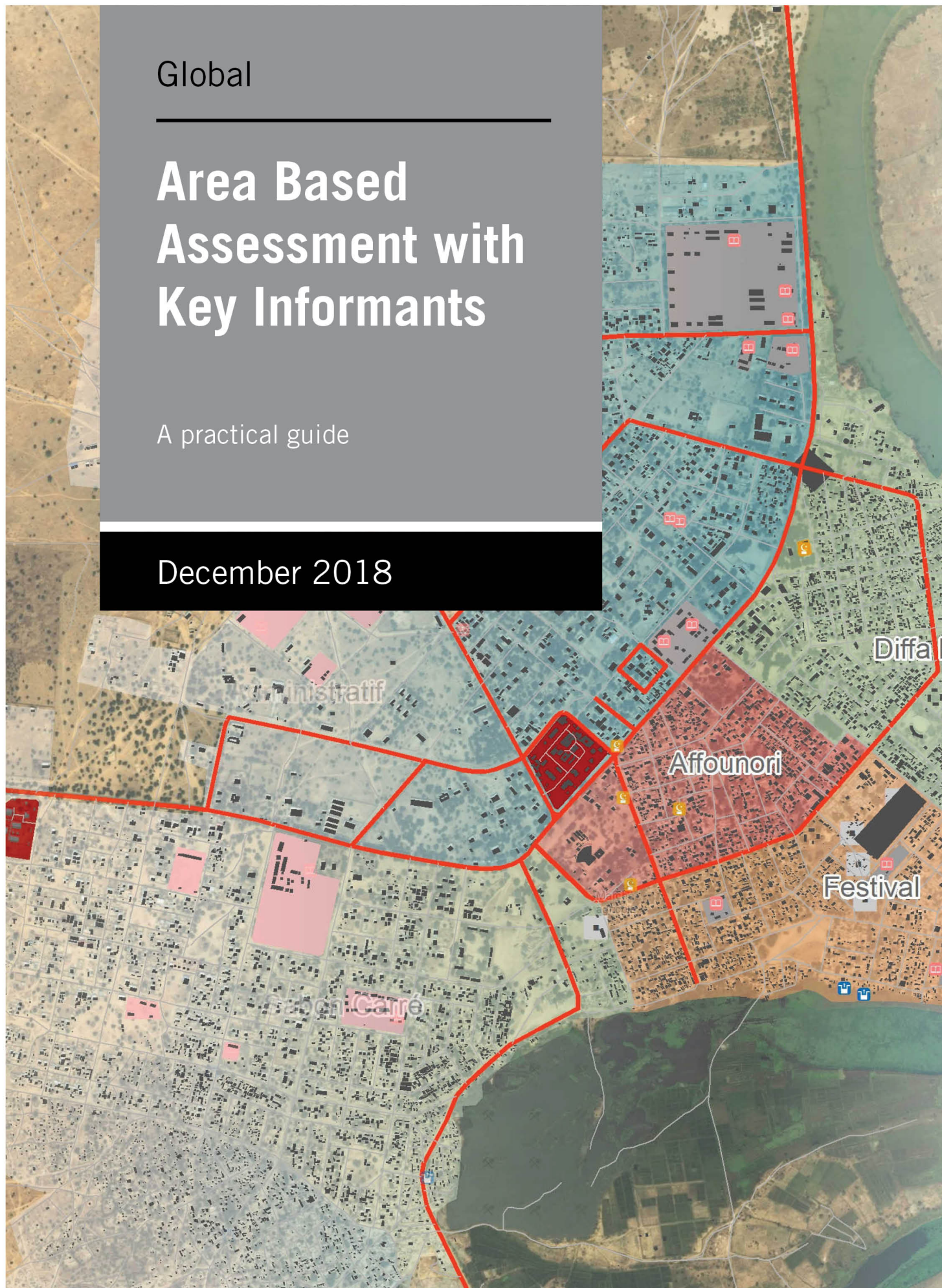


Global

Area Based Assessment with Key Informants

A practical guide

December 2018



Funded by the
United States
Government



UNHCR
The UN Refugee Agency

IMPACT

Shaping practices
Influencing policies
Impacting lives

About IMPACT

IMPACT Initiatives is a Geneva based think-and-do-tank, created in 2010. IMPACT is a member of the ACTED Group.

IMPACT's teams implement assessment, monitoring & evaluation and organisational capacity-building programmes in direct partnership with aid actors or through its inter-agency initiatives, REACH and Agora. Headquartered in Geneva, IMPACT has an established field presence in over 15 countries. IMPACT's team is composed of over 300 staff, including 60 full-time international experts, as well as a roster of consultants, who are currently implementing over 50 programmes across Africa, Middle East and North Africa, Central and South-East Asia, and Eastern Europe

Table of Contents

Figures & Tables	4
Abbreviations and Acronyms	4
Glossary	4
Introduction.....	7
Response efforts for out-of-camp displacement.....	7
Overview of the area-based approach.....	8
Purpose of the toolkit.....	11
Pilot testing for the toolkit	11
What the toolkit provides	13
Who should use the toolkit	13
How to use the toolkit	13
Section A: Identification of Area	15
Definition of an area	15
Identification of an area	16
Step 1: preparation	17
Step 2: Identification of participants.....	23
Step 3: Data collection.....	24
Step 4: Analyse and interpret MFGD data.....	29
Section B: Identification of Key informant	32
Challenges with using KIs for humanitarian response planning	32
Lessons learned from pilots.....	32
Proposed steps for KI selection: service and community areas	33
Step 1: Confirm research question	34
Step 2: Identify KI profiles.....	35
Step 3: Shortlist individuals for interview	37
Section C: Identification of Key informant with Network Data	38
What is Network Analysis?	38
Relevance of SNA	38

Proposed steps for network data collection	39
Step 4: Snowball to identify KIs	39
Step 5: Collect network data.....	40
Step 5: Select KIs.....	41
Section D: Collecting multisectoral data from KIs.....	46
Step 1: Considerations for KI multisectoral assessments.....	46
Step 2: Data collection.....	48
Step 3: Data interpretation.....	48
Conclusion	46
Annexes.....	51
<i>Annex 1: KI reliability measures</i>	<i>52</i>
<i>Annex 2: Complete MFGD questioning route used in Mafraq, Jordan</i>	<i>53</i>
<i>Annex 3: MFGD moderator and note-taker tips.....</i>	<i>57</i>
<i>Annex 4: Sample KI indicator list used in Mafraq, Jordan</i>	<i>59</i>
<i>Annex 5: Network related questionnaire.....</i>	<i>62</i>
<i>Annex 6: Example of accuracy of multisectoral indicators for a sample KI.....</i>	<i>63</i>
<i>Annex 7: Theoretical Framework (Social Network Analysis)</i>	<i>64</i>

Figures & Tables

Table 1: Extract from a KI matrix by REACH in support of 2015 HSOS	Error! Bookmark not defined.
Table 2: KI ranked based on network characteristics	43
Figure 1: Components of area based approach	9
Figure 2: Function of area based assessment.....	10
Figure 3: Pilot process for area-based assessment.....	12
Figure 4: Toolkit pathway.....	14
Figure 5: Key steps to conducting a MFGD	15
Figure 6: Example service and community areas.....	17
Figure 7: Example reference map for one MFGD in Arua, Uganda.....	22
Figure 8: Example of Community Areas in relation to official administrative levels	25
Figure 9: Example of a community mapping exercise in Arua, Uganda	27
Figure 10: Illustration of gaps or overlap in community boundaries.....	30
Figure 11: Example of a final community area map from Diffa, Niger	31
Figure 12: Overview of practical steps to shortlist potential KIs (without SNA)	34
Figure 13: Example of coverage area and communication pathways between community and service KIs.....	35
Figure 14: Overview of practical steps to use network data to select KIs (including SNA).....	39
Figure 15: Example of Adjimeri network in Diffa collected through snowball.....	41
Figure 16: same network with data collection based on a list.....	41
Figure 17: Example of closeness and authority.....	44
Figure 18: KI Network mapping from Diffa, Niger pilot.....	45

Abbreviations and Acronyms

ACAPS	Assessment Capacities Project
BPRM	U.S. Department of State, Bureau of Population, Refugees and Migration
CBO	Community-Based Organisation
CSO	Community Service Organisation
DRC	Democratic Republic of the Congo
FGD	Focus Group Discussion
GIS	Geographic Information System
HSOS	Humanitarian Situation Overview in Syria
IASC	Inter-Agency Standing Committee
IDP	Internally Displaced Person
IIED	International Institute for Environment and Development
INGO	International Non-Governmental Organisation
IOM	International Organisation for Migration
IRC	International Rescue Committee
KI	Key Informant
LAAB	Local Area Advisory Board
MFGD	Mapping Focus Group Discussion
NGO	Non-Governmental Organisation
PII	Personally Identifiable Information
OCHA	UN Office for the Coordination of Humanitarian Affairs
GBV	Gender-Based Violence
SDR	Secondary Data Review
SNA	Social Network Analysis
UNHCR	UN Refugee Agency
UN DESA	United Nations Department of Economic and Social Affairs

Glossary

Accuracy (key informant)	the average deviation of all or some of the key informant answers from the reference household level survey. For more detail on the method please see annex 1
Area-based approach	a geographically targeted, multi-sectoral, and participatory approach which may be applied in both urban and rural settings
Area-based assessment	the data preparation, collection, analysis, and interpretation components of area-based approaches
Area of knowledge	the geographic area or expertise about which each key informant is able to provide information
Community area	a socio-economic whole, comprised of people, resources, and common sets of values or identities – could be neighbourhoods in urban areas or rural villages
Edge	a term used in social network analysis to designate connections between nodes within a network – also known as links or ties
Equilibrium	an indication that, on the basis of the data collected or analysed hitherto, further data collection and/or analysis are unnecessary since it would be unlikely to meaningfully increase understanding (also referred to as saturation)
Key informant	an individual whose informal or formal position give them specific knowledge about other people, processes, or events that is more extensive, detailed, or privileged than other individuals
Node	a term used in social network analysis to designate individuals within a network – also known as vertices
Probability sampling	a sampling technique wherein the samples are gathered in a process that gives all the individuals in the population equal chances of being selected
Purposive sampling	a sampling technique wherein a researcher relies on subjectivity when choosing members of population to participate in the study
Questioning route	a pre-determined sequence for asking questions in the context of a focus group discussion

Secondary data review	a rigorous process of data collation, synthesis, and analysis building on a desk study of all relevant information available from different sources ¹
Service area	a catchment area for services and/or markets. These are often, but not necessarily, linked to administrative and/or geographical boundaries
Snowball sampling / technique	a sampling technique wherein households and individuals are selected according to recommendations from other informants. Each informant recommends the next set of informants.
Social Network Analysis	the process of investigating social structures through the use of networks and graph theory
Symbology	the set of conventions and rules defining how geographic features are represented with symbols on a map
Weighting	the process of emphasising the contribution of some aspects of a phenomenon (or of a set of data) to a final effect or result, often used during data analysis

¹ ACAPs, Technical Brief, May 2014; available online at: https://www.acaps.org/sites/acaps/files/resources/files/secondary_data_review-sudden_onset_natural_disasters_may_2014.pdf; accessed 19 April 2018

Introduction

People exposed to forced migration, such as refugees or Internally Displaced Persons (IDPs), are increasingly more likely to live in cities and towns – as well as rural out-of-camp contexts – rather than camp settings. The UN Refugee Agency (UNHCR) estimated that over half of all refugees lived in urban areas in 2017.² Growing alongside advances in policies and programmes tailored to out-of-camp displaced populations, assessments must also be designed to effectively and accurately discern the service gaps, the needs of displaced persons in out-of-camp settings to support response efforts.

Response efforts for out-of-camp displacement

Following the 2009 revision of its urban refugee policy,³ UNHCR acknowledged “that camps should be the exception and only a temporary measure in response to forced displacement”.⁴ By 2014, the agency formally committed to pursuing alternatives to camps, especially in situations of protracted displacement.⁵ Thus, signifying an important shift in displacement response to focus more intently on the unique needs of out-of-camp displaced populations, as well as the situational understanding for areas of displacement.

In 2010, the Inter-Agency Standing Committee (IASC) called for a “paradigm shift in humanitarian assistance in urban areas based on a district or community-based [approach], rather than an individual beneficiary approach so as to forge partnerships for assistance delivery and recovery with actors on the ground.”⁶ This shift posed as a considerable challenge for traditional project management tools and approaches designed for humanitarian settings.⁷ While prior out-of-camp responses were mainly designed for rural regions affected by floods or droughts or displacement across remote borders,⁸ today’s out-of-camp displacement is most often to urban areas in neighbouring countries - and these countries may already struggle to provide services to host populations. In its advocacy for community-based approaches, the IASC effectively encouraged agencies to better integrate contextual considerations that influence displacement outcomes. These considerations are vastly more complex in out-of-camp settings where services, livelihoods, and settlement are less defined and more dynamic than camp settings.

Despite the contextual complexity, out-of-camp contexts can foster a more impactful and cost effective response to displacement in the long-term.⁹ There is increasing recognition across both humanitarian and development communities that displacement outside of camps can be part of a long-term and even durable solution, particularly in cases of protracted displacement.¹⁰ Also now acknowledged is the need for an integrated approach alongside state, municipal, and other sub-national actors – one which clarifies, residency, movement, access to services, and employment entitlements for displaced people. However, there are still notable gaps in out-of-camp responses. Advances in area-based approaches may enable

² UNHCR, 2018, [Global Trends in Forced Displacement 2017](#).

³ UNHCR, 2009, [UNHCR Policy on Refugee Protection and Solutions in Urban Areas](#).

⁴ UNHCR, n.d., [Urban Refugees](#).

⁵ UNHCR, 2014, [UNHCR Policy on Alternatives to Camps](#).

⁶ IASC, 2010, [Strategy Meeting Humanitarian Challenges in Urban Areas](#).

⁷ Sanderson & Sitko, 2018, [Ten principles for area-based approaches in urban post-disaster recovery](#).

⁸ Earle, 2016, [Addressing urban crises: Bridging the humanitarian–development divide](#), International Review of the Red Cross.

⁹ ICRC, 2014, [Responding to Urban Displacement: What needs to be done](#).

¹⁰ BPRM, 2012, [PRM Principles for Refugee Protection in Urban Areas](#).

these gaps to be filled, especially when area-based assessments are used to inform how data are collected and analysed.

Data challenges for populations displaced to out-of-camp settings

Much remains unknown about displaced communities in out-of-camp areas as identification constraints hinder knowledge on the overall situation and preeminent needs of an area. When compared to regularly-monitored in-camp populations, less is known about the health, sanitation, livelihoods, food security, nutritional status, protection situation, and school attainment of out-of-camp populations.

There are a number of reasons why data on refugees and IDPs displaced by conflict or natural disasters who live in urban areas or other out-of-camp contexts remains limited in comparison to data of in-camp populations, these include:¹¹

- While camps constitute usually clearly demarcated areas of intervention with a well-defined actors and target population, out-of-camp settings are characterised by dispersed populations across a variety of settings and different degrees of coexistence between host communities and displaced people.
- There are large numbers of displaced people forced to live on peripheral peri-urban terrain unsuitable for residential development – lacking services and often exposed to risks of natural disasters – and may fear drawing attention to themselves, for they have nowhere else to go.
- Marginalised and stigmatised populations, such as Gender Based Violence (GBV) survivors or ethnic minorities, may be particularly vulnerable or hard-to-reach.

Overview of the area-based approach

As displaced people seek refuge outside of camps they tend to settle in marginal areas, creating or expanding informal settlements which are often not fitting officially designated neighbourhoods or jurisdictions of the sub-national government. These areas may not be recognised by the authorities and their inhabitants may not be given access to public services or provided with documentation. Hardly ever are such communities homogenous, and conventional methods of needs assessment have failed to distinguish between the chronic needs of poor host community households and the acute vulnerability resulting from an influx, or triggered by disaster or conflict. This is particularly critical when humanitarian assistance for displaced urban communities excludes host populations, thus may contribute to increasing social tensions.

Building on the call from the IASC's in 2010, there has been increasing uptake of area-based approaches to provide better responses in urban areas following a displacement. An area-based approach seeks to support populations in a specific location to transition effectively from relief to recovery, as well as strengthen resilience.¹² Guided by the core components of area-based approaches, available in *Figure*

¹¹ UNHCR, 2012, [State of the World's Refugees; In Search of Solidarity](#).

¹² Maynard, Parker, Yoseph-Paulus & Garcia, 2017, [Thinking bigger: area-based and urban planning approaches to humanitarian crises](#), IIED.

1, area-based assessments aim to provide quality data to inform planning and response for area-based approaches.

Figure 1: Components of area based approach



Area-based approaches, which can be applied for *assessment, planning, coordination, and programming*, generally include the following characteristics:¹³

- Focus on a defined location
- Include a multisectoral view of the needs of the population being analysed
- Participatory in nature, meaning that inclusion of affected populations living in the area is essential
- Rooted in strong collaboration between sectorial and key actors working in the area
- Adaptable to different contexts
- Work with existing local structures to ensure sustainability

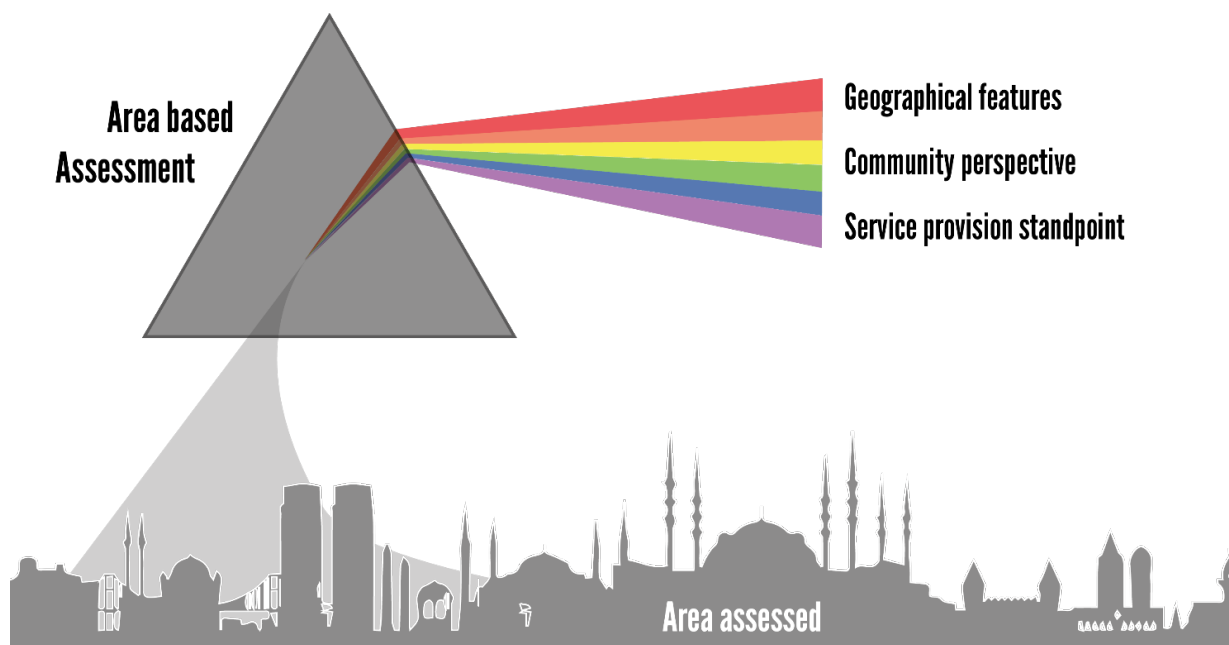
Despite the overarching characteristics, area-based approaches are interpreted in different ways by different agencies.¹⁴ There is, thus, an imperative need to improve shared understanding of the areas

¹³ Adapted from: Sanderson & Sitko, 2017, [Urban area-based approaches in post-disaster contexts. Guidance note for Humanitarian Practitioners](#), IIED.

¹⁴ Parker and Maynard, 2015, [Humanitarian response to urban crises: a review of area-based approaches](#), IIED.

where displaced people live and their interactions with others. This cannot be done without methodologies, including area-based assessments, to identify appropriate household needs assessments and area interventions. Area-based assessments incorporate a holistic approach to area distinction and Key Informant (KI) selection (see Figure 2).

Figure 2: Function of area based assessment



Challenges of common area-based approaches

There are several research methods to conduct area-based assessment which can be used in isolation or conjunction. The most common sampling approach, probability sampling, can be an accurate way to gather reliable data about the households and community needs within a given area.¹⁵ However, there may be challenges in the implementation of probabilistic sampling:

- There may be no accurate or geo-localised population numbers. Often outside camps the absence of reliable information about the displaced population might make it challenging or impossible to use probability sampling.
- When resources are limited, it may not be possible to ensure sufficient accuracy for probability sampling requires conducting a considerable number of household interviews to ensure representation of all the targeted groups.

Another approach, using KI methods to assess community needs, has sometimes been criticised because of respondent bias, for example excluding minorities and vulnerable groups, and lack of data accuracy. However, KIs are widely used by humanitarians because they:

- *Do not require accurate population numbers* – contingent on the area of knowledge being properly defined through area mapping

¹⁵ For more information, see Alexander & Cosgrave, 2014, [Representative sampling in humanitarian Evaluation](#), ALNAP.

- *Are generally quick* – enabling information gathering in a short time frame, for example after a natural disaster or rapid displacement flow
- *Utilise minimal resources compare to household assessments* – permitting recurrent data collection to monitor the area of interest over time
- *Are flexible* – supporting remote implementation in locations which humanitarian organisations cannot access

Purpose of the toolkit

This toolkit outlines how to facilitate an area-based assessment: inclusive of data collection, area mapping, and data analysis. This toolkit has been developed to support humanitarian response in out-of-camp displacement contexts and builds on existing literature, best practices, and lessons learned.

It should also be stressed that:

- This toolkit does not offer a one-size fits all approach. Rather, it provides general tools – augmented by extensive case studies and examples – which need to be adapted to the contexts being analysed.
- This toolkit may not be relevant in all kind of settings. Thus, implementing organisations should extensively review whether area-based assessments are appropriate and viable.
- The methodologies outlined in this toolkit complement – and certainly do not replace – the variety of methods¹⁶ which have been developed to gather data and communicate with host populations and communities.

