|  |  |  |
| --- | --- | --- |
| logo msf | Insert here the logo of the local partner institution | logo epicentre |

Retrospective mortality survey in [details where the survey takes place, i.e. name of refugee camp, district/province, region, country]

Survey protocol

[dd.mm.20XX]

[Version, i.e. first version or revision after comments from the co-investigators and/or ethical review board]

[Name of principal investigator, degree i.e. PhD/MSc]

This template was edited by Francesco Grandesso (Epicentre) from a previous version prepared by Sibylle Gerstl with input from Ruby Siddiqui, Jane Greig, Philipp du Cros and Bhargavi Rao (Manson Unit, MSF-OCA) and Annick Lenglet (Public Health Department, MSF-OCA). This work was informed by reference to several retrospective mortality survey protocols from MSF and Epicentre. This new version received comments from several field epidemiologists of MSF and Epicentre.

This template is to be used as a framework to edit your contextualized mortality survey protocol. It was pre-approved by the MSF Ethic Review Board. However, bear in mind the following:

1. The contextualized protocol must be approved by the Director of the Medical department of your MSF Operational Centre, who should confirm the exemption of ethics review.
2. The contextualized protocols should be presented in track changes vs the template protocol (or highlighted changes vs the standardized survey), to facilitate the task of those who should confirm the exemption of ethics review.
3. In case of doubts about whether any changes (to the methodology, data management, ethics aspects, modules, etc.) are substantial, the Ethic Review Board can be consulted for a rapid opinion on whether the exemption is applicable.
4. Enquire in advance whether the contextualized protocol must be approved by the country Ethics Committee.

While editing the contextualised protocol, please note:

* Blue text in square brackets [ ] indicates the places where you have to fill-in, add or remove context-related information for the survey
* Red and italic text – like the one you are reading – provides instructions on how to complete the protocol and is to be deleted before submission.

**Also remove this page entirely after having finalised the contextualized protocol.**

|  |  |
| --- | --- |
| **Version** | [Number and date of the version in DD.MM.20YY format]*Mark the version by adding here a comment, i.e. revised version after comments from the co-investigators and submitted to the ethical review board or version approved by ethical review board. You may also keep the track by listing here all the versions that the protocol went through with description of the steps.*  |
| **Survey design** | Retrospective survey with recall period |
| **Survey period** | [Month(s) of estimated field activity of the survey] |
| **Survey site** | [Detail where the survey takes place i.e. region, district/province, country] |
| **Principal investigator** | [Name of principal investigator (qualification)][Institution, Town, Country][email] |
| **Study coordinator/Senior investigator** | [Name][Institution, Town, Country][email]*Responsible for ensuring that research achieves its expected impacts and is appropriately disseminated and translated into materials**In some circumstances, the principal investigator may act as study coordinator*  |
| **Local Partner Institutions** | Often the Ministry of Health of the country or/and a local NOG where the survey takes place are local partner institutions. |
| **International Partner Institutions** | International Partner Institutions could be universities and academic institutions of the country where the survey takes place and local Field Epidemiology Training Programs. |
| **Co-investigators** | [Names of all co-investigators (qualification)][Institution, Town, Country][email]List one by one all persons who are involved in carrying out the survey. Usually co-investigators are:- the health advisor/Medical referent at MSF headquarters- the medical coordinator of the mission where the morality survey will be carried out, and- a contact in the Ministry of Health of the country where the survey will be carried out |
| **Protocol development and survey design** | List here the name of persons participating in writing the protocol with his/her affiliation(s).This is always the principal investigator and usually one to two more persons who contributed substantially to the writing of the protocol. |

**CONTENTS**

To update this table, right-click on the text and select “Update Field”

[List of abbreviations 7](#_Toc108448920)

[1 Introduction 8](#_Toc108448921)

[1.1 Context 8](#_Toc108448922)

[1.2 MSF presence in the country 8](#_Toc108448923)

[1.3 Justification for the survey 9](#_Toc108448924)

[2 Objectives 11](#_Toc108448925)

[2.1 Primary objective 11](#_Toc108448926)

[2.2 Secondary objectives 11](#_Toc108448927)

[3 Methods 13](#_Toc108448928)

[3.1 Survey design 13](#_Toc108448929)

[3.2 Survey area and recall period 13](#_Toc108448930)

[3.3 Survey population 13](#_Toc108448931)

[3.4 Definitions 14](#_Toc108448932)

[3.4.1 Definition of household 14](#_Toc108448933)

[3.4.2 Definition of head of household 14](#_Toc108448934)

[3.4.3 Definition of member of a household 15](#_Toc108448935)

[3.5 Recall period 15](#_Toc108448936)

[3.6 Inclusion and exclusion criteria 16](#_Toc108448937)

[3.7 Sampling 17](#_Toc108448938)

[3.7.1 Sample size 17](#_Toc108448939)

[3.7.2 Sampling procedures 21](#_Toc108448940)

[3.8 Procedures 26](#_Toc108448941)

[3.8.1 Introduction of the survey to the community 26](#_Toc108448942)

[3.8.2 Introduction of the survey to the households 27](#_Toc108448943)

[3.8.3 Interview of household members 28](#_Toc108448944)

[3.9 Data collection and management 29](#_Toc108448945)

[3.10 Data analysis 30](#_Toc108448946)

[4 Ethical considerations 34](#_Toc108448947)

[4.1 Ethical references 34](#_Toc108448948)

[4.2 Ethical approval 34](#_Toc108448949)

[4.3 Administrative endorsement 34](#_Toc108448950)

[4.4 Verbal consent 34](#_Toc108448951)

[4.5 Risks and benefits for participants and contingency plans 35](#_Toc108448952)

[4.6 Confidentiality 36](#_Toc108448953)

[5 Collaboration 36](#_Toc108448954)

[6 Dissemination of results 37](#_Toc108448955)

[7 Field implementation 37](#_Toc108448956)

[7.1 Selection and tasks of the interviewer teams 37](#_Toc108448957)

[7.2 Supervision 38](#_Toc108448958)

[7.3 MSF field support 39](#_Toc108448959)

[7.4 Training and pilot phase 39](#_Toc108448960)

[7.5 Safety and security of interviewers 40](#_Toc108448961)

[7.6 Timeframe in the field 41](#_Toc108448962)

[8 Budget 43](#_Toc108448963)

[9 References 44](#_Toc108448964)

[10 Appendixes 45](#_Toc108448965)

[10.1 Appendix 1: Information sheet 45](#_Toc108448966)

[10.2 Appendix 2: Interviewer consent declaration 48](#_Toc108448967)

[10.3 Appendix 3: Basic questionnaire: mandatory questions 49](#_Toc108448968)

[10.4 Appendix 4: Additional questions by topic 51](#_Toc108448969)

[10.5 Appendix 5: Interviewers training program 53](#_Toc108448970)

