

# Population estimation techniques

28 November 2023

## Key points

- Each emergency situation is different, and therefore the choice of method(s) for estimating populations must reflect the diversity of contexts and available resources. Prioritize your information needs, assess the local capacity and the context, and choose the most appropriate methodology(ies) to produce population estimates. Document how UNHCR will use the population estimates at country and regional levels
- Producing population estimates is usually a multi-agency effort. It is essential to identify and convene key actors to agree on a common approach and to review and agree on the estimates prior to their publication
- Methods for rapid estimation exercises do not usually generate population figures of good quality, because sample sizes are often small and techniques are not representative. If resources allow, develop a robust estimation plan to ensure that the data collected is interoperable with other data and can be used to support other activities, such as needs assessments
- Triangulate the available primary and secondary data; assess its quality, including its accuracy and timeliness, and the reliability of the data's source
- Validate your estimates and publish them following the CORE guidance. Population estimates must be reviewed and cleared by the Representative and, as appropriate, key stakeholders and coordination mechanisms (e.g., the regional or national information management working group), and the government

## 1. Overview

**Robust and interoperable** population estimates are essential to understand the situation, inform the planning of a response, deliver and adjust programmes, advocate for, and mobilize resources. As such, population estimates are of interest to the entire humanitarian response.

[Multi-agency coordination](#) of the population estimates is therefore essential so that they meet the needs of all organisations involved in the response effort, to create confidence in the methods used and the figures they generate, to effectively disseminate the resulting estimates, and to ensure their effective use for a variety of purposes.

The steps explained below will help operations and regional bureaux to gather, quality assure, calculate, and report population estimates that will inform the emergency response.

### **Resources and partnerships**

Population estimation is a collaborative process, led by an information management officer or equivalent function, working closely with protection colleagues, other thematic and sectoral focal points (e.g. Health, WASH, Nutrition, Education, Food/NFI distribution, and Supply) as well as the partner organisations. The Regional Bureau can support as needed, with additional support from relevant HQ entities if required.

## **2. Relevance for emergency operations**

Timely, accurate and reliable information about people displaced by a crisis is crucial for an effective and efficient emergency response. Information on the **magnitude, locations and composition** of population movements is required. This includes information about where they are and where they are from (two location aspects), who they are (profiles), how many (number), when / since when (time), and why (causes of displacement). Population estimates are **aggregate population figures drawn from the best available primary and secondary data sources**.

## **3. Main guidance**

### **Information required to produce population estimates**

Population estimates include both flow figures and stock figures:

- Flow figures describe changes in populations over time. For example, the number of individuals who have been forced to flee from specific location(s) and are arriving in transit areas during a given week.
- Stock figures: describe populations at a given moment in time. For example, the number of people at a particular point in time living in a specific village. Stock figures change via flows (e.g. arrivals and departures as well as in situ changes such, as births and deaths). The estimates should include sex and age disaggregation, which is vital to AGD-inclusive humanitarian assistance activities, notably strategic planning, programming resource mobilization, and advocacy.

The following information should be obtained to produce the population estimates:

- **Where?** Where are the key entry/exit/transit areas? Where does the displaced population come from? Where are they settling temporarily? Where are they going?

- **How many?** What is the magnitude of the population flow (i.e. the number of in/outflows by day/week; the number of forcibly displaced people in specific locations in a point in time)? This should be disaggregated to include the populations on the move through (flow) and presence (stock) in specific areas.
- **Who?** The broad age groups and sex of the displaced population. And what are their main causes of displacement? The latter is essential in mixed movements to assess if the target population is mainly refugees or a mixed movement with refugees and migrants.

To capture this information, use mixed methods at a geographical level that is sufficient to inform the response, as explained in the following steps.

Some population estimation methodologies require **personal data** to be collected and processed. You must take the necessary proactive steps to ensure that statistical [data is managed](#) in line with [UNHCR's personal data protection and privacy framework](#) (accessible to UNHCR staff only), as applicable. Critically, UNHCR does not further reuse personal data collected for statistical purposes for incompatible purposes and without an appropriate legitimate basis.

