I: DETAILED TIMELINE

	Sun	Mon	Tue	Wed	Thur	Fri	Sat
October	14	15	16 - Finalised draft of indicator lists - Finalised draft of ToR including sampling strategy shared with NPM/ACAPS/ISCG	17	18	19	20
	21 - Draft questionnaire for sector review	22	23 - ACAPS completes review of indicator list - ACAPS suggests social cohesion indicators for inclusion	24	25 - Deadline for sector inputs to questionnaire	26	27
November	28 - Questionnaire finalised and sent to Translators without Borders - SoP for joint data collection agreed with NPM - ISCG endorses indicator list and ToR, shares in meeting/with meeting minutes	29	30	31 - Questionnaire received from Translators withour Borders	1 - Questionnaire coded into Kobo - Data cleaning drafted	2	3
	4 - Enumerator training	5 - Enumerator training	6 - Enumerator training	7 - Pilot	8 - Pilot - Pilot debrief	9 - Questionnaire/SoP revisions	10 - Questionnaire/SoP revisions
	11 - Questionnaire/SoP revisions - Enumerator refresher post-pilot - Data cleaning SoP finalised	12 - Survey	13 - Survey	14 - Survey	15 - Survey - Data analysis plan for preliminary analysis finalised	16	17 - Survey
	18 - Survey	19 - Survey	20 - Survey	21 - Survey	22 - Survey - R script for data analysis based on data analysis plan finalised	23	24 - Survey
	25 - Survey	26 - Survey	27 - Survey	28 - Survey	29 - Survey	30	1 - Survey
December	2 - Data collection completed	3 - Finalisation of data cleaning	4 - Finalised clean dataset	5 - Preliminary results available	6	7	8

ANNEX II: GIS-BASED SAMPLING PROTOCOL

The sampling strategy for the Host Community MSNA is based on a combination of two datasets. First, the Worldpop raster dataset from 2015 estimates population density across the assessment area based on corrected 2011 census data. The methodology for the generation of this dataset is outlined here.[7] Second, the Open Street Map (OSM) building footprint provides polygons of structures visible from satellite images across the assessment area. The dataset is built by volunteers and does not distinguish between inhabited shelters and other structures such as shops, schools or latrines. In addition, shelters hidden by vegetation coverage cannot be mapped.

The first step of the sampling strategy involved creation of a density raster for each Union. From the OSM dataset, a square grid was created covering the area with a resolution equal to the WorldPop raster (0.0008333 degrees). Centroids were then created for each structure in the OSM dataset, and the number of points inside each square of the grid was counted. From the WorldPop dataset, a second square grid was created covering the area with a resolution equal to the WorldPop raster (0.0008333 degrees). The value of the number of people per square in the WorldPop raster was then transferred to this grid. Following this, the two grids were joined, and the mean of the population values between OSM and WorldPop grids was calculated in order to align the two datasets.

In order to exclude refugee populations living in camps from the sample, ISCG camp boundaries were used to remove these areas from the grid. In cases where camps and host communities coexist together (camps 8E, 9, 23, 24, 25, 26, 27), NPM's Mahjee block boundaries were used to exclude areas with high densities of refugees while including host community populations not covered by the Mahjee block system.

Following this, the final polygon grid was converted to a raster with the same resolution as the grid size, and separated by Union.

In order to generate sample points for the assessment, the number of points required according to the sample size were calculated, and tripled for the sake of redundancy in case the initial set of sample points did not yield the target number of interviews (e.g. due to a higher than anticipated non-response rate). Based on these figures, sample points were generated using an R script and randomly distributed across the grid, with distribution weighted according to the population value of each cell. Once all points were generated, an initial sub-group of sampling points equal in number to the original sample size was extracted to serve as the primary sample, with the remainder of points reserved as back-up. The primary set of sample points were then exported to KML files and uploaded to enumerator smartphones.

---

[7] Worldpop, "Methods." http://www.worldpop.org.uk/data/methods/ (accessed 4 November 2018).