USE CASE C

INCLUDING IDPS IN A SAMPLE SURVEY OF THE NATIONAL POPULATION, OR RUNNING A STAND-ALONE SURVEY OF IDPS
This is Use Case C from the Compiler’s Manual on Forced Displacement Statistics. The Use Case describes how Internally Displaced Persons (IDPs) can be included in a sample survey of the national population and how to run a stand-alone survey of IDPs.

The Compiler’s Manual is aimed primarily at technical personnel in National Statistical Systems who want to include displaced populations – refugees and / or IDPs – in official statistics. Each Use Case discusses a different scenario relevant to producing official statistics on refugees and IDPs, with a focus on the elements of statistical production cycles that are specific to refugee and IDP contexts. Spotlight examples of good practice in the production of refugee and IDP statistics are interwoven throughout the Use Case.

The Compiler’s Manual and its individual Use Cases are intended to be a ‘living document’ which will be amended and extended as the body of expertise and knowledge develops worldwide.

Note: Paragraph numbering is per the complete version of the Compilers Manual.

The Expert Group on Refugee, IDP and Statelessness Statistics

The Expert Group on Refugee, IDP and Statelessness Statistics is a UN Statistical Commission mandated, multi-stakeholder group that works with National Statistical Offices, international organizations and civil society to develop and support implementation of international standards and guidance to improve official statistics on forcibly displaced and stateless persons.

The Compilers’ Manual

The Compilers’ Manual offers clear operational instructions on producing official statistics on refugees, asylum seekers, IDPs and related populations. It complements the content of the International Recommendations on Refugee Statistics and the International Recommendations on Internally Displaced Persons Statistics by providing hands on guidance.
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample surveys – general background</td>
<td>4</td>
</tr>
<tr>
<td>Questionnaire design</td>
<td>5</td>
</tr>
<tr>
<td>Sampling considerations when surveying IDPs</td>
<td>12</td>
</tr>
<tr>
<td>Fieldwork planning</td>
<td>18</td>
</tr>
<tr>
<td>Enumerator training</td>
<td>21</td>
</tr>
<tr>
<td>Fieldwork in practice</td>
<td>22</td>
</tr>
<tr>
<td>Data quality</td>
<td>22</td>
</tr>
<tr>
<td>Data protection / confidentiality / disclosure</td>
<td>23</td>
</tr>
<tr>
<td>Analysis and dissemination</td>
<td>23</td>
</tr>
</tbody>
</table>
Sample surveys – general background

102. Sample surveys are an important data source for measuring the characteristics and living conditions of IDPs and IDP related populations. When planning a representative survey of IDPs, NSOs have the option of including IDPs in a broader survey of the national population (e.g. LFS, MICS, DHS, integrated living conditions survey, etc.), or to design a specialized survey for IDP populations. Both approaches have their merit: a specially designed IDP survey has the potential to collect broader and deeper information, whereas inclusion in a national multi-topic survey allows IDPs to be compared with other groups in the general population. The two options of course are not mutually exclusive, and countries may decide to implement a complementary mix of both options. (A detailed discussion of pros and cons of the two approaches, and of sample survey more broadly, is provided in IRIS paras 220-225 and 238-244).

103. The priorities when considering including IDPs in a sample survey are:

- In the design of the questionnaire, include recommended questions to identify IDPs.
- Identify which indicators of progress towards durable solutions are priority for data collection and analysis. If expanding an existing survey, establish whether any are already included in the data collection and add in question sets for any additional indicators.
- Establish a sampling frame. If possible, expand an existing sampling frame or develop an alternative sampling frame that includes IDPs, or the locations where IDPs reside, including camps and temporary settlements.
- Undertake field visits to test sampling strategies, identify potential issues or concerns and to build relations between the parties involved in fieldwork implementation.
- Ensure that enumerator training allows sufficient time to cover IDPs, including the relevant definitions and concepts, practical survey considerations and data protection.

104. Other than these special considerations during sample design and adjustments of the questionnaire and analysis, there are no further major technical implications from the statistical framework set out in IRIS for surveys of IDPs. All the usual principles and best practices for running sample surveys for official statistics apply - including on fieldwork security risk assessment and mitigation - and are not discussed here.

105. Many of the issues discussed in this Use Case are similar to the considerations when including refugees in a sample survey (Use Case B) and so are repeated here to ensure both Use Cases serve as stand-alone resources.
Questionnaire design

106. Designing an effective questionnaire is fundamental to the success of any survey. When considering including IDPs in an existing or new survey the key factors are to ensure appropriate identification and classification questions are included and the question flow in the sections covering places of residence and migration status works accurately given the additional complexity. Typically, questions on IDP status should be introduced in the flow of questions on migration status.

107. Wherever possible, when including IDPs, survey development timescales should allow additional time for questionnaire development, to provide an opportunity to bring together relevant stakeholders and seek their input to the survey design. Investing time to engage stakeholders (for example, relevant government agencies, key communities, organisations supporting IDPs, UNHCR) early in the project can help ensure the success of the survey by bringing together relevant knowledge and appropriately accounting for any sensitivities.

CASE STUDY: COORDINATING WITH STAKEHOLDERS ON QUESTIONNAIRE DESIGN IN SUDAN

In 2020, a large-scale analysis covering eight localities across Darfur’s five states was initiated to provide a shared evidence-base to support peacebuilding and durable solutions under the UN Peacebuilding Fund (PBF). JIPS was requested to provide technical support. The project targeted both displaced and non-displaced population groups to measure the progress towards durable solutions among IDPs.

The design of the household survey tool was a consultative process that involved all stakeholders in the project. This meant presenting and sharing the progress at each step and collecting input, questions and comments. It ensured that the final questionnaire had wide buy-in and support and could underpin a successful survey. JIPS led the design of methodology and tools and conducted the analysis and reporting. The process was guided by UNHCR and the other PBF agencies (UNDP, UNICEF, IOM, UN-Habitat and FAO), with technical guidance from the Durable Solutions Working Group in Sudan (DSWG). IOM collected the survey data and the Sudanese Development Initiative (SUDIA) undertook the qualitative area-level data collection.