Pilot testing for the toolkit

The methodologies and guidance included in this toolkit have been tested during field pilots undertaken by IMPACT in four (4) locations with out-of-camp displaced populations – Arua in Uganda,¹⁷ Diffa in Niger,¹⁸ Kabul in Afghanistan,¹⁹ and Mafrag in Jordan.²⁰ Lessons learned from the pilots were combined with extensive global and country-level consultations. In addition to the four pilots, additional area-based assessments by IMPACT Initiatives contributed to the lesson learned used for the toolkit.

Method

In each of the four pilot locations, mapping focus group discussions²¹ (MFGDs) were conducted to identify areas. KI were identified both through MFGDs and snowball techniques in order to recruit as many KIs as possible and have an exhaustive list of KIs in each community before collecting KI network data. Then, multisectoral assessments using KI and household data took place simultaneously with the intention to

¹⁶ Such as UNHCR, 2006, [Guidelines for Participatory Assessments](#) and UNHCR, 2017 [Needs Assessment Handbook](#).

¹⁷ IMPACT Initiatives, 2018, [Research terms of reference, Pilot study #4, Arua Uganda](#).

¹⁸ IMPACT Initiatives, 2017, [Research terms of reference, Pilot study #2, Diffa, Niger](#).

¹⁹ IMPACT Initiatives, 2017, [Research terms of reference, Pilot study #3, Kabul, Afghanistan](#).

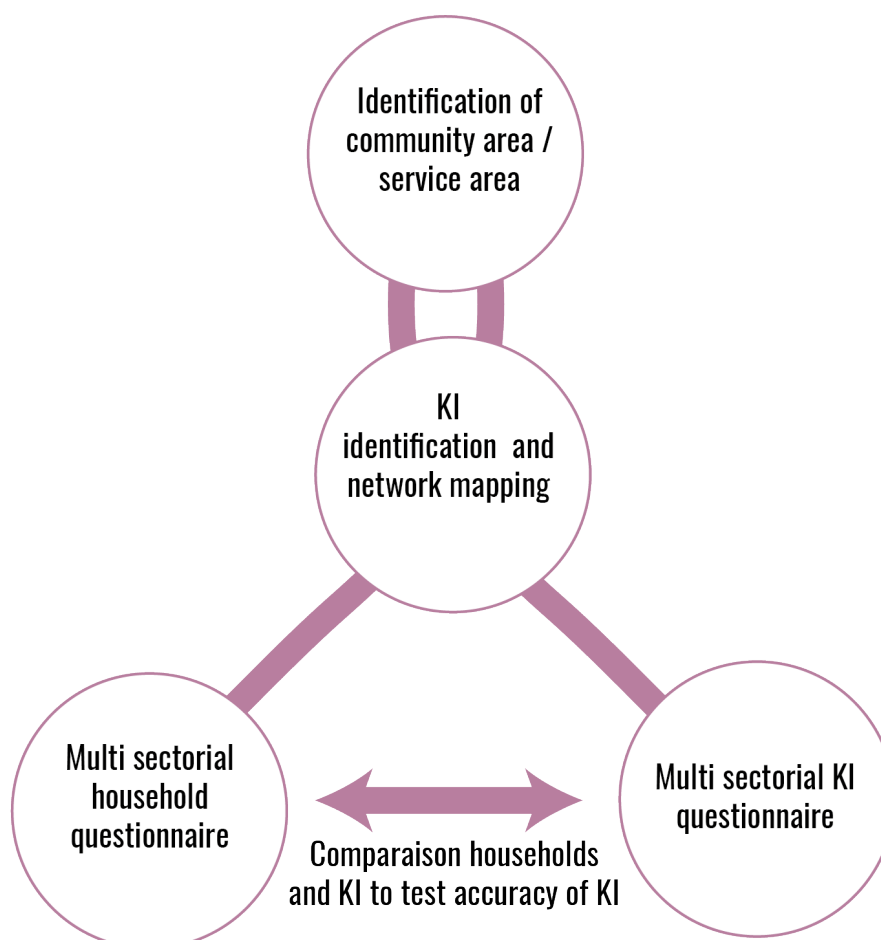
²⁰ IMPACT Initiatives, 2016, [Research terms of reference, Pilot study #1, Mafrag, Jordan](#).

²¹ See section B of this document and to IMPACT Initiatives, 2017, [Identifying areas for response planning & aid delivery](#)

study the reliability of KIs by comparing results.²² Probability sampling was used to select households. Comparable questions were asked of both KIs and household members to assess KI reliability.

The sampling of the household assessment gave results with 95 percent confidence and 10 percent maximum error margin at the community area.²³ For each KI, all the answers were compared with the household results aggregated at area level (determined during MFGDs).²⁴

Figure 3: Pilot process for area-based assessment



²² Simultaneous multisectorial assessments were conducted specifically to determine KI reliability for the pilots. It is not advised that simultaneous multisectorial assessments should be conducted for area-based assessments meant to inform future programming or response.

²³ In the course of the pilots, 2,249 KIs were interviewed and a total of 148,873 individuals' answers were compared with findings from 9,333 households.

²⁴ More information on reliability testing during the pilots is available in *Annex 1*.

What the toolkit provides

The toolkit aims to provide both the theoretical framework and the procedural guidance that are needed to implement an area-based assessment through KIs.

Guidance is presented using practical steps and instructions which are accompanied by explanations of the underlying theoretical framework. Contextual considerations and stakeholder consultation should be made before applying the theory and concepts within this toolkit.

The toolkit uses a combination of four components across four *Sections*:

- Section A. Defining areas:** designation the geographical distinctions of the assessment
- Section B. Identifying KIs:** determining which individuals can best inform community needs
- Section C. Identifying KIs using SNA:** incorporating Social Network Analysis (SNA) into KI selection
- Section D. Utilising KI data:** integrating KI data in two-way communication and analysis

Who should use the toolkit

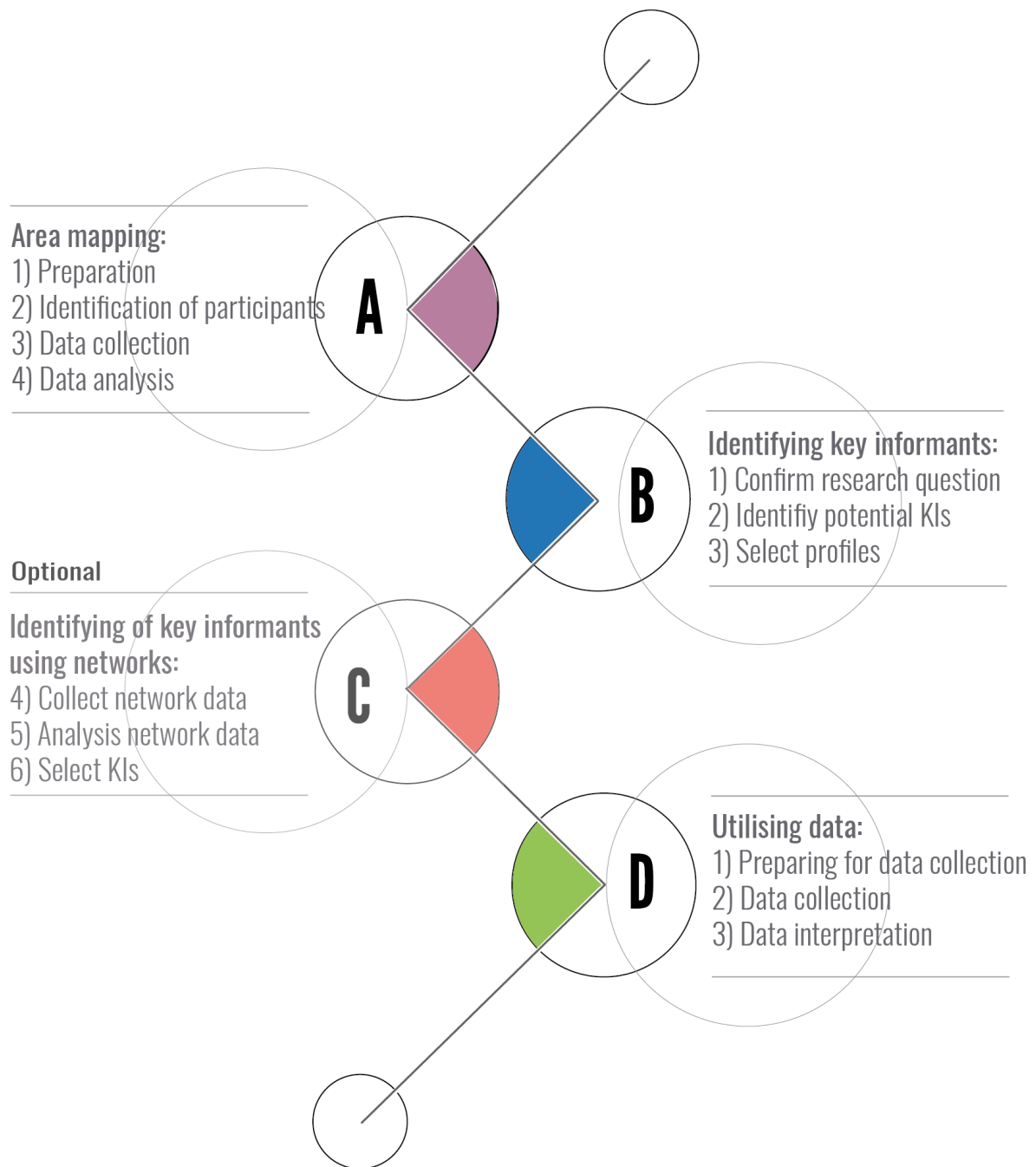
The toolkit is targeted to support:

- *Coordinators and programs staff* by informing on how the area-based assessment can feed an area-based response
- *Information managers* by providing support to design an area-based assessment with stepwise instruction for implementation

How to use the toolkit

There are two pathways included in this toolkit, depending on whether SNA is being conducted or not. For those interested in conducting area-based assessments with Social Network Analysis (SNA), this toolkit should be reviewed consecutively from *Section A* until *Section D*. However, if not conducting SNA, *Section C* of this toolkit may be skipped. The pathway for this toolkit is available in *Figure 4*.

Figure 4: Toolkit pathway

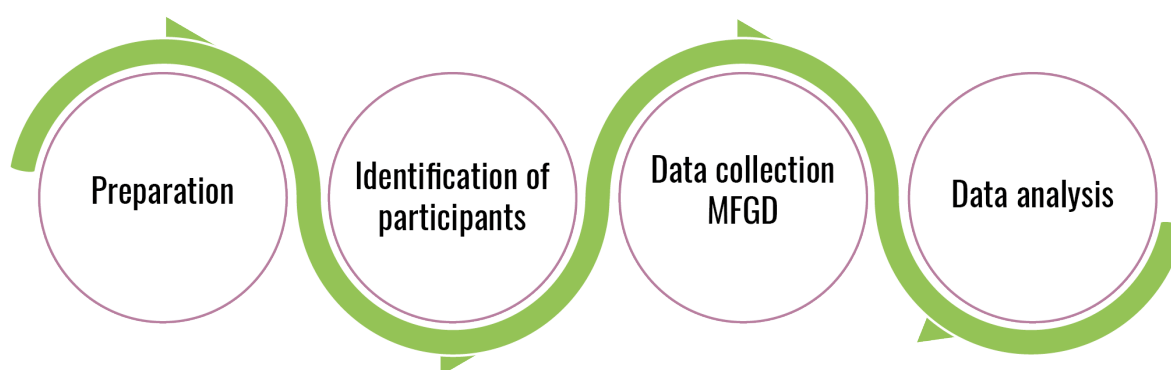


Section A: Identification of Area

Area-based assessments are geographically targeted, participatory, and multi-sectoral. They provide an effective method to understand needs and capacities at the local level and are dependent on an empirically based and collaboratively determined geographical area. This section investigates how areas are recognised and understood by the population, as well as how different areas are linked with the knowledge of KIs. Specific focus is made for two area catchments: service areas and community areas.

MFGDs should identify both service and community areas. A successful MFGD exercise needs preparatory work to ensure discussions are well-oriented, participants are chosen properly, and the overall objective is clear. There are four key steps to conducting a MFGD (see *Figure 5*).

Figure 5: Key steps to conducting a MFGD



Definition of an area

Area-based assessments build on the assumption that there are specific spatial features, physically or culturally defined, which are known to the affiliates or members of a given area or territory. The commonality of these known spatial features allows for clearly defined areas to be extracted from a larger geographic setting. For example, clustering of displaced persons within a specific location may support the designation of an area just as well as a river supports the designation of another area. Two common area typologies that are particularly relevant for area-based assessments due to their relevance for future programming and response are service areas and community areas. These areas, while distinct in their definition, are interrelated. Thus, the nuances of each area require specific attention in regards to assessments and response planning.

Service area

Service areas are catchments of service provision – legal, health, education, food, water, etc. – for a designated group of people across a specific geographical location. Depending on the type of information that is required for the intended area-based assessment, one or more service areas may be included in MFGDs. Often, but not necessarily, service areas are linked to clear administrative or geographical boundaries.

It is important to bear in mind that the area of one service may not necessarily coincide with the area of another service nor with the administrative boundaries. Nevertheless, delineating service areas should,

in most cases, be straightforward. Sufficiently precise information should be available from service providers, including municipalities and ministries, during background research and preparation phase then confirmed through MFGDs.

While the entire population within the geographical space of a service area may not be of interest when implementing programming, it is important to designate the extent of a service area as its relation to community areas as this may interact with future programming.

Community area

Community areas can be neighbourhoods “typically defined by social, economic, and physical features, which often serve as the basis for administrative and political recognition within larger jurisdictions.”²⁵ While community areas could be neighbourhoods in urban areas or rural villages, they may also be shaped by geographic features – such as valleys or waterways. Such community areas can be considered as singular units comprised of similar resources, socio-economics, values, and identities. They are often, but not always, smaller than service areas.

Community areas usually have their own history and culture shaped by shared customs, interests, values, and identity. These factors evolve over time and may influence the boundaries of a community area. Depending on its history and distinct characteristics, a community area may include a sense of belonging by its inhabitants that distinguishes it from neighbouring community areas. As such, a community area has its own identity that is formed and defined more by the local knowledge and perceptions of the area rather than by externally applied geographic demarcations – such as international frontiers or internal administrative boundaries. Ultimately an area is defined, named, and recognised endogenously (i.e. internally by the population inhabiting the area).

Example: Illustrations of community areas can be found across the globe and across cultures. For example, Berliners speak of a *kiez* – an informal word for a small neighbourhood – as a place defined by its inhabitants, where one lives. There are around 20 *kiezes* in and around the Berlin’s city centre. While the *kiezes* are acknowledged by residents and non-inhabitants alike, they are not delineated by the government.

Example: Piloting in Diffa, Niger revealed a layer of community area called a *lay* – a small unit, often comprising only a couple of households. Each *lay* is named for the first family that settled there or from the name of the first shop or other enterprise. Boundaries of *lays* do not accord with administrative boundaries used by Diffa local government. Similarly, these layers are called *mantaqa* in Mafrqa, Jordan and *parish* in Arua, Uganda.

Identification of an area

The information catchment for an area varies according to the type and method of information collection. The methodology suggested henceforward proposes to focus on identifying the two area typologies defined above:

1. *For supply side information* – the aim is to identify what this toolkit refers to as service areas. Depending on the type of information that is required, these can include one or more catchment

²⁵ USAID, 2011, [The “Neighborhood Approach”, a Means of Improving the Delivery of Humanitarian Assistance in Urban Areas](#)

areas for services, markets, and/or land use management. Often, but not necessarily, they are linked to clear administrative or geographical boundaries.

2. *For community information* – such as vulnerability, capacity, and demographics – the aim is to identify community areas. These are areas where people share key characteristics or, at a minimum, feelings of belonging. They can be neighbourhoods in cities, villages in rural areas, or separately identified areas by community members.

Shared geography of areas

While distinct in their functions, community and service areas may share some geographical locations. It was found in during the pilots implemented by IMPACT, that service areas tend to be larger than community areas and may include more than one or parts of other community areas. See *Figure 6* for an example of how community areas and service areas may present in relation to each other.

Figure 6: Example service and community areas



Step 1: assessment preparation

Before beginning of the assessment, the study team must identify the overall scope of interest and determine hypotheses around the geographical and non-geographical influences of a given area. While specific research questions regarding the situation of an area are not necessary at this stage, it is important to narrow the scope of interests to ensure practicality and to have a preliminary hypothesis about area delineations to ensure feasibility. Having a specific scope of interest is especially relevant when gathering information about service areas, as service areas may vary across service typologies, while community areas are a relatively static concept. While both the scope and hypotheses may be malleable or change across the course of the preparation phase, it is important that both are identified before beginning to liaise outside the study team.

Example: During the preparation of the pilot in Arua, Uganda, the team realised that the initial definition of a community area led to the delineation of community areas that were too small for effective

planning. It was then decided to use the *ward*, rather than community area, to gather insight from local authorities and community members in the determination of community area boundaries.

Connect with the local authorities and community members

After identifying the scope of interest, contact the authorities (where appropriate) and community members to inform them about the exercise. It is vital to inform both groups about the objectives of the exercise to ensure a collaborative approach, to refine the scope of interest, and avoid any misunderstanding of intentions. Ask the authorities to identify official boundaries and provide any existing maps of the area. This will enable understanding of the difference between an area of knowledge and official administrative boundaries. Connecting with local authorities and community members may also support the later application of MFGDs, as these individuals can provide as contacts when looking for participants in the MFGDs. Local authorities and community members may also provide secondary data to use for further refinement of hypotheses.

Setting up a participatory assessment forum, for example a Local Assessment Advisory Board (LAAB) in the community can be beneficial. This LAAB maybe composite of NGOs, CBOs, local government, community members or other local institutions which are identified during background research. Members of a LAAB can give advice on methods, sequencing of questions, as well as share information about the results and, more broadly, offer a forum for exchanging lessons learned on community outreach and mapping.

Conduct background research

A Secondary Data Review (SDR) enables identification of existing community and administrative boundaries, as well as information regarding available or utilised services, infrastructure, and markets. An SDR will also help to identify stakeholders for MFGDs – such as service providers, community members, NGOs, CBOs, and local government. Findings from the SDR provide the basis for building hypotheses on community and service areas. MFGDs – the use of which is explained in the following section – test their validity.

During the process of gaining background research, it is important to take note of characteristics that may indicate the thresholds for a service area or the shared knowledge about a community area. These noted characteristics for areas should then be presented during MFGDs to explore what boundaries have the strongest significance within an area.

Potential shared characteristics of a community area:

- Available services (such as schools, places of worship, and medical centres)
- Natural features (such as mountains, ravines, and rivers)
- Infrastructure (such as roads, bridges, and tunnels)
- Resources (such as water, and land)
- Common livelihoods (such as cultivation, trade, or provision of services)
- Common socio-cultural characteristics (such as ethnicity, faith, or nationality)

The area characteristics are rooted in their geographical locations and may differ depending on whether an area is in a rural or urban setting. For example, natural features may be defining characteristics in rural settings while shared livelihoods may demarcate areas in urban settings. Depending on the socio-political context, social factors may be more or less relevant than physical factors in demarcating community areas. These factors should be anticipated through background research before conducting MFGDs.

While it is important to pre-define the potential area to be targeted before initiating the MFGDs. Depending on the objective of data collection, these pre-defined areas could involve a cluster of informal settlements – as was the case during the pilot in Afghanistan – or a city-level analysis of neighbourhoods – as was done during in Jordan, Niger, and Uganda.

Questions to answer before pre-defining the service area:

- *What service-related information is useful for the planned response?* While there may be an abundance of varied service areas, it is important to narrow the scope of a service area to only those that are relevant to the overarching planned response that the data are meant to inform.
- *To what extent should service layers be aggregated?* For example: if there are three health areas in a small city, it may be decided to aggregate them.
- *Are the service layers overlapping?* Insofar as possible, layers of different services should be aligned to facilitate inter-sector comparison.
- *What disruptions may interfere with identification of the service area?* In conflict zones or besieged cities, service areas could be more difficult to identify, since most services could be disrupted or destroyed. In such cases, the identification of service areas could be based on areas controlled by particular combatants.

Questions to answer before pre-defining the community area:

- *Can this area be used for planning of aid delivery?* Similarly, it is important to identify if there are stakeholders in the area who can be identified as implementation partners.
- *Does the area provide sufficient granularity to implement area-based programming?* Keep note of what geographical information is used in the identification of a community area, even in relation to socioeconomic and sociopolitical constructs.
- *Is this area recognised as an area of knowledge?* Determine if this area is recognised by multiple, groups of people and is there consensus on its distinction.
- *Is this area an administrative boundary?* Keep note if an area aligns or contends with administrative distinctions, as this can influence the planned activity.

Clarify terminology

Based on the SDR, it is imperative to ensure that the vocabulary used during the assessment is clearly defined and understood by everyone. When collecting information during background research double check that there is clarity among the study team (including MFGD moderator and note-taker) about the type and size of the areas that are to be identified. Below are some sample questions to clarify with team:

- What do we mean by service area or community area?
- Do these correspond (roughly) to neighbourhoods or to smaller areas?

- What comprises a neighbourhood?
- What is the difference between a neighbourhood and settlement and what terms are used in local languages to refer to them?

Create the questioning route for MFGDs

Having a predefined question sequence is essential to ensure both comparability between the findings of each MFGDs and to ensure all discussion points are covered. It is important to include in MFGD questioning routes not only questions about the limits of a certain area but also why such limits may exist. In other words, it is important to ask, “What makes this area different from its neighbouring area?” Importantly, participants in MFGDs should also propose individuals who they perceive as having a good understanding of the area. These suggested individuals will inform later potential KI identification.

Example of MFGD topics from Mafraq, Jordan²⁶

- Identify the location(s) and recognise or add key landmarks to the map
- Indicate where they accessed such services as bakeries, barbers, schools or cafés
- Give the name of the neighbourhood (locally known as a *mantaqa*)
- Identify any differences between inhabitants of the neighbourhood – for example, with regard to income and ethnicity
- Demonstrate at what point on the map the *mantaqa* starts to be different
- Indicate until what point on the map they felt able to talk on behalf of neighbours about shared daily issues, needs, and opportunities.