# List of abbreviations

Here below an example of most frequently used abbreviations in such protocols. To be updated if needed

|  |  |
| --- | --- |
| 95% CI | 95% Confidence Interval |
| CMR | Crude Mortality Rate |
| IPTp | Intermittent preventive treatment [for malaria] in pregnancy |
| MoH | Ministry of Health |
| MSF | Médecins sans Frontières |
| MSF*-[OCX]* | Médecins sans Frontières – Operational Centre *[MSF section]* |
| U5MR | Under 5 Mortality Rate (mortality rate in children under 5 years of age) |
| WHO | World Health Organization |

# Introduction

## Context

This chapter is very context related; however, it usually should include:

* Special characteristics of the country relevant for the context of the survey, such as ongoing crisis and conflicts, availability of natural resources (e.g., oil, gold, diamonds), information on infrastructure (education, roads, health system), key players present in the country and region.
* Population figures of the country
* Demographic data, such as life expectancy at birth stratified by sex, maternal mortality ratio, crude and under-five mortality rate, and principal causes of death. You may have your own sources of data for the area that the survey may take place. Useful background information may be available on dedicated websites such as <https://ourworldindata.org>. Remember to cite the source and the date when the data were accessed in the format [accessed on dd-mm-yyyy].

Finalise this chapter with a map of the country. Useful maps can be found at:

* <http://www.un.org/Depts/Cartographic/english/htmain.htm>
* <https://gadm.org/maps.html>
* <https://mapcentre.msf.org>

You may want to create a specific map. Here some links:

* <https://gadm.org/data.html>
* [https://www.openstreetmap.org](https://www.openstreetmap.org/)
* <http://www.diva-gis.org/gdata>

Consider contacting the MSF GIS Unit based (<https://geo.msf.org/help/contact>) if a convenient map is already available of if the unit can make one for you.

Note that in some contexts the country borders may be under controversies.

|  |
| --- |
| Insert here the map image |
| **Figure 1.** [map Title] |

## MSF presence in the country

This chapter should include:

* The name of the region where MSF is operating in the country
* The year and month when MSF started its activities
* An overview of the different MSF programmes/projects in the region related to the planned mortality survey
* Statistics of the most recent “Monthly Medical Report” of the MSF project, such as
	+ Top diseases for admission to OPD and IPD usually stratified by persons aged 5 years and older and children under 5 years of age
	+ Inpatient mortality data (note that you cannot directly draw conclusions from the inpatient mortality to the population mortality in the community, as there are different contributing factors to these rates such as quality of healthcare; nevertheless, inpatient mortality is interesting to know in the planning of a community mortality survey)

Finalise this chapter with a map of the catchment area of the MSF project. Normally the catchment area is equivalent to the survey area chosen for the planned mortality survey.

|  |
| --- |
| Insert here the map image of the MSF catchment area |
| **Figure 2.** [Map title of MSF catchment area map] |

## Justification for the survey

This chapter usually should start by providing previous mortality estimate, if available:

* Previous estimates of crude mortality rate (CMR) and under-five mortality rates (U5MR)
	+ in the same region in the past, including their references/sources
	+ over the last 5 years carried out in the country, including their references/sources
* Comments of the results of these mortality studies and if they have exceeded emergency thresholds; there might be locally adapted reference rates for CMR and U5MR.

The emergency threshold should be established considering the mortality baseline: A doubling or more of the baseline CMR or U5CMR indicates a significant public health emergency and requires an immediate response. In cases where there is no reliable baseline for the region of interest, baselines from national level or neighbouring countries can be used (e.g. in displacement crisis the baseline of the hosting country could be used). In the complete absence of a known baseline, generic threshold references can be used (see table 1 below).

The interpretation of the table 1 is as follows:

* Being a baseline value not available, we consider that the expected CMR of a population in a refugee camp should be below 1 death per 10,000 persons per day.
* Therefore, a CMR equal or above 1 death per 10,000 persons per day indicates a population in a refugee camp being in an emergency.
* Similarly, a CMR twice or more the emergency threshold value (i.e., a CMR of 2 deaths per 10,000 persons per day or a U5CMR of 4 deaths per 10,000 children under 5 per day) indicates the population in a refugee camp being in a critical situation.

|  |
| --- |
| **Table 1.** Emergency threshold values for CMR and U5MR when local, national, or neighbouring country mortality references are not available |
|  | CMR \* | U5MR \* |
| Industrialised countries | 0.5 | - |
| Developing countries | 1.0 | 2.0 |
| Stable refugee camps | 1.0 | 2.0 |
| \* Number of deaths per 10,000 persons per daySource: MSF: Rapid health assessment of refugee or displaced populations. Médecins Sans Frontières, Paris, France, 2006; UNHCR (United Nations High Commissioner for Refugees): Handbook for Emergencies. 3rd edition. Geneva, Switzerland, 2007 |

Moreover, and most importantly, this chapter must provide a problem statement with justification for the need to carry out a mortality survey. Here below some suggestions.

[Currently there is no regular assessment of community health indicators in the MSF catchment area, such as mortality.]

[It is not clear whether the [insert name of the project] project is identifying and responding sufficiently to the key health issues in the area.]

[There is a lack of data in the area where MSF is working, which we want to address with the present survey.]

Finally, an explanation of the expected use of the results of the survey should be given. Explanations could be as follows:

[The planned survey will therefore give results on the scale of the health status in the population.]

[This survey will provide results on retrospective mortality and the likely causes of deaths.]

[It will further provide results on…]

Insert expected results according to supplementary chosen specific objectives (refer to chapter 2.2)

[This information will help MSF to better target its medical programmes and advocacy strategies to the major causes of death.]

[This survey will also obtain data for advocacy purposes]

# Objectives

## Primary objective

* To estimate retrospectively the crude mortality rate (CMR) for the total population and the mortality rate for children under five years of age (U5MR) in the catchment area during the defined recall period.

## Secondary objectives

* To describe the population surveyed according to sex and age
* To obtain an indication of the major causes of death, as well as the age and sex distribution of the deceased

You may consider adding an additional secondary objective to estimate CMR and U5MR after having split the recall period for example into rainy and dry season, harvest and hunger gap periods, malaria season, meningitis season.

This might be an important objective. However, you need to be aware that splitting the recall period adds a layer of complexity in the survey, similar to the stratification. In particular, it is important to understand that splitting the recall period into two or more periods (1) will require the questionnaire to add questions on when deaths occurred, and (2) will lead to more imprecise estimates per split period. This second point is particularly important when one split of the recall period is very short. The shorter the recall period, the worse the precision of the estimates will be. As a rule of thumb, do not split the recall period into periods of less than 2 months. If in your survey precise estimates in a short recall period is an important objective, you should consider increasing the sample size. In this case, the estimation of the mortality in different periods is to become a primary objective if the survey.

* To obtain an indication on whether mortality is more associated with [rainy or dry season, harvest, or hunger gap periods] TO CHOSE ONE CATEGORY.