The large-scale sample-based survey was combined with extensive in-depth qualitative data and community consultation sessions. These were led by the Sudanese Development Initiative (SUDIA), a Darfur-based non-governmental organization, with support from UNHCR and funding from UNDP. SUDIA’s well-established expertise in conflict and displacement dynamics in the region, paired with a solid network of local authorities and stakeholders, were key in this context. Piloted in Tawila, the approach was then rolled out to the remaining seven localities and all population groups in focus of the PBF exercise. Consultations were pivotal not only to ground-truth the analysis, but also to increase transparency on how the data collected was going to be utilised, and to enable communities’ co-ownership of their locality action plan.
IDP identification and classificatory questions

108. IDP statistics can be separated into two distinct population groups: IDPs themselves, and IDP-related populations that are potentially of interest. Identification questions may focus just on IDPs or also cover some or all IDP-related populations.

- IDPs - Persons who have displacement-related protection needs:
  - IDPs in locations of displacement
  - IDPs in location of return
  - IDPs in other settlement locations;
- IDP-related populations:
  - Children born after displacement to at least one IDP parent;
  - Other non-displaced family members of IDPs;
  - Those who have overcome key displacement-related vulnerabilities;
(See IRIS Figure 3.1 page 28).

109. IRIS recommends adding a set of screening questions to national household surveys similar to those recommended for population census (IRIS paragraph 241) where the sample design is suitable for obtaining reliable estimates. It also advises that camps, reception centres, informal settlements and other institutional accommodation should be included in the sampling frame.

110. While self-identification of IDPs is relatively widespread in current national survey practice, IRIS acknowledges self-identification / self-declaration merely as an exceptional last resort¹. This is largely because of the risk of over or under reporting, due to differences in how individuals interpret or define the concept of forced displacement and the potential for individuals to misreport in cases where being identified as an IDP is linked to assistance provision, or where it evokes fear of social stigma, discrimination or persecution. It should therefore not be considered best practice, and – if at all – be used only to complement a more objective measurement practice.

111. The IDP identification question set recommended for censuses (Use Case A, paragraph 37) constitutes a "bare minimum" approach aimed primarily at census contexts where the potential for expanding the questionnaire is minimal. While this set attempts to capture all of the necessary criteria for being an IDP, it does not cover the full extent of complexity: for example, the place of usual residence before displacement may not be properly established and no information is collected on the causing event or multiple displacements, so it is not possible to accurately identify the 3 sub-populations of IDPs. The census questions also do not allow identification of IDP-related populations, i.e. the children of at least one IDP parent, where these children do not live in the same household as their IDP parent.

¹ See IRIS, paragraph 107.
112. EGRISS has published a methodological paper, *Towards a standardized approach to identify IDPs, refugees and related populations in household surveys* that expands on the IRIS recommendation and suggests survey question flows for use in household surveys. The paper discusses the questions that need to be included in order to address each element of identification of IDPs and related populations; the limitations and potential sources of inclusion and exclusion error; and the practical considerations around administering the questions within a survey. It also provides question flows for inclusion in MICS and DHS.

**Capturing indicators of progress towards durable solutions, and priority SDG indicators**

113. Beyond basic socio-demographic variables (age, sex) and the identification and classificatory questions discussed in the previous section, IRIS also proposes a composite measure for assessing if IDPs have overcome their key displacement-related vulnerabilities. The *Criteria and sub-criteria included in the Composite Measure of Overcoming Displacement-Related Vulnerabilities* is explained in the IRIS Chapter 4 and is illustrated in Table C.1 slightly expanded set of criteria and sub-criteria is also recommended in IRIS to comprehensively measure progress towards solutions.

<table>
<thead>
<tr>
<th>Table C.1: Criteria and sub-criteria included in the composite measure of overcoming displacement-related vulnerabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>1. Safety and security</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2. Adequate standard of living</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>3. Access to livelihoods</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>4. Restoration of housing, land and property</td>
</tr>
<tr>
<td>5. Access to documentation</td>
</tr>
</tbody>
</table>

Source: IRIS, 2018 Table 4.3

114. There is currently no agreed list of indicators for measuring the sub-criteria, but EGRISS is developing a prioritised list of relevant indicators that can be obtained through surveys.

---

2 EGRISS Methodological Paper Series (2023) *Towards a standardized approach to identify IDPs, refugees and related populations in household surveys*
115. In the meantime, for each sub-criterion, there are many different indicators which can be chosen. The indicators selected in each national context should, as far as possible, be aligned with already tested and standardized indicators. The Durable Solutions Indicator Library provides a thorough overview of potential options (see IRIS, 2020 paragraphs 127-128). Indicators selected should be carefully chosen to ensure they are relevant for all three sub-stocks of the displaced population: IDPs in locations of displacement, IDPs in locations of return and IDPs in other settlement locations. For a full discussion of this topic refer to IRIS, 2020 paragraphs 135 – 137.

116. In addition to the list of indicators being developed by EGRiSS, the IAEG-SDGs in collaboration with EGRiSS has highlighted 12 SDG indicators that are of particular relevance to displaced people. For survey practitioners, this implies that these indicators should be captured in any survey including IDPs (whether this is a national survey that includes IDPs or an IDP-specific survey) if applicable and feasible within the scope of the planned survey. EGRiSS does not itself promulgate definitions of these SDG indicators but rather refers to the relevant SDG metadata sheets and statistical standards for definitions and questionnaire specifications. Table C.2 provides references to such relevant international standards and specifications. Measurement of these SDG indicators for IDPs should be aligned wherever possible with the national statistical practice for capturing these indicators among the wider national population, to ensure comparability.

117. To make the assessment of the progress of the IDP and IDP-related population a comparator population is needed, this may be the general population or sub-national population in the vicinity of the initial displacement who have not themselves suffered displacement. These variables should also be available in the comparator population where relevant to enable comparisons to be made.