Create reference map for MFGDs

The map is the centrepiece of the MFGD and thus forms part of the questioning route. The map orientation should be clear and without ambiguity, and potentially defining characteristics should be included. It is therefore imperative that the team is familiar with the map and the community surroundings before starting the session.

Important considerations when creating a reference map:

- Prepare several versions of the reference – with and without satellite imagery – before determining the most appropriate.
- Print separate maps for service area and community mapping, even within the same group of MFGD participants.
- Do not display the neighbourhood boundaries on the map for the community area mapping exercise to limit the bias of participants toward the community boundaries.
- Include as many points of interests or landmarks (such as hospitals, other public buildings, schools, other types of infrastructure and road names) as possible on the maps that will be used for the MFGDs. This makes it easier for the participants to orient themselves; however, the map should not be unclear or confusing because of these points of interest.
- Adjust the points of interest on the map depending on the MFGD participants.

²⁶ For the complete questioning route used in Mafraq, see Annex 2.

- Ensure that the map is wide enough in case the area is bigger than expected. If the conversation extends the printed area, take notes.
- Ensure that field teams are aware of the geographic coverage of the maps and know the larger location well. They will be map translators if it is found individuals lack map interpretation capacities.

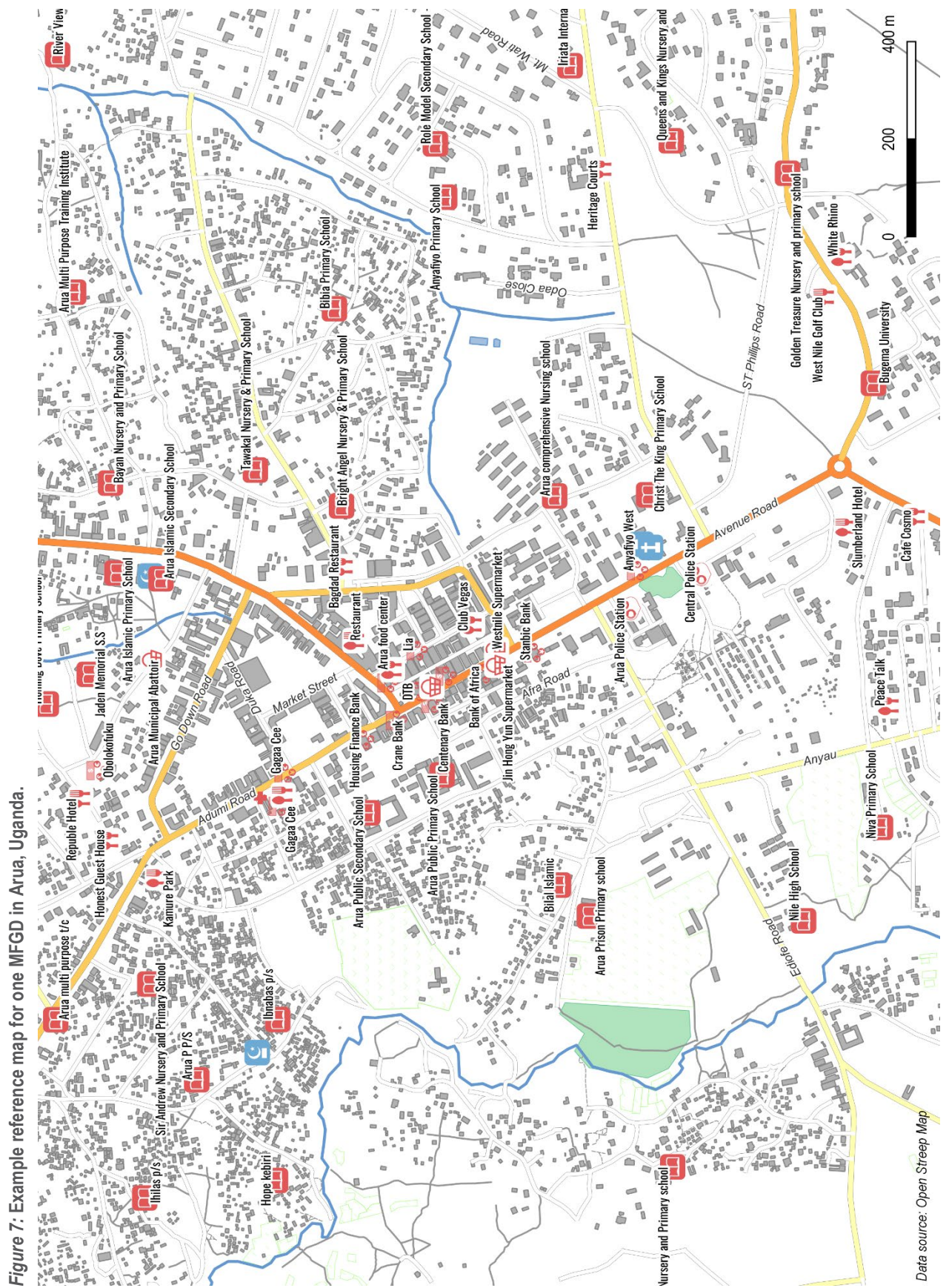


Figure 7: Example reference map for one MFGD in Arua, Uganda.

Data source: Open Street Map

Step 2: Identification of participants

Each MFGD should include a minimum of four (4) and no more than eight (8) participants. As much as possible, participants should not live in the same block, street, or quarter. Rather, participants should represent different locations within the area of focus for the MFGD. Neither should participants be engaged in the same livelihoods. Both host and displaced community members should also be represented in MFGDs, but in different MFGD groups.

While it is important to ensure a relatively heterogeneous mix of participants to capture differences in perception about the community, it is important to be aware of barriers that may inhibit certain individuals from participating in a MFGDs. Social norms in the community may discourage some participants from speaking up in mixed-groups. Thus, it may be prudent to organise segregated MFGDs. A common example in certain contexts of this is gender-segregated MFGDs.

CASE STUDY CONSIDERATIONS: DIFFA, NIGER

In Diffa, it was found during background research that there were two different levels of an area, the *lay* and the *quartier*. Because of the small size of the *lay* and lack of common understanding around the *lay* boundaries, the MFGDs were designed to investigate the *quartier* level. Discussion participants were selected from different *lays*. While the *lay* was not the focus of the MFGDs, during the discussion community members were asked to demarcate the *lay* inside the *quartier*.

Multiple MFGDs should be conducted and triangulated, though the exact number of MFGDs is dependent on the hypothesis on the selected area, additional MFGDs should be held in each area until the data for the topic meet equilibrium.²⁷ At least one MFGD should be conducted for each specified group of interest in an area. At a minimum, at least four MFGDs should be conducted per community area (i.e. displaced females, displaced males, host females, host males). However, further specification of groups may be necessary if there are specific perspectives needed by a different group (ex. ethnic minorities).

Considerations for service area MFGDs

MFGD participants should be selected using purposive sampling²⁸ method per each service to be mapped. Conversations and findings during SDR should inform participant selection for service area MFGDs. Selected individuals should have technical or practical knowledge of a specific service across the area. They must be able to describe the operating status of services and infrastructure, as well as primary service access points in service catchment areas. The participants need to represent the different perspectives of the service delivery and utilisation. Documenting the function and community role is important to factor in the bias in answers during the analysis phase.

Example: When assessing the service area of education in Diffa, head-teacher, school directors, parents, municipal services in charge of education were selected for the MFGDs.

²⁷ See *glossary*: An indication that, on the basis of the data collected or analysed hitherto, further data collection and/or analysis are unnecessary since it would be unlikely to meaningfully increase understanding

²⁸ See *glossary*: A sampling technique wherein a researcher relies on subjectivity when choosing members of population to participate in the study.

Considerations for community area MFGDs

Participants must have deep knowledge of specific communities, as well as being able to read maps of the area. Individuals involved in community area mapping differ from those involved with service areas in that they do not necessarily have the service-specific technical expertise to describe the status or capacity of infrastructure. However, they should be able to identify the boundaries of a community and describe the characteristics of its inhabitants. They may also be asked to map key service delivery points in their communities such as water access points, health centres, public latrines, and schools.

Step 3: Data collection

Materials and arrangement

The MFGD should be conducted by a moderator who will help facilitate the discussion and supported by a note-taker who will take notes of key topics.²⁹ MFGDs should take place in a safe and friendly space with minimal distractions, preferably in a commonly known setting such as school, municipal building, or community centre. Participants should be arranged in a circle around the detailed reference map – or with clear sight of the projection area – along with the facilitator. However, the note-taker's location – either outside or within the circle – may depend on the group dynamics. Drinks, snacks and name labels may help participants feel comfortable and energised during the discussion.

Equipment needed for the MFGD includes:

- A recorder to capture the discussion and enable retrospective checks (if the context allows)³⁰
- A detailed, large reference map of the immediate area within which the participatory mapping is conducted. A projector may also be used to project the map onto walls
- A clear legend on how to annotate the map
- A less detailed, smaller reference map of the overall area on which adjacent community areas can be identified by the participants
- Marker pens in several colours for participants to mark boundaries, as well as key infrastructure and services.

Moderation and documentation

Good facilitation and documentation of MFGDs depends on the combined efforts of both the moderator and the note-taker. It is important to work as a team and agree upon the set roles. Prior to starting the MFGD, the moderator and note-taker should discuss procedures (such as consent and turning on the recorder), expectations, and techniques to assist each other. These techniques may include agreement on non-verbal silent ways to communicate during the MFGD, such as use of hand signals, when important issues arise. To ensure the highest possible data quality, the team should debrief on all materials from the MFGD immediately after the MFGD to capture the key information and flag any missing key points from the discussion to the maps or questionnaires.

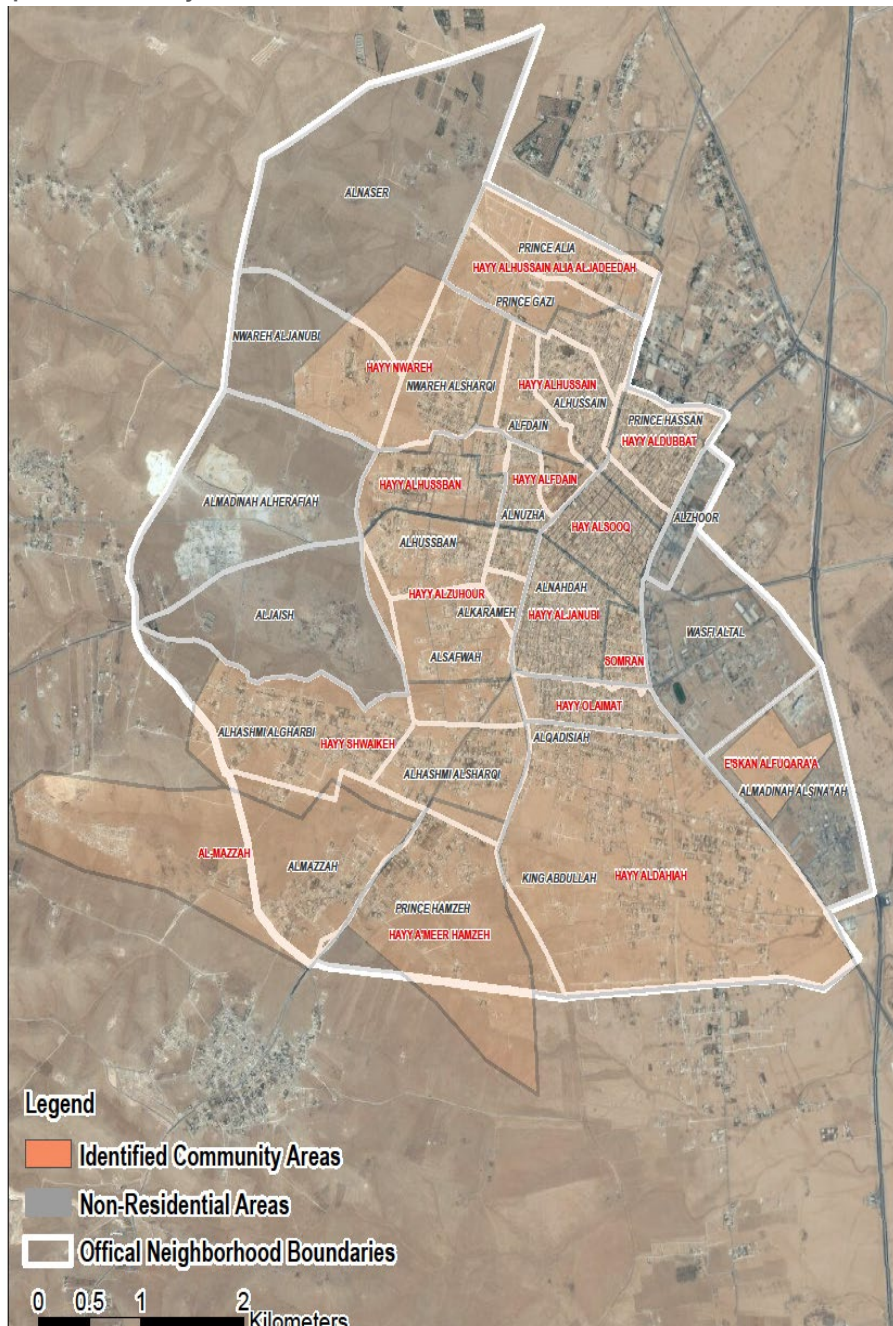
²⁹ The objective is not to record the discussion verbatim (i.e. not transcribing). See more information on MFGD moderation and note-taking in Annex 3.

³⁰ Consent is necessary for all aspects of data collection, and MFGD consent forms need to include reference to any audio data collection.

Preliminary crosscheck

To help the data collection team guide the participants in identifying the geographic areas of interest, the defined service area or community area can be cross-checked against the other borders identified and, ideally, adjacent to the defined area. Depending on the type of key shared characteristics that determines an area, the delineation above and below may differ in relation to official administrative levels. An example of community areas in relation to administrative boundaries is included in *Figure 8*.

Figure 8: Example of Community Areas in relation to official administrative levels



It is important to note that community areas may not be neatly nested within official boundaries or service areas. Instead, these areas may intersect, depending on which types of shared characteristics delineate each area.

Informed consent

Informed consent should be collected throughout all stages of data collection, and agencies should refer to existing standards regarding informed consent.³¹ However, particular considerations should be made regarding informed consent when audio is recorded (i.e. MFGDs). Audio recorded can be considered Personally Identifiable Information (PII) and is, therefore, in need of specific protections and considerations. Field staff should be extensively trained to ensure that the notion of informed consent is understood and respected. No data collection, especially that which includes PII, should be collected until the informed consent process is fully completed and consent is provided.³²

Once collecting informed consent and the accompanying data, agencies should take special care to safeguard audio recordings and accompanying transcripts. This safeguarding should be inclusive of having clear documentation of how such data will be transported, stored, and utilised, as well as who will have access to these data. Access to audio data should be reduced to a minimal number of persons. While redacted transcripts may be shared more broadly, specifications regarding who can access the redacted transcripts should also be determined. In all instances, MFGD participants should be provided general information about who will have access to which data and how these data will be used. All audio recordings and transcriptions with PII should be secured in a locked and stable storage device for physical documents, while soft copies should be guarded under password protection.

Orientation

After gaining consent and creating consensus on the mapping rules, it is important to ensure that all participants can orient themselves on the maps to enable accurate mapping. Map literacy can be a challenge in many contexts. It is, therefore, essential to ensure prior to the actual discussions that each MFGD participant knows how to orient themselves within their community area. In some cases, it may be helpful to familiarise participants with map reading. As much as possible, turn the map so that it corresponds with actual cardinal directions – thus the northern edge of the map should face north. Participants should be asked to provide any landmarks that have not been identified in the reference map.

CASE STUDY CONSIDERATIONS: GARBEK, SOUTH SUDAN

In a similar mapping exercise in South Sudan³³, none of the MFGD participants were able to read a map. Instead of presenting maps, data collectors therefore read the names of local places out-loud from a prepared list and asked questions in reference to the community characteristics around each place to discern where divergent characteristics existed. It is important to have a contingency plan which enables the use of place names, street names, or landmarks as a means of completing the participatory mapping when participants have limited map literacy.

³¹ ICRC. 2018. [Professional Standards for Protection Work](#).

³² *Consent* signifies the approval by the participant for the information to be used as explained. ICRC. *Idib*.

³³ REACH Initiative, 2018, [Research Methodology Note, Service Access Gaps Assessment in Garbek; South Sudan](#):

Labelling

Consistent ways of marking and labelling are needed throughout the mapping exercise to enable interpretation of the maps in the analysis stage. Two approaches can be used to mark landmarks and boundaries on the maps – either pre-defined symbology or numbering. Map scale and number of layers are the key factors in choosing whether to use predefined symbology or numbering.

Comparison of symbology and numbering:

- *Symbology* is a term used in cartography to describe graphical techniques to represent geographic information on a map. The advantage of using symbology is that it reads clearly on a map and can be more quickly digitalised. However, when multiple layers of information are involved or a smaller area is mapped, symbols can quickly become cluttered, thus reducing legibility.
- *Numbering*, on the other hand, offers the opportunity to put a lot of detail on a participatory map while preserving legibility. It is crucial to record a detailed legend as numbering runs the risk that if a number is not properly annotated its meaning will be completely lost.

Initiate mapping

Once all participants are oriented and landmarks are indicated on the map, participants may be asked to identify where they live on the map and mark all locations with an 'X'. Once this is done, participants may be asked to identify the location of key features, depending on whether it is a community area MFGD or service area MFGD.

Figure 9: Example of a community mapping exercise in Arua, Uganda



Sequence of service and community area mapping

It is advantageous to conduct service area MFGDs before community area MFGDs, as much of the service information can inform community area information. With a clear idea of the geographic context – including services, land use, and resources, as well as locally used terminologies for geographic categories and boundaries – participants in community area MFGDs can be asked to identify their community area boundaries. It is important to keep the definition of community relatively open at this stage to encourage discussion around what particular factors describe it. At the same time, it is vital for the MFGD facilitators to have a good understanding of the local context and to be able to use and understand common local terminologies in order to direct and probe participants to accurately delineate their community areas.

Probing

A starting point can be to ask participants to outline the immediate area occupied by a community, where there is a group that share services, resources, and socio-cultural characteristics. The moderator should probe where service locations, such as schools, fall outside of the suggested boundaries in order to understand the rationale underpinning the community area.

Similarly, where suggested boundaries do not correspond with natural features such as rivers or where arable land cultivated by community members is outside of the boundaries, it is again important to probe. Depending on the context, ethnic, religious, or other socio-cultural characteristics may underpin community area boundaries. If this is not made clear by participants, the moderator should seek to probe for areas inhabited by different groups to cross-check against the boundaries of the suggested community area. Boundaries may be a sensitive topic and must be well researched during *Step 1* to ensure the data collection team have a clear strategy on how to approach the topic during the MFGD.

Once the area has been defined, the MFGD participants should be asked to identify the approximate number of individuals and/or households that reside within the community area. It is important to bear in mind that it is usually challenging to get reliable population numbers from MFGD participants. Thus, population estimates should be used cautiously and confirmed with local authorities, whenever possible.

Preliminary identification of KIs in the area

Before ending the MFGD session, participants should be asked to suggest individuals in the community who are knowledgeable about the community area or service area. Participants can be probed to suggest different individuals based on specific services or other topics that are of interest. Finally, participants can be asked to indicate adjacent community areas to help identify subsequent locations for follow-up MFGDs. These could be recorded on another reference map that displays a wider area. If they have knowledge, participants could also help identify contacts and even introduce the data collection teams to the adjacent community members. The moderator should conclude the session by thanking all participants for their time, provide an opportunity for asking questions, and outline the next steps for the process. Teams should have the following information at the end of all MFGDs:

- Transcript of the discussion with clear criteria's that were used to delineate the area.
- Information about each area

- Field map with the drawn community or service area
- A list of potential KI and service stakeholder in the community

Step 4: Analyse and interpret MFGD data

Analysis of MFGDs rely on basic analytical skills in both mapping and qualitative information. For example, analytical skills in mapping are necessary to interpret the parallels of labelling across reference maps and analytical skills with qualitative data are necessary to interpret how MFGD participants define the characteristics of a community. Both sets of analytical skills are then necessary to determine why an area is where it is and why other geographical locations are not part of the area.

Triangulation

During the MFGD data collection phase, discrepancies across MFGD sessions need to be duly recorded and investigated. In particular, it is important to cross-check how far community area and service area demarcations differ between groups (ex. displaced vs. host communities, male vs. female). Both discrepancies and convergences need to be recorded and cross-analysed until there is data equilibrium or an understanding of the difference.

The transcripts of MFGDs will be the basis for analysis and triangulation. Perceptions of the different groups reflected in the transcript should inform analysis of how differences in boundaries are perceived by different groups within the same community or whether they are simply entry mistakes. For example, in Niger some differences of community boundaries were found in areas to which displaced people had recently arrived. MFGD groups with displaced populations tended to include a larger area on the periphery. It was found through the script of the MFGDs and by field verification that the areas included by displaced groups were in fact peripheral areas where the newly arrived population had settled.