### **Step 1: Prepare by gathering baseline population data and assessing the context**

Building on the Emergency Information Management strategy and [Guidance on Emergency Preparedness](#), gather baseline population data and assess the context. The following actions will help you determine the most appropriate methodology/ies to produce population estimates:

- [Common operational datasets \(CODs\)](#): Review the available geographic data for the operation [in UNHCR's geodatabase](#) or humanitarian [CODs](#). If these are not up-to-date, compile a list of locations and names of villages, existing service facilities (e.g., hospitals, health centres) in areas that are likely to receive refugees and/or IDPs.
- *Actors and informants*: identify possible key informants, including local/national authorities (e.g. migration directorates, municipalities, civil protection / emergency institutes), drawing on active humanitarian services and partners' presence in areas likely to receive refugees and/or IDPs.
- *Historic population flow and stock data*: If not already available, identify reliable data sources on the number and composition of people moving through relevant border or transit points between the country of origin and the host country (cross-border situations), or within the country (internal displacement). Try to review at least the previous 12 months to assess typical changes in movements.
- *Contextual data*: Consult the contextual information that has been gathered to inform the emergency response (and supplement it if needed), notably to understand the local security context. This is important to understand trends and to anticipate potential access constraints in implementing certain types of estimation methodologies. If not yet gathered, obtain recent conflict, access and armed groups presence data (e.g. [Armed Conflict Location & Event Data Project \(ACLED\)](#) and [ACAPS](#) data) for the country of origin, the country of displacement (if different), and the host country (for cross-border situations).

### **Step 2: Agree "Who" should be counted by defining the target population group(s), together with partner organisations**

- The definitions of the target population(s) for the estimation should reflect the information

needs of UNHCR and the wider community of actors responding to the emergency (including the government), and align with international standards and recommendations on forcibly displaced or stateless population groups (notably the Expert Group on Refugee, IDP and Statelessness Statistics (EGRISS) guidance on international statistics on [refugees](#), [internally displaced people](#) and [stateless people](#), as well as the [Refugee Data Finder](#) and [UNHCR's Master Glossary of Terms](#)), and guidance on international migrants ([IOM key definitions](#), UNDESA).

- When elaborating the definitions of the target population groups, include the following dimensions: the population groups (e.g., refugees, IDPs, stateless – and whether the displaced population is also stateless), the temporality of the movement (e.g. what is the minimum and/or maximum length of displacement), its geographic scope (countries of origin, area of origin, country(ies) / region(s) of destination of the population groups to be estimated), and relevant AGD profiles (e.g., ethnicity, socio-economic status, etc.).

### **Step 3: Identify information gaps required to produce the population estimates**

Conduct a desk review to identify and assess the available data to determine what information is currently missing:

- Identify reliable sources that may provide the data you need. Remember to consider sources such as traditional media reports and news articles in the country of origin, country of displacement and host country. Explore the existing systems that are in place to capture information on population movements, including those implemented by external partners (e.g. border monitoring). Where possible, prioritise corporate tools (e.g. in a refugee emergency, use emergency registration if possible) or systems that are implemented by the authorities.
- Quality assure and triangulate the secondary data you compile.
- Organise the data by category, format, theme, date, source, reliability, accuracy, geographic and time coverage.

The secondary data review will determine if there is still a gap between the information you need and the information you have. If you need to collect primary data consider the methodology(ies) in the next sections.

In many cases, the available data will be insufficient, but to collect additional primary data:

- *time* will be too limited,
- relevant *resources* (human and financial) will not be available,
- the locations hosting the displaced population will not be *accessible* (e.g. due to physical constraints or the security situation) or
- it may be difficult to find a suitable methodology for specific populations (e.g. due to social profiles, fast changing patterns of movements).

These are all considerations to bear in mind if you need to collect primary data to generate the estimate. Note that in many cases, the estimate will be generated by combining the existing (secondary data) and new data that has been collected.

### **Step 4A: Collect primary data (if needed) - methods and techniques requiring access**

## **to the displaced population**

If you need to collect new data to meet your information gaps, bear in mind that this will require time, money and human resources. Taking account of the context (see the section on *considerations for a practical implementation* below), the following methodologies can be applied:

**A.1. Consultations with key actors/stakeholders:** consult the key stakeholders that were identified during the preparedness activities in step 1 to map the main displacement areas, routes, access constraints, and /or entry points at national, sub-national or local levels. The information can be gathered remotely (i.e. phone interviews, emails, online questionnaires) or in person.

**A.2. Direct observation** can be used to rapidly collect the information required to estimate populations on the move as well as those currently staying at specific locations that are in scope for the estimate. Conduct visual habitation counts and/or static crowd estimations. Drive or walk through the locality and record your observations on the location (its nature, size), population stocks, and population movements. Try to identify possible key informants, and places where enumerators could potentially conduct surveys with the target population.

**A.3. Set up an alert system:** if the context permits, regular updates can be organised to allow notification of new arrivals or new internal displacements. This requires a level of engagement of the authorities and/or humanitarian partners (or other stakeholders identified in *step 1*) to help set up the alert systems. The modalities of this alert system will depend on the context and should be defined at the outset and agreed with all members of the network. This would typically include a short questionnaire that members would update with information on the size, locations and nature of areas receiving the target population(s).