118. It is also important to consider collecting GIS data where possible, to enable locational analysis.

Table C.2: SDG priority indicators on forced displacement

<table>
<thead>
<tr>
<th>SDG Priority indicator (with links to metadata)</th>
<th>Details</th>
<th>Relevant references</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1 Prevalence of stunting among children under 5 years of age</td>
<td>Measure height of children of age 0-5 years; stunting indicated by height-for-age &lt; -2 standard deviation from the median of the WHO Child Growth Standards</td>
<td>SDG indicator metadata sheet. In terms of survey questionnaire design, only fields for age/DOD and height are required. However, it must be noted that height measurement is non-trivial. See guidance on anthropometric data collection from DHS (Link), MICS (Link) and the FANTA Project (Link).</td>
</tr>
<tr>
<td>3.1.2 Proportion of births attended by skilled health personnel</td>
<td>Measure attendance of births by skilled health personnel (e.g. doctors, nurses or midwives) during delivery</td>
<td>SDG indicator metadata sheet</td>
</tr>
</tbody>
</table>
### 6.1.1 Proportion of population using safely managed drinking water services

Measure improved basic drinking water source which is located on premises, available when needed and free of faecal and priority chemical contamination; where

*Improved drinking water sources* mean piped water into dwelling, yard or plot; public taps or standpipes; boreholes or tubewells; protected dug wells; protected springs; packaged water; delivered water and rainwater

*Located on premises* means if the point of collection is within the dwelling, yard, or plot

*Available when needed* means households are able to access sufficient quantities of water when needed

*Free from faecal and priority chemical contamination* means water complies with relevant national or local standards (or follow the WHO Guidelines for Drinking Water Quality)

[SDG indicator metadata sheet](#). For survey questionnaire specification, see table 1 and “core questions” W1-W5 of the JMP recommendations ([Link](#)).

### 11.1.1 Proportion of urban population living in slums, informal settlements or inadequate housing

Measure the urban population living in slums or informal settlements that are defined by a lack of at least one of the following:

*Access to improved water source* (see definition under indicator 6.1.1)

*Access to improved sanitation facilities* (access flush/pour-flush toilets or latrines connected to a sewer, septic tank or pit; ventilated improved pit latrine, pit latrine with a slab or platform and composting toilets/latrines)

*Sufficient living area* (not more than three people share the same habitable room)

*Housing durability* (permanent building materials for walls, roofs, and floor; not in need of major repair or in a dilapidated state; complies with local building standards; location is not near toxic waste, in a flood plain, on a steep slope, or close to rail, highway, airport or power lines)

*Security of tenure* (evidence of documentation to prove secure tenure status, de facto or perceived protection from forced evictions)

Measure inadequate housing, where inadequate housing is defined as housing that is not affordable to the household (i.e. the net monthly expenditure on its cost exceeds 30% of the total monthly income of the household)

[SDG indicator metadata sheet](#)
<table>
<thead>
<tr>
<th>1.2.1 Proportion of population living below the national poverty line, by sex and age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute a comprehensive estimate of total household consumption or income (with consumption being the preferred welfare indicator) and construct a correctly weighted aggregate of consumption or income per person. Compare this aggregate with the national poverty line. To measure consumption, cover questions on food and non-food expenditure as well as food consumed from own production (either through recall questions using lists of consumption items or through diaries).</td>
</tr>
</tbody>
</table>

**SDG indicator metadata sheet.** Poverty measurement is a complex undertaking. No internationally standardized questionnaire module for capturing consumption expenditure exists (since food and non-food items consumed vary from country to country). When comparing consumption aggregates against a national poverty line, it is crucial to capture consumption in a way that aligns with the primary poverty survey instrument used in the country from which the poverty line was derived. This paper documents best practices around survey-based consumption and poverty measurement.

<table>
<thead>
<tr>
<th>4.1.1 Proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure the minimum proficiency level in reading and mathematics for a. Grade 2/3 b. End of primary education c. End of lower secondary education Proficiency levels are established via the Minimum Proficiency Levels (MPLs)</td>
</tr>
</tbody>
</table>

**SDG indicator metadata sheet.** Collecting data for this indicator requires detailed learning assessments at the three levels specified in the indicator, which in most countries are administered through participation in various national, regional and/or international school-based assessments (as opposed to household-based surveys, which often will not sample sufficient numbers of children at a required age and education level). Hence, inclusion of forcibly displaced children in these learning assessments is the preferred data source for this indicator. Countries wishing to nevertheless include a learning assessment in a household survey covering forcibly displaced children should align the assessment method in the survey with the relevant school-based assessments that a country carries out. If not applicable, the MICS (Link, see Foundational Learning: module) and EGRA EGMA may also provide useful survey-based approaches relevant to the lower schooling levels.

<table>
<thead>
<tr>
<th>7.1.1 Proportion of the population with access to electricity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure access to electricity where access is only considered if the primary source of lighting is the local electricity provider, solar systems, mini-grids and stand-alone systems. Sources such as generators, candles, batteries, etc., are not considered due to their limited working capacities and their function as backup sources for lighting.</td>
</tr>
</tbody>
</table>

**SDG indicator metadata sheet.** A single survey question on primary source of lighting is sufficient to capture this indicator. See question HL2 in the broader set of energy-related questions recommended by WHO and the Word Bank (Link).
### 8.3.1 Proportion of informal employment in total employment, by sex

Measure all people of working age who, during a short reference period (one week) are engaged in any activity to produce goods or provide services for pay/profit.

Measure informal employment which refers to people that in their main or secondary job were in one of the following categories:

- a. Own-account workers, employers, members of producers’ cooperatives employed in their own informal sector enterprise
- b. Own-account workers engaged in goods production for their own final use
- c. Contributing family workers (no explicit written contract of employment, not subject to labour legislation or social security)
- d. Employees holding informal jobs

[SDG indicator metadata sheet](#) See ILO’s model questionnaires ([Link](#)) for labour force surveys which include the required questions for capturing informality, as well as UNHCR’s guidance to labour measurement in forced displacement contexts ([Link](#)) for a more specific discussion.