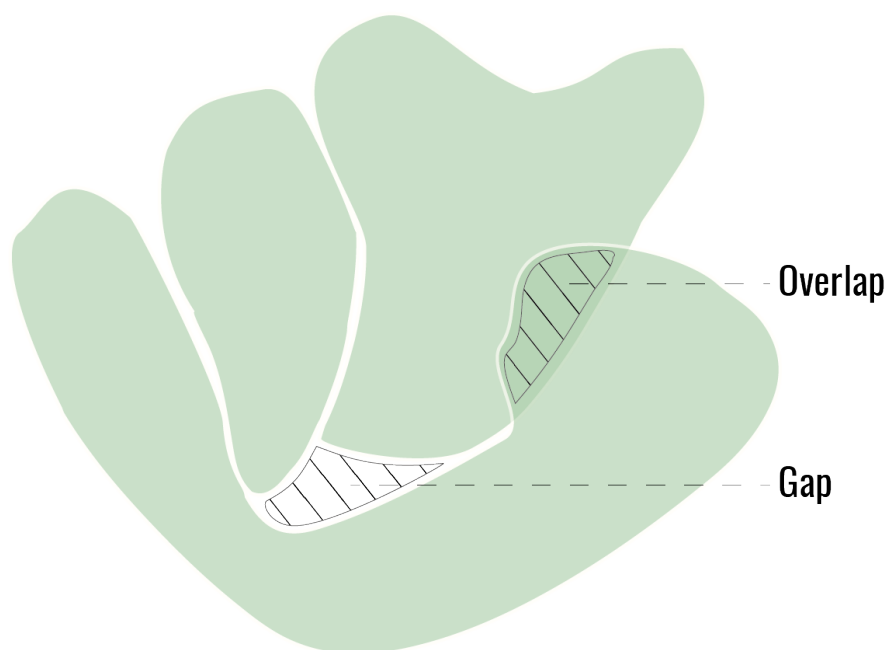
Determining final boundaries

After triangulation of data among different groups in the same area it is beneficial to organise a workshop with field teams and the LAAB to brainstorm the final boundaries of an area. Before arranging the workshop, the study team should propose boundaries and prepare the workshop in relation to the interpretation of the data collected. This workshop will also provide an opportunity to discuss questions that have arisen during the triangulation process and promote a collaborative process.

Dealing with inconsistencies

During analysis, it may become apparent that there are some gaps or overlap between the areas digitised. It is important to investigate the cause and gain insight from the LAAB. In most cases the investigation results allow resolution of inconsistencies and make it possible to attribute the gap or the overlap to a community. See *Figure 10* to review an example of community area gaps and overlapping.

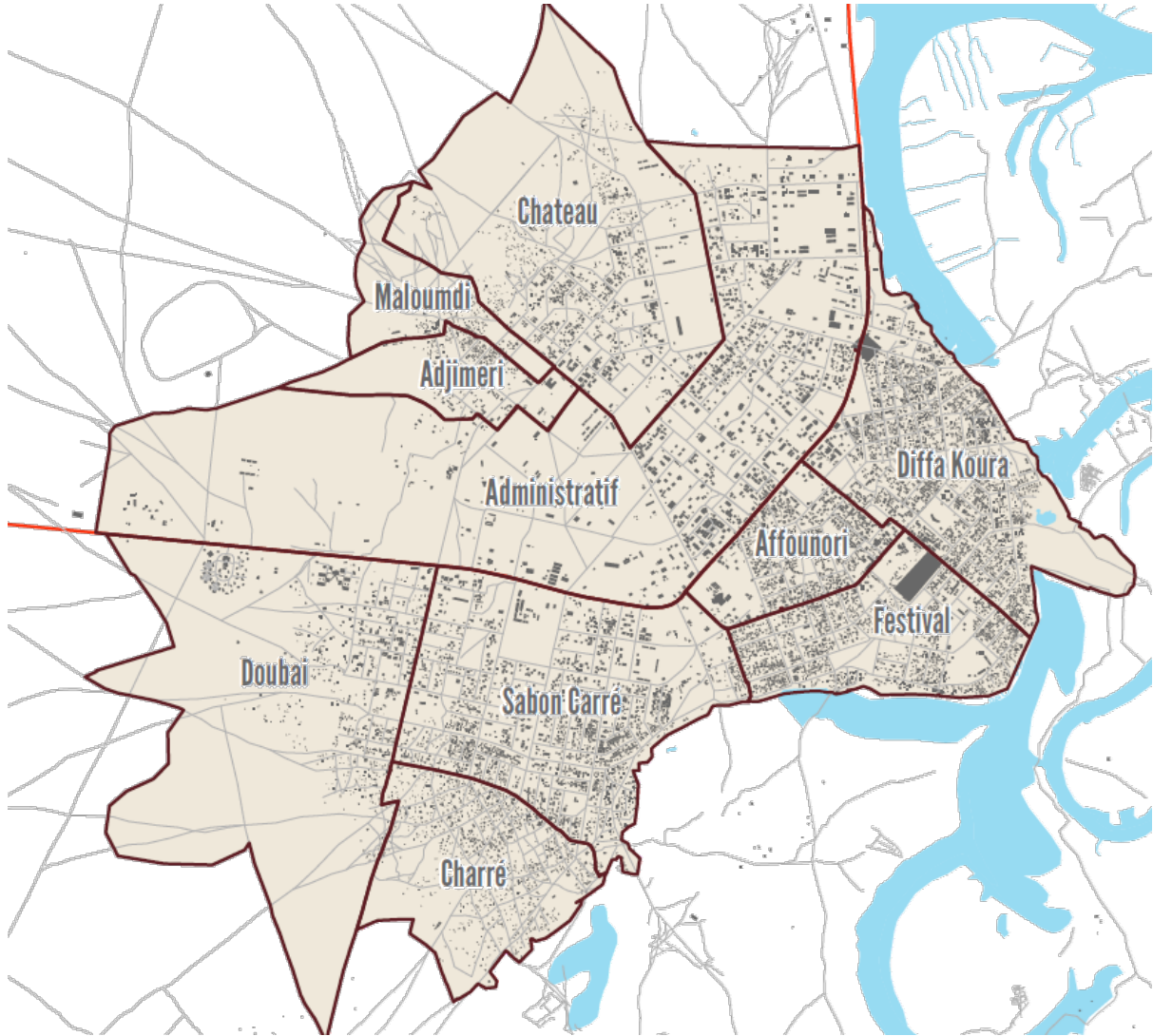
Figure 10: Illustration of gaps or overlap in community boundaries



Questions to raise when encountering inconsistencies

- *Is that an entry mistake?* Digitisation of maps can have led to mistakes. In such cases it is important to double check the paper maps, investigate with the data collection teams and conduct additional MFGDs with respondents from this overlapping / gap area so as to clarify where the population feel they belong to.
- *Is that related to land use?* In some case the gap in the mapping coverage can be related to land use, a forest, a park, or inhabited area. This was the case in Arua municipality where it was found that the local golf course was omitted from the community area mapping, since it was not considered as being part of the community. Investigation can include double checking with the mapping team, verification via land use data (if available), and trips to the location to investigate further.
- *Are there differences in perception?* Different groups, especially host communities and displaced populations, may differ in how they define a given community area. In such instances, it is both important to mitigate any conflicting definitions as well as confirm who the future programming will inform. In some instances, the study team may have to defer to the defined community area of a particular group with opposing definitions from another group. That was the case in the Arua pilot where refugees participants tends to recognised smaller area of the community, equivalent to the cell.
- *Is the unit of analysis appropriate?* There are risks in choosing a unit of analysis that is too large or small. If the unit is too large – as was discovered during the piloting phase in Uganda – participants may provide conflicting information because the area is too broad for the participant to provide specific and accurate information about the population. Alternatively, if the unit is too small – as was discovered during the piloting phase in Niger, it may be difficult to gain consensus on the geographic distinctions of an area from MFGD participants.

Figure 11: Example of a final community area map from Diffa, Niger



Section B: Identification of Key informant

Key Informants (KIs) are individuals who provide information about a larger group of individuals or a space or a service. KIs can have lasting impact on response efforts when used as main information sources. In the case of this toolkit, a KI is a proxy for her or his associates within a community or organisation.³⁴ Thus, KIs should be individuals who have reliable information about key topics related to the population of interest – either specific services within an area or for a particular community area.

Challenges with using KIs for humanitarian response planning

In humanitarian contexts, KIs are increasingly utilised to gather multisectoral information to be used in response planning. Accurate and practical area definition, in turn, enable accurate and practical KI data collection. Unfortunately, it is often the case that the area on which KIs report information is not properly defined. As the quality of KI data is strongly dependent on the area definition, improper area definition can lead to inaccurate and even misleading data.

Often times, KIs are expected to provide information for an area bearing the same name as an official jurisdiction or referring to a wider area which extends beyond or fall within its boundaries. However, these designations are often too broadly and inconsistently defined for any KI to provide representative and accurate information. When this happens, inaccurate definition of the area being reported on can have major impacts on data accuracy. An organisation using official boundaries as a reference might make wrong decision based on the reporting of a KIs referencing a community area with different boundaries.

Further complicating the utilisation of KIs data, organisations may ineffectively select community leaders or self-proclaimed leaders as KIs, instead of undergoing a critical selection process to ensure the accurate data possible. While a one-size-fits-all approach to KI selection may not exist or be appropriate, it is still important to create a replicable process that can be applied to different contexts, thus allowing for an objective definition of criteria used for KI selection.

Lessons learned from pilots

The four pilot studies that informed this toolkit investigated the accuracy of KI selection methods and provided lessons to improve selection criteria of KIs for future responses in out-of-camp displacement contexts. There were six primary lessons learned in relation to KI selection:

- **The importance of the area of knowledge** - During some community area MFGDs, participants had different perceptions of the community boundaries. When comparing the results between KIs and households, the KIs results were, on average, more accurate when MFGDs findings were consistent between the MFGDs groups in a given area. This finding highlights the importance of having a properly defined area of knowledge in order to collect reliable data. MFGDs findings should meet equilibrium³⁵ before continuing to area definition. Otherwise, area distinctions may be improperly defined and corresponding area-based information will be inconsistent.

³⁴ Lavrakas, 2008, Encyclopedia of Survey Research Methods, [Key Informant](#). SAGE Publications.

³⁵ an indication that, on the basis of the data collected or analysed hitherto, further data collection and/or analysis are unnecessary since it would be unlikely to meaningfully increase understanding (also referred to as saturation)

- **The question of a multisectoral community KI** - It was highlighted during the pilots that no KI is able to provide accurate information on all topics of a multi-sector assessment. Accuracy of KIs on the complete multisector assessment are in average low. However, when KIs are targeted for their specific knowledge on a subject, for example health profile for health questions, the accuracy on subject-specific topics improves greatly and bias reduces. Thus, community KIs should also be selected based on topic-specific knowledge and no KI should be expected to provide reliable information on all topics in a multi-sector assessment.
- **The use of community leaders as KIs** - While implementing an assessment in a community, it is indispensable to engage with leaders for project buy-in and support. However, using community leader as community KIs poses as a challenge. It was found during the pilot that, in average, the community leaders are not more accurate in reporting community situation than KIs of different profiles. Thus, community leaders should only be selected as KIs if there is strong reason to believe they have access to information representative of the community, not just outlier information.
- **The use of host and displaced populations** - In contexts where the community is populated by host and displaced individuals, KIs are more likely to give reliable information on the host and displaced population separately than if referring to the community as a whole. In addition, members of each group give more reliable information when responding about his or her own population. This finding highlights the importance of having KI from the different population groups.
- **The value of referrals for KI selection for community areas** - During the pilot studies, KIs identified through snowball technic were asked each to identify the person who in their opinion was the most informed member of their and about the community. No difference in accuracy was found when compared with KIs identified by referral and KIs identified by snowball. Referrals often led to community leaders who could not provide reliable information on the broader community. Thus, referral may not be the most reliable KI identification method.
- **The value of referrals for KI selection for service areas** - The same identification process was tested for service area KIs and community area KIs. It was found that referred service area KIs had, on average, higher accuracy that the service area KIs identified by snowball. This finding indicates that selection of informed KIs on the basis of referral can works at the condition that the referral is linked to a specific topic or service.

Proposed steps for KI selection: service and community areas

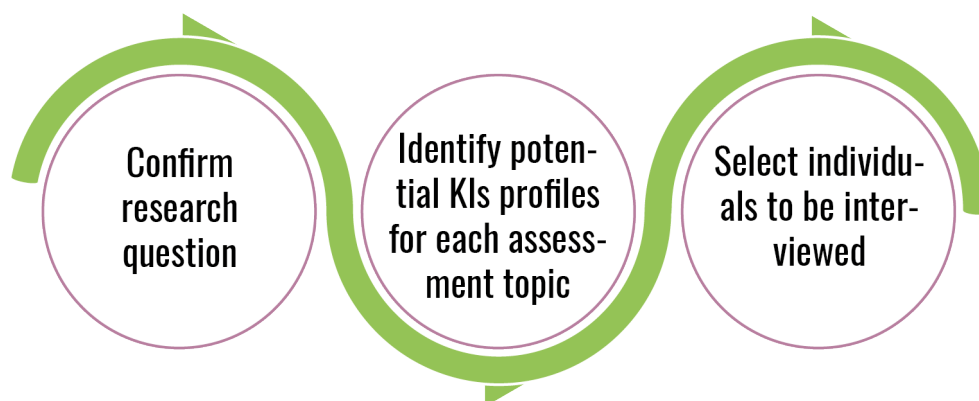
Following the designation of the area typologies focused on here (services and community areas), KIs can be categorised as either service or community area KIs. Service and community KIs are not expected to have the same expertise or knowledge. For Service area, KI expertise is specific and service-linked while community area KIs have more general expertise around issues to do with community needs, interests, and capacities but the same steps for selection apply.

Once community or service areas have been identified, KIs can be selected for the corresponding geographic area and the type of services.

A three-step process is proposed when selecting KIs in out-of-camps contexts:

- *Step 1: Define KI specific research questions based on information needs* – A clear research question is essential and a list of indicators should be drafted – both for service and community KIs. The indicators should reflect the information needs of the assessment in the area assessed.
- *Step 2: Identify potential KIs profiles by topic* – Develop based on local knowledge a list of potential KIs profiles for the area for each knowledge topic. Displaced and host groups need to be represented in the list of KIs selected. It is necessary to target specific groups in the community that would have the knowledge on each topic (for example, teachers for questions related to access to education or women for women access to services focus on women).
- *Step 3: Identify KIs with the desired profile and asked them to refer the most knowledgeable KI* – For each topic, identify individuals that fits the profile at the community level. Ask the identify individuals to refer you to the most knowledgeable individuals about the knowledge topic.

Figure 12: Overview of practical steps to shortlist potential KIs (without SNA)



Step 1: Confirm research question

Without a clearly defined research question, studies run the risk of being too broad in scope and vulnerable to derailment. In such case, since the objective is not defined a lot of effort will be spend collecting information having no or only very limited use in the study. Having a well-defined research question is important in regards to efficiency of data collection and quality assurance of data. The research question(s) should be informed through collaboration with the LAAB, relevant to area definitions from *Section A*, and guided by the ultimate interest of response efforts.

Corresponding topics and indicators

For both service and community KIs, a list of topics (for example WASH or Health) and indicators should be drafted prior to data collection. These should relate to the underlining research question and reflect the informational needs of the assessment. Thus, the list of topic and indicators will inform the KI selection.

Example: A research question in Arua was “What is the education situation, regarding access, for displaced and host household children in the community?”. The topic the research question is focussing on is *education*.

Preparation of indicators should, ideally, be done in close collaboration with the implementing programme partner, the organisation planning the humanitarian response and if setup, the LAAB. That will ensure that the information collected later during the assessment will be useful to partner to support an effective response in the area. Background research to identify service and community areas may also help to disclose informational gaps and narrow down the research topics of interest within the target area. *Annex 4* contains an example indicator list to help inform displacement trends, access to services, and humanitarian priorities in a service area. These indicators are important to eventually develop the questionnaire, and thus ensure that the questionnaire is fully in line with and relevant to the informational needs of the assessment.

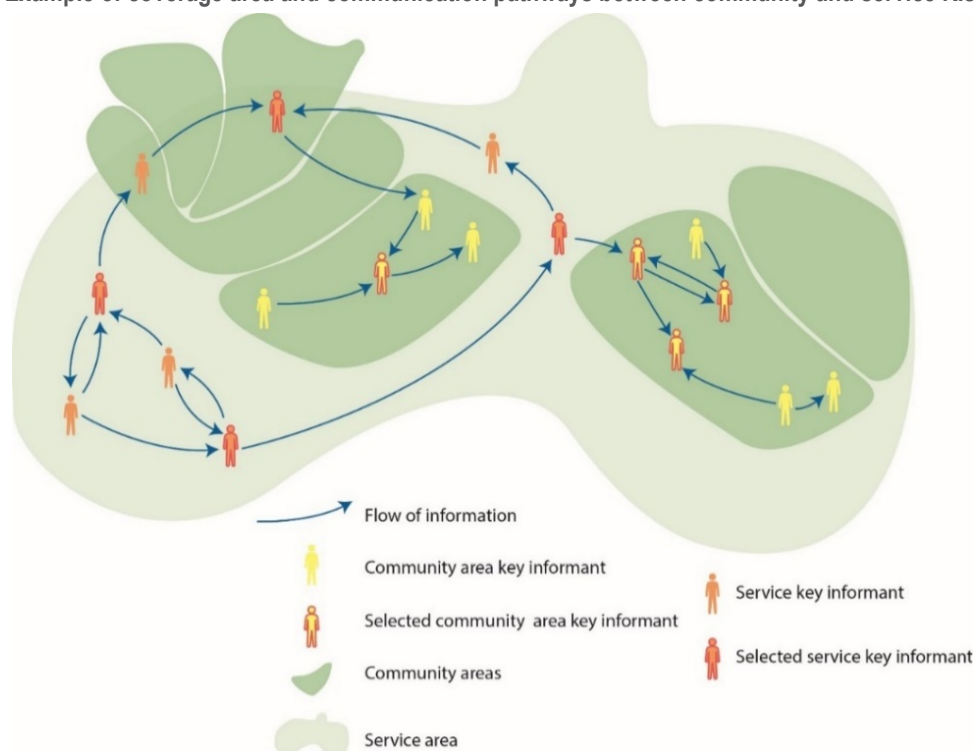
Set of education indicators during the Arua, Uganda pilot:

- Percentage of school-age children (i.e. 6-18 years old) enrolled in formal education
- Percentage of enrolled children regularly attending formal education
- Percentage of school-age children out of school for over a year
- Average months spent out of school for children aged 6-18 over previous five years
- Top three priority needs to enhance access to, and quality, of education in Arua.
- Percentage of children attending schools outside Arua

Step 2: Identify KI profiles

In this step, a KI profile list should be created for each topic of the assessment, in consultation with community members, local organisation and local government. This profile list should indicate what optimum characteristics (eg. age, gender, profession) for KIs, as this list will guide KI selection. In this

Figure 13: Example of coverage area and communication pathways between community and service KIs



Service KIs

Service area KIs need to have a general understanding of services within their sector of expertise, including the availability, access, strengths, gaps, and connectivity of the given service. They must be able to describe the operating status of services and infrastructure, as well as key places where services are accessed. Service KIs should be an individual highly knowledgeable about the service provision itself, as well as someone who interacts with a high number of community members who need the particular service. Additionally, service KIs may be able to provide information relevant to more than one community area or only a sub-set of a community area (or vice versa for community KIs). See *Figure 13* for a representation of how a network of community KIs and service KIs may be interconnected.

Example: For the education sector, KIs may be teachers, school directors, secondary school students, or members of parent associations.

Community KIs

Community KIs need to understand the needs and capacities of the inhabitants of the targeted community area. They need to be well-connected to individuals and services providers and tend to be representative community on a particular topic or population group. When selecting community KIs, it is assumed that communities are not homogenous; therefore, each group should have its own KIs. As much as possible, characteristics of KIs – such as gender, age, migration status, profession, and community position – should be documented for each community KI. Having this kind of documentation may help highlight if there are gaps in representation.

Example: During the pilot in Kabul, some specific questions focused on the access to health services of women and children. Additional efforts were made to recruit female KIs. The results of the accuracy study about those questions, thus, show that women were more reliable than men on these topics.

Define KI profiles for each information topic

Conversations with community members and relevant stakeholders will be helpful in determining profiles of persons who may be particularly well-informed. This can be done in conjunction with the LAAB or it can also be a specific exercise with the community and a field team. It can be more efficient to realise this during the MFGDs. At the end of the mapping exercise, the assessment topic along with specific questions can be presented to MFGD participants. A brainstorming session with MFGD participants can produce a list of KI knowledgeable profiles at service and community areas for each topic of the assessment.

Example of KI profiles selection: The table below shows an example of the work done by REACH since 2015 for the Humanitarian Assessment Overview Syria (HSOS) to target KIs in villages inside Syria. It ranks the desirability of particular KI profiles so as to get specific information on specific topics (seen in the top row). For example, a hospital doctor would be more desirable than a carpenter to get information about the health situation in a community. The table uses a ranking scale 1-3, whereby 1 is not desirable and 3 is highly desirable. It represents the level of confidence analysts can have that those with certain profiles are well-informed about specific topics. This matrix may be based on certain paradigms but is, nevertheless, a useful field tool to

guide study teams in the initial selection of relevant KIs as well as for triangulation i.e. if two different KIs are providing conflicting information we can go back to these scores can be used to determine which information is more reliable.

Table 1: Extract from a KI matrix by REACH in support of 2015 Humanitarian Assessment Overview Syria (HSOS) ³⁶

KI profile	Knowledge score³⁷ Topic: displacement
School managers	1
Teachers	1
Students	2
Community worker	3
Doctors	2
Local charity workers	2
Local council representatives	3

Step 3: Shortlist individuals for interview

After determining the individuals that fit the pre-defined profiles related to topic at each area level (service and community), a shortlist of KIs should be created. Ideally, the LAAB would support the creation of this shortlist after reviewing the inputs and suggestions from MFGDs. In order to create distinct and practical shortlists, separate lists should be created for each topic in each area (service in community). If not incorporating network analysis, KIs can be then selected from this shortlist to be included in KI data collection (see *Section D*). Otherwise, this shortlist of KIs can be used to guide network analysis for the final selection of KIs.

A referral exercise can be conducted to improve the KI selection. Individual identified for each topic in each area was asked to refer to the person that knows best about the topic in the area. The same questions were asked to people across waves, thus creating a hierarchy of referrals. KIs referred the most often will be selected as KIs While this exercise may seem tedious to conduct, it was made simpler and cheaper when the context allowed the collection of information by phone.

³⁶ REACH Initiative, 2015, [Area of Origin – Jordan, The Kurdistan Region of Iraq, Lebanon & Turkey, Terms of Reference](#).

