Estimation of population denominators for the humanitarian health sector

Guidance for humanitarian coordination mechanisms

15 November 2018

Nada Abdelmagid MBBS MSc, Francesco Checchi PhD
Department of Infectious Disease Epidemiology
Faculty of Epidemiology and Population Health
London School of Hygiene and Tropical Medicine
Nada.abdelmagid@lshtm.ac.uk , Francesco.checchi@lshtm.ac.uk

Table of contents

Acronyms

AIDS	Acquired Immuno-Deficiency Syndrome
ART	Antiretroviral treatment
BEmONC	Basic Emergency Obstetric and Neonatal Care
CEmONC	Comprehensive Emergency Obstetric and Neonatal Care
d	Days
DHS	Demographic and Health Survey
GIS	Geographic Information System
GPS	Geographic Positioning System
HeRAMS	Health Resources Availability Monitoring System
HIV	Human Immunodeficiency Virus
HMIS	Health Management Information System
HNO	Humanitarian Needs Overview
HRP	Humanitarian Response Plan
IOM	International Organisation for Migration
MICS	Multiple Indicator Cluster Survey
MIRA	Multi-Sector Initial Rapid Assessment
mo	Month
NGO	Non-Governmental Organisation
PMTCT	Prevention of Mother-To-Child Transmission of HIV
SGBV	Sexual and Gender-Based Violence
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNFPA	United Nations Population Fund
UNOCHA	United Nations Office for Coordination of Humanitarian Affairs
WASH	Water, Sanitation And Hygiene
WHO	World Health Organisation
y	Year

Acknowledgements

Nada Abdelmagid was funded by UK Research and Innovation as part of the RECAP project, as well as by Save the Children UK. Francesco Checchi was funded by WHO Agreement to Perform Work WHE/EMO/HIM/001 (“Technical support to WHO Health Emergencies Programme”).

1 Introduction

1.1 Background

A key prerequisite for any effective humanitarian response is the availability of timely, reliable and robust information.[1] Sound decision-making for humanitarian health interventions requires robust public health information about the health status and risks faced by the affected population, the availability and functionality of health resources and services, and the performance of the health system.

One of the critical pieces of information needed from the very onset of a crisis is an accurate estimation of the size and composition of the affected population. In addition to an overall estimate, different population denominators are needed for planning and monitoring, such as the total number of people in a given location, the population within the catchment area of specific health services or projects, the number of children aged less than 5 years or the number of pregnant women.

Obtaining accurate estimates of population denominators is challenging in most crises. Some of these challenges are inherent to the nature of the crisis (e.g. large-scale population displacement), while others result from the lack of robust pre-war population estimates. As a result, conflicting estimates of crisis-affected population size and/or considerable uncertainty about population denominators are commonly encountered.

1.2 Purpose of this document

This guidance note accompanies the Global Health Cluster's Standards for Public Health Information Services in Activated Health Clusters and Other Humanitarian Health Coordination Mechanisms.[2] It specifically:

- provides an overview of the applications of population denominators in the humanitarian health sector;
- briefly introduces methods for estimation of population size in emergencies;
- demonstrates how to appraise and triangulate existing estimates;
- summarises calculations and parameters required to estimate the size of different vulnerable population groups;
- describes a systematic approach for estimating health service catchment populations; and
- suggests a simple method for defining populations in need of humanitarian health services.

This guidance note does not contain detailed methods or software/data collection instruments for primary data collection to estimate population denominators. Where appropriate, key methodological references and provided in the bibliography.

1.3 Target audience

This guidance note is intended for:

- Staff working in activated Health Clusters or other sectoral coordination mechanisms, including Health Custer Coordinators, Information Management Officers and Public Health Officers;
- Epidemiologists who may be deployed to support humanitarian health responses, and who may require population denominators to design surveys or analyse / interpret findings;
- Monitoring and evaluation and other health project managers within individual humanitarian sector agencies.

2 Applications of population denominators across the humanitarian project cycle

2.1 Needs assessment and analysis

At the outset of a crisis, population denominators are needed to construct essential indicators to quantify health status, public health threats faced by the affected population and the availability and functionality of health resources and services (see examples in Table 1). While the overall population is sometimes the appropriate denominator, many indicators instead require either or both of the following:

- the number of people belonging to relatively more vulnerable sub-groups, e.g. pregnant and lactating women, children less 5 years old, people living with a given condition (see Chapter 4);
- the estimated catchment population, i.e. people served by a health programme or facility, as delineated based on factors such as population distribution, natural geographic boundaries, and transportation accessibility. Chapter 0 suggests methods for quantifying this figure.

Whenever health indicators are inaccurate, needs analysis becomes unreliable and may be called into question, reducing the credibility of resulting resource mobilisation efforts. This may negatively affect the attention a humanitarian crisis receives with regards to both advocacy and resource allocation, and/or skew the response towards specific sub-populations or areas of service, resulting in inequity and areas of neglect. For example, a comparison of mortality surveillance data (reliant on population denominators) and mortality survey estimates (not reliant on population data) in Eastern Chad (2006–2010) showed discordant estimates.[3] Population overestimation was thought to explain the low death rates recorded by the surveillance system, compared to survey results. As seen in this example, inaccuracy in population estimates is not always visible – in this case, only a comparison of two mortality estimation methods allowed for the problem to be spotted.

2.2 Strategic response planning and resource mobilisation

The total affected population, and the population in need of humanitarian health services (see Chapter 0), commonly feature in Flash Appeals, joint inter-sectoral documents such as Humanitarian Response Plans (HRP), monitoring reports and humanitarian dashboards. These estimates are usually drawn up during a Multi-Sector Initial Rapid Assessment (MIRA) in the acute emergency phase, or periodic Humanitarian Needs Overview (HNO) updates in protracted crises.

Such estimates are essential for resource mobilisation, as they provide a crude metric of the size of a crisis and (should) underpin health sector budgets. They are also critical for various aspects of response planning: for example, the number of people in need of treatment continuation (e.g. for HIV: see Table 1) is needed to project drug supply requirements, while the number of children aged 6 months to 14 years is the basis for planning measles mass vaccination. If these target population denominators are under-estimated, excess HIV/AIDS mortality and transmission of HIV will occur, and a vaccination campaign will not achieve sufficiently high coverage to interrupt measles transmission in the population; vice versa, over-estimation will lead to unnecessary resources and funds being used that could instead have been directed to other public health priorities.

More generally, population denominators are needed to plan for adequate service availability. For example, according to the Sphere Project[4], minimum adequate standards for the availability of different services are:

- ≥ 80% of the population can access primary healthcare within one hour's walk from dwelling;
- ≥ 18 inpatient beds per 10,000 people;
- ≥ 5 health facilities offering basic emergency obstetric and neonatal care (BEmONC) per 500,000 people;
- ≥ 1 health facility offering comprehensive obstetric and neonatal care (CEmONC) per 500,000 population.

2.3 Response monitoring

As the response gets underway, population denominators become critical for monitoring health system performance through relevant indicators (see example in Table 1), as well as detecting emerging gaps (e.g. an increase in the population in need of secondary care due to the withdrawal of one or more actors) and threats (e.g. an unusual trend in malaria incidence, suggesting a possible epidemic). For example, the Sphere standards and other guidelines suggest benchmarks of acceptable coverage, e.g. measles vaccination coverage should be > 95% among children between 6mo and 15y old to prevent an outbreak.

Such indicators rely heavily on accurate population estimates: inaccuracy in the latter carries serious implications for decisions made during the humanitarian health response, and may result in forms of malpractice, as important emerging threats, gaps and inadequacies may be overlooked. Ultimately, robust population estimation is critical for accountability to beneficiaries and donors.

Table 1. Examples of public health indicators relevant to the humanitarian response.