**A.4. Key informant interviews:** Design a short questionnaire to gather estimates of population flows at entry, transit/settlement and/or exit locations in the areas of interest through local key informants:

- Key informants can include district or village/residential areas authorities (in urban settings), host community leaders; service providers, humanitarian workers, religious leaders, refugee or IDP leaders and education or health staff. They can also include the key stakeholders you identified at Step 1.
- The information can be gathered by enumerators and through a range of modalities, including face-to-face interviews or by phone, email, surveys, and SMS. Estimates provided by key informants will be less accurate than by direct assessments, but they are particularly useful in situations where access to the population is limited.

**A.5. Questionnaires with the target population(s):** where resources are available, design a short questionnaire to collect data about the target population at entry, transit/settlement and exit locations in the areas of interest via enumerators. Enumerators conduct short interviews with each person/family and collect the data noted above. This method is generally more applicable in border crossing contexts, where population flows are manageable. Where resources and capacity are limited, prioritize household/group interviews over individual interviews to ensure greater data collection coverage:

- Quantitative data on population flows gathered at household or group level will provide information on the scale of the movement, coupled with:
- Sampled individual data can be used to characterize the population and understand population flow trends locally.

Note: A statistical approach may also be inappropriate for the context, for example because it would be too time consuming, or the required level of expertise to design the approach and interpret the results is not available.

**A.6. Conduct registration, enrolment or head counts:** count the population through fixing exercises, rapid registration, enrolment or head counts. See [UNHCR's guidance on registration and identity management](#).

#### **Step 4B. Collect primary data (if needed) - Alternative data collection methods**

In addition to the methodologies explained above, it can be necessary to explore other data collection methods that are presented below, with regional or HQ technical support if needed. They have limitations, involve additional processing time and require technical resources - therefore, their implementation should be assessed on a case-by-case basis. That said, they can be pragmatic and relatively cost-effective options to triangulate secondary data and/or primary data collected through the methods above. They can also be used to provide information on past, current, and potential future trends of population movements. Here are some examples of methodologies that have been implemented by operations and partner organisations when the emergency context did not permit the primary data collection methods listed above in Step 4a:

**B.1. Earth observations:** if available and relevant, use data derived from UNOSAT [satellite imagery](#), such as NASA's [night-time lights](#) imagery to indicate population movements through changes in the levels of electricity luminosity in particular areas.

**B.2. Call Data Records (CDRs):** If there is *mobile connectivity*, displacement within countries can be assessed from analysing CDRs. Work with other UN Agencies or organisations such as [Flowminder](#) and [GSMA](#) to source and [clean](#) the data responsibly.

**B.3. Social media and mobile application usage statistics:** consider [usage statistics from social media](#) (e.g. [X/Twitter](#) or [Meta data for good](#)) if the internet penetration rate is greater than 70% (see [World Bank statistics](#)), following UNHCR [guidance on social media for community-based protection](#). Otherwise, consider using [mobile application usage data](#) such as the opt in use of locations services.

**B.4. Proxy datasets:** Explore other datasets that could be *proxies* for displacement such as trends in market prices and financial transactions/cash assistance data.

**B.5. Mine radio broadcasts:** to [understand movement and protection incidents](#).

**B.6. Explore Google data:** including [Google trends](#) (e.g. search for terms in the local language like “how to seek asylum in”, or “how to migrate to”, or “how to go to X country” or “what documentation or vaccines is required in this country”), and the [Google distance matrix API](#).

**B.7. Data science techniques:** can be applied if sufficient historical data is available (e.g. daily or weekly data for at least a year, or two to three years of monthly data). Estimates can be produced using nowcasting or forecasting approaches. Multivariate regression analysis helps to explore correlations and relationships between variables. Strive to corroborate the findings of data modelling by triangulating this with other data.

### **Step 5. Presenting the population estimates**

- Use standard CORE templates and guidance to ensure population estimates are consistently presented. Produce both static and dynamic dashboards / reports, to support the needs of a broader set of users.
- Share those population estimates and information products through systems such as UNHCR's Operational Data Portal.

For more information, please read the entry "Emergency Information Products".