³⁷ Knowledge score from 1 = not particularly knowledgeable to 3 = very knowledgeable

Section C: Identification of Key informant with Network Data

Utilisation of Social Network Analysis (SNA) techniques are emerging in humanitarian contexts and can be useful for the identification of KIs for programme development and two-way communication, that include using the network for information diffusion and data gathering. Thus, the following section builds directly on the information from *Section B* and includes specific ways that network data can support the identification of KIs. To identify KIs with the most strategic positions within their networks, this toolkit proposes using specific measures to assess the quality, scope, and strength of communication between KIs. A more technical description of the SNA can be found in the academic literature,³⁸ for the framework and measure used in this toolbox please refer to annex 7.

What is Network Analysis?³⁹

Network Analysis is the broad representation of relationships among a set of items. Networks exist in all systems – societal, biological, financial, trade – but are often researched through the frame of social networks. These social networks are exemplified consciously and subconsciously through in-person, online, and telephone communication. Interpreting the structure of a network can answer questions about complex phenomena or hypothetical concepts.

SNA falls under the umbrella of Network Analysis. For the purposes of this document, SNA is defined as the multi-disciplinary application of network theory to the modelling and analysis of social systems of individuals, groups, or stakeholders. Guided by graph and map theories, SNA represents a theoretical approach to understanding relationships and influences with a social network. SNA analyses relationships, through an explicit focus on dyads (pairs), for the purpose of interpreting complex interactions.

Relevance of SNA

The objective of using SNA in this toolkit is to identify KIs who are connected in a manner that enables them to provide information about their respective area of knowledge.⁴⁰ Implementing SNA methods for KIs identification may be beneficial with the aim of establishing longer term communication with the community by opposition to an ephemeral data collection exercise. In that context, the network position of a KI is less important for the information passively gathered from their networks than for the potential reach the KI can have – enabling the KI to gather information needed as well as play the role of mediator with the community. Thus, KIs selected through SNA can facilitate two-way communication between humanitarian agencies and communities.

Identification of KIs through the use of SNA needs more investment of time, effort, and costs than identification of KI in *Section B*. In fact, collection of KI network ties can be considered an assessment by itself. Therefore, utilisation of SNA for KI selection should be critically considered alongside the broader aims and interests of the project, as well as the resources available to the study team.

³⁸ Jackson, 2008, [Social and Economic Networks](#).

³⁹ IMPACT Initiatives, 2017, [Social Network Analysis and Targeting of Assistance for Urban Refugees](#)

⁴⁰ Depending on the context, cross-sector KI networks may be more helpful to gather the information required. For example, informing a needs assessment for a planned child protection project, an approach could be to combine KIs from across various services, such as teachers, doctors, community workers, etc. in one and the same KI network to identify the most informative KIs on child protection related indicators.

Rationale of SNA for the pilot

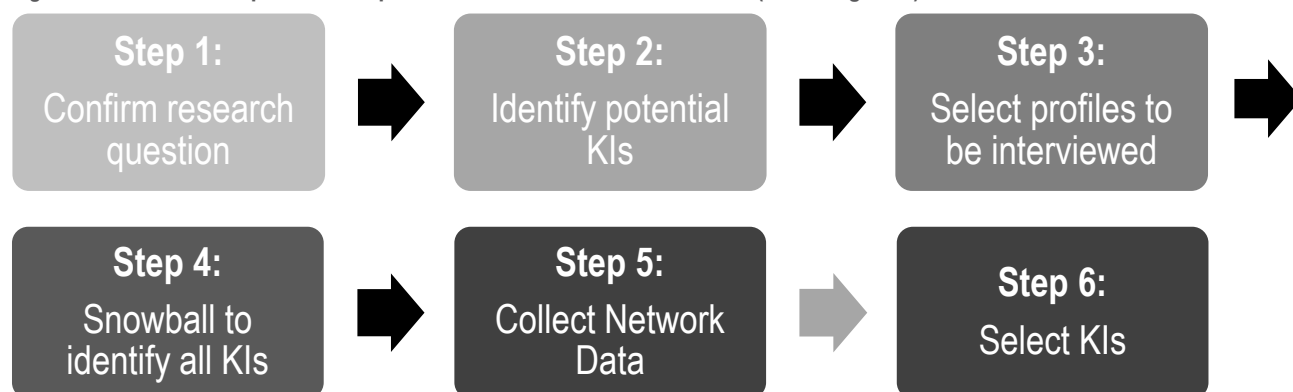
The theoretical framework of the pilots was based on the assumption that KI connected to each other exchange information. It was found during the pilot that KIs are more likely to have similar answers if they are connected to each other than if they are not. In other words, KIs who are closely tied are more likely to have related knowledge. On average, the probability of providing similar answers increases by 40 percent if KIs are directly connected. This finding supports the notion that information may be directly transmitted between KIs and highlights that KIs can be used as a way to get but also to spread information in the community.

Proposed steps for network data collection

The steps involved in selecting KIs through SNA are listed below:

- Step 1: Confirm research question – see *Section B*
- Step 2: Identify potential KIs – see *Section B*
- Step 3: Select profiles to be interviewed in relation to topics – see *Section B*. These profiles will be considered as “seeds⁴¹” for snowball methods.
- Step 4: Snowball from each seed – use snowball technique to identify all the knowledgeable KIs in the area.
- Step 5: Collect network data – At a minimum, data collection should include network information on the relationships between KIs. On the basis of the data obtained through this survey, a network can be drawn and use for analysis.
- Step 6: Select KIs – Through analysing KI network data, a final list of KIs should be selected to provide multi-sectoral information (See *Section D*) and maintain two-way communication.

Figure 14: Overview of practical steps to use network data to select KIs (including SNA)



Step 4: Snowball to identify KIs

After selecting the seeds⁴², snowball techniques can be used to get an exhaustive list of KIs in each community. Identification of KIs should stop after reaching equilibrium. Equilibrium is an indication that, on the basis of the data collected or analysed hitherto, further data collection is unnecessary since it

⁴¹ Seeds, when referring to a snowball method, are the starting points for recruitment of additional elements through snowball

would be unlikely to meaningfully increase understanding. When collecting KI network data, equilibrium is evident when the same KIs are being referenced and no additional network nodes are provided through the snowball.

Snowball techniques can be implemented through field work with face to face interview but can also be conducted by phone if the condition allows. The use of phone will allow a quicker and cheaper data collection.

Step 5: Collect network data

A minimum criterion for conducting SNA to select KIs requires collecting network data from each KI identified in the *Step 4*. The network data survey should be structured in nature, include the same questions for all KIs, and include specific network questions from which SNA can be applied. A suggested questionnaire to enable KI informant network mapping and subsequent network analysis is available in *Annex 5*.

Each KI is surveyed about all other KIs identified in the area during *Step 4* – it is important that network data is not collected before snowballing finishes. KIs need to be asked ties to all area KIs, not only the ones to which they have a shared experience (i.e. host communities or displaced populations). As an example, refugee KIs will have to be asked about their links to refugee KIs within their community area, but also about their ties to host community KIs in the same community area. A KI from the education sector will, similarly, be asked about their links to other education KIs as well as health or shelter KIs.

The KI survey questionnaire should contain, at a minimum, four main questions that are asked to each KI:

1. Which KIs from within the area are you exchanging with directly? (strength of exchange)
2. Do you both receive and provide information through this exchange? (direction of exchange)
3. How often do you communicate with each other? (frequency of exchange)
4. How reliable the information received is? (trust of exchange)

The detailed network questionnaire is available in *Annex 5*. In order to obtain a measure for perceived trust, which is necessary for subsequent network analysis, KIs were asked during the pilot to rate the information provided by each KI he is receiving information from. Response options comprise: “Very reliable”, “Reliable”, “Somehow reliable”, “Unreliable”, “Very unreliable”. For the pilot study, every link not perceived as reliable were removed from the analysis.

Equally, KIs were asked how often they interact, or more specifically, how often they receive information from a given key informant on a given sector. Response options depends on the sector specific information and its weight. Education specific information, for example, may be less quickly outdated than displacement related information, hence KIs would have to exchange information on displacement in shorter time periods than information on education related issues to stay up to date. Frequency of exchange is measured on a scale of five different time periods that should be adapted to context and type of information. An example could be: “daily”, “weekly”, “bi-weekly”, “monthly” and “less than once a month”.

These questions can be translated into weights indicating the strength of a connection between two KIs and are a prerequisite for the subsequent network analysis. See *Annex 7* for more information on scoring, including weights.

CASE STUDY CONSIDERATIONS: DIFFA, NIGER

While it may be tempting to survey the KIs concurrent with the snowball or referral process, KIs should not be surveyed until the final and cumulative list of KIs is created. Having the final list can help with KI recollection of their exchanges with other KIs about the community or the service provision. For example, *Figure 15* and *16* show the networks for the same community. *Figure 15* shows the network constructed by identifying KIs and their ties (Steps 4 and 5) at the same time, while *Figure 16* shows the same community, but the network constructed with Step 4 and 5 conducted in a stepwise approach. *Figure 16* is a stronger network representation as it demonstrates the closed nature of the network, as well as demonstrating exhaustive network ties.

Figure 15: Example of Adjimeri network in Diffa, collected through snowball

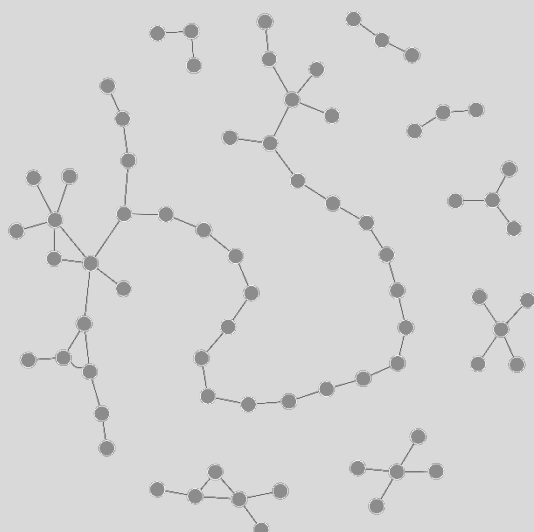


Figure 16: same network with data collection based on a list



Step 5: Select KIs

Two type of approach will be presented in this section to select KIs:

- 1) A more qualitative approach that examines the place of a KI in its network (community or service area) in the perspective of mid- to long-term engagement with the community.
- 2) A more quantitative approach (still in development) that aims to extrapolate the pilots finding to other context in order to improve KIs selection in the framework of short-term assessments.

Qualitative approach

In the context of mid- to long- term information gathering needs and in order to create link with the communities or the service provision, SNA allows to identify the most *important* measure of centrality within a network.⁴³ The most *important* can refer to the:

- *Central node*: the KI closest to all the other KIs in the network – this individual can reach directly or through other node a large number of other individual in the network.
- *Intermediate*: a bridge in between two part of the network – this individual is an intermediate – he can moderate the exchanges between two network part.
- *Most consulted*: the KI with a lot of outward connection – this individual provides information to a large number of KIs in the networks and, therefore, can influence the KI connected to him.

Central node

The central nodes refer to KIs that have high closeness centrality. The closeness centrality is defined as the average length of shorter path⁴⁴ (distance) between a KIs and all the others. See *Annex 7* for more details. In short it measures how easily a KI can obtain information from other KIs in a given network. The rationale behind this is that a KI may not receive information directly from many other key informants, but could still be well connected to the rest of the network through intermediates, allowing access to wider knowledge. In sum, a key informant who is closer to the rest of the network than another key informant is expected to have access to a larger proportion of information that is shared within their network and disposing of reliable channels to access information.

While we demonstrated in *Section B*, that multisector community KIs do not provide reliable information, a node with a high closeness centrality has the potential to be a good key informant at community or service level. In the context of slower pace assessment, a KI with a high closeness, while not having all the information himself, can use its network to get the information on each assessment topic. In other term, if the KI is willing to, we can give him some time in order to contact other KIs that would be knowledge about specific part of the questionnaire.

Intermediate

The intermediate nodes refer to KI that have high betweenness. This measure quantifies how often a key informant acts as an intermediate in passing on information (on shortest paths) within the network. As such, it indicates the importance of a given key informant in facilitating exchange of information within a key informant network.

As such, KI with a high betweenness can be in a position to moderate the exchanges between two sub-networks for example (see case study from Diffa below) or between to group of population. Understanding where the intermediate is will allow actors to 1) Understand potential blockage in the spread of information 2) Remove the blockage by for example promoting communication between other actors.

Most consulted

The consulted node refers to a KI with a high outward degree of centrality. This measures how many other key informants a key informant gives information to on a defined area of interest. More technically

⁴³ See Hollstein (2014) "Qualitative Approaches" in Scott & Carrington (2014) "The SAGE Handbook of Social Network Analysis"

⁴⁴ Freeman, 1978, [Centrality in Social Networks: Conceptual Clarification](#).

speaking, the degree of centrality of a key informant within their network is the weighted sum of all “outgoing” links that connect them directly with other key informants of their network. See annex 7 for more information. In smaller network the most consulted could be seen as the main source of information and hence has influence on the network knowledge.

Selecting KIs

Each key informant is ranked according to the three measures. This produces rankings, one *closeness* ranking and one *betweenness* ranking, for each network. In other words, for each network, there are as many rankings as employed centrality measures that allow to target the right KI depending of the usage.

Table 2: KI ranked based on network characteristics

KI	Degree	Rank	KI	Closeness	Rank	KI	Betweenness	Rank
KI1	1.9	1	KI2	1.20	1	KI4	0.62	1
KI4	1.8	2	KI3	0.13	2	KI3	0.40	2
KI3	1.7	3	KI5	0.12	3	KI2	0.39	3
KI9	1.3	4	KI1	0.12	4	KI6	0.23	4
KI5	1.2	5	KI8	0.11	5	KI8	0.22	5
KI8	1.1	6	KI6	0.11	6	KI5	0.13	6
KI7	1	7	KI4	0.11	7	KI1	0.08	7
KI6	0.3	8	KI7	0.10	8	KI9	0.04	8
KI2	0.3	9	KI9	0.00	9	KI10	0.00	9
KI10	0.2	10	KI10	0.00	10	KI7	0.00	10

It is also important to note that networks change over time, as KIs may leave or arrive in the network. Thus, it is important to make sure the KIs composing the initially identified network are contacted and assessed regularly.

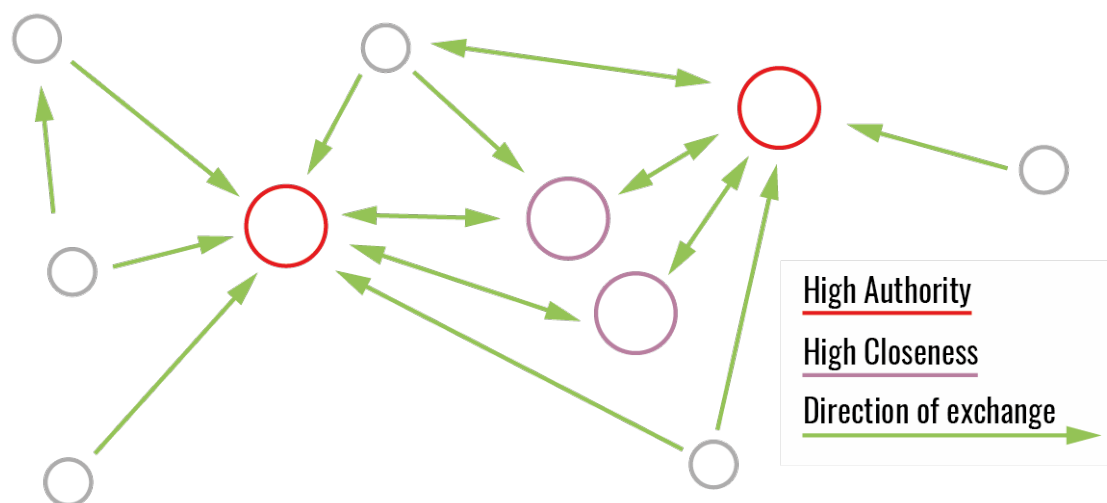
Quantitative approach

Selecting KIs could also be informed by more quantitative analytical techniques to determine their function as part of the wider KI network. Two node level network measures (authority and closeness centrality), when applied in conjunction with contextual deliberations such as discussions with local stakeholders and triangulation with secondary data, yielded an efficient and effective selection approach for KIs in the pilots. This analytical technique is still under development.

Closeness centrality

As seen previously, closeness centrality ranks each KI based on their comparative proximity to each KI within the network in regard to network connectivity (based on information gathered during network data collection). The two purple circles in the middle of *Figure 17* represent high closeness centrality because they are close to the other circles and centrally located within the network.

Figure 17: Example of closeness and authority



Authority

Authority measures the extent that a KI is a provider of information. Authority is calculated per KI and is related to the number and direction of connections of the other KIs directly connected to each KI. Technically, a KI is authority if its inward connections are from KIs that refers to many others⁴⁵. The two red circles in *Figure 17* represent high authority because (large circle size) but have the greatest numbers of inwardly directed communication flows from a diversity of well-connected sources.

Network analysis and the pilot

While there are a number of network analytics that can be applied to measure the relationship between KI and the network, the four pilot studies identified two primary analysis measures to identify reliable KIs – closeness centrality and authority. Controlling for country of data collection and area (service or community), the pilot demonstrated that KIs with strong closeness centrality and authority provided responses were more likely to give accurate responses (less than 25 percent deviation from the household survey replies) about the community or the service. The pilot's modelling of centrality and authority yielded an 85 percent accurate prediction of which KIs are reliable when compared to household data.

Other common SNA measures such as degrees, eigen centrality, prestige, and betweenness were tested during the pilot stage but not found to be strong predictors. There was also no significant relationship between the displacement status, gender, or community position of KIs and accuracy of information. Therefore, while sociodemographic characteristics should be collected for each KI, these characteristics could not inform global trends on KI selection in regards to ensuring reliable information.

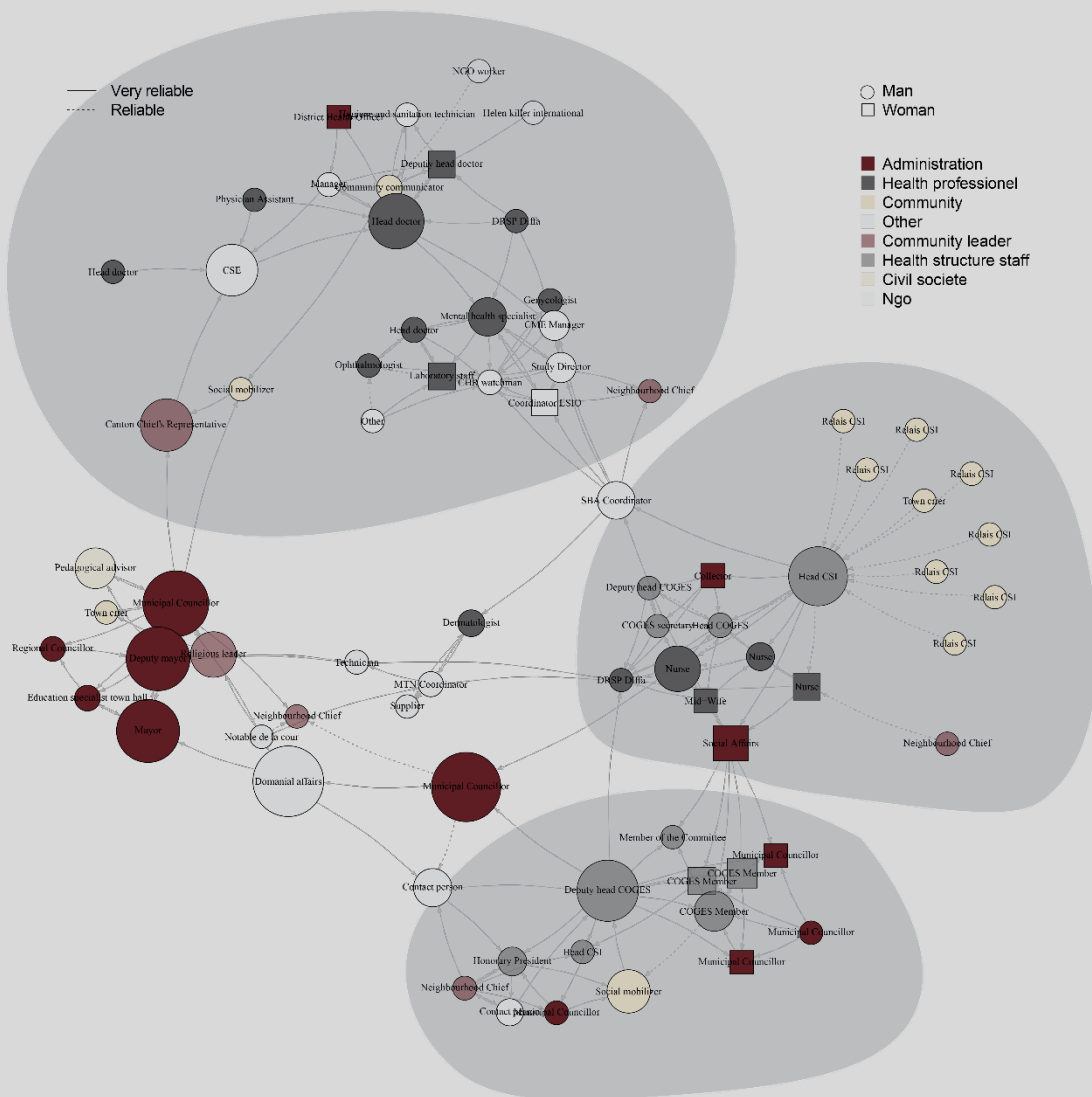
The results from the pilot based on these two analytical techniques were identified as the strongest 'predictor' for KI reliability. However, this result should be carefully considered until further validated in other contexts.

⁴⁵ John M. Kleinberg, [Authoritative Sources in a Hyperlinked Environment](#)

CASE STUDY CONSIDERATIONS: DIFFA, NIGER

This pilot was specific to health service provision. Thus, all the identified and interviewed KIs were selected based on their area of knowledge in the health sector. The map exemplifies the three primary clusters within the network: two integrated health clinics and the primary hospital. The size of the nodes represent the betweenness. Through visualising the exchange pathways between KIs and the betweenness measures, it is clear that the municipality officials are linked to the three clusters – thus acting as potential intermediaries. Not only does this map provide a critical picture of the exchange of information, but it may also support the understanding of critical areas of entry to support health service delivery. The SNA map can highlight overarching trends and connectivity. Through selecting KIs in each of the primary clusters, as well as critical intermediaries and local government representatives, targeted information about health provisions could be collected from this network.