### 8.5.2 Unemployment rate, by sex, age and persons with disabilities

Measure unemployed people, which are defined as those of working age who were not in employment (in a short reference period), carried out activities to seek employment and were currently available to take up employment given a job opportunity.

[SDG indicator metadata sheet](#) See ILO’s model questionnaires ([Link](#)) for labour force surveys which include the required questions for capturing unemployment, as well as UNHCR’s guidance to labour measurement in forced displacement contexts ([Link](#)) for a more specific discussion.

### 1.4.2 Proportion of total adult population with secure tenure rights to land, with legally recognized documentation and who perceive their rights to land as secure, by sex and by type of tenure

Measure secure tenure rights to all types of land use (residential, commercial, agricultural, forestry, grazing, wetlands), where secure tenure rights are compromised of two sub-components:

- a. Legally recognized documentation (the landholder reports having the right to bequeath the land)
- b. Perceived security of tenure (the landholder does not report fear of involuntary loss of land through disputed land ownership, inability to use land, fear of eviction or dispossession)

[SDG indicator metadata sheet](#)

### 16.14 Proportion of the population that feel safe walking alone around the area they live

Measure the feeling of fear of crime in a context outside the house; suggested formulation is "How safe do you feel walking alone in your area/neighbourhood? (Very safe/fairly safe/bit unsafe/very unsafe/ I never walk alone/don’t know). Feeling safe refers to those respondents who feel "very safe" and "fairly safe".

[SDG indicator metadata sheet](#) Survey questionnaire specification provided directly in the “methodology” section of the metadata sheet.
16.9.1 Proportion of children under 5 years of age whose births have been registered with a civil authority, by age

Capture whether the birth of children under 5 has been registered with a civil authority, where birth registration is understood as a universal recording of the birth, a birth certificate is issued as a record that documents the circumstances of the birth and a civil authority is an officially authorized body for registration of births.

SDG indicator metadata sheet

119. Examples of questionnaires from other IDP surveys are available in the UNHCR microdata library3.

Sampling considerations when surveying IDPs

120. In theory, sampling in forced displacement contexts is no different from sampling elsewhere, and the same general principles and methods apply4. However, in practice the selection from and application of these methods is often faced with specific challenges when it comes to IDPs (IRIS paras 226-231).

121. Forcibly displaced persons may be hard to reach, either because they are difficult to identify, not willing to be identified, on the move or live in areas difficult to access. Moreover, displaced people often constitute a small proportion of the wider population and are not evenly distributed around a country or region – making it difficult or impossible to rely on sampling frames which don’t allow for the direct identification of forcibly displaced households, or at least of small geographical areas where they are predominantly located. IDPs can live either in dedicated camps or they can live in houses among the general population, either as part of another household or as an independent household.

Identifying a sampling frame

122. The first question that guides the decision on sampling from IDP populations is how one can construct a frame from which to sample IDPs in a targeted manner. In the simplest case, this can either be a list of the wider population which allows identifying forcibly displaced, or a list solely describing the IDP population. Such a list can often come from administrative records, either from relevant government agencies or operational data from international agencies such as UNHCR or the International Organisation for Migration. In order to gain access to such sources it is likely that a data sharing agreement will be required, which can take some time to establish and should set out how data privacy will be protected - for example, by basing sample selection on anonymised data, so that only the personal details of those invited to take part in the survey are shared.

3 UNHCR Microdata Library
4 Useful reading on survey sampling in general can be found in UNSD’s “Designing Household Survey Samples: Practical Guidelines”, “Household Sample Surveys in Developing and Transition Countries”, and “Sampling Rare and Elusive Population”.
CASE STUDY: PROFILING OF INTERNAL DISPLACEMENT DUE TO VIOLENCE IN EL SALVADOR

The Government of El Salvador, with the technical leadership of the General Directorate of Statistics and Census and the Latin American Faculty of Social Sciences, and the technical support of UNHCR and JIPS, conducted a profiling of IDPs between 2016 and 2018. The objectives of the exercise were to identify the number of families displaced by violence in El Salvador in recent years, and to better understand the living conditions and specific needs faced by the victims of displacement.

There was no register of internally displaced persons or households available and it was not possible to complete a full enumeration of the whole country, so a sampling approach was needed that would allow estimates of the entire IDP population on a country level and also ensure that the sample was representative of the target population.

Geographic areas that were receiving flows of IDPs were identified through a data review and mapping exercise using interviews and workshops with key informants. This produced a list of 202 priority study areas containing a total of 648 census segments. For the displaced target population, a stratified random cluster sampling approach was chosen. For the non-displaced target population, a two-stage sampling approach was chosen. The process of selecting the two samples can be summed up in the following way:

**Stratification** → **Disproportionate selection of clusters** → **Enumeration of sampled clusters** → **Inclusion of all identified IDP households in final IDP sample** → **Selection of households for inclusion in non-IDP sample**


123. If one does not have an existing list of possible respondents, it needs to be created. The following methods are commonly used:

a. If the locations in which IDPs reside are known, or can be determined, and the National Statistics Office maintains a list of Primary Sampling Units (PSUs) for the country, the two can be used to identify those PSUs where IDPs reside. The first stage of sampling can then include either only those PSUs, or over-sampling from those PSUs. Relevant PSUs are selected randomly or systematically, and then a household listing is carried out or updated in the selected PSUs, capturing household’s IDP status.
CASE STUDY: SAMPLING FOR THE 2018 MULTIPLE INDICATOR CLUSTER SURVEY IN GEORGIA

The 2018 Georgia Multiple Indicator Cluster Survey (MICS) was carried out by the National Statistics Office of Georgia in collaboration with United Nations Children’s Fund (UNICEF) and National Center for disease Control and Public Health (NCDC), as part of the Global MICS Programme. The primary objective of the sample design was to produce statistically reliable estimates of most indicators, at the national level, for urban and rural areas, and for the ten regions of the country and country-level estimates for the IDP population.