To avoid affecting the quality of the data collected on the main objective, it is recommended to add only a few specific objectives (and thus a few questions in the questionnaire) that go beyond the mortality part of the survey. Moreover, the selection of questions required to respond to the additional objectives should be discussed with the field and headquarters to see whether they fit in this mortality survey.

Common additional topics include proportion of malaria prevention during pregnancies, measles vaccination coverage, coverage and utilization of bed nets and health seeking behaviour. Modules on nutrition, violence, health seeking behaviour and malaria with a list of possible questions are provided in Appendix 4.

The following specific objectives could be added:

* [To determine the proportion of pregnant women (in the 2nd and 3rd trimester, age 15-49 years), in the survey population who received intermittent preventive treatment (IPTp) for malaria]
* [To determine the proportion of children (6-59 months) in the survey population who are vaccinated against measles]
* [To determine the proportion of persons (stratified by children under 5 years of age and pregnant women) in a household sleeping under long-lasting insecticide-treated bed nets (LLITNs)]
* [To determine the coverage of LLITNs in the households]
* [To facilitate improvement of the MSF activities in the catchment area by basing our intervention on the findings of the survey]

**A REMINDER:**

Around 20 minutes is an acceptable duration for an interview. Make sure you do not overload the mortality questionnaire with too many questions. The longer the questionnaire is, the lesser the quality of the data will be.[[1]](#footnote-2)

Refrain from including additional topics in the mortality survey, if separate nutritional and/or vaccine coverage surveys are planned.

If questions on violence are added, they need to be thoroughly discussed with the community representative, so that they are formulated in a form accepted by the interviewees.

# Methods

## Survey design

This is a retrospective mortality survey of a representative sample of the reference population.

The population sample will be selected using [simple random sampling] [systematic sampling] [spatial sampling] [one/two stage cluster sampling] [modified Epi-Method sampling].

Mention here briefly the methods. More details will be given in chapter 3.7.2. A combination of methods may be used. In rare contexts different methods need to be used according to the information available and logistic constraints.

If you opted to stratify the population before sampling, here is the place where you should mention how the stratification will be done.

## Survey area and recall period

The survey area will be the entire catchment area of [name of refugee camp, district/province, region, country] repeat here what was written in the title.

Indicate also if some areas of the MSF catchment area will be excluded and give the reason for exclusion, such as “inaccessible due to security problems”.

You can further insert here a quick overview of the typical climate of the region including annual average temperatures. If it is relevant for the results of the mortality survey, please indicate the months of rainy and dry season(s).

The recall period will include [months or seasons of the mortality survey] (see chapter 3.5)

## Survey population

The survey population will consist of all people living, or have lived, in the survey area during the defined recall period.

As a matter of clarity, if the survey population is a mix of resident and refugee/displaced population, you may add here above a sentence that states that the survey population will include the refugees or the displaced population that lived or is living together with the local/host population in the survey area.

Information on current population estimates is based on [source, usually the last national population census] data and an estimated annual growth rate of [insert percentage of growth rate for the country].

Population estimates are usually of limited availability or outdated. Very often the last population census on national level was carried out years before the time of the planned survey. Therefore, the population figures obtained may not entirely reflect the current population in the catchment area - a limitation of the survey that should be mentioned in the discussion chapter of the survey report.

Always try to get the most recent population figures of the survey area and add the annual growth rate for the country. If no annual population growth rate is available for the country, use standard annual growth rates for regions and continents, which can be found at: <https://www.unfpa.org/world-population-trends>.

If you opted to stratify the population or to use a cluster sampling (see chapter 3.7), you need to obtain break-down population figures at cluster level. If the required population figures are not available, you may consider to estimate them. If there are no up-to-date data available, you may consider to estimate them considering a similar context

You may also assess whether partial population estimates, or count were carried out through public health activities such as Expanded Programme of Immunization. If the under 5-year population estimate is available, the total population estimates might be extrapolated by multiplying by 5 the estimate of children under 5 years of age.

The total population was [population figures of the latest available census or any other population figure source] in [year of latest available census] [reference in brackets]. With an estimated population growth of [growth rate], the current survey population is estimated at around [current population figures].

It is important to mention in this chapter if there are groups of people (i.e., nomads) or people that should be in the sampling frame but are not accessible to the survey interviewers due to insecurity or logistic constraints (i.e., limited access due to poor roads or weather, or not reachable due to cultural and local constrains, etc.). Beside mentioning the reason for exclusion, explain here the consequences that such exclusion entails in the survey outcomes.

## Definitions

### Definition of household

A household will be defined as a group of people who [are under the responsibility of one person or head of household] [slept under the same roof the previous night] [ate together the previous night].

Select one definition only, in the square brackets. Other definitions may be considered depending on the context and country culture.

### Definition of head of household

The head of household is defined as follows:

* Adult household member [age that corresponds to local laws or cultural norms e.g., >18 years]*, and*
* Can give accurate information on all demographic and mortality issues in his/her household (can describe with reasonable accuracy the events that occurred during the recall period), or can authorise some other household member to provide the information, *and*
* Has lived in the household the entire recall period, *and*
* Is present at the time of the survey

A household will be excluded from the survey if none of the household members fulfil all these criteria.

### Definition of member of a household

A member of the household is defined as a person who is or was part of the household according to the household definition during the recall period. This includes members of the household who died or moved out but was a member of the household at any time during the recall period.

## Recall period

The recall period comprises from [date] to [date] corresponding to [number of days]days.

The recall period is a critical aspect of the survey because the persons interviewed will have to accurately place in time some event such as arrivals departure births and deaths relative to the recall period of interest. Therefore, the start of the recall period should be an easily remembered point in time, such as a large festival (e.g., Christmas, Easter, beginning or end of Ramadan) or holiday (e.g., Independence Day) or a date on which something memorable occurred in the survey area[[2]](#footnote-3).

Most populations may not be oriented enough to a western calendar to easily think of a date that occurred weeks or months ago. Therefore, you should not use arbitrary time periods (months, weeks), but work with events of which everybody in the survey area is aware. It is very helpful to generate an events calendar for the chosen recall period in order to determine more accurately the time of the occurred deaths. An events calendar may also help to avoid biases in the estimates because people generally tend to remember traumatic events as having occurred more recently than they really did.