Figure 18: KI Network mapping from Diffa, Niger pilot.



Section D: Collecting multisectoral data from KIs

After determining the various areas of knowledge of each KI within a network, multisectoral or sectoral data can be collected from selected KIs. Periodic (yet targeted) data collection and engagement with KIs supports the dual functions of improving information sharing to and from an area. Identical data collection methods are used for KI data collection in both service and community areas.

While similar methodologies are employed, KI multisectoral data collection is not meant to replace household needs assessments. Instead, multisectoral KI data can be collected as a preliminary or a complementary approach to household needs assessments.⁴⁶ Data collection of multisectoral information from KIs can be used at the preliminary stage of programme development to determine key areas of need or to further identify vulnerabilities within a community. Then a more tailored household needs assessment may be conducted. More information on household needs assessments is available in UNHCR's "Needs Assessment Handbook".

Step 1: Considerations for KI multisectoral assessments

KI data cannot replace data from household needs assessments. However, KI data can often be collected more efficiently than representative household information and KI data can fit specific roles that household data cannot. As an example, service area KIs are better able to provide specific information on the administrative components of a sector (ex. availability of and accessibility to certain goods) and community area KIs are able to provide specific information about their identified group (ex. host vs. displaced community).

The four pilots shed light on the conditions that can support collection of reliable data with a KI methodology for service areas and community areas. However, even after carefully selecting KIs, there are some perspectives that no KI can predictably provide. Most notably, KIs should not be utilised to discern questions related to normative thoughts or behaviours. For example, "Over the past 12 months do you think community satisfaction with the education system has increased, stayed the same or decreased?" Instead, household needs assessments are best placed to provide information on norms, while KIs can provide more objective interpretations of community needs and services. Thus, research objectives may be similar for multisectoral KI and household assessments but, before collecting multisectoral data from KIs, underlining KI indicators and questions should be reviewed for appropriateness.

Questions

The level of measurement included in each question of a multisectoral KI survey should be explored, as this may vary based on the type of input sought. The four primary types of variables – nominal, ordinal, interval, and ratio – are commonly in household needs assessments but may not be appropriate for multisectoral KI surveys.

- *Nominal*: Nominal questions account for categorical distinctions, where there is no inherent hierarchy in the initial ordering of answer options. However, the questions may include an interest

⁴⁶ UNHCR, 2017, [Needs Assessment Handbook](#).

in determining hierarchy. For example, *“What is the most important common shelter-related issue in your community?”*

- *Ordinal*: Ordinal questions account for distinctions in hierarchy without objective differences in categories. For ordinal measure, the “order” matters. For example, *“On a scale of 1-10, how satisfied are your community members with their access to water.”* Numerical questions may be made into ordinal questions through further categorising options into ranges. For example, *“Approximately what percentage of people in your area needed to access medical facilities in the previous six months?”* with the answer options of “None”, “1%-25%”, “26-50%”, “51%-75%”, “76% - 99%”, and “All”
- *Interval & Ratio*: Interval and ratio questions include a specific numeric response. However, the two levels of measurement vary as ratio questions include a true zero, while interval questions do not. This is important because the potential levels of analysis for each vary. Interval questions only allow for addition and subtraction comparisons, whereas ratio questions allow for subtraction and division. Ratio questions are more common in multisectoral surveys. These questions can be asked in regard to specific numbers, such as *“How much do you think refugee community households pay, on average, for water per month?”* Or these questions can be asked in regard to percentages, such as *“What proportion of households in your community have sufficient access to drinking water?”*

Evidence from the pilots showed that the accuracy of KI responses differs according to the type of question. A high level of inaccuracy was found when KIs were asked to provide reliable numerical estimations (not percentage). Answers were more accurate when KIs were asked to provide ordinal input (for example 1-10%; 11-20%), or a percentage or to identify most prevalent situations. It is thus advisable to avoid questions requiring respondents to give a number, rather than a percentage.

In order to get accurate information about the community, it is important to understand which questions can be measured through KI interview methods with significant accuracy.⁴⁷ Some questions could not be measured through responses provided by KIs to any degree of reliability from KIs during the pilot. These examples are available below:

- Average time from place of residence to reach the nearest community water source KIs
- Health access challenges
- Intentions of displaced persons in regard to returning to places of origin
- Percentage of households in the community with members separated in other locations

Each KI may be provided the same extensive multisectoral survey to support comparability. However, it is also possible to only ask select questions to each KI, depending on their area of knowledge (sector or community area).

⁴⁷ During the pilots, KI accuracy was determined by comparing KI answers to household level responses.

Step 2: Data collection

Data collection of KIs for multisectoral assessments should follow the same procedures of regular multisectoral needs assessments.⁴⁸ However, maintaining the database of KI contact details is central to enable recurring KI data collection. Once this is compiled, data collectors may arrange periodic interviews with each KI.

As previously noted, it is essential in all out-of-camp displacement contexts to recruit KIs from both host and displaced populations. It was found during the pilots that accuracy was higher when KIs were asked questions specifically about their identified group (either host or displaced). Conversely, questions concerning the broader community were less accurate for both types of KIs. This finding echoes the importance of predetermining what areas of knowledge each KI has.

Step 3: Data interpretation

Analysis

Aside from aggregating the sample, the techniques for analysing multisectoral KI data are the same as analysis applied for household needs assessments (i.e. basic statistical processes, such as frequencies). More information on analysing household needs assessments is available in UNHCR's "Needs Assessment Handbook".⁴⁹

However, in addition to basic statistical procedures, analysing multisectoral KI data are dependent on triangulation with available data. Once data collection is conducted with KIs for each service and community area, the information they have provided must be triangulated to obtain the best estimation for each indicator assessed. Examples of triangulation include conferring with LAAB, reviewing secondary data, and discussing with community members.

The area-based assessment process is cumulative by nature⁵⁰; thus, earlier data (such as MFGDs) should be used to corroborate or inform analysis of multisectoral KI data. Collating information from each step of an area-based assessment also highlights different perspectives (ex. secondary data, LAAB, community members, KIs, etc.). The MFGD provides unique information regarding the composition of areas and can provide insight for how community areas (as well as service areas) are linked. MFGDs for service areas can be used to identify the overall situation and existing gaps for service provision in a given area, as well as how the situation relates across communities. Whereas, multisectoral KI and household data can provide a more detailed look into the situation for community members. KI network data can, then, highlight critical areas for efficient and effective community engagement. When used in tandem, these data sources can provide a holistic view of the needs of a given area across sectors.

Utilisation

There are two types of multisectoral KI data: formative and complementary.⁵¹ *Formative* KI data is data utilised to inform on the baseline situation of an area. This data is collected before any household needs

⁴⁸ UNHCR, 2017, [Needs Assessment Handbook](#).

⁴⁹ UNHCR, *ibid*.

⁵⁰ Area-based assessments are cumulative, as opposed to a purely iterative process.

⁵¹ Area-based KI data collection and household data collection should never be conducted in tandem.

assessments are conducted and is meant to provide an overarching understanding of a given area. *Complementary* KI data is data collected on an iterative basis after household needs assessment are conducted. This data is meant to provide periodic updates on the situation of an area, without requiring a full household needs assessment. Complementary KI data can support ongoing monitoring efforts, as well as opening two-way communication with communities.

Conclusion

Concurrent with the growing inclination of displacement to out-of-camp settings, programmes are also shifting in toward area-based approaches to inform and respond to the nuanced needs in these out-of-camp displacement settings. Alongside the application of area-based approaches in displacement contexts, particularly in humanitarian settings, area-based assessments are suited to inform the situation and needs within areas in an integrated, effective, and cost-efficient way.

Thus, the primary aim of this toolkit is to provide generalised guidance on how to collect and utilise area-based assessments in out-of-camp displacement contexts. Through a stepwise approach, the toolkit provides a cumulative method and detailed guidance on how to conduct area-based assessments with KIs and SNA to inform ongoing and future programming in out-of-camp displacement contexts. Additionally, this toolkit outlines how to effectively integrate SNA to area-based assessments. As a result, this toolkit allows for comparability and application of area-based assessments with or without SNA in order to provide relevant guidance across humanitarian actors interested in area-based approaches and, more specifically, area-based assessments.

The stepwise sections from this toolkit are 1) Identifying an area, 2) Selecting KIs, 3) Selecting KIs with SNA, and 4) Collecting multisectoral KI data. Each section builds off earlier sections and informs future sections. Thus, area-based assessments are unique in the cumulative integration of previous findings *and* data across steps. This is in contrast to iterative assessments that solely build from findings from prior points of data collection (such as household needs assessments). **Area-based assessments can be inclusive of data from MFGDs, KI network data, KI multisectoral data, and household multisectoral data.**

Given the wide range of data sources and analytical techniques, there are unique skills needed for an agency to conduct area-based assessments. Teams should be well-trained and experienced in data collection, and above-average analytical capacities are necessary if conducting SNA. Since these analytical capacities are not commonly applied in displacement contexts, **teams should be trained on basic SNA techniques if collecting network data.**

If conducted well, the utility of area-based assessments is far reaching. Important to area-based approaches, carefully selected KIs can provide as connectors to an area and ease communication with the population, as well as provide a situational overview of an area. Therefore, area-based assessments can directly foster two-way communication with areas of interest. Additionally, KI-informed area-based assessments are more efficient and less costly than exclusively relying on standard household needs assessments. Area-based assessments may also reduce survey fatigue within an area since KIs are the primary providers of information.

With the expansion of area-based approaches and corresponding standards, guidelines to area-based assessments should also evolve. This toolkit provides as the first set of guidelines for area-based assessments, and **contextual considerations and local advisement should always be taken before conducting an area-based assessment.**

Annexes

<i>Annex 1: KI reliability measures</i>	52
<i>Annex 2: Complete MFGD questioning route used in Mafraq, Jordan</i>	53
<i>Annex 3: MFGD moderator and note-taker tips</i>	57
<i>Annex 4: Sample KI indicator list used in Mafraq, Jordan</i>	59
<i>Annex 5: Network related questionnaire</i>	62
<i>Annex 6: Example of accuracy of multisectoral indicators for a sample KI</i>	63
<i>Annex 7: Theoretical Framework (Social Network Analysis)</i>	64

Annex 1: KI reliability measures

In order to assess reliability of KIs, IMPACT developed a methodology to compare KI data to the aggregated answers of representative household survey. To ensure that comparison held, both KI and household surveys were developed to measure the same indicators, on the same time frame, on the same area.

The *accuracy* is the absolute score for each KI answer compared to the equivalent household survey result. The extent of deviation between a KIs answer and household confidence intervals yielded the outcome accuracy score. Any deviation greater than 25 percent yielded a score of 1 (i.e. completely inaccurate). The 25 percent threshold was put in place due to stakeholder consultation indicating that a deviation from the household data of more than 25 percent is considered misleading. The exact threshold applied (ex. 10, 20, or 25 percent) can vary and should be confirmed during the planning stages. Note that the threshold is applied on either end of the confidence interval. Therefore, programming should prioritise tighter accuracy thresholds, as much as possible.

The *bias* was the direction of the deviation from the households' data for each question, either overestimated (+) or underestimated (-). Table 1 includes an example of KI reliability outcomes from the pilot in Arua, Uganda.

Example KI Reliability Outcomes: Percentage of households able to access a health facility in the previous six months in Awindara community (Arua, Uganda)

code KI	Accuracy score	Bias	KI answer	Household result - Confidence Interval
uga_ki_.1001	1	(-)	30%	[75.38% ; 84.91%]
uga_ki_.1015	0.05	(-)	70%	[75.38% ; 84.91%]
uga_ki_.1016	0	None	80%	[75.38% ; 84.91%]
uga_ki_.1018	0.1	(+)	95%	[75.38% ; 84.91%]

Example takeaways from Table 1:

- The KI code UGA_KI_1001 indicates that 25 percent of households could access a health facility in the previous six months. This is an underestimation of more than 30 percent compared to the household findings: the accuracy score is -1 (i.e. completely inaccurate).
- UGA_KI_1015 tells us that 70 percent of households could access a health facility in the previous six months. This is a deviation of +5 percent compared to household findings. Thus, the accuracy score is -0.05.
- UGA_KI_1016 indicates that 80 percent of households could access a health facility in the previous six months. That is within the range of household data and thus the accuracy score is 0.

Annex 2: Complete MFGD questioning route used in Mafrag, Jordan

INTRODUCTION

A. Facilitator's welcome, introduction and instructions to participants [5 minutes]

- *Facilitator completes an ODK form for each participant, recording the FGD code, location, number of participants, and start and end times of the discussion. Facilitator assists each participant in filling out their portion of the ODK form - age, gender, size of household, household demographics, type of residence, and number of years lived in Jordan.*
- Welcome and thank you for volunteering to take part in this discussion. You have been asked to participate as your point of view is important. I appreciate your time.
- This discussion is designed to understand how people in Mafrag define their community and the geographic space associated with this community, as well as to look into what factors influence people's identification with a certain community. To do this, the discussion will include some participatory mapping exercises.
- **Anonymity:** I would like to assure you that the discussion will be anonymous. We would appreciate it if you would refrain from discussing the comments of other group members outside of this session. If there are any questions or discussions that you do not wish to answer or participate in, you do not have to do so; however please try to answer and be as involved as possible.
- The discussion will take no more than 1-1.5 hours. We will have a quick break in between.

B. Ground rules [2 minutes]

- The most important rule is that only one person speaks at a time. There may be a temptation to jump in when someone is talking but please wait until they have finished. There are no right or wrong answers.
- You do not have to speak in any particular order.
- When you do have something to say, please do so. There are many of you in the group and it is important that I obtain the views of each of you.
- You do not have to agree with the views of other people in the group.
- Does anyone have any questions? (*answers*)
- With this in mind, may I tape the discussion to facilitate its recollection? (*if yes, switch on the recorder*)
- OK, let's begin.

QUESTION ROUTE (60-90 minutes)

Stage 1: ESTABLISH FAMILIARITY WITH MAP AND GEOGRAPHIC CONTEXT [20-30 minutes]

We are interested in understanding how you define your community and what are the reasons influencing this definition. To do this, we will engage in some participatory mapping activities which will help us identify the boundaries for your community and understand what are the factors influencing your identification of these boundaries.

1. **[Introduction to Mapping]** This is a map of your area within _____ (*Introduce participants to map and mark map with date, time and location*). Which location within this area do you live in? (*Mark each location with X*)
 - **[Probing Questions]** *Direct participants to the map and ask them to identify the following in their area. Mark each response using the symbol/ number assigned in the end of document. If a response is mentioned that is not part of the list, write down the response next to where it is marked.*
 - a. What are the main landmarks and infrastructure within this area? (*This could include main roads; public buildings such as schools, mosques; industries; dams, natural features such as hills, ridges, rivers, etc.*)
 - b. Within this area, what are the main public services being used by you and your family on a regular basis? Where are they located? (*This could include water points, schools, medical centres, public transport stops, public gardens, leisure centres, etc.*)
 - c. If agriculture is important for this area, and you or other residents are engaging in agriculture, what is the area of the land under cultivation? (*Use green marker stripes to mark this area*)

Stage 2: IDENTIFY COMMUNITY BOUNDARIES [30-40 minutes]

2. Now that we have identified the key features and landmarks in this area, we will move onto discussing more specifically about your community. What, in your opinion, defines a 'community'? What do you consider to be your 'community'? (**Note:** *Definition of 'community' should be kept relatively open at this stage to encourage discussion around what particular factors participants think define a community. Prior understandings of what defines and delineates a community area should be used as probes towards the second half of the discussion.*)
 - **[Probing questions]**
 - a. Do you consider e.g. your neighbourhood/village/[insert other possible community boundaries that you have identified during your background research] as your community?
 - b. Why do you consider this to be your community? (*Probe into what specific characteristics participants consider to be defining their community. For example, is it their village or neighbourhood? Is it the area which shares the same services and/ or resources? Or is it the group of people who share similar socio-cultural characteristics such as religion, tribe, nationality, etc.?*)
 - c. What is the immediate area occupied by your community? (*Mark area boundary with red marker*)
 - **[Additional questions]**
 - d. How many individuals and households are residing within this community?

Stage 3: IDENTIFY ADJACENT COMMUNITY AREAS [20 minutes]

3. If you are aware of other communities within this area, what are these communities and where are they located on this map? (**Note:** *Use smaller reference to mark adjacent community area boundaries with a blue marker pen*)
 - **[Probing questions]**

- a. What differentiates this community/communities from yours?
- **[Additional questions]**
- b. If you know individuals within this community/ communities who can introduce us to similar focus groups, could you share their details with us? (*Note down contact and other details as available*)

Stage 3: IDENTIFY COMMUNITY KI PROFILES [20 minutes]

4. Among the members of your community which we have just defined, who do you consider to be well-informed about the following issues? (**Note:** *Stress that we are not only asking about names of persons who they know personally but also about individuals in general they consider to be well-informed, such as a nurse in a hospital*)
 - a. Who, in your opinion, is well-informed about the availability and quality of education and healthcare services accessed by your community?
 - b. Who, in your opinion, is well-informed about the availability of and issues related to shelter and housing within your community?
 - c. Who, in your opinion, is well-informed about community members' involvement in economic and income-generating activities, including their employment patterns?
 - d. Who, in your opinion, is well-informed about displacement patterns within your community, especially the arrival and departure of refugees since the onset of the crisis in Syria in 2012?
 - **[Probing questions]**
 - e. *Use specific indicators from the indicator list/questions from the KI survey and probe participants: Do you think these people are able to give accurate information on [insert specific survey question] ?*
 - f. Why do you consider these individuals to be well-informed about these issues?
 - **[Additional questions]**
 - g. If you know specific individuals within your community who can inform us on the above, could you share their details with us? (*Note down contact and other details as available*)

CONCLUSION [5 MINUTES]

- Thank you for participating. This has been a very successful discussion.
- We hope you found it interesting. Your opinions will be a valuable asset to the study
- Do you have any questions before we conclude?
- I would like to remind you that any comments featuring from this discussion will be anonymous.
- Before you leave, please ensure you have completed the personal details questionnaire

MAPPING LEGEND

'X': Location of residence

Service Mapping (Numbering)

1. Schools
2. Public universities
3. Hospital
4. Clinics
5. Water points
6. Public parks/ gardens
7. Leisure facilities
8. Public transportation stops

Key infrastructure Mapping (Numbering)

- A. Mosque
- B. Industries
- C. Government building
- D. Other public building
- E. Dam
- F. Main roads and highways (highlight with black marker in addition if necessary)
- G. Bridges

Natural Features Symbology



Hills



Forest



River/ lake



Canyon (Wadi)

Green marker pen boundary with stripes: Land in use/ cultivation

Red marker pen boundary: Community area boundary

Blue marker pen boundary: Neighbouring community/ communities

Annex 3: MFGD moderator and note-taker tips

Tips for conducting MFGDs

The moderator should warmly welcome participants as they arrive and establish the MFGD space as a safe and friendly one in which participants may feel comfortable freely expressing themselves. The note-taker should complete the participant detail section in the MFGD questionnaire with each participant. If name labels are used, they should be prepared at this stage before the discussion begins.

Once participants have settled, the moderator should read the welcome script from the questionnaire. It is important to stress that information is confidential and that any views expressed during the discussion will not affect any future receipt of assistance. Finally, it is important to check if any of the participants have questions and ask for permission to turn on the recorder.

To facilitate analysis, it is helpful to save all digital files (such as recordings and, questionnaires) with the same file name which can also be recorded on all paper files – for example, reference maps. This is to ensure all materials from one MFGD can be matched correctly. The more communities and mapping sessions are conducted, the more important it is to establish an efficient filing system.

The moderator starts the recorder by clearly stating his/her name, date, community name, name of the official administrative unit within which the community is located, followed by information about the group (for example, host or displaced population, male or female). The note-taker records the file name at the top of each reference map that will be used during the MFGD.

Word-for-word translation of the discussion is not the objective. Instead, the note-taker should record detailed summaries of the discussion, making as many annotations on the notes on the maps as possible.

Tips for Moderators

Probing is the practice of eliciting good follow-up comments. A good probe:

- is specific
- is connected to the discussion underway
- can sometimes just be silence or waiting for more comments
- can be a listening response – repeating what the speaker is saying so that the speaker knows you are listening
- is neither judgmental nor a statement of opinion

Verbal signs of active listening include

- **Reflection**, repeating or paraphrasing what the speaker has said so as to show comprehension and generate further discussion.
- **Questioning**. The listener can demonstrate they have been paying attention by asking relevant probing questions.
- **Remembering**. Recalling a few key points, or even the name of the speaker, can help to reinforce that messages have been received and understood, - in other words, that listening has

been successful. Remembering details, ideas and concepts from previous comments proves that attention was kept and is likely to encourage the speaker to continue.

- **Summarising.** Like reflecting, it involves focusing on key points a participant has made. Particularly useful if a participant has made a complicated point which may not have been understood by the rest of the participants.

Non-verbal signs of active listening include:

- **Eye contact.** Participants should usually be encouraged to look at the speaker. However, eye contact can be intimidating, especially for the timid so gauge how much is contextually appropriate. Combine eye contact with smiles and other non-verbal messages to encourage the speaker.
- **Posture.** Attentive listeners tend to lean slightly forward or sideways whilst sitting. Other signs of active listening may include a slight slant of the head or resting the head on a hand.
- **Distraction.** An active listener will not be distracted and therefore will refrain from such things as fidgeting, slouching, checking the time, doodling, playing with their hair or picking their nails.