Step in the project cycle	Information domain	Indicator	Typical expression	Formula (population denominator)
Needs assessment	Health status and public health threats	Coverage of three doses of pentavalent vaccine (routine)	Percentage	(Number of infants who receive the third dose of pentavalent vaccine before age 12mo, during a given period / Number of infants who turn 12mo old during the same period) x 100
	Availability and functionality of health services	Availability of primary health services	Percentage	(Total population that lives within 1h walk from a primary healthcare facility / Total population) x 100
Strategic response planning	Availability and functionality of health services	People in need of HIV/antiretroviral treatment (ART) continuation	Number	Number of people living with HIV x (ART coverage pre-crisis - estimated ART coverage now)
Response monitoring	Health status and public health threats	Crude death rate	Deaths per 10,000 person-days	(Total number of deaths during time period x 10,000 persons) / (Mid-period population x Number of days in time period)
	Health system performance	Outpatient utilisation rate	Consultations per person-year	(Number of curative consultations due to any cause over a given time period of data collection x t †) / (Outpatient health facility catchment population x time period of data collection)

† If the time period is expressed in units of days, t = 365; if weeks, t = 52; if months, t = 12; etc.

2.4 Challenges with estimating population size in crises

Estimating accurate and updated population denominators throughout the various phases of a crisis is hampered by several challenges:

- Crises disproportionately occur in countries with weak vital registration and other public information systems.[1] Accordingly, pre-crisis census data may be outdated and local health maps not reflective of recent changes in service availability or internal migration (meaning catchment populations may be outdated or altogether unavailable);
- Population displacement is a common feature in many crises, with populations displaced multiple times in some occasions. Moreover, migration rates may differ by age and gender, altering the pre-crisis population age-sex distribution (for example, many refugee populations feature a higher proportion of children than in the country of origin[5]). While camp-based populations are relatively amenable to estimation, most displaced people live outside camps and are harder to identify and enumerate. Furthermore, in acute scenarios the high rate of displacement quickly renders any point-in-time estimates obsolete;
- In crises where physical access to affected populations is hindered by infrastructural damage or insecurity, data collection becomes extremely difficult, and established methods, such as census counts, become unfeasible;
- Crises can affect populations across administrative boundaries – for example, the population affected by flooding might be that living in the low-lying areas of several districts or villages. This complicates the use of pre-existing population estimates defined by administrative area;

- Over-reporting of population size is known to occur if households or community informants perceive that population counts are associated with food or other relief good distributions, and/or if people maintain multiple residences (e.g. in a camp and in a nearby host community) to maximise their access to distributions[6];
- Among humanitarian agencies, few health actors have the requisite expertise and resources to implement rigorous methods for population estimation or apply demographic methods for population forward-projection.

3 Estimating population size in humanitarian health responses

3.1 Sources of population estimates

Methods to estimate population vary in terms of their feasibility (access and resources required) and robustness. A brief overview of the main methodological approaches is provided below, noting that not all have been tested for feasibility and accuracy in crisis settings. Further information on different methods and how to appraise the robustness of estimates derived from them is provided in the Annex and elsewhere.[1,7]

3.1.1 Projection from census or large-scale surveys

A **census** requiring exhaustive enumeration of the population provides the number and demographic characteristics of people in the affected area before the crisis. A census is an exhaustive enumeration of every resident within the national territory of any country within a specific period, and is a gold standard but is rarely, if ever, conducted in a crisis setting.[8] A recent**, large-scale household survey**, such as Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS), may also have collected useful data on the number of households in different parts of the country, as part of the sampling process, and, within its questionnaire, may collect data useful for estimating important demographic parameters such as fertility, birth, death and migration rates.

The number of people affected by a crisis can be derived from the above baseline figures by forward-projecting them to account for natural population growth (or decline, i.e. the difference between births and deaths) and population movement/migration. An appropriate rate of annual population growth is usually selected based on what is known about fertility and mortality trends (for example, in Sub-Saharan African settings 2-3% annual growth is typical). Accurate displacement data, particularly for internally displaced persons and broken down by location, may be more difficult to come by, and this is an inherent limitation of the method (see Annex). The International Organisation for Migration (IOM) offers an increasing range of data on displacement through its Displacement Tracking Matrix.

3.1.2 Registration, enumeration and habitation counts

In camp settings or similarly accessible settlements where sufficient resources are available, a **registration** of residents is usually conducted to establish services. Following baseline registration, records are updated to reflect new arrivals, births, deaths and people permanently leaving the settlement. Where registration is not possible, an **enumeration**, or simple headcount, can be done. Population estimates from exhaustive registration and enumeration can be highly accurate, although important sources of bias may arise:

- affected people may choose to register in more than one location to maximise their opportunities to receive assistance (in this respect, simple enumeration carries less potential for bias);
- affected people may not all be present on site when enumeration or registration take place.

Alternative methods include a **habitation (or residential structure) count**, which is then multiplied by an estimated **average habitation occupancy** to obtain a population estimate. The habitation count may be done on the ground, through aerial photography or by analysing very high-resolution satellite imagery. At ground level, walkabout or drive through processes can be used to validate estimates based on aerial/satellite images.[9] An average occupancy may already be available through past surveys or registration efforts; alternatively, it may be estimated through a rapid sample survey.

Lastly, populations moving across national borders or along major transportation roads can be estimated through **flow monitoring**, conducted at specific points along the migration route, although this may be of limited accuracy.

3.1.3 Area and distance sampling methods

Alternative sampling methods have been used, particularly in camp settings, to estimate key parameters that can be multiplied together to compute a population estimate. A necessary ingredient for each of these methods is an **estimate of the total surface area** in which the crisis-affected population is present, obtained through mapping software and/or Geographic Information System (GIS) analyses of satellite imagery. All of these methods are rapid but technically complex to implement. As they rely on a sample, they produce a point estimate as well as a confidence interval whose width cannot easily be predicted a priori and may hamper confident interpretation.

The **quadrat method** is particularly useful in very large camps or well-delineated settlements. It consists of drawing a spatially representative sample of squares of known size (quadrats) within the perimeter of the settlement, and counting people living within each of these squares: this yields an estimate of population density, which is then multiplied by surface area.[10] Similarly, the **transect method** counts people within a 'transect area' (i.e. a straight-line cross-section of the settlement). The transect is meant to capture the heterogeneity of population density across the settlement.

The **T-square method**[11] starts with a selection of random points within the surface area. Distances from these points to the nearest habitation, and from that habitation to the nearest next habitation, are measured. The number of people living in each sampled habitation is also collected. The method accordingly estimates the mean area per habitation and the mean habitation occupancy, which are then combined with surface area. This method provides relatively accurate estimates and can be done rapidly, but > 50 sampling points are needed to ensure accuracy. A variant, the **spatial interpolation method**, collects data from small circles around randomly sampled GPS points (usually > 50), including distance from the centre of the circle to each habitation, and habitation occupancy. Spatial statistical methods are then used to estimate population density across the settlement, and the latter is multiplied by surface area.[12]

3.1.4 Qualitative and convenience methods

The **Delphi method** brings a group of experts (e.g. people with strong local knowledge or relevant quantitative expertise) together to reach a consensual opinion about a situation. The panel of experts exchanges several rounds of anonymous information to make an educated guess on the population size.[13]

Community estimates can be provided by community leaders based on visual assessments or by compiling initial reports (which also provide numbers of vulnerable sub-groups). The reports are checked and verified at random and updated regularly. Estimates are gathered from more than one source and triangulated to determine reliability. These may provide a useful source of number of affected people in specific locations.[14]

Service data from systematic screening or mass vaccination campaigns targeting children under 5y old can be combined with an assumption about the proportion of under-5y olds in the population (usually 15-20% in Sub-Saharan Africa; see Chapter 4). In anthropometric screening, children who fall below a certain height (e.g. 110 cm) are assumed to be under 5y old. After a vaccination campaign, data on children vaccination can be divided by estimated campaign coverage (ideally provided by a survey) to estimate the total number of children. The number of children under 5y is then divided by the proportion of under 5y to estimate total population. This method will be unreliable if the proportion under 5y is unusually high or low due to the nature of the crisis, or if service data are inaccurate.