Here below the main principles on how to decide the length of the recall period:

* The length of the recall period must be carefully defined so that to find the best trade-off between two factors: the accuracy of the information and the statistical precision of the outcomes.
	+ The accuracy of the information relies on the capacity of the interviewees to place the events (the deaths of a family member) into the correct point in time. In this scope, information of a short recall period is more accurate. Conversely, a long recall period is prone to recall bias (also known as telescoping effect); interviewees recall recent events as more remote than they were, and remote events as more recent. Interviewees may place therefore events within the recall period while they were not, or vice versa.
	+ Because the denominator of mortality rates is person-time, the longer the recall period, the more person-time units are in the denominator of the mortality rate. In this scope, a longer the recall period, therefore, provides a better statistical precision of the estimates if the same number of households are selected for the survey sample (see also chapter 9).
* In most acute humanitarian emergencies, a morality estimate over a period of around the last 3-5 months is much more useful than a long or an older one.
* Investigators might opt for a recall period longer that 3-5 months; however, they must provide a sound reason of a longer recall period and keep in mind the risk of the telescoping effect (events recalled being occurred within the recall period while they were not, or vice versa). In any case it is highly recommended to set a recall period no longer than 12 months.
* The outcome of the survey is the average mortality rate during the recall period. If the mortality is expected to vary within the recall period, investigators may consider to divide the recall period into 2 or 3 segments separated by memorable event dates. Investigators, however, are to be aware that this will complexify the questionnaire and the interviewing procedures and therefore, should provide sound reasons to do so.

After the decision of the length of the recall period for the survey is made, some explanations should be given. Possible explanations are:

This recall period comprises both the rainy season ([months of season, number of days] days) and the dry seasons ([months of season, number of days] days).

With a chosen recall period we will be able to compare CMR of previous studies carried out in [year in which prior mortality studies were carried out].

The exact beginning of the recall period will be discussed with the team in the field, especially taking into account the experience of the national staff. The end of the recall period will be the day before the start of the interviews in the field.

Together with the field team an events calendar will be generated for the chosen recall period to determine more accurately the dates the deaths occurred.

Ideally the exact date of death should be recorded. However, in many survey areas it is not feasible to record the day of death, especially when a death occurred at home and not in a health facility. Nevertheless, it should always be tried to estimate the month in which the death occurred.

## Inclusion and exclusion criteria

A person will be included in the survey if s/he satisfies all the following criteria:

* Is or was a member of the selected household during the recall period, *and*
* Informed consent has been given by the head of the household (see chapter 4.4 for details on the informed consent process) and by the person being interviewed.

A person will be excluded from the survey if s/he satisfies one of the following criteria:

* Refusal to participate in the survey, *or*
* Inability to locate the participant after two attempts to trace him/her, if no one in the household was entitled to answer on his/her behalf.

Specific situation when we want to survey only certain groups: For example, if we want to determine mortality only among displaced population, we should exclude native population.

There definition of head of the household is context/cultural related and is an issue that is to be discussed during the interviewers training. Related discussion will be on how to deal situations when more than one adult is qualified as head of the household, and how to proceed when the head of the household is not of major age (see chapter 7.4 for more instructions).

## Sampling

### Sample size

Sample size was calculated with the help of [name of software]*.*

Examples of useful software include the online support OpenEpi (1) and ENA for SMART 2020" freeware(2). In the “Planning” page of the ENA software there is a section called “Sample size calculation”.

The criteria listed in Table 1 were taken into consideration for the calculation of the sample size.

|  |
| --- |
| **Table 1.** Criteria for the calculation of the sample size, [name of the MSF project, region, country, year] |
| **Criteria** | **1st period** | **2nd period** |
|  | **CMR** | **U5MR** | **CMR** | **U5MR** |
| Expected mortality of 10 000/day |  |  |  |  |
| Precision of 10 000/day |  |  |  |  |
| Recall period in days |  |  |  |  |
| Design effect |  |  |  |  |
|  |  |  |  |  |
| Total population to be sampled |  |  |  |  |
| Total households to be sampled |  |  |  |  |

The table above considers the recall period split into two periods. Remove the columns of the second period, if there is a unique recall period, or include more columns if the recall period is split in more than two periods.

Rename the headings 1st and 2nd period with more meaningful terms, e.g., rainy season of XX numbers of days (1st period) and dry season of YY numbers of days (2nd period).

Four basic parameters are needed to estimate on the sample size for the mortality survey:

* The expected point estimates of the mortality rates (one for CMR and one for U5MR)
	+ The point estimates (or the expected value) can be based on mortality studies carried out in the country prior to the planned survey and/or on general reference rates for CMR and U5MR (see Table 1)
* The desired precision around the point estimates
	+ The precision defines the distance of the lower and the upper limits of the confidence interval for a given alpha error (see point below). For instance: if the expected CMR point estimate is set at 0.5 deaths per 10000/day, and the precision at +/- 0.3 per 10000/day, the lower and the upper limits of the confidence interval are 0.2 and 0.8 deaths per 10000/day.
	+ If the current survey will compare results from previous studies, it is very important that the expected confidence intervals do not overlap.
* The desired alpha error (or desired level of confidence around the estimates). It is commonly accepted to set an alpha error or 0.05, which means a z of 1.96.
* The recall period in days
	+ Refer to chapter 3.5 on how to decide on recall periods.
	+ However, it is important here to understand the effect of the length of the recall period on the sample size. The denominator for the estimates is in person/days. This means that, for a given sample size of number of persons, the number of person/days will big if the recall period is long, and, conversely, little if the recall period is short. For instance, for a sample size or 400 persons and a recall period of 180 days, the denominator will be 400\*180 = 72,000 person/days. For the same sample size, and a recall period of 60 days, the denominator will be 400\*60 = 24,000 person/days.

All the above factors have an impact on the sample size. Here below indications on the effects of these factors on the sample size:

* The alpha error and the precision are the two factors that greatly drive the size of the sample.
* While the alpha error is normally set to 0.05, investigators should work out on the best choice for the precision and the recall period.
* Even a slightly better precision (for instance from +/- 0.3 to +/- 0.2) will increase considerably the sample size.
* If the recall period (or a split of the recall period) is short, investigators need to increase the sample size or accept a lower precision.

**Here the formula for estimating the sample size in terms of person/days**

$$Total sample size (in person/days)=\frac{\frac{espected number of deaths per day}{10000}}{\left(\frac{desired precision}{\frac{10000}{z}}\right)^{2}}$$

To calculate the number of households needed for the survey, the number of the population to be sampled must be divided by the average number of members per household and/or the estimated number of children under 5 years of age per household (if only U5MR is being estimated). Software like ENA does these calculations automatically. Use local estimates if available. The national or local offices of statistics have often these estimates; if not, you may need to use your best guess.

Usual estimates in African developing countries are:

* 5-6 members per household
* 1 (or, to be more conservative, 0.8) child under 5 years of age per household

At this stage it is also important to consider that some households selected will not be available or willing to respond. Also consider that, if you are using a spatial sampling to select roof, some roof may not correspond to a household (a house where a family lives, but rather a shop or a school, etc. The sample size needs to be increased of the expected percentage of non-respondents or roofs not corresponding to a household. This value is in general 10-20%.