Be aware of non-verbal signs of agreement/disagreement. These may include other participants shaking their heads or showing signs of disagreement while somebody is making a point. Once the participant has finished her/his point, ask the others whether they disagree and why

Tips for Note-takers

- Look out for non-verbal cues from participants such as head nodding, laughter, discomfort or pauses.
- Use a new paragraph for each new point raised and look out for key themes for each question. These will likely be identified by several different participants. Sometimes they are said only once but with emphasis.
- Write down a new opinion the first time it is shared. When another person voices agreement with that opinion, put a check-mark next to it rather than writing down a similar statement. Note only the extra details or opinions offered.
- Record details of key relevant stories or experiences discussed by the group that inform their opinions.
- Record any contradictory opinions among the group, detailing roughly which proportion of the group holds which opinion. Do not generalise if not everybody agrees.
- Note follow-up questions that could be asked. The moderator is busy directing the discussion and may miss the importance of a particular follow-up probe.
- Record any big ideas, hunches or thoughts. For the note-taker may think of ideas that will be helpful in later analysis. Make a note of them on the form.
- Develop short-hand codes to write sentences, concepts, or words more quickly. For example, arrows, symbols, and abbreviations can be very useful to quickly capture ideas.
- Note the time displayed on the recorder when key quotes are said so you can go back and listen to capture the full comment.

Annex 4: Sample KI indicator list used in Mafraq, Jordan

Sector	Topic	Indicator
EDUCATION	Education needs	% of children of school-going age i.e. 6-18 years enrolled in formal education (by gender, nationality and age)
		% of enrolled children regularly attending formal education (by gender, nationality and age)
		% of children aged 6-18 that have been out of school for over one year (by gender, nationality, age and reason)
		Average time in months spent out of school for children aged 6-18 over the past five years (by gender, nationality and age)
		Top 3 priority needs to enhance access to and quality of education in BSU
	Access to education	% of children attending schools outside BSU, by reason
		Average distance in meters to school of attendance from HHs
	Quality of education	% of children with access to adequate learning materials, from any source
		Average class size by number of students (per primary/secondary schools)
	Impact	% of HHs reporting changes to the availability, access and quality of educational services within the last five years
	Sector	Topic
HEALTH	Healthcare needs	% of HHs with a member who suffered from health issues in the past 6 months, by type of health issue
		% of HHs with at least one member with a disability
		% of HHs with at least one member with a chronic illness, by type of illness
	Access to healthcare	% of HHs with a member who suffered from health issues in the past 6 months able to access required healthcare, by type of facility accessed
		Average distance in meters to nearest healthcare facility from HHs
		% of HHs reporting challenges in accessing healthcare in the past 6 months, by type of challenge
		Top 3 priority needs to enhance access to and quality of healthcare services in BSU

	Health expenditure	% of HHs covered by type of health insurance
		Top 3 alternative means to cover healthcare costs if not insured
	Impact	% of HHs reporting changes to the availability, access and quality of healthcare within the last five years
Sector	Topic	Indicator
LIVELIHOODS	HH income	Top 3 sources by share of HH income
		Top 3 sources by share of HH income before 2011
	Employment	% of adults within community in employment (by gender, nationality, and employment type)
		% of HHs by sector of employment
		% of HHs by type of employment (part-time/ full-time and temporary/ permanent)
	Work Permits	% of refugees in community who were able to acquire work permit authorisation since the change in Ministry of Labour regulations
	HH expenditure	Top 3 items by share of HH expenditure
	Challenges	% of HHs facing challenges in maintaining livelihoods, by type of challenge
	Coping strategies	% of HHs adopting strategies to cope with challenges faced in maintaining livelihoods, by strategy type
	External assistance	% of HHs receiving humanitarian aid, by type of assistance received
	Impact	% of HHs reporting changes to livelihood stability within the last five years
Sector	Topic	Indicator
SHELTER	Accommodation	% of HHs by type of current accommodation arrangement
		% of HHs by type of shelter
	Shelter Needs	% of HHs with inadequate housing conditions, by type of inadequacy
	Access to shelter	Average monthly rental costs in community
		% of HH who have experienced challenges accessing shelter in the past six months, by type of access challenge

	Shelter Expenditure	Average expenditure per household per month on shelter needs, by type of expense
		% of HH cash assistance spent on shelter and housing
	Impact	% of HHs reporting changes to availability and costs of housing within the past 5 years
Sector	Topic	Indicator
COMMUNITY AREA SPECIFIC ("DEMAND")	Household demographics	Average household size
		Average dependency ratio per household
		% of households headed by males/females
	Household arrival	% of families who have lived in neighbourhood for less than 6 months
		% of families who have lived in neighbourhood for 6 months to 1 year
		% of families who have lived in neighbourhood for 1 year to 3 years
		% of families who have lived in the neighbourhood for 3 years to 5 years
		% of families who have lived in neighbourhood for more than 5 years
	Main sources of income	Top 3 sources of household income in past month
	External assistance	% of HHs receiving humanitarian aid in past month, by type of assistance received
	Livelihoods challenges	% of HHs facing challenges in maintaining livelihoods in past month, by type of challenge
	Coping strategies	% of HHs adopting strategies to cope with challenges faced in maintaining livelihoods in past month, by strategy type
	Access to education	Top 3 challenges in accessing education
	Access to healthcare	Top 3 challenges in accessing healthcare in the last month
Shelter issues	% of HHs with inadequate housing conditions, by type of inadequacy	

Annex 5: Network related questionnaire

Do you know anyone from the list of KI provided?

[For each known KI]:

We identified {Name} as someone who is well informed about the needs and characteristics of your ward. Would you agree with that?

Yes No

Do you personally receive information from {Name} about your community?

Yes No

What is your relationship with {Name}?

Family ties Friendship ties Professional ties Prefer not to say Other

How do you usually receive information from {Name}?

Face-to-face By phone Through social media In writing Prefer not to say
 Other

Do you personally receive information from {Name} on these topics within {community}?

Our community WASH Healthcare Education Shelter situation
 Displacement Livelihood and income Other Not sure

In general, how reliable do you find the information provided by {Name}?

Very reliable Reliable Somehow reliable Unreliable Very unreliable

How often do you receive information by {Name} about your community?

Daily or almost daily (4-7 times a week) At least once a week (1-3 times a week) At least once a month (1-3 times a month) At least quarterly(1-2 times every three months)

During the last 3 months, how would you describe your communication with {Name} on {community}?

Passive transmission information: You send or receive information from this person via a non-personal way (email, sms, WhatsApp) Active transmission information: You send or receive information directly from this person Discussions: You have conversations and discussions with this person where issues are identified, discussed and clarified. Hypothesis exploration: You engage in conversations that lead to a new understanding of issues between the two of you. Other
 Prefer not to answer

Annex 6: Example of accuracy of topic from KI during pilots

Indicator	Value	Lower confidence interval	Upper confidence interval
Breadwinner	0.63	0.57	0.69
Distance water source	0.52	0.45	0.59
Drainage issue	0.52	0.50	0.55
Enough water	0.63	0.61	0.64
Health access	0.59	0.58	0.61
Health access challenges	0.58	0.56	0.60
Household size	0.55	0.53	0.57
Livelihood challenge	0.45	0.42	0.48
Monthly income	0.39	0.31	0.46
Priority wash	0.27	0.23	0.32
Public private water source	0.10	0.06	0.14
Shelter issue	0.42	0.39	0.46
Solid waste disposal	0.63	0.61	0.65
Toilet type	0.42	0.41	0.44
Water concern	0.58	0.51	0.64
Assistance source	0.30	0.23	0.36
Enough food	0.54	0.53	0.55
Expenditure food	0.34	0.26	0.41
Food source	0.36	0.35	0.37
Household number	0.51	0.46	0.56
Income source	0.10	0.07	0.12
Leave reason	0.08	0.08	0.09
Livelihood coping strategy	0.42	0.36	0.49
Member separated	0.71	0.54	0.87
Mental disability	0.03	0.03	0.03
Physical disability	0.03	0.03	0.03
Shelter type	0.29	0.29	0.30
Formal education	0.20	0.19	0.22
Intention duration	0.47	0.40	0.54
Female head of household	0.13	0.12	0.14
Number children	1.00	1.00	1.00
Water price	0.71	0.66	0.75
Documentation children	0.43	0.34	0.51
Hosting number	0.64	0.61	0.66
Improvement education	0.69	0.64	0.74
Intention return	0.62	0.55	0.69
Toilet location	0.34	0.32	0.37
Water source	0.39	0.38	0.40
Education access challenges	0.30	0.28	0.32

Annex 7: Theoretical Framework (Social Network Analysis)

The theoretic framework employed here builds on the assumption that KI networks are social networks comprised of informed stakeholders, or key informants, who share interpersonal ties between each other. These ties allow KIs to receive, disseminate and exchange information on a given topic with one another. Some ties may be stronger, some weaker, some stakeholders may not even have a direct tie to another stakeholder in their network because they do not know each other, yet they may indirectly receive information through mutual contacts. A key informant network may be thought of as a web of interpersonal links that allows for a flow of information on a given topic. Depending on a key informant's position within such a network, she may have broader or more limited access to better or less reliable information

This section develops the theoretical framework on the basis of which the Social Network Analysis is implemented. It will first explain different centrality measures, that are necessary to describe a key informant's position within a network, and then briefly introduce network characteristics, which may further inform the Social Network Analysis. In how far network characteristics will have to be taken into account in employing the network analysis as presented in this approach, will have to be seen after analysing pilot results.

Centrality Measures

The theoretical framework for the Social Network Analysis as it is employed in this approach comprises three different centrality measures, namely degree, closeness and betweenness. These are developed in the following.

Degree of Centrality

As the most basic of the three *centrality* measures employed in this approach, the *degree of centrality* measures how many other key informants a key informant receives information from on a defined area of interest. More technically speaking, the *degree of centrality* of a key informant within her network is the weighted sum of all "incoming" links that connect her directly with other key informants of her network. The degree centrality of Key Informant (i) can be expressed mathematically as follows:

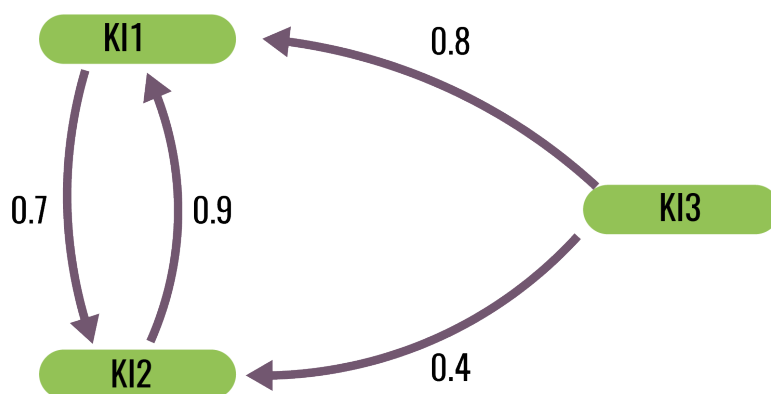
Formula 1: Degree Centrality of a Key Informant i

$$C_D(i) = \sum_{j \neq i} w_{j,i}$$

With: i = Key informant for which the degree centrality is calculated; j = Key informants that directly inform i, i.e. without going through any intermediaries; $w_{i,j}$ = the weight of the link that goes from j to i.

Example A1: Calculating the degree of centrality for a given key informant

An example of a Weighted Directed Network of key informants



Tying in with Example 2 this example illustrates how to actually calculate scores for both Sam's and KI2's degree of centrality within the network presented there. Using the formula above, Sam's *degree of centrality* score is:

$$C_D(KI1) = \sum_{j \neq \text{Sam}} w_{j,KI1} = w_{KI2,KI1} + w_{KI3,KI1} = 0.9 + 0.8 = 1.7$$

While KI2's score is:

$$C_D(KI2) = \sum_{j \neq \text{Sarah}} w_{j,KI2} = w_{KI1,KI2} + w_{KI3,KI2} = 0.7 + 0.4 = 1.1$$

In accordance with the conclusion made in Example 2, these scores confirm that KI1 has a higher degree of centrality in this network than KI2 does, as his score (1.7) is higher than hers (1.1).

Closeness Centrality

The centrality of a KI in a network is only partially described by her degree of centrality (see Example 3). Other measures of centrality are thus necessary to conduct a thorough Social Network Analysis. This section develops and explains the components that feed into the measure of closeness centrality. The last part of this section then presents the mathematical definition of *closeness centrality*.

The *closeness centrality* measure is composed of two components:

1. A notion of **distance** between key informants in a network
2. The number of key informants from which a given key informant can receive information through both direct and indirect links, the **scope of inward connections**.

In order to understand the *closeness centrality* measure, it is helpful to think of key informant networks as webs that allow information to flow and travel throughout. Information is not usually inherent to one person, but is passed on and disseminated to other peers. In a sense this allows information picked up on one end of the network to "travel" to the other end. In doing so, it can be passed on and disseminated through different "paths", that is through different intermediates, depending on how interconnected the network is. The *closeness* measure, as used here, is based on two assumptions:

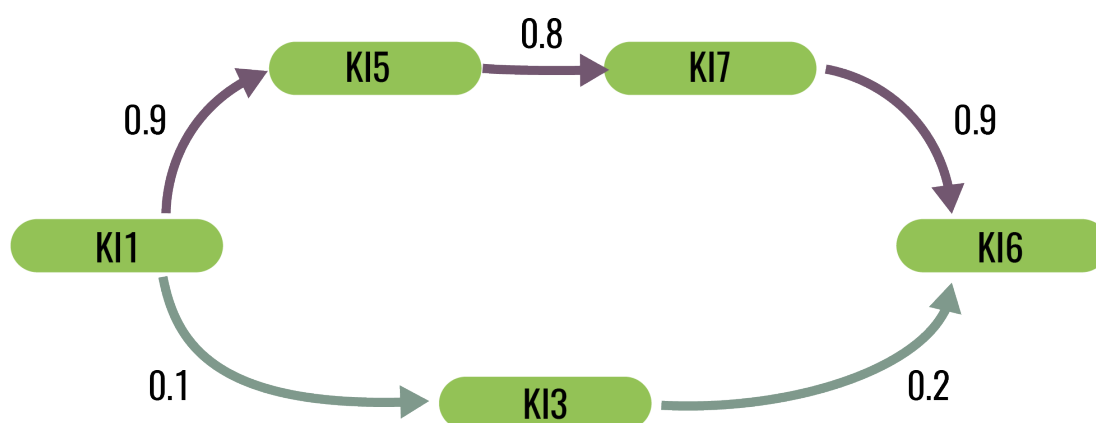
1. The shorter the path information travels on, meaning the fewer the intermediates information has to go through before reaching a given key informant, the more accurate and complete the information received on the other end

- The higher the level of trust and frequency (i.e. the higher the weight) for each link through which information is passed on, the more reliable the flow of information, and again the more accurate and complete the information received on the other end

From these assumptions, it follows that the ideal path through which information can “travel” between two key informants would be as short as possible, meaning going via the fewest intermediates possible, while comprising the most trustworthy and frequently used links possible.

Example A2: The most reliable path to receive information

Flow of information in a key informant network



This figure shows an example of a weighted directed network in which information is passed on between key informants. Given the links between key informants in this network, it is likely that information picked up by K11 will at some point reach K16, albeit indirectly because K16 and K11 do not directly exchange information. Although information passed on through the green path only has to go through K13 to reach K16 it may be less reliable and accurate than had it been passed through the blue path. This is because the links between K11 and K13 as well as between K13 and K16 are not as strong as the links between key informants on the blue path. The weights suggest that K13 only so often receives information from K11 and that this information is most of the time inaccurate (for whatever reason). The same goes for information passed on from K13 to K16. Therefore, information that was initially passed on by K11 and ultimately reaches K16 through K17 and K15 is expected to be more accurate than information that travelled through the green path. The concept of closeness centrality takes this into account using a measure of “distance” between key informants that builds on both the number and the strength of links.

The notion of ‘distance’ within a key informant network⁵²

Within this theoretic framework, the notion of *distance* is, loosely speaking, understood as the distance information has to “travel” between two key informants in a given network. As such, it takes into account the number of intermediaries information passes through to get from one key informant to another. However, as outlined in Example A2 above, a “short” path between two key informants within a network is not necessarily one that goes through the fewest intermediaries. “Shortness” is also a matter of going through stronger links, i.e. links that are characterized by higher frequency of exchange and higher levels of trust. Furthermore, direction of exchange needs to be taken into account as well. It makes little sense to talk about *distance* between two key informants in a scenario where information cannot possibly reach the key informant on the other end. In this example:

⁵² Adapted from Newman, M. E. J., 2001. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E* 64, 016132.

$$K11 \rightarrow K16 \rightarrow K13 \leftarrow K14$$

one could talk about the *distance* between K11 and K13, but not between K11 and K14, because K14, in this setting, would never possibly receive information that initiated from K11, as he doesn't receive information from K13.

Based on these considerations, **distance $d(i,j)$** between two key informants is defined as the sum of inverse weights of all links that information has to pass through to travel from Key Informant j to Key Informant i:

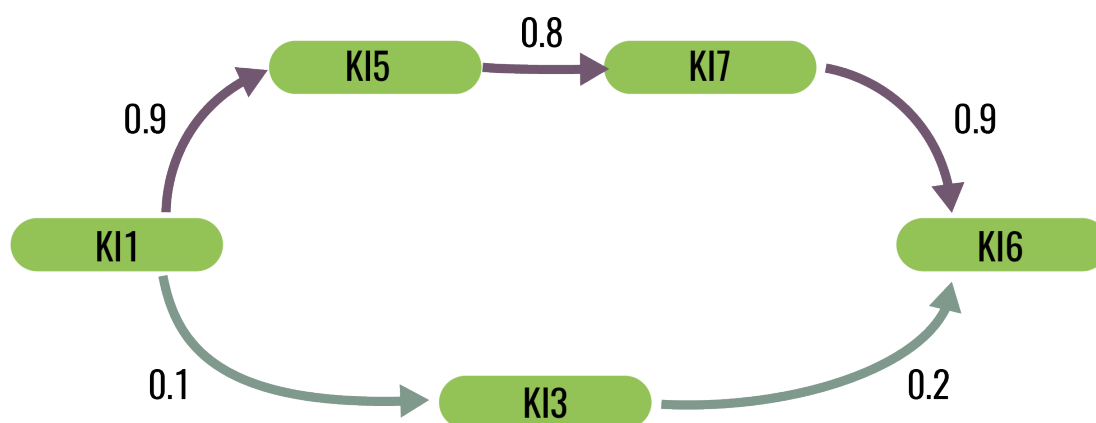
Formula 2: Distance between key informants i and j

$$d(i,j) = \sum_{k \neq l, k,l \in M} \frac{1}{w_{k,l}}$$

With: M the set that includes all intermediate key informants between Key Informant i and Key Informant j; $w_{k,l}$ the strength of the directed link between Key Informant k and Key Informant l, on which information travels on its way to reach Key Informant j.

Example A3: Calculating the distance between two key informants

Flow of information in a key informant network



This is the network on the basis of which in Example A2 above the potentially most reliable, or *shortest*, path between K11 and K16 was discussed. With the above definition of *distance* the most reliable path between K11 and K16 can be determined by calculating the *shortest* path between the two.

The distance between K16 and K11 on the green path is calculated as follows:

$$d(K16, K11) = \frac{1}{0.1} + \frac{1}{0.2} = 15$$

While the distance on the blue path:

$$d(K16, K11) = \frac{1}{0.9} + \frac{1}{0.8} + \frac{1}{0.9} = 3.47$$

Hence, information from K11 travels a “shorter” distance to K16 on the blue path than on the green one. In other words, information coming through the blue path is expected to be more accurate and reliable than information coming through the green path.

The sum of shortest distances

With a measure of distance between two key informants, as described above, it is easy to determine the shortest path between key informants (see Example A3). It is important to note that there might be more than one shortest path between two key informants, depending on the characteristics of the network. The *closeness centrality* requires to calculate the sum of shortest paths between a given key informant and all other key informants from whom he or she can directly or indirectly receive information. In this case, **only one shortest path between two key informants should be accounted for**.⁵³ In line with the definition of distance between two key informants, let $l(i,j)$ denote the shortest path between Key Informant i and Key Informant j , or more precisely **from** Key Informant j **to** Key Informant i . Then the sum of shortest paths for Key Informant i can be written as:

Formula 3: The sum of shortest paths for Key Informant i

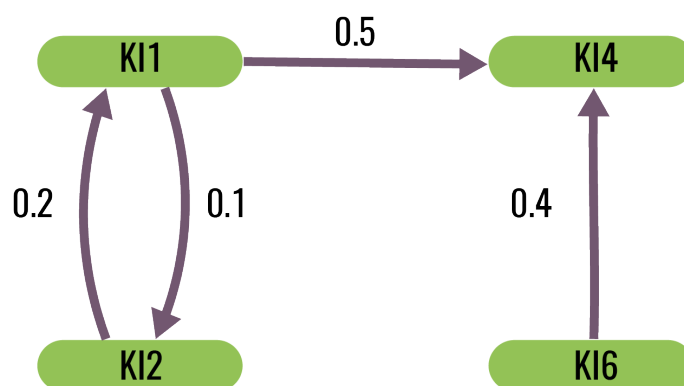
$$\text{Sum of shortest paths } (i) = \sum_{j \neq i} l(i,j)$$

With: j all key informants within the network key informant i can possibly receive information from, be it directly or indirectly

Example A4 shows, how the sum of shortest paths is calculated based on the above formula.

Example A4: Calculating the sum of shortest paths for a given key informant

Figure 1: Shortest paths within a key informant network



Based on the definition of distance between two key informants, calculating the sum of shortest paths between a given key informant and all other key informants from whom he or she can directly or indirectly receive information is straight forward.

First, all shortest (inward) paths need to be identified, then their distances can be summed up. Given the simplicity of this example, this can be done in one step. For Sam, the sum of shortest paths calculates as follows:

$$\text{Sum of shortest paths}(KI1) = \sum_{j \neq KI1} l(KI1,j) = l(KI1, KI2) = \frac{1}{0.2} = 5$$

⁵³ Else, as can be seen from the formula for *closeness centrality* (see page 40), a key informant would be penalized in their *closeness* score for having more than one shortest path with another key informant, which shouldn't be considered a bad thing in terms of centrality within a network.

As there is only one other key informant K11 can receive information from within this network, namely K12, the sum of shortest paths equals the distance of the link from K12 to K11.