A multi-stage, stratified cluster sampling approach was used for the selection of the survey sample. The sampling frame was based on the 2014 General Population Census of Georgia. The primary sampling units (PSUs) selected at the first stage were the enumeration areas (EAs) defined for the census enumeration. A listing of households was conducted in each sample EA, and a sample of households was selected at the second stage.

The urban and rural areas in each of the ten regions were defined as the sampling strata. Each major stratum (Region by Urban/Rural) was divided into IDP and Non-IDP sub-strata. Since some of the PSUs have no IDPs, or the proportion of IDPs was very low, only the PSUs where the IDP population is more than 48 percent of the PSU population were included in the IDP strata.


b. A similar approach can be applied if the locations of IDPs are unknown, by drawing a random selection of PSUs and undertaking a listing exercise, using a short set of questions to elicit IDP status. In those PSUs with higher concentrations of IDPs – above a pre-determined threshold such as 10% - the household listing is extended to the neighbouring PSUs (or one nearest neighbour PSU). This process continues extending until PSUs no longer cross the threshold. In this way, PSUs with the highest concentration of IDPs are identified and the listing can be used as the sampling frame. This adaptive cluster sampling approach has been used successfully for refugees in urban areas with high concentrations of refugees and should be applicable to IDPs.
c. If a list of PSUs of reasonable quality does not exist, alternative ways to identify smaller geographical areas to sample from exist. Satellite, aerial or drone photos of areas in which IDPs are known to reside can be used to divide geographical areas into sampling units to be sampled and the households existing there listed to be sampled in a second stage.

d. If IDP locations are (partially) unknown, Mobile Phone Tracking may be an option. Digital trace data generated by IDP’s mobile phones can help identify the areas in which IDPs are located and thus where to list/sample. This is not the same as identifying IDP’s phones and then sampling them for a phone survey (see section (e) for comparison). Approaches such as these naturally require careful consideration around data privacy. Further discussion of non-traditional data sources is included in Use Case F.

e. Moving away from area-based sampling approaches, it may be possible to undertake a phone survey, assembling a list frame of phone numbers of the IDP population, by Mobile Phone Tracking or via Random Digit Dialling. If launching a face-to-face IDP survey it may also be worth collecting phone numbers and consent for future phone-based panel surveys at the same time.

124. Sometimes a combination of different approaches is required. For example, if expanding a national survey to include IDPs living in camps and outside camps, it is possible that PSUs will be available for out-of-camp populations, but that the PSU frame exclude the camps. In this case, it may be necessary to use registration lists of IDPs in the camps as a complementary sampling frame.

Drawing a sample

125. Constructing a sampling frame of IDPs allows a survey to apply probability-based sampling methods. This ensures that every unit in the population has a known, non-zero chance of being selected into the sample, and its probability of selection can be accurately determined. This makes it possible to produce unbiased estimates of population totals, by weighting sampled units according to their probability of selection. It also allows for estimates of uncertainty, like confidence intervals. Probability-based sampling is considered best practice when sampling for surveys and ought to be the first choice - whenever it is possible to construct a sampling frame, i.e. a list of units to draw a sample from.

126. Including IDPs in a general household survey is likely to require a boost to the sample ratio for IDPs to ensure that a sufficient number are reached through the survey – because IDPs tend to be a small proportion of the overall population. The size of a boost will depend on the analysis requirements – for example the disaggregations and level of precision that are expected.

---

The usual risks of sample under-coverage resulting from phone ownership and network connectivity rates among the target population need to be taken into account.
CASE STUDY: BOOSTING THE HARMONIZED HOUSEHOLD LIVING CONDITIONS SURVEY IN CENTRAL AFRICAN REPUBLIC

The National Statistical Institute of CAR carries out the Harmonized Household Living Conditions Survey (Enquête Harmonisé sur les Conditions de Vie des Ménages). The 2021 survey will cover a total of 6,000 households nationally. Nearly 15% of the total population in CAR are IDPs (an estimated 702,000 as of February 2020) but detailed data on this vulnerable population is lacking. Alongside the IDPs identified through the regular sampling strategy, the sample size in IDP camps will be increased by 600 households. The expanded sample in IDP camps will allow more detailed analysis of the welfare and living conditions of IDPs in both IDP camps and in the general population.


127. If sampling from a highly concentrated population such as IDPs living in camps, it is generally not necessary to consider clustering the sample. Clustering is a technique that one uses to bring down sample dispersion and make travel during fieldwork manageable, but it comes at the cost of losing precision and driving up design effects. When sampling from a highly concentrated population there is no real cost saving from clustering but there is still a loss of precision, so it is preferable to obtain a registration list for the camp (if it exists and is sufficiently up to date) and use it as a complementary sampling frame from which to draw a systematic and unclustered selection of dwellings.

128. If carrying out a dedicated survey of IDPs, it is also important to identify an appropriate “host community” to include in the sample as a comparator population. There is no standard definition of host community, so at present this will need to be locally determined. Examples include the wider population that lives within a certain radius of an IDP camp, or households within the same administrative area.

129. It should be noted that responses to questions on attitudes, intentions to move, and decision-making power in the household may differ substantially by gender. When feasible and appropriate, using a random selection protocol of female and male respondents within a sampled household, or interviewing more than one member of each household, are ways to ensure more reliable data.
Non-probability sampling methods

130. When constructing a sampling frame is not feasible, reverting to non-probability-based sampling is sometimes considered. Commonly used non-probability sampling techniques are respondent driven sampling (RDS) or snowball sampling. If IDPs are difficult to find, either because they are few in numbers, are hiding, are homeless, on the move or live in an area difficult to access, such methods may be considered as an option of last resort. They should however be used with caution, as they do not allow for statistically representative, unbiased inference from the sample to the population. Application of non-probability sampling methods in official statistics production is uncommon.