If the cluster design sampling was chosen, two additional parameters are required:

* The number of clusters
	+ It is highly advised to opt for no less than 30 clusters (3)
* The design effect (or DEFF)
	+ The DEFF is an effect that occurs when another procedure than the simple random sampling is used. In a cluster sampling, the DEFF considers the fact that individuals (or households) living close to each other are more similar than individuals (or households) living far away on a random distance.
	+ Taking the example of DEFF in a mortality survey, it is possible that the population in some clusters experienced a higher mortality than the population in other clusters. This may occurs among displaced populations, or when a conflict affected only some villages and not others(4).
	+ Technically speaking, the DEFF is a measure of the heterogeneity of the estimates between clusters. The DEFF is measured during the analysis and is to be used to adjust the precision of the estimates. A DEFF higher than 1 leads to loss of precision in the estimates. Therefore, to keep the desired precision, it’s necessary to anticipate the expected DEFF to recalculate the sample size.
	+ The expected DEFF is a multiplying factor to calculate the sample size. For instance, for a calculated sample size of 400 persons and a DEFF of 1.5, the sample size including DEFF will be 400\*1.5 = 600 persons.
	+ For cluster survey, an estimated DEFF of 2 has been often used (5), but may be set at 3 or 4, or even higher, if investigators expect a high heterogeneity of the estimates between clusters.
	+ To note that, following the logic, a DEFF of 1 means no loss of precision as for a simple random sampling.

The number of clusters and the DEFF are connected. A sample with a higher number of clusters is frequently (though not always) expected to result to a smaller DEFF. The gain in the reduction of the DEFF, however, fades progressively with the increasing of number of clusters, so that at one point there is no more effect on the DEFF.

In the two-stage cluster sampling the increase of the number of clusters will result automatically to the decrease of number of units per cluster. For instance, for a given sample size of 900 units (individuals/households), the investigator may opt to set the sample in 30 clusters of 30 units or in 60 clusters of 15 units.

**Example of sample size calculation (expected mortality of 0.5/10000/day, precision ±0.3, alpha error of 0.05 (z = 1.96), recall period of 90 days, average number of 5 members per household, 30 clusters with design effect of 2)**

**Basic formula:**

$$Total sample population=\frac{\frac{0.5}{10000}}{\left(\frac{0.3}{\frac{10000}{1.96}}\right)^{2}}=213 422 person/days$$

**With a recall period of 90 days (3 months):**

$$Total population sample= \frac{213422 person/days}{90 days}=2371 persons to be interviewed$$

**With an average number of members per household of 5:**

$$Total household sample=\frac{2371 persons to be inteviewed}{5 persons per household}=474 households $$

**With an expected 10% of non-respondents:**

$$Final households sample=474\*1.1=522 households$$

--------------------------------------------------------------------------------------------------------

**For a cluster design sampling, with an expected design effect of 2 in 30 clusters:**

$$Total households sample=522\*2= 1044 households$$

$$Number of households per cluster=\frac{1044}{30}=35 households per cluster$$

Based on an expected CMR of [expected number of deaths]/10,000/day, a precision around the estimate of [desired precision] with 95% confidence interval (95% CI; alpha error = 0.05), a number of [number of person-days] person-days were estimated to be required to deduce CMR. With the length of [number of days of the recall period] days for the recall period and an average number of [average number of household members] members per household [or the estimated number of children under 5 years of age per household use this last sentence if only U5MR is being estimated], and an expected [percent of non-respondents] % of non-respondents, a total of [number of households] households are to be sampled.

Add the sentence below if a cluster design sampling was selected.

With a design effect of [design effect value] for a total of [number of clusters] clusters, the total number of households to be sampled raised to [number of households with design effect] households; this means [number of households per cluster] households per cluster.

### Sampling procedures

The section below lists all recommended sampling procedures. The choice of the procedures depends on many factors including the information available, the human resources deployable, the logistic constraints and the timing. In any case the choice should always fall for the procedures that allow the most representative sample of the population.

You may opt for two or more stages procedure with different sampling procedures in each stage. It is very popular, for instance, the two-stage cluster sampling.

In rare circumstances (for example in a stratified sampling when different information for each stratum were available) more than one sampling procedure may be chosen.

The list below is ordered on the theoretical capacity to provide the most representative sample of the population.

1. **Simple random sampling of households**

All households within the camp/village [are/will be] marked with a unique identification number. From this list, households/shelters will be chosen randomly through a random number generator. Because household/shelters are marked clearly, the survey teams can easily find their assigned houses.

Simple random sampling requires the list of all units (sampling frame) from which sample of units is randomly selected. It is the best choice, when feasible; in complex emergencies, however, there is often no comprehensive list of individuals or households, so this method is rarely feasible. It may be used in small areas with clear geographic boundaries where all households or individuals may realistically be identified. For this, it could be the choice for the second stage of a two-stage sampling.

1. **Systematic sampling of households**

Sampling units are arranged in rows, households will be selected starting from one point (chosen at random) at one end of the village/camp and then systematically choosing subsequent households at regular intervals (sampling interval) until the end of the village/camp.

Systematic sampling is adapted to well defined, organized, ordered populations; it is feasible if a local guide with good knowledge of the place is available. Useful method once population size increases but population is still ordered/organized. However, the number of total eligible units needs to be reliable (to calculate the sampling interval) and to avoid that otherwise some areas are over- or under-represented.

As for the simple random sampling, the systematic sampling could be used in the second stage of a two-stage sampling.

1. **Spatial sampling**

This method uses spatial data (i.e., satellite imagery or aerial images and GPS coordinates) and may be applied in four different procedures, according to the information available or the capacity to collect the necessary information on due time.

* 1. ***The roof simple-random sampling***

The sampling procedures will be based on a sampling frame corresponding to the list of coordinates of all roofs in the area of interest taken from a recent [satellite | aerial] image. A sample of roofs will be randomly selected from the coordinates sampling frame. The source of [satellite | areal] images was [source of images i.e. Google Earth, Bing …] and the images were dated from [earliest date] to [most recent date].

Selected roofs will be found with the help of a GPS device.

All roofs need to be pinpointed beforehand to create the sampling frame, and a sample of roofs is randomly selected. Be aware that pinpointing all roofs may be time consuming. Coordinates might have been already collected for other purposes (see OpenStreeMap <https://www.openstreetmap.org/>).

With recent images and with the condition that a roof corresponds to a unique sampling unit (a household) and vice versa, this method is comparable to the simple random sampling. This and the rooftop points option (see below) are the two options that best respect the geographical distribution of the population.

This method requires recent satellite images of survey area. Moreover, you may need to verify any recent population movement that may not be captured in the most recent satellite images.

Depending on the context you may need to consider specific situations: many households may be empty, in camp situation or areas with apartment buildings more than one household may live under one roof and this information should be obtained for further weighting, some roofs might belong to business and not households etc.

* 1. ***The rooftop points option***

The sampling procedures will require a preparatory step to define the area of interest with one or more geo-referenced polygons. Random coordinates will be then drawn within the polygon/s and superposed to remote sensing [satellite or areal] images. The coordinates corresponding to a roof will be kept and the households living under those roofs are those selected. Coordinates not corresponding to a roof will be discarded. The source of [satellite or areal] images was [source of images i.e. Google Earth, Bing …] and the images were dated from [earliest date] to [most recent date].