For K14, it is slightly more complicated, as he receives information from K11 and K16, but can potentially receive information also from K12. His sum of shortest paths is thus an actual sum:

$$\begin{aligned}
 \text{Sum of shortest paths}(K14) &= \sum_{j \neq K11} l(K14, j) = l(K14, K11) + l(K14, K16) + l(K14, K12) \\
 &= \frac{1}{0.5} + \frac{1}{0.4} + \left(\frac{1}{0.2} + \frac{1}{0.5} \right) = 11.5
 \end{aligned}$$

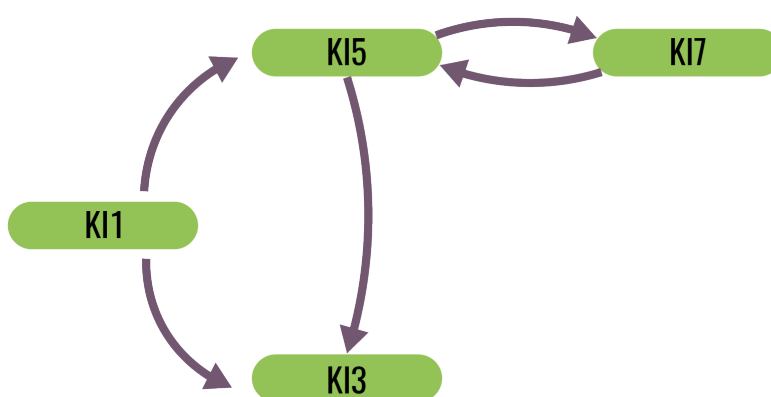
Intuitively, the smaller the sum of shortest paths for a given key informant, the “closer” connected a key informant to the rest of the network because the smaller, on average, the distance information has to travel to reach this person from across the network. However, the sum of smallest distances does not take into account how many key informants in a network a person can potentially receive information from. Say a key informant is only connected to one other key informant in a network, her sum of shortest paths would hence be comparatively small as there is only one possible path that feeds into the equation above. The sum of shortest paths therefore is, taken on its own, an insufficient measure for centrality. For this reason, the *closeness centrality* measure also takes into account the number of key informants from which a given key informant can potentially receive information: the *scope of inward connections*.

The Scope of Inward Connections

Aside from *distance*, the *closeness* measure includes the number of key informants from which a given key informant can potentially receive information. This number is indicative of the amount of knowledge within a network a given key informant potentially has access to. This number is called here the *scope of inward connections*.

Example A5: The scope of inward connections

Figure 2: Scope of inward connections in a key informant network



Shows a key informant network with links indicating the direction of information flow between key informants. Based on this information, it is easy to determine the *scope of inward connections* for each key informant:

$$\text{Scope}(K15) = 2; \text{Scope}(K11) = 0; \text{Scope}(K13) = 3 \text{ and } \text{Scope}(K17) = 2$$

K11's scope of inward connections, for example, is 0 because no information reaches him from within the network (all his links are outward). K13, on the other hand, does directly receive information from K11 and K15, and through K15 can potentially receive K17' information, too. Hence, his scope of inward connections is 3. K13, in that sense, has potentially access to a much wider share of knowledge within the network than K11 does.

Calculating the Closeness Centrality of a Key Informant

The *closeness centrality* of a key informant is calculated by her scope of inward connections and the sum of shortest paths between her and all other key informants within her scope. The measure is constructed in a way that it produces higher scores for key informants that benefit from a larger scope of inward connections, while being positioned in the network in a way that distances to other key informants are small. Hence the **closeness centrality** can be expressed as:

Formula 4: The Closeness Centrality of a Key Informant

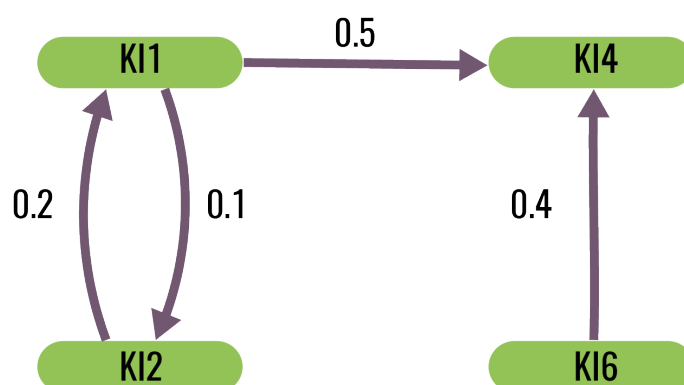
$$C_c(i) = \frac{Scope(i)}{\sum_{j \neq i, j \in M} l(j, i)}$$

With: M the set comprised of all key informants from which key informant i can potentially receive information, $l(i, j)$ = shortest path going **from** key informant j **to** key informant i, $\sum_{j \neq i, j \in M} l(i, j)$, the sum of shortest paths **to** key informant i, and $Scope(i)$ the scope of inward connections for key informant i.

A key informant who scores higher on *closeness centrality* than another key informant is then expected to have access to a larger share of knowledge that exists within their network, while disposing of more reliable channels to tap this information.

Example A6: Calculating the closeness centrality of a key informant

Figure 3: Shortest paths within a key informant network



Picking up from example A3 above in which the sum of smallest paths was calculated for K11 and K14, this example looks at the *closeness centrality* of these two key informants within their key informant network.

Based on the definition of *closeness centrality* and given the sum of smallest paths, calculated in the example above, K11's *closeness centrality* score is as follows:

$$C_c(Sam) = \frac{Scope(KI1)}{\sum_{j \neq Sam, j \in M} l(KI1, j)} = \frac{1}{l(KI1, Ki2)} = \frac{1}{5} = 0.2$$

While K14's *closeness centrality* score is:

$$C_c(Hazim) = \frac{Scope(KI4)}{\sum_{j \neq Sam, j \in M} l(KI4, j)} = \frac{3}{l(KI4, KI2) + l(KI4, KI1) + l(KI4, KI6)} = \frac{3}{11.5} = 0.26$$

This means K14 is considered to be more "closely" connected to the overall network than K11, as his *closeness centrality* score is higher, if by just a little, than K11's. This result seems, at first sight, contrary to what one may expect, as there are more links (inward and outward) that connect K11 to the rest of the network (three in total)

than there are for KI4 (only two). It is important to keep in mind, though, that what really counts here are the “inward” connections, that is the links through which a key informant can receive information from other key informants. Seen from this perspective, KI4 has access to a larger share of information within the network than KI1 does, as he receives information from KI6, KI1 and indirectly potentially from KI2 as well, while KI1 only receives information from KI2. KI4 is thus “closer” to other key informants in the network than KI1.

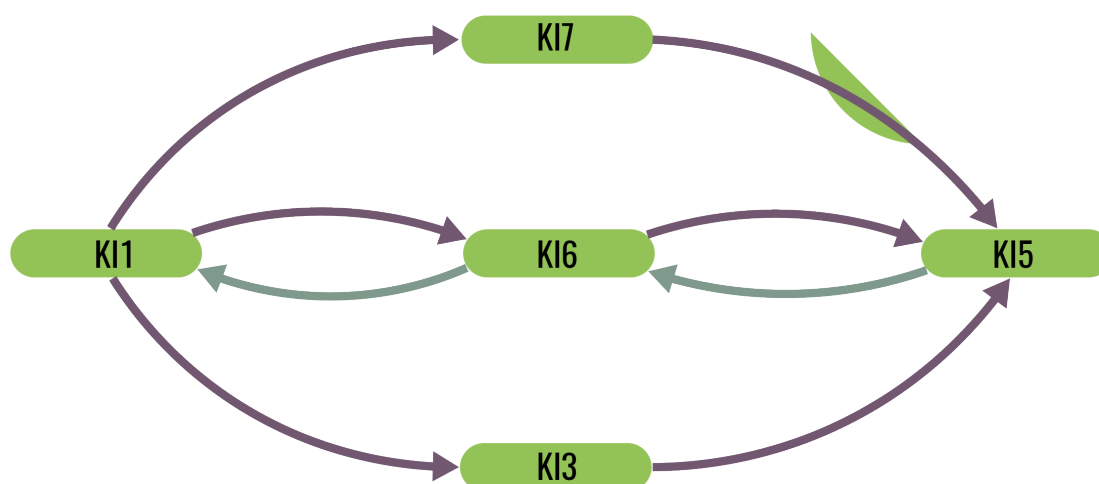
Betweenness Centrality

As discussed in the section on betweenness centrality complements the degree and closeness measures to determine a key informant’s position within a network. This measure technically quantifies how often a key informant acts as an intermediate in passing on information (on shortest paths) within the network. As such, it indicates the importance of a given key informant in facilitating exchange of information within a key informant network.

The betweenness centrality measure is composed of the number of shortest paths that go through a given key informant (see Example A7) and the overall number of shortest paths between each pair within the network (see Example A8). The ratio between the two is a measure of how well positioned a key informant is to receive and disseminate information within the network.

Example A7: The number of shortest paths between key informants

Shortest paths between two key informants



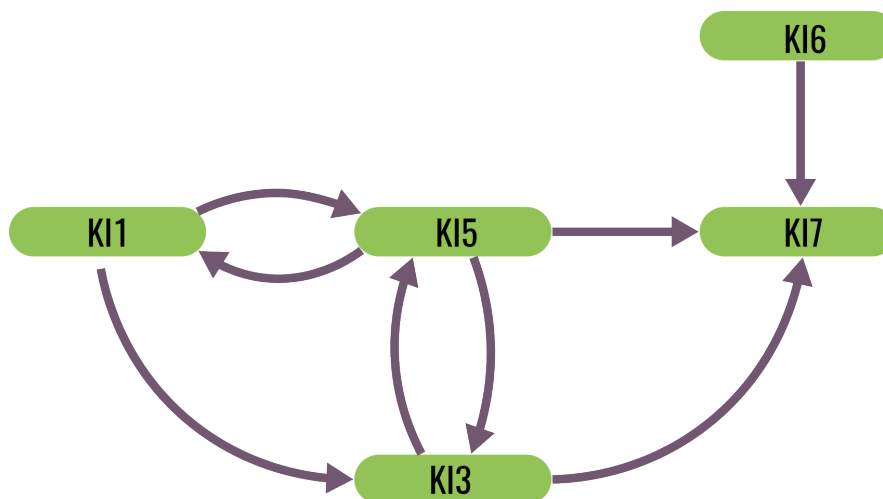
Example A7, shows a key informant network with directed links between key informants, indicating the flow of information. For reasons of simplicity, this example does not include weights, meaning all links are regarded to be equally strong. The number of shortest paths between two key informants depends on their direction. Rather than “the number of shortest paths between A and B” one should thus refer to “the number of shortest paths from A to B” and “the number of shortest paths from B to A”. The latter two may not be the same.

In this example, the number of shortest paths from KI1 to KI5 is 3. Disregarding weights, each blue path is equally “long”, as each of them goes through only one intermediate. The number of shortest paths from KI5 to KI1, on the other hand, is 1, the green path being the only, hence the shortest path from KI5 to KI1.

While the example above considered the number of shortest paths between two given Key informants, the following example examines the number of all shortest paths within a network.

Example A8: The sum of shortest paths in a network

Shortest paths within a key informant network



Similarly to example A7, this example shows a key informant network with equally weighted directed links. In order to calculate the sum of shortest paths within this network, all shortest paths between each pair of key informants have to be determined. In a network as simple as the above, this is done very easily without the help of software. These are the shortest paths between each pair of key informants:

1. KI1 → KI5
2. KI1 → KI3
3. KI1 → KI5 → KI7
4. KI1 → KI3 → KI7
5. KI5 → KI1
6. KI5 → KI3
7. KI5 → KI7
8. KI3 → KI5
9. KI3 → KI7
10. KI6 → KI7

Note that there are two shortest paths in this example from KI1 to KI7; both need to be accounted for to determine the sum of shortest paths in this network. As mentioned in Example A7, direction of paths also matters: there is a shortest path from KI1 to KI5, for example, and a shortest path from KI5 to KI1, equally for other pairs. Taken these considerations into account, the sum of shortest paths within the network in the figure is 10.

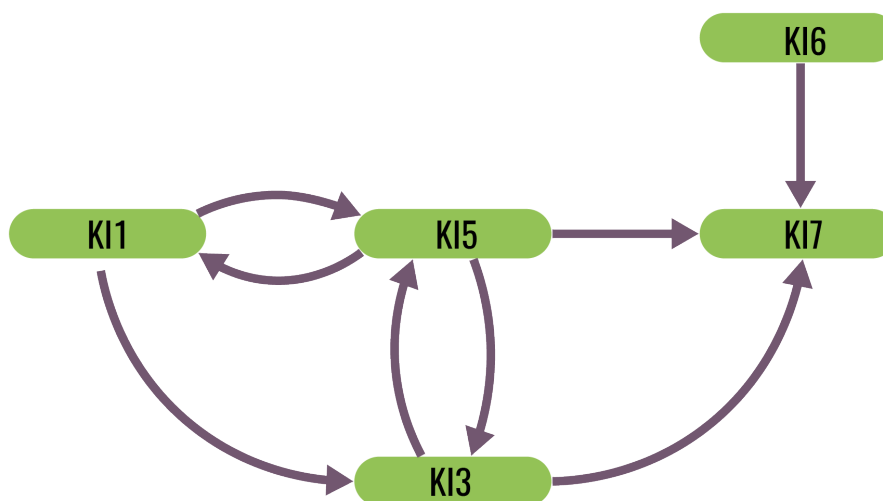
With the number of shortest paths between two key informants and the sum of shortest paths within a network the measure of betweenness centrality can be constructed.

Calculating the Betweenness Centrality of a Key Informant

As mentioned above, the *betweenness centrality* is the ratio between the overall number of shortest paths between each pair of key informants within the network and the number of shortest paths that go through a given key informant. This example illustrates how a key informant's *betweenness centrality* is determined.

Example A9: Betweenness centrality of a given key informant

Shortest paths within a key informant network



Picking up from the example above, the overall number of shortest paths within the network has been found to be 10. In order to determine KI5's *betweenness centrality*, the number of shortest paths that go via him, needs to be determined. Simple as the network in this example is, this can be done manually. In fact, there are only two shortest paths in this network that go via intermediate key informants:

1. K11 → K15 → K17
2. K11 → K13 → K17

However, only one of these goes via KI5. The number of shortest paths in this network that go via KI5 is thus 1. KI5 is therefore an intermediate on 1 out of 10 shortest paths within the network. In other words, 10% of all shortest paths in this network go via him. This percentage corresponds to his *betweenness centrality*. It is easy to see that the same goes for KI3, while no key informant other than the two functions as an intermediate in this network. KI5's and KI3's *betweenness centrality* is therefore higher than anyone else's in this network.

As for the other two centrality measures outlined above, there is a mathematical description to calculate a key informant's *betweenness centrality*. Let $P_i(k, j)$ denote the number of shortest paths that go from Key Informant k to Key informant j and pass through Key Informant i , and $P(k, j)$ the number of shortest paths that go from Key Informant k to Key Informant j (regardless of who these pass through), then *betweenness centrality* of Key Informant i can be expressed as:

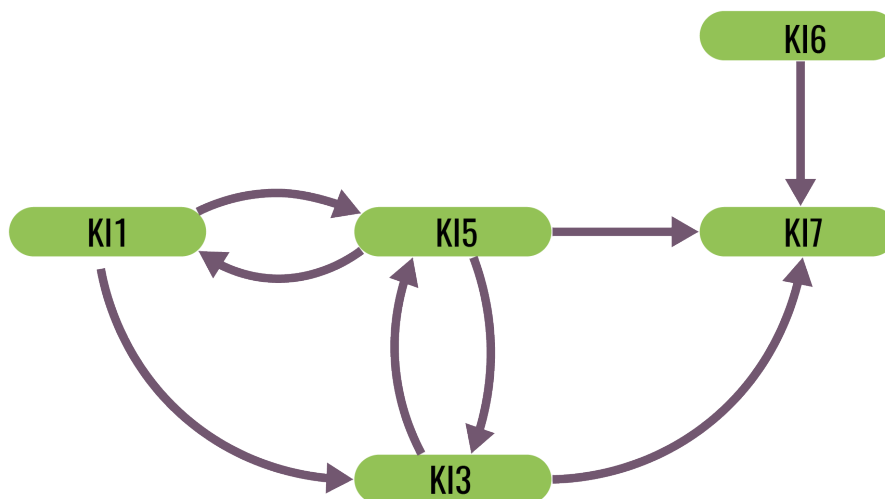
Formula 5: Betweenness Centrality of a Key Informant i

$$C_B(i) = \sum_{k \neq j; i \in (k, j)} \frac{P_i(k, j)}{P(k, j)}$$

The following example shows how this formula is applied.

Example A10: Calculating the *betweenness centrality* of a given key informant

Shortest paths within a key informant network



This is the network presented in Example A9 above to illustrate how to determine the *betweenness centrality* of a key informant. It was found that K15 and K13 each function as intermediates on 10% of all shortest paths that exist in this network. This example reproduces these findings using the above formula to calculate K15's and K13's scores of *betweenness centrality*:

$$C_B(Max) = \sum_{k \neq j; Max \notin (k,j)} \frac{P_{Max}(k,j)}{P(k,j)}$$

With $\sum_{k \neq j; Max \notin (k,j)} P_{Max}(k,j)$ being the sum of all shortest paths within the network that pass through K15, and $\sum_{k \neq j; Max \notin (k,j)} P(k,j)$ the overall sum of all shortest paths within the network.

$$C_B(Max) = \sum_{k \neq j; Max \notin (k,j)} \frac{P_{Max}(k,j)}{P(k,j)} = \sum_{k \neq j; Max \notin (k,j)} \frac{P_{Max}(K11, K17)}{P(k,j)} = \frac{1}{10} = 0.1$$

Similarly for $C_B(Woody)$.

With the *betweenness centrality* formula at hand, it is easy to see that everyone else's *betweenness centrality* score is 0 because the numerator is 0, for there are no shortest paths with intermediates other than K15 and K13 in this network.

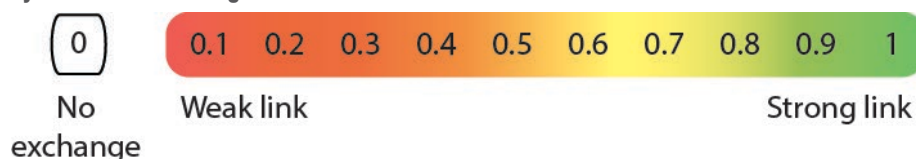
Constructing Weights

Weights of links between two KIs are an indication for the strength of their connection. Based on the KI network survey, scores for frequency of exchange and perceived trust, which both range from 0 to 5, are determined. The weight of a link between two key informants is then constructed as follows:

$$w = \frac{\text{perceived trust score} + \text{frequency score}}{10}$$

The score for perceived trust and frequency of exchange will be summed up, and the sum divided by ten. This will produce weights between 0 and 1, 1 being the strongest possible link and 0.1 the weakest possible, while 0 can be considered as no interaction between two key informants.

The scale of key informant link weights



As discussed, links are actually directed, meaning there is a difference between the link that goes from, say, Key Informant *i* to Key Informant *j* and the link that goes the other direction, from Key Informant *j* to Key Informant *i*. More precisely, the weight of the link going from Key Informant *j* to Key Informant *i* depends on *i*'s perceived trust with regards to information she obtains from Key Informant *j* and the frequency she reports, hence:

$$w_{j,i} = \frac{\text{perceived trust score } (i) + \text{frequency score } (i)}{10}$$

Example A11: Strength of Links

Within her community area, Leila was identified as a potential key informant on health-related issues. She reports to receive health-related information from Laith, another identified key informant in her community area, on a daily basis. Leila further states that the health-related information she receives from Laith is often accurate. Based on this information, the weight of the link between Leila and Laith within this health-related key informant network is calculated as follows:

Laith's scores: "daily" = 5 and "often" = 4

$$\text{Weight} = \frac{5+4}{10} = 0.9$$

With a weight of 0.9 out of 1, this can be considered a very strong link. Note that the direction of this link goes from Laith to Leila. We do not know, in this example, about the flow of information in the other direction, from Leila to Laith. In a weighted directed network this will be taken into account.

The **Strength of links** outlined above is used to construct a weighted directed network. The weights are assigned based on the strength of the links between each pair of informants exchanging information (including those cases where there is no exchange but a one-directional flow of information between two key informants).

Network characteristics

Additional factors that may feed into the Social Network Analysis concern network characteristics. Characterization of a network include size, density and reciprocity within the network. These measures indicate the level of network efficiency in spreading or collecting information in a specific community. Depending on size, structure and reciprocity of the network, different strategies might be used to identify the most reliable key informants for each area of interest. Based on pilot results, this may have to be explored further. Suffice it for now to briefly introduce network characteristics without incorporating these into the analysis approach as developed above.

Size and density of the network – can indicate how fast information can spread through the network:

Formula 6: Density of a network

$$\text{dens} = \frac{n(n-1)}{d(g)}$$

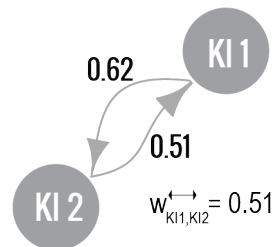
With: n = total number of key informants (“size of the network”); d (g) = Number of links in the network

Reciprocity – indicates the rate of two-way reciprocal connections between key informants within the network:

Reciprocity of a network

For each connection between KIs, the minimal weight (based on **Strength of exchange**) is calculated

$$w_{i,j}^{\leftrightarrow} = \min[w_{i,j}, w_{j,i}]$$



For the entire network, the sum of all weights between two nodes w and the sum of the minimal weights between 2 nodes w^{\leftrightarrow} are calculated as follows:

$$w = \sum_i \sum_{j \neq i} w_{i,j}$$

and

$$w^{\leftrightarrow} = \sum_i \sum_{j \neq i} w_{i,j}^{\leftrightarrow}$$

The reciprocity is the ratio of the sum of minimal weights between two nodes and the sum of all weights between two nodes:

Formula 7: Reciprocity of a network:

$$r = \frac{w^{\leftrightarrow}}{w}$$