### **Considerations for practical implementation**

The following tips and best practice should be considered when implementing the methodologies listed above:

- Population estimation is *not* protection monitoring or needs assessment: population estimation generates information about the number and characteristics of a population in a specific place and time. By contrast, protection monitoring generates information about trends in the protection situation of specific population groups and/or specific locations. Needs assessments identifies the needs, conditions and capacities of specific groups and/or specific locations, in order to determine gaps between a current situation and agreed-upon standards.
- *Multiple approaches* or methods may be employed to generate the population estimates in different areas or at different moments in the emergency response, according to gaps in data needs and operational contexts.
- *Key stakeholders* can include law enforcement officials (border, police, security), civilian government authorities, representatives of the civil society/refugee-led or IDP-led organisations, researchers, experts, humanitarian, development and peace organisations) at national and sub-national level. To assess which organisations are present, review the available 3W information.
- Review *the target locations* for which you are producing population estimates. Record them following [corporate location standards](#) to improve the comparability and re-usability of the data collected over time. Minimise overlap between those locations or areas by assessing their geographical coverage and ensuring they are well known by the key informants and enumerators. Avoid gathering information in geographically close locations (e.g. on each side of a border, on the same road etc.).
- Improve *temporal coverage* by conducting time-location sampling (selecting location/time period combinations at random e.g. different days of the week and hours of the day). The magnitude and composition of flows can fluctuate, and for example, mornings are often busier.
- Improve analysis by collecting both *stock* and *flow* data, e.g. in key informant interviews in

camps or host community settings. Collect a) stock data on the target population present in the area at the moment of the data collection, b) flow data on new incoming and outgoing individuals/families over a certain period of time in that same area. Do not try to produce stock data from flow data or vice versa, as the data will not be reliable.

- Be transparent about *data coverage* and limitations when disseminating data to partners, highlighting changes in the access and data collection coverage between rounds of data collection.
- Define the required *duration* of the data collection to manage expectations of how long it will take to generate the estimate.
- If the geographic scale or magnitude of the displacement exceed what is possible to assess with the available resources, define a sampling strategy: where possible, define *sampling frames* based on probabilistic methods to enhance statistical representativeness, such as network scale-up, the quadrant method, capture-recapture, and multi-stage and stratified samples. If probabilistic sampling is not possible, consider using respondent-driven sampling. Seek guidance from a regional or HQ statistical expert.

## Annexes

[ACAPS, Rapid Estimation of Affected Population Figures, 2012](#)

[IASC, Guidelines on the Humanitarian Profile Common Operational Dataset, 2011](#)

## 4. Learning and field practices

Tips for applying the methodologies described in steps 4A and 4B in common operational contexts:

- For all scenarios: consult key actors/stakeholders (method A1).
- **Scenario 1: Measuring displaced population flows in cross-border situations:** Use method A4 to source data from border officials, supporting a border monitoring system in coordination with them, if required. If the security situation permits, triangulate with data from method A2, or if resources and time permits, method A5. If most displaced people cross the border at less accessible situations, prioritize data collection in nearby transit locations (e.g. nearby villages and bus stations), as described in scenario 2.
- **Scenario 2: Estimating displaced populations located in host communities:** If key informant networks are active, use method A3. If not, conduct a field mission to identify key informants who have the capacity and willingness to monitor flows and stocks and can provide regular updates in the coming weeks or months.
- **Scenario 3: Estimating displaced populations located in camp-like settings:** Always use method A2 to triangulate other methodologies. Where resources and available time are limited use method A3. Explore opportunities for proxy indicators as described in method B4, such as analysing total water consumption by area/block in a site. Where greater time and resources are available, use methods A5 or A6.
- **Scenario 4: Estimating large population flows in rapidly evolving situations in**



**transit locations:** If there is access to the locations, use method A2 and if resources are available, method A5. In mixed movement situations, knowing the reasons for fleeing is important to help assess how many people are likely in need of international protection. People fleeing persecution will be underreported in group responses, therefore ensure to conduct individual surveys to better assess their reasons for moving. If there is no access to the transit locations, consider the methods described in step 4B.

## 5. Links

[ACAPS, Secondary Data Review, Technical Brief, 2014 ACLED early warning hub Google data IASC Humanitarian Profile Support Guidance, 2016IFRC Community mapping MMC - 4MI Methodology - Longitudinal Surveys UNHCR Master Glossary of Terms UNHCR Refugee Statistics International Recommendations on Internally Displaced Persons Statistics \(IRIS\) Refugee Statistics \(IRRS\) and Statelessness Statistics \(IROSS\) IOM, Key Migration Terms VIEWS violence early warning](#)

## 6. Main contacts

Global Data Service (GDS) - Statistics and Demographics Section: [hqcs00@unhcr.org](mailto:hqcs00@unhcr.org)

Innovation Service: [hqinlab@unhcr.org](mailto:hqinlab@unhcr.org)