Selected households will be found with the help of a GPS device.

The process of keeping coordinates that correspond to a roof, and discarding the others, allows to respect the spatial distribution of sampling units. With recent image and with the condition that a roof corresponds to a unique sampling unit and vice versa, this method is comparable to the simple random sampling. However, it may be time-consuming to set-up the unit selection, as many coordinates randomly drawn need to be discarded because not corresponding to a roof.

As for the point census option, this method requires recent satellite images of survey area. Moreover, you may need to verify any recent population movement that may not be captured in the most recent satellite images.

* 1. ***The*** ***rooftop points with radius option***

The sampling procedures will require a preparatory step to define the area of interest with one or more geo-referenced polygons. Random coordinates will be then drawn within the polygon/s and superposed to remote sensing [satellite or areal] images. The coordinates at [distance in meters] meters maximum distance from a roof will be kept and the households living under those roofs are selected. The other points will be discarded.

Selected households will be found with the help of a GPS device.

This method is similar to the rooftop points option. The only difference is that random coordinates are kept if they fall within a predefined distance from a roof. This option, however, may introduce a selection bias, by giving sparse roofs/shelters higher chance to be selected than shelters close to each other; the longer the radius, the higher the chance to select isolated roofs/shelters.

The length of the radius depends on the context and especially on the dispersion heterogeneity of roofs/shelters. The higher the dispersion heterogeneity, the smaller the radius should be. Once the length of the radius is decided, it cannot to be changed.

* 1. ***The any points option***

The sampling procedures will require a preparatory step to define the area of interest with one or more geo-referenced polygons. Random coordinates will be then drawn within the polygon/s regardless their distance to any household/shelter.

Selected coordinates will be found with the help of a GPS device.

This can be considered the extreme variant of the rooftop points option as all randomly drawn coordinates are kept regardless their distance from a roof. This may be the choice when recent remote sensing images are not available, or when there is no possibility to verify the points by others means, or when you do not want to charge the interviewers of the task to measure the distance of the selected coordinate to the closest household/shelter (see below the any points with radius option). If the population is geographically unequally distributed (i.e. areas with different population densities), there is an important selection bias toward the isolated units. For this reason, this is to be considered as the least of the spatial sampling options.

Conversely it could be the choice at the only rare condition that all sampling units are equally distant to each other, and large empty spaces were removed from the sampling polygons.

* 1. ***The any points with radius option***

The sampling procedures will require a preparatory step to define the area of interest with one or more geo-referenced polygons. Random coordinates will be then drawn within the polygon/s regardless their distance to any household/shelter.

Selected coordinates will be found with the help of a GPS device.

During the field activity, interviewers will then look for the closest households/shelter at a maximum distance of [units in metres or steps] [meters | steps] from the coordinates. Interviewers will discard coordinates for which no household/shelter is within the maximum pre-set distance.

This is a variant of the any points option with the difference that the interviewers are in charge during the field activity to measure the distance of the selected coordinates to the closest household/shelter. If there is one household/shelter within the pre-set distance, the household is interviewed. If there is no household/shelter within the pre-set distance, the coordinates are discarded. This may be the case when recent remote sensing images are not available, but there is the possibility to verify the points in the field.

This option, although better than the any points option, bares the same drawbacks of the rooftop points with radius option. It may introduce a selection bias, by giving sparse roofs/shelters higher chance to be selected than shelters close to each other; the longer the pre-set distance, the higher the chance to select isolated roofs/shelters. The length of the pre-set distance depends on the context and especially on the dispersion heterogeneity of roofs/shelters. The higher the dispersion heterogeneity, the smaller the pre-set distance should be.

1. **Cluster sampling**

The cluster sampling refers when the target population at some point in the selection process is divided into groups or clusters (e.g., camp sectors, villages, neighbourhoods of a town, schools, etc.) and a random sample of clusters is selected. Within the selected clusters there is then the final selection of the sample units, household/shelters.

The cluster sampling design is considered more efficient than other random sampling designs particularly when the population is dispersed, as it allows grouping several respondents within relatively small areas (the sampled clusters). Nevertheless, because of clustering, there is a loss of precision in the estimates; a design effect, therefore, need to be anticipated and the sample size increased accordingly.

* 1. ***One-stage cluster sampling***

The population will be divided in clusters ([villages] [camp sectors] [others]). Cluster will be randomly selected and all households/individuals in the selected clusters will be interviewed.

If for unforeseen reasons a selected cluster cannot be visited, the cluster will not be replaced.

The procedure is defined as a “one-stage” cluster sampling when ALL units (i.e., households or individuals) in each selected cluster are sampled. This procedure is well adapted when the total number of units in each cluster is equal. If the total number is unequal, a weighting is to be used during the analysis, which is a common situation.

Most of the time, however, a fixed number of units are sampled. The process is defined as a “two-stage” cluster sampling.

* 1. ***Two-stage cluster sampling***

In the first stage the population will be divided in primary sampling units (PSUs) ([villages | camp sectors | others]). From the list of PSUs a number of them will be randomly selected with probability proportional to population size. In the second stage, within the selected PSUs, a sample of units is selected from the list of secondary sampling units (SSUs) ([households | individuals]). The SSUs selected for the interviews is what we will call cluster sample. The size of the cluster sample therefore ([number of households | individuals]) will be the same everywhere. The SSUs will be randomly selected using a [simple random sampling] [systematic sampling] [spatial sampling] procedure. If for unforeseen reasons a selected unit in the cluster cannot be visited or interviewed, the unit will not be replaced.

The two-stage cluster sampling design is a good option for large, non-organized, non-ordered populations.

The selection of the PSUs according to the probability proportional to population size allows to a self-weighting so that all SSUs have equal, non-zero probability to be sampled. This method, therefore, relies on accurate population breakdown estimates at the PSU level.

The fact that the sample is self-weighted means that, in theory, there is no need to use weighting during the analysis. Nevertheless, it frequently happens that some sampled units in a cluster end up not being interviewed, leading eventually to unequal cluster sizes. If this issue is thought to bias the results, the investigator should consider applying a weighting corresponding to the inverse of the actual cluster size during the analysis. In case of doubt, apply the weighting.

In the second stage sampling, another method like the simple random, systematic, spatial sampling, or, as a last resort, the modified EPI method can be used (see below).

* 1. **The 2005 standard WHO/****EPI cluster sampling procedure (****the modified EPI method)**

The 2005 standard WHO/EPI sampling procedure is carried out directly in the field and follows the following steps:

1. Identify the geographical centroid of the village/sector selected and choose a random direction by spinning a pen (or any object like a bottle that allows to indicate a direction)
2. Walk in a straight line until the boundary of the village/sector is reached
3. Spin the pen/bottle once again. Walk along this second direction, and tag households/shelters/tents on both sides (left and right) with an incremental number
4. Draw a random number between 1 and the number of households/shelters/tents counted
5. The households/shelters/tents tagged with the drawn number is the first household to survey
6. Continue to next household on the right until the total number of units in the cluster is completed.