Other service data, such as on water usage, food consumption or mobile network usage[15] have been used to estimate the total population size. These methods, however, rely on sufficient coverage of the service and accurate service data collection.

3.2 Evaluating the accuracy of existing estimates

Before adopting an existing estimate, it is important to assess its strength.

Table 2 proposes a procedure for grading the quality of population estimates based on any available documentation or verbal report of how the estimate was obtained. The procedure considers seven criteria, each of which is attributed a sub-score. The criteria are weighted unequally to reflect their relative importance. The individual criterion sub-scores and weights are then averaged to come up with a summary quality score from 0 to 1. This semi-quantitative score in turn helps to:

- interpret the estimate with due caution;
- know which estimate to give more/less weight (credence) to when making sense of multiple estimates;
- make a case for investing in rigorous estimation of population size in situations where only low-quality estimates are available.

Table 2. Criteria for grading the quality of a population estimate.

	Criterion	Sub-score (S)	Weight (W)
1	**Choice of method**		0.25
	The method used is a gold standard (census)	1.0	
	The method used is theoretically highly accurate (census projections; registration/enumeration followed by prospective demographic surveillance; Quadrat method; T-square method)	0.8	
	The method used has moderate theoretical accuracy (Delphi method; spatial interpolation method; habitation count using ground, aerial or satellite imagery combined with data on habitation occupancy)	0.6	
	The method used has low theoretical accuracy (use of service data; key informant guesstimates by community, authorities or humanitarian actors; flow monitoring for moving populations only; transect method)	0.4	
	Other method (not mentioned above)	0.2	
	The method used is unknown	0.0	
2	**Precision**		0.10
	The estimate is from (or projected from) an exhaustive count (census, registration, enumeration, etc...)	1.0	
	The estimate is from (or projected from) a representative sample, with a confidence interval that is narrow enough to allow for confident decision-making	0.8	
	The estimate is from (or projected from) a representative sample, with a confidence interval that is too wide to allow for confident decision-making	0.6	
	The estimate consists of a guesstimated range	0.4	
	The estimate is a single figure and comes from a guesstimate	0.2	
	Other / unknown (e.g. the estimate seems to have been drawn from a sample, but no confidence interval is provided)	0.0	
3	**Bias** (for further guidance see Annex, 'How to interpret the result' column)		0.15
	All potential sources of bias have been accounted for in implementation of the method	1.0	
	Some potential sources of bias have been accounted for in implementation of the method while others have not	0.8	
	The method consists of a guesstimate by experts/respondents with no plausible motives for exaggerating, under-reporting, or distorting their answers	0.4	
	No potential sources of bias have been accounted for in implementation of the method OR The method is a guesstimate by experts/respondents with plausible motives for exaggerating, under-reporting, or distorting their answers	0.2	
	Other / unknown	0.0	
4	**Expertise / credibility**		0.10
	The estimate was performed by a reputable, independent institution/expert, with a track record of similar work	1.0	
	The source is not an expert individual or institution, but is not known to have previously provided biased or inaccurate information	0.6	
	The source (institution/expert/informant) is previously known for having provided biased or inaccurate information	0.2	
	Other / unknown	0.0	
5	**Timing of data collection**		0.15
	Data were collected during the current crisis; if any displacement has occurred, the estimate was obtained after the last wave of displacement AND/OR (for scenarios of ongoing population movement) within the last 2mo	1.0	
	Data were collected during the current crisis; if any displacement has occurred, the estimate was obtained before the last wave of displacement OR (for scenarios of ongoing population movement) > 2mo ago	0.8	
	Data were collected during the immediate pre-crisis period (within the last 2y)	0.6	
	Data were collected during the distant pre-crisis period (> 2y ago)	0.4	
	Other / unknown	0.0	

	Criterion	Sub-score (S)	Weight (W)
6	**Population included in estimation**		0.15
	The estimate only includes the population/population group of interest for the intended application of the estimate AND excludes all other populations/population groups (e.g. people not affected by the crisis, or not in need of specific services)	1.0	
	The estimate only includes part of the population/population group of interest AND excludes all other populations/population groups	0.6	
	The estimate includes all or part of the population/population group of interest AND all or part of other populations/population groups (e.g. neighbouring administrative areas; non-affected populations in the same area; other vulnerable groups)	0.4	
	Other / unknown	0.0	
7	**Plausibility**		0.10
	The estimate appears plausible, given what is known about the current crisis	1.0	
	The estimate appears plausible, given what is known about previous crises in the same population / similar crises	0.8	
	The estimate appears plausible, given what is known about the pre-crisis population OR similar crises in other populations	0.6	
	The estimate does not appear plausible, given what is known about the pre-crisis population / the current crisis / previous crises in the same population / similar crises	0.4	
	Other / unknown	0.0	
Summary Quality Score = $(S_1 \times W_1) + (S_2 \times W_2) + (S_3 \times W_3) + (S_4 \times W_4) + (S_5 \times W_5) + (S_6 \times W_6) + (S_7 \times W_7)$			

Panel 1 demonstrates how to compute quality scores in three hypothetical scenarios. Further guidance on how to attribute sub-scores depending on the method used is found in the Annex.

Panel 1. Example of how to calculate summary quality scores for existing population estimates.

In the wake of a typhoon, a town has suffered major infrastructural damage and most (though not all) residents have fled to an informal settlement 10km from the town centre.

Three population estimates are available:

a. A United Nations Population Fund (UNFPA) projection from a census carried out 15y ago. According to this projection, the town's population is currently 39,103;
b. A rapid needs assessment carried out in the settlement 2 days after the typhoon struck by an NGO. The data collectors were told by community leaders that there were 28,000 people in the settlement, but that they were expecting more families to arrive over the coming few days;
c. One week after the typhoon struck, the World Food Programme started food distribution in the informal settlement, after interviewing heads of households on the size of their households and compiling a food distribution list. The total number of people registered for food distribution was 54,522.

Corresponding quality scores would be as follows:

	Data Source	Summary score calculation	Notes
a	UNFPA population projection (39,103)	S_1 x W_1 = 0.8 x 0.25 = 0.20 S_2 x W_2 = 1.0 x 0.10 = 0.10 S_3 x W_3 = 0.8 x 0.15 = 0.12† S_4 x W_4 = 1.0 x 0.10 = 0.10 S_5 x W_5 = 0.4 x 0.15 = 0.06 S_6 x W_6 = 0.4 x 0.15 = 0.06† S_7 x W_7 = 0.6 x 0.10 = 0.06 **Summary Quality Score = 0.70**	†The method does not distinguish those who were displaced from those left behind in the town.
b	Community estimates (28,000)	S_1 x W_1 = 0.4 x 0.25 = 0.10 S_2 x W_2 = 0.2 x 0.10 = 0.02 S_3 x W_3 = 0.4 x 0.15 = 0.06 S_4 x W_4 = 0.6 x 0.10 = 0.06 S_5 x W_5 = 1.0 x 0.15 = 0.15 S_6 x W_6 = 1.0 x 0.15 = 0.15 S_7 x W_7 = 1.0 x 0.10 = 0.10 **Summary Quality Score = 0.64**	
c	Food distribution data (54,522)	S_1 x W_1 = 0.4 x 0.25 = 0.10 S_2 x W_2 = 1.0 x 0.10 = 0.10 S_3 x W_3 = 0.2 x 0.15 = 0.03‡ S_4 x W_4 = 1.0 x 0.10 = 0.10 S_5 x W_5 = 1.0 x 0.15 = 0.15 S_6 x W_6 = 1.0 x 0.15 = 0.15 S_7 x W_7 = 0.4 x 0.10 = 0.04# **Summary Quality Score = 0.67**	‡Inflation of household size by respondents is a likely, but apparently unaccounted-for source of bias. #The estimate seems implausible given pre-crisis census figures for the entire town (not just those displaced).