CASE STUDY: NON-PROBABILISTIC APPROACH IN IDP PROFILING SURVEY IN THE STATE OF CHIHUAHUA, MEXICO

Residents in the state of Chihuahua in northern Mexico have suffered from violence-related displacement due to organised crime in the state, with populations displaced to cities or border regions. In 2022, the National Institute of Statistics and Geography (INEGI) worked with the Joint IDP Profiling Service and UNHCR to plan a survey to understand more about the displaced populations. This is a population that is difficult to locate for probabilistic survey purposes, so a non-probabilistic approach was identified as the most appropriate sampling strategy.

INEGI identified three routes to reach IDPs in Chihuahua.

- An existing crime survey had asked about the measures respondents had taken to protect themselves from crimes, including a response option for relocation. Those respondents were targeted for the IDP survey.
- The Executive Commission for Attention to Victims of the State of Chihuahua (CEAVE) provides shelters and facilities for victims of crime. Interviewers visited CEAVE locations to conduct interviews.
- A snowball approach was attempted to identify other potential displaced people; however, this was not as successful as expected due to the sensitivity of the topic (the underlying reason for displacement being organised crime).

In total, 1,255 interviews were completed out of a total of 1,388 planned, approximately 70% at CEAVE locations and 30% from the crime survey. By focusing on the population that received attention from CEAVE, there is a bias in the sample frame (because it is not representative of all displaced persons) which needs to be taken into consideration during the analysis of the results.
Limitations of registration systems as sampling frames

131. A frequent challenge with sampling frames derived from IDP registration systems is that the registers are only focused on a sub-set of IDPs receiving assistance; or can be incomplete, for example if areas of the country are inaccessible to registration systems, or the address/contact information can be outdated.

Box C.1

Guidance material on sampling in forced displacement contexts

Two dedicated resources on sampling in forced displacement contexts have recently been published and may be considered for further details and examples:

- Stephanie Eckman and Kristen Himelein (forthcoming): Innovative Sample Designs for Studies of Refugees and Internally Displaced Persons

Fieldwork planning

132. Before any survey fieldwork commences, it is advisable to conduct one or more field visits to fully understand and plan for the locations and practicalities involved in reaching IDPs. This can be important in overcoming the additional challenges that surveying IDPs can entail. Ideally one or more of the most challenging locations / scenarios should be visited, with the following objectives:

- Observe the situation on the ground;
- Test whether the sampled households can in practice be located and accessed and make contingency plans if not, for example through engaging local community leaders, local authorities or landlords for access;
- Build relations between the different agencies and communities involved in implementing the survey – trust and rapport can be essential to success of the survey;
- Identify any language barriers and how best to address them (questionnaire translations, availability of interpreters);
- Test the questionnaire, both in terms of content and flow. Do the questions lead to correct identification and classification of refugees; are there any issues with translations; are there any sections of the questionnaire that are seen as more sensitive and potentially might result in non-response.

133. There may be additional costs for extended duration of fieldwork compared with a standard survey, reflecting the challenges of identifying IDPs, and potentially greater use of supervisors to train, monitor and support enumerators.
Surveys may involve a communication plan at a national or local level, to ensure that residents are aware the survey is taking place and primed to take part at the correct time. It is important to ensure that displaced people are considered as part of the communication plan, if good coverage is to be achieved. There are no fixed prescriptions that will work everywhere, but experience from various countries should be observed and evaluated. The main messages to convey include:

- That the survey will include displaced people, refugees and IDPs and that their responses are important.
- How the data will be used (positive messages) and clarity that it is not linked to allocating assistance.
- That confidentiality and data security are assured.
- Any special arrangements that have been made to include displaced people (such as availability of interpreters).
- What opportunities there will be for respondents to hear about the survey results following data collection. Depending on the scale of the survey, it may be appropriate to offer some dissemination aimed specifically at the target population.

The communication plan may also consider raising awareness of any existing referral mechanisms available to the community, for feedback or protection concerns, according to the principles of “do no harm”.

**CASE STUDY: PILOT VISIT DURING IDP PROFILING SURVEY IN THE STATE OF CHIHUAHUA, MEXICO**

As part of its 2022 survey of IDPs in Chihuahua, the National Institute of Statistics and Geography (INEGI) planned to conduct interviews at the office locations of the Executive Commission for Attention to Victims of the State of Chihuahua (CEAVE), which provides shelters and facilities for victims of crime. Residents in Chihuahua have suffered from violence-related displacement due to organised crime in the state, with populations displaced to cities or border regions.

To ensure the success of this element of the survey, INEGI organised a pilot visit aiming to address the following key issues:

- Logistics of interviewing at CEAVE offices – finding a suitable location (privacy concerns)
- Communicating the importance of the survey, to engage CEAVE officials and potential respondents.
- Testing the questionnaire through 20-25 test interviews

As a result of the pilot visit, enumerators were provided with additional training to improve their success in obtaining complete interviews, given the challenging topics covered. In total, the IDP survey achieved 1,255 complete interviews, of which 70% were at CEAVE locations.
To reach displaced people, communication at local and community level is often the most effective. Examples of communication strategies include:

- Leaflets or visits to relevant institutions (refugee camps, etc) in the lead up to Census, with interpreters available if language is likely to be a barrier.
- Identification of and engagement with community leaders before the survey, to establish the most effective way of reaching displaced populations and to gain their buy-in and support.
- Acknowledge the potentially sensitive nature of the question topics for displaced populations and explain the purpose of collecting the data and principles of data protection.

CASE STUDY: COMMUNITY ENGAGEMENT IN DATA PROCESSES IN SYRIA

The Urban Analysis Network Syria (UrbAN-S) project, implemented in 2018-2019, produced area-based urban profiles that provided a snapshot of the physical and social conditions from a conflict and displacement context. The project was funded by the European Union and implemented through a collaborative approach, led by iMMAP with JIPS, Mercy Corps’ Humanitarian Action Team, and the European Commission Joint Research Centre.

Due to restrictions on international organisations’ work in Syria in 2018-2019, it was difficult to access communities to conduct sample-based household surveys. The project therefore used an alternative approach that relied on data collection from ‘community focal points’. Transparent criteria were defined to identify which community members would be included in the data collection, based on chain referral (snow-balling) and assessing the individual’s connections, reputation, influence, motives and potential biases, length of time living in the area and aspects of their understanding of the situation of the wider community and geographical location.