If the size of the cluster is smaller than the number of sample units required, the interviewer team will continue by selecting the geographically closest village/camp sector.

This procedure was used often in the past as the second stage of a two-stage sampling process. Nevertheless, nowadays it’s to be considered as a last resort when other methods cannot be applied.

This procedure may be considered for use for large non-organized populations and lack of foreknowledge of the set-up of the geographic area (i.e., perimeter not defined in advance). It requires less “office preparation” for the investigator, but it is more demanding for the field teams, as this procedure is entirely carried out directly in the villages or camp sectors.

This procedure suffers of two major drawbacks. The first drawback is that it relies on the judgment of interviewers (or other persons in the field team) to identify the random starting direction and ultimately to choose the first and all the other units. The second and most relevant drawback is that by using the proximity criterion (“Continue to next household to the right”) the units selected are ultimately a cluster – a group of households close to each other – within the cluster. The selected units may therefore not be a good representation of the cluster.

These are the reasons why the modified EPI method should be reserved as a last resort when simple random, systematic, or spatial sampling procedures are not possible.

## Procedures

### Introduction of the survey to the community

During the preparation, the [position of person/s responsible for informing villages] will share information and discuss the study with key authorities and community leaders/ groups to ensure mutual understanding and the smooth and respectful implementation of the study. The [position of person/s responsible for informing villages] will also inform the administrative authorities and community leaders/representatives of the visit of the interviewer team of the selected administrative zone/village/camp sectors/etc. one or two days in advance.

Sharing information and discussing the study with key authorities and community leaders/ groups is key to ensure mutual understanding and the smooth, sensitive, and respectful implementation of the study. This will include:

* explaining in detail the study aims, processes, risks and benefits.
* giving people the opportunity to ask questions, discuss concerns, and provide feedback and suggestions for implementation.
* discussing initial plans for validating and disseminating findings.
* establishing platforms or channels for ongoing engagement during the study, as well as ways for people to share feedback or raise concerns during study implementation (including complaints and ‘red flags’ signalling the study should be put on hold).
* discussing plans for piloting of study tools, and opportunities of input/ adaptation.

To ensure that the survey population in the selected villages is present at the day of the planned interviews, it is highly recommended to inform the head of the village and members of the village at least one or two days in advance the arrival of the survey team. This can be easily organised with the help of motorcycle drivers bringing information letters to the head of the village, by phone depending on context. It is recommended to ask the motorcycle drivers for a confirmation of the reception of the letter to the head of the village. In the planning stage, efforts will be made to limit the time persons have to stay at home as much as possible and ideally to a maximum of one day, to reduce the potential personal costs.

The purpose of the survey will be explained with the help of an information notice to the administrative authorities and a [verbal/written] authorisation will be obtained before starting the interviews. The [position of person/s responsible for informing villages] will clearly inform the administrative authorities and community leaders/representatives that they are freely allowed to decline the participation without any consequences or penalty.

A written authorisation from the administrative authorities should be sought, unless unfeasible.

During the analysis you will need to mention any refusal (from the head of a village or the head of the household) as a limitation of the survey in the discussion chapter of the survey report. Therefore, investigators should take care to document any refusals. To know that administrative authorities as well as any eligible person to be interviewed may refuse to participate without giving any reason for their decision.

### Introduction of the survey to the households

In each selected household, the interviewer teams will explain the purpose of the survey to the head of the household in the language he/she is familiar with and [written or verbal]consent will be obtained to conduct the interviews (the consent form is presented in chapter 4.4)*.*

If the head of the household declines to participate, the interviewer team will record this information, thank the head of the household, and proceed with the following household selected.

If the head of the household accept to participate, the interviewer team will agree with the head of the household to identify a place where the interview will be carried out with an adequate confidentiality. The head of the household may propose another household member to respond to the question on behalf of all household members. The interviewer team will encourage the head of the household to discuss with the other adult members whether they agree he/she gives consent on their behalf, and whether anyone would like to answer to the questions by themselves. In any case, a parent/caregiver will respond on behalf of children. The interviewer team will then explain which information will be collected for each member of the household.

### Interview of household members

Interviewers will fill a questionnaire form for each eligible members in selected households.

If a household member will not be present at the time of the survey, and the head of the household will not want to answer to questions his/her behalf and will not delegate to another household member to do so, interviewers will return once later that day to interview the absentee.

If a selected household will not be available after two visit attempts (morning and afternoon), or is not willing to respond, that household will not be replaced.

The interview will be based on a standardised questionnaire that consists of the following sections:

* Age and sex of household members who had arrived, had left, were born, or died in the household during the recall period of the survey
* Educational status of these members
* Cause of deaths and time of deaths (e.g., rainy or dry season) for all deceased members of the household during the recall period
* Other additional information that was added in the secondary objectives (see chapter 2.2)

A questionnaire template is provided in Appendix 3.

[Modules on nutrition, violence, health seeking behaviour and malaria diarrhoea/cholera with a list of possible questions are provided in Appendix 4]

A specific training of the interviewer teams will be carried out to allow the interviewers to get familiar with the questions, the way to ask the questions and the mode to record them into the [paper or electronic device] questionnaire form.

Information related to mortality is sensitive. Interviewers must be aware that asking questions on this topic can provoke difficult conversations with an emotional burden for participants, with a risk of re-traumatisation. Interviewers themselves might be at risk. An understanding of the context and dynamics related to mortality is essential to understand how best to engage with communities to prepare for and implement the survey to minimise the risk of harm or distress for participants, the risk of recall and desirability bias, and optimise the accuracy of data.

Before starting with the individual questionnaires, it is extremely important to draw up the list of all household members for which a questionnaire is to be filled, to ensure that no member will be missed. When listing the members of the household, take special care to list:

* Persons that are not anymore household members, but were so at one point during the recall period; these persons may be:
	+ Someone who quit or disappeared
	+ Someone who died, including a child that was born alive and then died even the same day of birth.
* Stillbirths (a child that was dead at birth) are to be excluded

If relevant in the context, questions on socio-economic status, ethnicity and religion may be added. Risks and benefits should be carefully considered if these issues are sensitive in the local community.

No personal identified data (name, telephone number) will be collected during the survey, reducing the risk that participants will be identifiable after the survey has been completed. The address and/or the geographical coordinates that were used to sample the household will be kept and managed separately from the survey questionnaire (more details are given in chapter 4.6).