3.3 Triangulating different estimates

When multiple estimates are available for the same population, it is useful to triangulate these into a single set of estimates, comprising of:

- a weighted average point estimate (i.e. the most likely figure); and
- a low to high range.

3.3.1 Calculating a weighted average point estimate

One may use the quality scoring approach proposed above to calculate a weighted average (mean) of different estimates. The summary quality score may be directly used as the 'weight' for each estimate, i.e. relatively how much it counts towards the average. Accordingly, a 'point estimate' (i.e. most likely value) of population size combining all available sources can be computed using the following formula, where N_1, N_2, N_3 etc. are the available estimates (up to m different estimates available) and Q_1, Q_2, Q_3 etc. their corresponding summary quality scores:

$$\text{weighted point estimate} = \frac{N_1 \times Q_1 + N_2 \times Q_2 + N_3 \times Q_3 + \ldots + N_m \times Q_m}{Q_1 + Q_2 + Q_3 + \ldots + Q_m}$$

A worked-out example is provided in Panel 2.

Panel 2. Example of how to compute a weighted average point estimate of population size.

Given the three estimates provided in Panel 1, a triangulated population point estimate will be calculated as follows:

Source	Population estimate	Summary Quality Score
a	39,103	0.70
b	28,000	0.64
c	54,522	0.67

$$\text{Weighted point estimate} = \frac{(39{,}103 \times 0.70) + (28{,}000 \times 0.64) + (54{,}522 \times 0.67)}{(0.70 + 0.64 + 0.67)} = \mathbf{40{,}707}$$

3.3.2 Composing a low to high range

The lowest and highest among estimates available can be adopted as the low and high range values around the weighted point estimate. This will make the amount of uncertainty around population figures more readily visible.

For the example above, the low to high range would be **28,000 to 54,522**.

3.4 What to do when there are no robust estimates

In situations where no robust estimates of population size are available (for example, only a single emergency 'guesstimate' from the community is available), the following steps are recommended:

- Advocate for prioritisation of population size estimation activities, specifically highlighting examples of how better data could strengthen the response and possibly reduce inefficiency (i.e. pay for themselves);
- Acknowledge uncertainty around population denominators whenever presenting indicators that are based on these and/or communicating estimates of populations in need, resources required or unmet gaps;
- For planning purposes, adopt a **'no regrets' approach** and use the highest estimate available for procurement of supplies (particularly medicines and vaccines), definition of the number of health facilities and staff required to cover the population adequately, etc. If only a single figure is available, and there is no evidence to suspect it is an overestimate, it may be prudent to inflate it by a sensible percentage, e.g. 30% (inflated population estimate = 1.3 x population estimate), so as to reduce the risk of acting based on unrealistically low denominators;
- Even if no expertise is immediately available locally to undertake more robust estimation, assess the plausibility of the available estimate(s) by talking to a range of community and humanitarian informants and, if the population is based in camps or other confined settlements, physically visiting locations to get a rough sense of the most likely direction of bias in the reported denominator.

4 Estimating population by age and other characteristics

Many health indicators should be disaggregated by age group (e.g. under 5y crude death rate, outpatient utilisation rate among people aged 50-59y, etc.) or other population characteristics (e.g. coverage of antenatal care among pregnant women). Such disaggregation is needed to inform age- or group-specific interventions.

The humanitarian community does not use a standard set of age or other group characteristic cut-offs. Table 3 provides reference values for the expected percentage of the population falling within different groups, as commonly used by humanitarian health and nutrition actors. These largely apply to **stable Sub-Saharan African settings only**. Generally, it is best to refer to local demographic estimates, e.g. from census, MICS or DHS, with the most recent data being most reliable unless information suggesting otherwise is available: if not online, these data will typically be available from district government offices, the central bureau of statistics or the population department of local authorities. The UN World Population Prospects and the United States Census Bureau also provide credible country-wide projections.

Table 3. Typical values for the proportion of various population groups in stable Sub-Saharan African settings.†

Group	Denomination	Proportion of the population that belongs to the group	Notes
Age			
0 to 28d	Neonates	0.3 to 0.5%	Best computed from the birth rate: see Table 4
0 to 6mo		1.5% to 3.0%	
6 to 11mo		2% to 4%	
0 to 11mo	Infants	5% to 8%	
12 to 23mo	Toddlers	4% to 7%	
0 to 59mo	Under-five	14% to 17%	
5 to 14y		25% to 30%	
0 to 14y		40% to 45%	Typical age range for vaccination campaigns
0 to 17y	Children	45% to 50%	
10 to 19y	Adolescents	20% to 25%	
15 to 49y		45% to 50%	
≥ 60y	Older persons	3% to 7%	
Other characteristics			
Women of reproductive age (15 to 49y)		22% to 25%	
Pregnant women		2.0% to 4.5%	Best computed from the birth rate: see Table 4
Lactating women		2.5% to 5.5%	Best computed from the birth rate: see Table 4
Pregnant and lactating women		4.5% to 10.0%	Best computed from the birth rate: see Table 4

† Approximate ranges adapted from the United Nations World Population Prospects, 2017, except where noted.

4.1 Age groups

Available age distribution values will mostly refer to pre-crisis conditions: crisis-affected and especially displaced populations may feature a somewhat altered age distribution, e.g. camps may house relatively more young children and women. Camp registration data or surveys that have collected age data as part of a representative household sample may yield more locally appropriate information.

Generally, **children** are defined as those aged less than 18y. Neonates are up to 28d (or 1mo if accurate information on the birth date is unavailable), infants are < 12mo old and adolescents 10-19y old. Planning figures (e.g. for vaccination) require estimates of the size of specific age groups within childhood; certain indicators also require such disaggregation (e.g. the prevalence of global acute malnutrition is usually estimated among children 6-59mo old, but is further disaggregated into the 6-23mo and 24-59mo age groups, e.g. because a high prevalence among older children indicates very poor nutritional status; proportional morbidity among children seen in outpatient services is disaggregated into the age groups <5y and ≥5y, reflecting the different spectrums of disease with increasing age; when analysing the incidence of sexually transmitted infections among adolescents, it is useful to observe patterns among young adolescents aged 10-14y; etc.). Information on these finer cut-offs is best sourced locally, i.e. from census, MICS, DHS or other credible data collection exercises.

4.2 Pregnancy, birth, lactation, infancy

The age group 15-49y is generally used to define women of reproductive age. However, in some settings 15-44y is used instead. The prevalence of contraceptive use is an example indicator requiring this denominator.

There are specific programmes and indicators that require information on pregnant women (e.g. the coverage of antenatal care), lactating women (e.g. infant and young child feeding), or both groups (e.g. for financing calculations when considering user fee exemptions or cash transfers). Furthermore, the expected number of births and pregnancy outcomes are needed to plan the availability of safe and assisted delivery, BEmONC and CEmONC services.

As shown in Table 4, most of these denominators can be worked out if the overall population and crude birth rate (expressed as live births per 1000 people per y) are known, and a few key assumptions on pregnancy outcomes are made (the latter should be verified and if needed modified locally).

Table 4. Projecting pregnancy outcomes and related populations in need.