The engagement with community members during data collection and the efforts put in to identify representative members to speak on behalf of the wider communities made space for a two-way information flow between the research teams and the communities.

## Enumerator training

136. Displaced people are often harder to reach and can need support in order to take part in a survey and this requires specific content in the enumerator training course. Allowing sufficient time to address these issues in the enumerator training is key to delivering a successful survey. It may take between one and two days to cover the appropriate material.

137. Topics to cover in enumerator training include:

- The definitions and concepts used in statistics about displacement, which may be unfamiliar to many experienced field staff. For example, IDP concepts such as habitual residence, usual residence, migration concepts, refugee categories, meaning of displacement, meaning of refugee and IDP, and reasons for displacement and migration.
- Accurate recording of all members of the household, including temporary members who may be displaced. This is important as displaced people may fail the interviewers standard understanding of ‘usual members of the household’, leading to them being inaccurately treated as visitors and excluded from household rosters.
- The sensitivities around interviewing IDPs and the data protection implications. When conducting surveys, the displaced population has special needs or conditions of interview, and the ‘no-harm’ principles and UN ethical standards on data collection and data protection must be applied. Enumerators must be aware of the existing referral mechanisms for community feedback or protection concerns, should any issues present themselves during the course of the data collection.
- Building trust and rapport, and skills in interviewing vulnerable people, such as children and disabled people. People who have been displaced have suffered trauma and any interview may ask people to relive painful experiences.

Strategies for managing personal well-being for the enumerators, if they are likely to be exposed to traumatic testimonies during the data collection – which can become emotionally exhausting.
On the whole, provided fieldwork planning and enumerator training has taken full account of the need to cover displaced people, there are few issues specific to IDPs in terms of the actual fieldwork. However, it is worth noting that effective feedback from enumerators during any survey can be crucial in identifying any emerging issues in reaching IDPs, which might be addressed by continued or improved communication and community engagement.

If budget allows, it is advisable to undertake more monitoring of fieldwork than for a standard survey, reflecting the challenges associated with identifying IDPs and potential issues overcoming sensitivities around data collection. Increased availability of supervisors can help to monitor the quality and consistency of data collection and play a facilitating role, communicating the purpose of the survey with more authority, to gain trust and avoid non-response.

There are few data quality issues that are specific to surveys including IDPs, and so data quality assurance should follow standard processes, such as monitoring response rates and monitoring the completeness and distributions of key indicators. The only specific concern of note is the risk that survey respondents may believe that their responses will influence the provision of humanitarian aid and will thus misreport consumption in an attempt to influence its distribution.
141. Often it is possible to triangulate survey data with data from other sources, such as administrative data from other government agencies or operational data produced during the regular activities of humanitarian organisations, to make sure that high level trends match between the various available datasets without any major deviation. For example, if the variation between two stocks is not significant, but administrative data identifies a high number of smaller, shorter term movements (e.g. in the case of preventative evacuations leading to quick returns, or mass movements leading to cross border returns) then this should help to verify or correct official statistics on flows or alert the need for better or more frequent data collection. Similarly, through comparing with cross border data, if movements lead to significant cross-border flight or returns, the datasets on both sides of the border should correspond with each other to ensure validity. Further discussion of administrative data, operational data and non-traditional data sources is included in Use Cases D, E and F respectively.

142. Even if variations in operational flow data from incidents do not match variations in stocks, the differences between the different stock types should be proportional. However, an increase in the number of displacements does not mean that a similar increase will appear in stocks, as those displacements could relate to the same people being displaced on multiple occasions. For example, if the numbers of IDPs in an area decreases, the numbers corresponding to that decrease should be reflected in another stock, e.g. there are 10,000 less IDPs in Borno State in Nigeria, because there are 10,000 more Nigerian refugees in Cameroon.

### Data protection / confidentiality / disclosure

143. Due consideration should be given to privacy in accordance with national and international laws. The UN Fundamental Principles of Official Statistics states that "individual data collected by statistical agencies for statistical compilation ... are to be strictly confidential and used exclusively for statistical purposes". Disclosure should also include considerations of interoperability: how datasets on different population types talk to each other, their ability to share information on those populations, or even on specific individuals, their characteristics and their needs.

### Analysis and dissemination

144. The process of analysing and presenting the survey data collected on IDPs is no different from that for any other data and should be planned to meet the needs of users. This section briefly describes the outputs that are commonly produced to meet users’ needs, with references for further information if required.

145. Data collected on identification and classification of IDPs is used to produce analysis of stocks and flows for the different categories of IDPs and IDP related populations, presented as counts and proportions cross-tabulated by sex and age group, details are in IRIS paras 114 – 125.
146. Some analysis will be required to identify the stocks and flows of those who have overcome key displacement-related vulnerabilities, in particular to derive the variable in the databases which enables the analyst to identify those who have overcome those vulnerabilities, and to remove them statistically from stock. It should be noted that the full methodology for determining IDPs to be removed from stock is in its early stages. A more developed methodology will be added to future editions of this manual.

Stocks

147. Analysis of stocks is recommended for all categories of IDPs and IDP-related populations where possible.

a. Counts of each type of IDP and IDP related populations, by sex and age, or age group, place of usual residence, place of habitual residence, and years since initial displacement. If age group is used, then it is important to distinguish children from adults aged 18 years and over and analysis should include the total number of IDP-related unaccompanied and separated children under 18, by sex and age. For some purposes it may also be helpful to identify those under 16 years of age.

b. Proportions of IDPs by the length of time the IDP was displaced; the location of the IDPs and if they have returned to their habitual places of residence, been resettled or are in locations of displacement; and proportions of women and children among the IDPs. IRIS elaborates on these, drawing out a recommended set of indicators – see Box C.2.