Parameter	Symbol	Value	Source	Notes
Total population (all ages, both genders)	N	Context-specific	See Chapter 0	May need to instead use catchment population (see Chapter 0)
Crude birth rate (per 1000 people per y)	r	Context-specific	Census, DHS, MICS, population projections	May decline during crises or among displaced populations (but hard to measure locally).
Proportion of pregnancies that end in stillbirths or miscarriages	c_0	0.15 (15%)	Assumed[16]	If robust data are available locally, this assumption may be modified
Proportion of pregnancies with complications that can be managed through BEmONC services	c_1	0.15 (15%)	Assumed[16]	If robust data are available locally, this assumption may be modified
Proportion of pregnancies with complications that require CEmONC services	c_2	0.05-0.15 (5-15%)	Assumed[16]	This assumption should be verified and modified using any locally available data
Mean gestation period	G	280d	Assumed[16]	Accounts for pre-term pregnancy outcomes
Mean lactation period	L	365d	Assumed[16]	If robust data are available locally, this assumption may be modified
Time unit of interest	t	365 for d, 12 for mo, 1 for y		

Denominator	Symbol	Formula	Notes
Number of pregnant women	P	(N x r x G) / (365 x 1000)	At any given time
Number of lactating women		(N x r x L) / (365 x 1000)	At any given time
Number of live births per time unit	B	(N x r) / (t x 1000)	
Number of women requiring safe and assisted delivery (live or stillbirths) per time unit	D	[(N x r) / (t x 1000)] x (1 + c_0)	
Number of women requiring BEmONC services per time unit		D x c_1	
Number of women requiring CEmONC services per time unit		D x c_2	
Number of antenatal visits expected per time unit		P x 4 x (365 / t)	Assuming all women should attend 4 antenatal visits; if more visits are recommended, replace 4 with the appropriate number
Number of postnatal visits expected per time unit		P x 1 x (365 / t)	Assuming all neonates should attend 1 antenatal visit
Number of neonates		(N x r x 28) / (365 x 1000)	At any given time
Number of children turning 12mo per time unit		B	Used as denominator for the coverage of routine vaccination. Assumes constant birth rate and no infant mortality.

4.3 People living with HIV and other conditions

The population of people living with HIV can be calculated by multiplying the total population by the HIV prevalence (%) in the general population, for example as reported in the latest Joint United Nations Programme on HIV/AIDS (UNAIDS) country reports. At the outset of an emergency, estimating the number in need of antiretroviral treatment (ART) continuation may be important if widespread health system disruptions have occurred, and can also be done if UNAIDS data on people receiving ART pre-crisis are combined with appropriately conservative assumptions on what proportion of these are likely to have lost access to ART as a result of the crisis. Moreover, the number of pregnant women in need of prevention of mother-to-child HIV transmission (PMTCT) may be computed by multiplying the number of expected live births (B in Table 4) by the prevalence of HIV infection among pregnant women, which is usually provided by UNAIDS alongside general population prevalence.

A similar approach may be taken to estimate the number of people in need of tuberculosis treatment continuation or care for key non-communicable diseases (e.g. hypertension, diabetes).

Estimating the number of people in need of mental health services is more complex since a range of mental health problems with varying severity and needs will occur depending on the crisis. To obtain accurate estimates for the prevalence of specific mental health problems, surveys must be conducted using validated instruments. However, WHO recommends focussing on assessment of local perceptions, resources and coping factors, not prevalence estimation.[1] In general, humanitarian mental health services should be made available in any crisis. As a rough indication, around 12mo into the timeline of a crisis about 3-4% of the general population may be expected to have severe disorders such as psychosis and very disabling anxiety and depression, while some 15-20% may have mild or moderate forms of anxiety, depression and post-traumatic stress disorder; a large percentage will experience some distress and ‘normal’ psychological reactions to the crisis.[17]

5 Estimating catchment populations

5.1 Applications of catchment populations

In crisis settings, the pre-crisis catchment populations of health facilities may change due to population displacement, obstacles to access such as insecurity, health facilities being damaged or going out of service, and/or humanitarian actors opening new or upgrading existing health facilities.

Estimates of catchment population have two main applications:

- Accurate quantification of resources needed by the facility (for example, staff of different levels; drugs or drug kits);
- Calculation of **service coverage** indicators needed to monitor the performance and utilisation of health services. For example, a very common coverage indicator, the outpatient utilisation rate (or consultation rate), usually expressed as the number of outpatient curative visits per person per year, requires as the denominator the catchment population of the facility of interest, or, if looking at a whole geographical area (e.g. a province, a district, a camp or the entire crisis-affected area), the total catchment population of all facilities that are functional in that area.

By contrast, when calculating **service availability** indicators, e.g. the number of basic emergency obstetric and neonatal care facilities per population, one should refer to the total population in need of services (see Chapter 6), irrespective of whether they fall within the catchment of available services.

Under- or overestimates of catchment population can seriously affect humanitarian decision-making and population health outcomes, e.g. by resulting in drug stock-outs, resource wastage, biased data on service utilisation, etc.

5.2 Methods for estimating catchment populations

Three methods for defining catchment populations are presented below, one of which should be selected depending on resources and information available. These methods will generate either of the following two types of estimate:

- (preferable) the **'natural' or 'effective' catchment population**, i.e. as defined empirically based on patient origin data;
- the **'theoretical' or 'geographical' catchment population**, defined based on distance or travel time between given locations and the facility, and arbitrary cut-offs for maximum distance or travel time.

All of the above methods are greatly enhanced by any available health maps (including hand-drawn versions) of the area: for example, local district health offices or routine vaccination facilities often display maps of health facilities and villages within the catchment of each. These maps should always be sought before setting out to define catchment populations.

5.2.1 Method 1: Patient origin analysis

This method provides an estimate of the 'effective' catchment population and requires the following data:

- The number of consultations (or other services depending on what one is interested in defining the catchment of) dispensed at the facility, overall and by place of residence of patients (e.g. village or urban neighbourhood, or a higher administrative level for hospitals or other referral centres catering to a large geographic area); these figures can typically be extracted from patient registers, if they are not already available; the data should be sourced for a reasonable period in the past (at least 1-4 weeks or at least the last 100 service utilisations);
- The theoretical catchment of the facility, as per local health maps: this refers to the villages or other administrative units that, for health service planning purposes, 'belong' to the health facility (see above);
- The total population of the administrative units within this theoretical catchment, if necessary triangulated as discussed in Chapter 0.

Let U be the total number of service uses (e.g. consultations), U_c the number of service uses by patients living within the theoretical facility catchment and N_c the total estimated population within this theoretical catchment. The effective catchment population may be estimated as

$$N = N_c \frac{U}{U_c}$$

The above equation attempts to augment the effective catchment population to account for the proportion of service uses from outside the theoretical catchment. A worked example is provided in Panel 3. The method is useful when it is suspected that the effective catchment area is in fact greater than the theoretical, but less so in the opposite scenario.

Panel 3. Example of the patient origin analysis method for defining catchment populations.

An outpatient health facility theoretically serves villages V_1, V_2, V_3 and V_4. Following torrential rains and floods, neighbouring communities and districts have suffered damage to their health facilities. It is suspected that many users from these neighbouring communities are now using the health facility (caseload is higher than average and essential drugs have unexpectedly run out earlier than forecasted). These observations suggest a need to better define the facility's catchment population.

The following data are extracted from available sources:

Population of the theoretical catchment: 5000 (V_1) + 9000 (V_2) + 7000 (V_3) + 4000 (V_4) = 25,000

Number of outpatient consultations over the last month in the health facility: 960

Number of consultations by patients resident within the theoretical catchment: 643 (67%)

Therefore, the estimated effective catchment population = 25,000 x (960 / 643) ≈ **37,000**.

A **variant** of patient origin analysis sets an arbitrary cut-off for the proportion (or percentage) of patients that must come from a given location, in order for that location to be included in the catchment population of the facility. For example, after recording the village of origin of the last 300 patients seen in an outpatient facility, any village that accounts for ≥ 10% of the patient load (≥ 30 patients) might be included, while all other villages would be excluded. The actual cut-off should be defined locally, depending on how many administrative units / locations contribute patients; practically it may be easiest to first collect patient origin data and tabulate percentages by location, sorting locations from most to least patients – a clear distinction between locations that mostly rely on the facility and those that use it only occasionally may be graphically visible, and make the choice of most appropriate cut-off obvious.

When defining the catchment population of **hospitals**, it may be impractical to work with small geographical units (e.g. villages), as these may be too numerous; instead, one may consider larger geographical units for the analysis. At the district level, cut-offs between 10% and 50% have been proposed.[18,19]

This variant requires population estimates to be available for geographical units not included in the theoretical catchment area and may thus be less feasible. Furthermore, the choice of cut-off may result in under- or overestimation of the catchment population: generally, an appropriately conservative choice should be made so as to avoid insufficient allocation of staff and resources.

5.2.2 Method 2: Distance or travel time

This method considers one of the following:

- the population residing within a circle whose radius is an arbitrary maximum **straight-line distance** from the health facility; this is least accurate in settings with sparse/poor road networks or geographical barriers such as hills, rivers or coastline, i.e. where straight-line distances do not reflect actual routes and travel times;
- the population within the area bounded by an arbitrary maximum **travel distance** to the health facility;

- the population within the area bounded by an arbitrary maximum **travel time** (by car or, depending on the main transport mode available in the area, walking, cycling, public transportation, etc.) to the health facility.

This method has not been validated in crisis settings, but versions of it have been used to estimate catchment and coverage of specialised services such as B/CEmONC.[20,21] It assumes that people will visit the closest facility, which may not apply in many crisis settings – for example, the closest facility may not be functioning, or people may prefer a more distant facility because of perceived quality of care.

The World Health Organization (WHO) recommends using travel time rather than distance.[22] Cut-offs of 5Km for distance or 1 hour for travel time are commonly used for primary-level facilities, and 2 hours travel-time for secondary-level facilities (e.g. hospitals, CEmONC).[23,24] However, these cut-offs can be aligned with local standards or tailored to the context.

Whether distance or travel time are used, measurements required to construct the catchment area can be taken using traditional ground mapping methods. However, Geographic Information System (GIS) technology is increasingly being used for such exercises. To measure straight-line or road travel distances, open source web-based tools and calculators are available, including the WHO's [AccessMod [version 5.0]]. The software comes with user guides, sample data and tutorials. Many local health offices hold maps of catchment areas for facilities, and GPS coordinates may already be available or can easily be recorded by actors on the ground, e.g. with standard smartphones. However, GPS coordinates, population estimates and maps/information on local topography or crisis-related access blockages (e.g. a checkpoint) for communities around the facility are also required.

5.2.3 Method 3: Participatory mapping

This method uses patient origin data and population size estimates, but adds an additional layer of information by interviewing key informants[25] (usually health facility service providers but also community representatives or civil society actors familiar with the area and involved in health service delivery). The output is a local or district-level map of the catchment areas of each facility.

As above, the method requires population estimates and, if GIS is being used, GPS coordinates. Producing maps to scale using software applications has manifold uses, and should be attempted if possible; however, even rough sketch maps drawn by community informants are better than no participatory data collection at all.

This method has not been validated in crisis settings, but versions of it have been used to estimate catchment and coverage of HIV[25] and routine vaccination[26] services.

5.3 Other issues with catchment populations

In a fast-evolving setting (e.g. with high rates of in- or out-migration), it is recommended that catchment populations be reviewed and updated on a quarterly basis. Elsewhere, updates every semester are probably sufficient. In a large crisis, the coordination team may not have enough staff resources and geographical access to personally implement one of the above methods everywhere. Instead, the team should motivate, train and remotely support cluster partners and other humanitarian actors to carry out this estimation work for the facilities that they support. The coordination team should also maintain and make public a master database of health facilities and corresponding catchment populations (note that such a database is also required for implementation of a Health Resources Availability Monitoring System or HeRAMS, a key Public Health Information Service for activated health coordination mechanisms).

Note that a single health facility may have different catchment populations depending on the range of services it provides: for example, the same secondary hospital might have a (relatively large) catchment population for its inpatient services, a (smaller and more local) catchment population for outpatient care, a catchment population of people living with HIV if it provides HIV care, and a catchment population of children aged below 2y if it is also a routine vaccination point. As such, it is helpful to not conflate catchment populations for different levels and areas of health service prevision, but rather keep them separate: this will make calculation of service coverage indicators (e.g. routine vaccination coverage) straightforward.

6 Estimating populations in need for the health sector

The denominator of people in need of humanitarian health services is a key contribution of the health sector to inter-sectoral planning and resource mobilisation documents including the HRP and the HNO. Such a figure, if combined with total funding requirements or actual funds disbursed for health, can also provide a basis for roughly assessing the per capita aid allocation, relative to other crises but also to what is known about the true cost of appropriate health service packages.

6.1 Estimation of overall population in need

A general approach for estimating populations in need of humanitarian health services is provided in Table 5. The premise of this approach is that a humanitarian health response primarily aims to reduce excess (i.e. crisis-attributable) mortality and disability (including mental health problems), and therefore that any population that risks experiencing such an excess in the absence of humanitarian support should be included in the overall denominator in need. Otherwise put, sound public health planning considers that the people in need of preventive or curative health services are not just those who will ultimately utilise any available services, but anyone who may require them; moreover, some services, like vaccination or management of common mental health problems, have an indirect benefit on people in the household and the community by preventing onward transmission or improving psychosocial wellbeing, livelihoods, feeding and care practices etc.

Table 5. Recommended steps for estimating the population in need of humanitarian health services.

Step	Computation	Criteria and notes
1	Divide the affected region into appropriate geographical units; come up with total population estimates for each unit, if necessary applying triangulation as described in Chapter 0.	Geographical units should ideally (but not necessarily) follow administrative boundaries, at a level consistent with the granularity of available information (e.g. administrative level 1 if data are not broken down further, or level 2 otherwise); a more granular level (e.g. county, zone) may be used wherever information is available (i.e. the level needn't be the same across the affected country or region). Population estimates should as far as possible be consistent with those used by the, as this will promote a balanced attribution of resources among sectors.
2	Decide for each geographical unit whether any humanitarian health services will be required during some or all of the planning period for which a figure is needed.	Humanitarian health services will be required if excess mortality and/or disability are expected due to the crisis, in the absence of such services. Review needs analysis products and any other available information, including informal reports, to systematically decide whether any of the following risk factors are likely to be present in a given geographical unit: ▪ Increasing acute malnutrition ▪ Worsening feeding and care practices ▪ Overcrowding ▪ Insufficient vaccination coverage ▪ Inadequate shelter ▪ Inadequate water, sanitation and hygiene (WASH) services ▪ Increased sexual and gender-based violence (SGBV) frequency ▪ Disruption and/or reduced access to health services If at least one of the above risk factors is present, assume that humanitarian health services are required.
3	Compute the total population in need by summing together the populations of any units where humanitarian health services are required.	Round to the nearest thousand. Estimate vulnerable populations (e.g. children) as needed, and as shown in Chapter 4.
4	Document the estimation thoroughly.	At a minimum, prepare and safely store a dated document (max. 2-3 pages) in which figures and decisions made in steps 1 and 2 are detailed clearly. State areas of uncertainty (e.g. geographical units for which a highly uncertain classification was made). It is highly recommended to make this document public.

The above analysis should be done by geographic unit (e.g. districts or provinces), and the resulting population in need should be summed up. Particularly at the outset of an emergency, information scarcity will likely mean that conservative, no-regrets assumptions need to be made about conditions in parts of the country or the affected region for which sufficiently granular information is not available. For example, even in the absence of comprehensive situational awareness, credible reports from a given district of serious damage to health facilities in the aftermath of a flooding disaster could be taken as sufficient evidence to assume that excess morbidity and mortality will occur unless humanitarian health services are provided.