Box C.2

Basic indicators for stock of IDPs (from IRIS, 2020)

i. Percentage of the total population who are IDPs.*#
ii. Proportion of women among all IDPs.* #
iii. Proportion of children among all IDPs.
iv. Proportion of IDPs who were first displaced in the last 5 years.
v. Proportion of IDPs who were first displaced in the last 10 years.
vi. Proportion of IDPs who were first displaced over 10 years ago.
vii. Proportion of IDPs who were displaced more than once, whose first displacement was during the last 5 years.
viii. Proportion of IDPs who have returned to their habitual place of residence and have not yet overcome key displacement-related vulnerabilities (IDPs in locations of return)#
ix. Proportion of IDPs who have settled elsewhere in the country and have not yet overcome key displacement-related vulnerabilities (IDPs in other settlement locations)#
x. Proportion of IDPs in locations of displacement and who have not yet overcome key displacement-related vulnerabilities (IDPs in locations of displacement).#

Key: # suitable for IDP-related stocks; * suitable for stocks of those who have overcome key displacement related vulnerabilities

Flows

148. Flows are inherently more difficult to collect through surveys and in practice the recommended indicators for in-flows are likely to need to come from administrative data, while measuring the outflows (also included in Box C.3) is a demanding statistical exercise for which the advice is still a work in progress.
Where analysis of flows is possible, it is recommended for all categories of IDPs and IDP-related populations, with in-flows and out-flows cross-tabulated by sex, age, reason for displacement, and place of usual / habitual residence. The full list of recommended tabulations from IRIS is replicated in Box C.3.

**Box C.3**

**Basic IDP flow tabulations (from IRIS, 2020)**

**Basic inflow statistics for IDPs**

a. Total number of IDPs in a country who were forcibly displaced for the first time, during a specified period of time, by sex and age

b. Total number of persons in a country who were forcibly displaced for the first time, during a specified period of time, by sex and reason for displacement

c. Total number of persons in a country who were forcibly displaced for the first time, during a specified period of time, by sex and by current place of current usual residence (province/district)

d. Total number of persons in a country who were forcibly displaced for the first time, during a specified period of time, by sex and by place of habitual residence (province/district)

**Basic inflow statistics for IDP-related persons**

a. Total number of children born to at least one IDP-parent after the parents’ last displacement, during a specified period of time, by sex and age

b. Total number of children born to at least one IDP-parent after the parents’ last displacement, during a specified period of time, by sex and parents’ place of habitual residence (province/district)

c. Total number of children born to at least one IDP-parent after the parents’ last displacement, during a specified period of time, by sex and current place of usual residence (province/district)

**Basic outflow statistics for IDPs**

a. Total number of IDPs who have died or emigrated during a specified period of time, by sex and age.

b. Total number of IDPs who have overcome all key displacement vulnerabilities during a specified period of time, by sex, age and by current place of usual residence.

c. Total number of IDPs who have overcome all key displacement-related vulnerabilities during a specified period of time by habitual place of residence and current place of usual residence and main reason for initial displacement.

**Basic outflow statistics for IDP-related persons**

a. Total number of children born to at least one IDP-parent after the parents’ last displacement, who have died or emigrated during a specified period of time, by sex and age.

b. Total number of children born to at least one IDP-parent after the parents’ last displacement, who have overcome all key displacement vulnerabilities during a specified period of time, by sex, age and by current place of usual residence.

c. Total number of children born to at least one IDP-parent after the parents’ last displacement, who have overcome all key displacement-related vulnerabilities during a specified period of time by parents’ habitual place of residence and current place of usual residence and main reason for parents’ initial displacement.

Geospatial analysis is also recommended, to visualise the stocks and flows.
In 2020, a large-scale analysis covering eight localities across Darfur’s five states was initiated to provide a shared evidence-base to support peacebuilding and durable solutions under the UN Peacebuilding Fund. The key insights from the resulting locality studies have been condensed into five thematic briefs.

The best examples of data dissemination are often through online resources, which enable dynamic access to the data and can tell the data story in an engaging and interactive way. This Case Study captures one image from an online resource, to illustrate and inspire. The full details are available in the following source material: [http://dswgsudan.org/pbfdarfur](http://dswgsudan.org/pbfdarfur).
Socio-economic indicators and basic progress tabulations

151. Statistics on the progress made by IDPs in overcoming their displacement related vulnerabilities should focus on the total number of IDPs who have achieved the defined target for each of the different durable solutions criteria and related sub-criteria during a specified period of time, by current place of usual residence (district/province). This should be measured against a reference population, either the national or regional average, or the host community.

152. The two measures on progress which can be used to produce official statistics on internal displacement are:

- The durable solutions progress measure can be used to show the change in the share of IDPs who have overcome vulnerabilities linked to the criteria of the IASC Framework on Durable Solutions for IDPs over time; and
- The composite measure can be used to specify whether all key displacement-related vulnerabilities have been overcome and thus, whether or not persons can be taken out of the total IDP stock.

153. Over time, as new data are collected and more testing conducted in different national contexts, the recommended progress and composite measures can be further refined, including a list of agreed indicators and other relevant guidance. This will be included in future versions of this manual.

154. There are no further specific recommendations relevant to IDPs on how these or other data on socioeconomic indicators (including SDGs) should be analysed and presented. Typical tabulations could include cross-tabulations by category of IDP, sex and age group.

155. In planning how the survey analysis will be disseminated, if part of a wider survey it is preferable for IDP analysis to be disseminated in the main survey report as a standard disaggregation but it may also be impactful to produce a separate report focusing just on IDPs. There is no right answer here; it will depend how users’ needs can best be met in the specific national context. It is also important to consider how survey findings can be shared with the IDP communities themselves. This feedback loop is important: as well as meeting the needs of a key user community it demonstrates what the survey is for and promotes the importance of the survey, encouraging future engagement and participation.
This is an excerpt from a statistical brief that presents the results from the third round of High Frequency Phone Surveys on Internally Displaced People (IDPs) in Burkina Faso, conducted between June 28 and July 20, 2021. The survey was designed to assess the socioeconomic experience of IDP households during the Covid-19 pandemic.