6.2 Population in need of specific services

The above approach is necessarily crude and does not take into account different levels and types of need across geographical units or crises. It is mainly meant to provide a top-line denominator for high-level strategic processes. Table 6 suggests ways to estimate the population in need of specific health services, on the basis of whether different crisis-emergent risk factors are occurring in the geographical unit being assessed, and additional assumptions to project expected service users. Such an assessment is best made in conjunction with the Public Health Situation Analysis service[2] of activated health coordination mechanisms.

Table 6. Estimates of the number of people in need of specific health services.

Service	Humanitarian health services are needed if any of the following factors are occurring:								Assumptions
	Increasing acute malnutrition	Worsening feeding and care practices	Overcrowding	Insufficient vaccination coverage	Inadequate shelter	Inadequate WASH services	Increased SGBV frequency	Disruption and/or reduced access to health services	
Outpatient curative care	X	X	X	X	X	X		X	Plan for 2-4 consultations per person-year in the acute emergency phase and ≈1 in the protracted phase
Inpatient curative care	X	X	X	X	X	X		X	10% of the population will require inpatient admission per year
Vaccination	X	X	X	X		X		X	Depends on vaccination strategy
Antenatal care	X	X		X		X		X	See Section 4.2
Safe assisted delivery	X	X		X		X		X	See Section 4.2
BEmONC	X			X		X		X	See Section 4.2
CEmONC	X			X		X		X	See Section 4.2
Clinical management of sexual and gender-based violence							X		Assume that SGBV, and in particular sexual violence, is taking place, regardless of the presence or absence of concrete and reliable evidence on prevalence. For planning purposes, provide services in all health facilities. Even if prevalence estimate is available, always assume it is underestimated as GBV is underreported almost everywhere in the world.
Management of common mental health problems	X	X	X	X	X	X	X	X	See Section 4.3
Continuation of care and control for HIV, tuberculosis, non-communicable diseases and/or locally important neglected tropical diseases	X							X	See Section 4.3

7 References

1. Checchi F, Warsame A, Treacy-Wong V, Polonsky J, van Ommeren M, Prudhon C. Public health information in crisis-affected populations: a review of methods and their use for advocacy and action. *Lancet* 2017; **390**(10109): 2297-313.
2. Global Health Cluster. Standards for Public Health Information Services in Activated Health Clusters and Other Humanitarian Health Coordination Mechanisms. Geneva.: Global Health Cluster, 2017. http://www.who.int/health-cluster/resources/publications/PHIS-standards.pdf?ua=1, accessed 28 May 2018
3. Bowden S, Braker K, Checchi F, Wong S. Implementation and utilisation of community-based mortality surveillance: a case study from Chad. *Confl Health* 2012; **6**(1): 11.
4. The Sphere Project. Sphere Handbook: Humanitarian Charter and Minimum Standards in Disaster Response. Geneva: Sphere Project, 2018. https://handbook.spherestandards.org/ (accessed 10 November 2018).
5. United Nations High Commissioner for Refugees. Global trends: Forced displacement in 2016. Geneva: UNHCR, 2016. http://www.unhcr.org/5943e8a34.pdf (accessed 8 November 2018).
6. Harrell-Bond B, Voutira E, Leopold M. Counting the Refugees: Gifts, Givers, Patrons and Clients. *Journal of Refugee Studies* 1992; **5**(3-4): 205-25.
7. Koedam A. Rapid estimation of affected population figures: desk review. Geneva: ACAPS, 2012. https://www.acaps.org/sites/acaps/files/resources/files/rapid_estimation_of_affected_population_figures_may_2012.pdf (accessed 10 July 2018).
8. United Nations Population Fund. Guidelines on Data Issues in Humanitarian Crisis Situations. Geneva: UNFPA, 2010. https://www.unfpa.org/sites/default/files/pub-pdf/guidlines_dataissues.pdf (accessed 5 November 2018).
9. Robinson C, Rowley E, Bolton P, Weiss W, Singh K. Demographic Methods in Emergency Assessment. A guide for practitioners. Baltimore: Centre for International Emergency, Disasters and Refugee Studies, Johns Hopkins University Bloomsberg School of Public Health, 2003
10. Grais RF, Coulombier D, Ampuero J, et al. Are rapid population estimates accurate? A field trial of two different assessment methods. *Disasters* 2006; **30**(3): 364-76.
11. World Food Programme. Emergency Food Security Assessments (EFSAs) Technical guidance sheet no. 11. Using the T-square sampling method to estimate population size, demographics and other characteristics in emergency food security assessments (EFSAs). Rome: WFP, 2009. https://www.wfp.org/content/technical-guidance-sheet-no11-using-t-square-sampling-estimate-pop-size-demographics-efsa (accessed 15 July 2018).
12. Pinto A, Brown V, Chan K, et al. Estimating Population Size Using Spatial Analysis Methods. In: Lai PC, Mak ASH, eds. GIS for Health and the Environment. Berlin: Springer; 2007: 271-87.
13. World Food Programme. Desk Review: Estimating population size in emergencies. Rome: WFP, 2006. http://reliefweb.int/sites/reliefweb.int/files/resources/1A1CB3E6D037A387C12572BF00363C1F-wfp-coordination-Dec06.pdf (accessed 15 September 2018).
14. United Nations High Commissioner for Refugees. Information Management Toolkit. Chapter 3: Rapid Population Estimation Methods. 2018. http://data.unhcr.org/imtoolkit/.
15. Bengtsson L, Lu X, Thorson A, Garfield R, von Schreeb J. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. *PLoS Med* 2011; **8**(8): e1001083.
16. Inter-Agency Working Group on Reproductive Health in Crises. Inter-agency field manual on reproductive health in humanitarian settings: 2010 revision for field-testing. Geneva, 2010. https://iawg.net/resource/inter-agency-field-manual-on-reproductive-health-in-humanitarian-settings-2010/ (accessed 14 November 2018).
17. World Health Organization, United Nations High Commissioner for Refugees. Assessing Mental Health and Psychosocial Needs and Resources: Toolkit for Humanitarian Settings. Geneva: WHO, 2012. http://www.who.int/mental_health/resources/toolkit_mh_emergencies/en/ (accessed 23 October 2018).
18. Goody B. Defining rural hospital markets. *Health Serv Res* 1993; **28**(2): 183-200.
19. Shortt NK, Moore A, Coombes M, Wymer C. Defining regions for locality health care planning: a multidimensional approach. *Soc Sci Med* 2005; **60**(12): 2715-27.

20. Chowdhury ME, Biswas TK, Rahman M, Pasha K, Hossain MA. Use of a geographic information system to assess accessibility to health facilities providing emergency obstetric and newborn care in Bangladesh. *Int J Gynaecol Obstet* 2017; **138**(2): 164-70.
21. World Health Organization. Investing the Marginal Dollar for Maternal and Newborn Health (MNH): Accessibility to EmOC services in Burkina Faso. Geneva: WHO, 2015.http://apps.who.int/iris/bitstream/10665/250269/1/WHO-HIS-HGF-GIS-2016.1-eng.pdf (accessed 12 November 2018).
22. Huerta Munoz U, Kallestal C. Geographical accessibility and spatial coverage modeling of the primary health care network in the Western Province of Rwanda. *Int J Health Geogr* 2012; **11**: 40.
23. World Health Organization. Surgical care systems strengthening: developing national surgical, obstetric and anaesthesia plans. Geneva: WHO, 2017.http://apps.who.int/iris/handle/10665/255566 (accessed 14 November 2018).
24. World Health Organization, United Nations Population Fund, United Nations Children's Fund, Mailman School of Public Health. Monitoring Emergency Obstetric Care: A Handbook. Geneva: WHO, 2009.http://www.who.int/reproductivehealth/publications/monitoring/9789241547734/en/ (accessed 10 November 2018).
25. Cunningham M, LaMois P, Mapala Y, Inglis A. Using Participatory Mapping to Assess Service Catchment and Coverage. Guidance from the Iringa Participatory Mapping Exercise, 2014.https://www.measureevaluation.org/resources/publications/sr-14-109 (accessed 10 November 2018).
26. Basics I. Health Facility Catchment Area Mapping – A Practical Guide: John Snow, Inc., 2008.www.immunizationbasics.jsi.com/Docs/IMMbasics_Nigeria_Practical_HFCAM_Dec2008.pdf (accessed 10 November 2018).
27. Checchi F, Stewart BT, Palmer JJ, Grundy C. Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations. *Int J Health Geogr* 2013; **12**: 4.
28. Brown V, Jacquier G, Coulombier D, Balandine S, Belanger F, Legros D. Rapid assessment of population size by area sampling in disaster situations. *Disasters* 2001; **25**(2): 164-71.
29. Bostoen K, Chalabi Z, Grais RF. Optimisation of the T-square sampling method to estimate population sizes. *Emerg Themes Epidemiol* 2007; **4**: 7.
30. Henderson A. Using the Delphi method to estimate population size and demographics in emergency food security assessments. Emergency Food Security Assessments (EFSAs) technical guidance sheet no. 10. Rome: World Food Programme, 2009. http://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp197300.pdf (accessed 10 November 2018).
31. Telford J, Gibbons LK. Counting and Identification of beneficiary populations in emergency operations: registration and its alternatives. London: Overseas Development Institute, 1997

8 Annex: Further guidance on methods for population estimation

Type of method	Feasibility in crisis settings	Validity in crisis settings	Notes on interpretation and bias assessment
Projection from census or large-scale surveys	Feasible as desk-based exercise, but requires basic demographic expertise in population projections.	Depends highly on the accuracy of population displacement figures, and of course of baseline census and survey estimates. The greater the time elapsed since the last census or survey, the higher the risk of inaccuracy.[8]	Requires expertise in population projections (you must know the underlying assumptions and methods used to be able to gauge the quality, validity and robustness of the estimates) If there is uncertainty in demographic assumptions (e.g. alternative data on displacement), the projections can be presented for each geographic unit as a range including low, medium (most likely) and high values. The low scenario might for example assume high uptake of family planning measures and/or low rates of out-migration, and vice versa for the high scenario.
Registration, enumeration and habitation counts	Mainly feasible in camp or other well-defined, concentrated settlements. <u>Registration and enumeration</u> require sustained access to the settlements (at least one week). <u>Ground habitation counts</u> can be very rapid and do not require specific expertise, but require access (at least 1-2 days). <u>Aerial or satellite imagery-based counts</u> can be done remotely if occupancy estimates are already available. However, they are technically complex. Cloud cover and bureaucratic impediments may hamper implementation.	Satellite-based imagery counts have been validated in crisis settings.[27], but have limitations in settings with multi-level housing units, extensive recent out-migration or extensive vegetation. They are best if complemented by ground truthing.	When interpreting the result, consider the following to appraise the robustness of the estimate: ▪ Is the population in a clearly demarcated and distinct geographical area, separate from unaffected population groups (this is necessary to prevent overestimation of population size, e.g. when displaced people settle within host populations)? ▪ If a habitation count is performed, how was the estimate of habitation occupancy derived? Ideally, this should be done through a ground sample survey, however rapid. Moreover, if habitations (structures) are counted, occupancy must also refer to people living within habitations, which may not be the same as household size, particularly in multi-occupancy structures.[27]
Area or distance sampling	Only feasible in camp or other well-defined, concentrated settlements. Requires access to affected population for about 1-3 days. Requires some technical expertise. Software is available for sampling and analysis.	Validated in crisis settings, albeit in very few settings.[10,12,28,29] The effect of heterogeneity in settlement patterns (i.e. sub-areas with very high or low density) on the estimates is not yet fully explored, but is likely to be large, affecting the accuracy of these methods.	When interpreting the result, consider the following to appraise the robustness of the estimate: ▪ What method was used to estimate surface area (GPS is more accurate compared to other methods[28])? ▪ How were the sampling points selected?

Type of method	Feasibility in crisis settings	Validity in crisis settings	Notes on interpretation and bias assessment
	Spatial interpolation requires expertise in GIS analysis.		▪ How were distances measured? (physical measurements such as footsteps, tape measure or rulers may be inaccurate in challenging physical environments – trigonometry may be better in such situations[10]) ▪ How many sampling points were measured? (>40 sampling points appear necessary[10])
Qualitative and convenience methods	These methods should be considered only if the above methods are all unfeasible. Most of these methods are highly feasible and can be implemented rapidly, with minimal access to the settlement, provided informants can be contacted or service data are available. The Delphi method requires a sufficient number (15-20) of experts. Mobile phone network analysis can be done remotely, but has many pre-requisites including niche expertise and reasonable network coverage and use.	The Delphi method may be most valid in complex situations where other methods are unlikely to cope well with potential sources of bias.[30] Mobile phone analysis has been validated in a post-earthquake scenario[15], and can be highly accurate if pre-requisites are met. However, it is most useful for tracking population displacement from/to specific areas, rather than for static populations. Key informant estimates may be most accurate in the initial stages of a crisis, before vested interests can influence reports.[31] Walkabouts, community mapping, and purposive sampling can all be used to improve the identification of key informants.[9] Service data on usage/consumption assume an arbitrary pre-defined average level of consumption per capita (e.g. L per person-day), which may not be accurate. Food distribution/consumption data are very prone to inaccuracy as rations can be diverted or sold, i.e. one ration may not equal one individual.	When interpreting the result, consider the following to appraise the robustness of the estimate: Delphi method ▪ Was Delphi the appropriate method to choose in this context? ▪ Did all experts complete the Delphi exercise? (the results could be biased if experts drop out before completing the exercise) ▪ Was the Delphi exercise executed effectively? (consistent anonymity maintained throughout, effective and unbiased facilitation, sufficient time allowed for experts to debate and reach consensus) Mobile phone analysis ▪ What assumption is made about movement patterns for groups with traditionally low mobile use (e.g. elderly, children)? (If these groups have substantially different movement patterns than groups with high mobile use, results will be biased) ▪ Is this method appropriate for the context? Were the pre-requisites met? (e.g. reasonably high mobile phone use, network coverage) Key informant estimates ▪ At which stage of the crisis was the estimate obtained? (in the earlier stages of the crisis estimates from key informants are likely to be less biased) ▪ How many key informants were used to arrive at the final estimates? (the larger the number of key informants the better the reliability of the estimate, or at least the more realistic the range provided) ▪ Who were the key informants asked to provide estimates? (a more diverse range of key informants may reduce bias, e.g. members of affected populations, local authorities, NGO actors familiar with the area, etc.)

Type of method	Feasibility in crisis settings	Validity in crisis settings	Notes on interpretation and bias assessment
			▪ How different are the estimates provided by different informants? (some estimates could differ by 100 percent or more, making triangulation difficult and reducing the reliability of the final estimate) <u>Analysis of service data:</u> ▪ Is it likely that the ratio of under-fives to the general population has remained the same as before the crisis? ▪ How have children under-5 been identified? (vaccination campaigns using a door-to-door approach would be the most exhaustive and accurate method, as compared to community mobilisation for height measurement or malnutrition screening) ▪ How was the age of children determined? (confirmation of age by birth certificate or vaccination card is more accurate than caregiver report or height measurement) ▪ How many sources of service data were used to estimate the population size? (a combination of sources would increase robustness by allowing for triangulation)