## Data collection and management

This section should always include explanation on how the following issues related to data collection and management will be addressed:

* In what format will questionnaire data be captured (e.g. paper, electronic)
* How will data security of the questionnaire be assured
* How will the data quality of the questionnaire be assured
* How will the questionnaire data be transposed into a database (from a paper form directly to an electronic format)
* What roles are involved in data protection (specify all roles from field to office)
* How will data security and integrity of the data base be maintained during the survey (e.g. password-protected database with access limited to coinvestigators and the medical coordinator)
* How will data security and integrity of the questionnaire forms and database be maintained after survey completion (e.g., paper questionnaire stored for 5 years under lock and key, electronic questionnaire stored password-protected in high-security MSF servers)

Two examples of text are presented below.

If a digital data capture is used instead of paper, use the following sentence.

[Data will be digitally recorded directly into an electronic support (tablet or telephone) by the interviewer team. A dedicated electronic template on [[Name and version of the software]] currently most used software are ODK (https://getodk.org), KoboCollect (<https://www.kobotoolbox.org>) and REDCap (https://www.project-redcap.org) will be prepared and tested during the pilot phase of the survey. Data stored in the electronic device are encrypted and transferred directly to a dedicated server located in [[name of the institution/organization/company that host the server]] premises in [[town and country where the server is physically located]]. The server in place is already set to protect the electronic records, and provision for technical support during the survey will be ensured.]

If a paper form is used for data capture, use the following sentence.

[Data will be captured by the interviewer teams into a dedicated paper form. Data will be then transferred into an electronic database using [[Name and version of the software]]. Paper questionnaires will be stored in a locked box during the time of data collection and later on locked cabinets in the office. Currently most used software are EpiData (<https://www.epidata.dk>), and REDCap (<https://www.project-redcap.org>). If possible don’t use Excel.

Data electronically stored in a computer will be protected with a password by the investigator.]

If data are stored in a server, replace the sentence above with the following

[Data will be electronically stored in a server located in [[name of the institution/organization/company that host the server]] premises in [[town and country where the server is physically located]]. The server in place is already set to protect the electronic records, and provision for technical support during the survey will be ensured.]

[Once completed the survey, the filled paper forms and the electronic database will be stored at MSF headquarter or country management level for 5 years. After 5 years the paper copies of all the questionnaire will be destroyed.]

Access to the electronic database [and paper form] will be restricted to the co-investigators of the survey and the Medical Coordinator. [After 5 years the paper copies of all the questionnaire will be destroyed.]

Beside the principal investigator and partners, deidentified datasets can also be accessed by the country Ministry of Health even if it is not a partner institution. Other local research partners may have access after the approval of the principal and partner institutions.

## Data analysis

Before any analysis, data will be checked for inconsistencies in data entry and responses.

Data analysis will be conducted using [name and version of the software] ([insert company and country that provides the software]).

Current commercially available data analysis software includes STATA (StataCorp, College Station, TX, USA, <https://www.stata.com>) and SPSS for Windows (IBM SPSS Statistics, IBM Corporation, USA, <https://www.ibm.com/analytics/spss-statistics-software>).

Free data analysis software includes EpiData Analysis (<http://www.epidata.dk>) and R (<https://cran.r-project.org>). In the <https://r4epis.netlify.app/surveys/> website there is a R template for mortality surveys. In addition, the ENA software mortality section may be used (<https://smartmethodology.org/survey-planning-tools/smart-emergency-nutrition-assessment/>).

If you opted for a cluster sampling, you must mention also how the cluster design will be taken into account. Choose one of the sentences below according to the software choice.

For STATA

[The cluster design will be taken into account by using the SVY prefix command with the cluster variable as the primary sampling unit]

For R

[The cluster design will be taken into account by using the survey package with the cluster variable as the primary sampling unit]

Please, note that EpiData has no functions to deal with cluster design and therefore is not to be used for cluster design surveys.

If you did not opt for a cluster sampling, delete the sentences above.

Categorical variable indicators (i.e. sex) will be presented as proportions with 95% confidence intervals (95%CI). Continuous variables (i.e. age) will be presented as means with associated standard deviations, if normally distributed, or as medians with associated 25th and 75th interquartile ranges (IQR), if not normally distributed. Some continuous variables may be summarised in categories (i.e. age-groups) and presented accordingly.

Estimates of actual design effect will also be calculated for each variable and those with effects greater than 1 will be reported. Where appropriate, differences in proportions will be tested using the Pearson χ2 test, differences in means will be tested using the Student t-test if normally distributed or the Wilcoxon rank-sum test if not normally distributed. Associated p-values (p) will be also presented.

Mortality rates during the defined recall period will be calculated using the following formula:

$$Mortality=\frac{total number of deaths during the defined period}{total number of person-times during the defined period}\*k$$

For the CMR the total number of deaths and person-times of the entire population will be used. For the U5MR only deaths and person-times of this specific age-group will be used in the formula. Ninety-five percent confidence intervals will be calculated and adjusted for the design effect.

Person-times can be expressed in days, months or years. The k is a conversion factor to better illustrate the estimates and is often related to the unit with which the person-time is expressed. Some common formats for mortality rates are reported in the table below:

|  |  |  |  |
| --- | --- | --- | --- |
| **Person-time unit** | **k** | **Expression** | **Comments** |
| Years | 1,000 | Number of deaths per 1,000 person-years | Most used by demographers to express annual mortality rates in stable populations |
| Months | 1,000 | Number of deaths per 1,000 person-months | Used in displaced populations after the situation has stabilized and mortality rates are calculated less frequently |
| Days | 10,000 | Number of deaths per 10,000 person-daysORNumber of deaths per 10,000 persons per day | Used in acute humanitarian emergencies when mortality is changing rapidly, and mortality rates are calculated daily or weekly |

The person-times will be expressed as [one of the expressions in the table above] and will be calculated with one of the two methods described here below:

1. By summing the exact number of time units that each household member passed during the recall period.
For the members who were present at the beginning and are still present at the end of the recall period the total time units of the recall period will be accounted.
For the members who were not present either at the beginning or at the end of the recall period (i.e. new-borns, new arrivals, deaths and departures) a fine calculation of the time units passed in the household will be counted (for example the person-times of a new born will be the number of days computed as the date of the survey minus the date of birth).
2. By calculating the average person-times over the entire recall period, using the following formula:

$$\begin{matrix}total\\person-time unit\end{matrix}=\left(\begin{matrix}total persons at the beginning of the recall period\\+ 0.5 of persons who joined \left[new arrivals+new borns\right]\\- 0.5 of persons who departed \left[departures+deaths\right]\end{matrix}\right)\*\begin{matrix}total\\time\\units\end{matrix}$$

The first method is highly recommended as it allows a more precise computation of the total person-times. Use the second method only when the first one cannot be computed. Be aware that the second method assumes that events (i.e. deaths) were evenly distributed over the recall period. If this is not so, the total person-times might be under- or over-estimated. You need to describe in the report which method was used.

If you opted for a cluster sampling or a stratified sampling design, consider the following during the setting of the analysis.

* You must instruct the software which sampling design was used. This can be done by a software specific function:
	+ STATA: svyset (<https://www.stata.com/manuals/svy.pdf>) for instructing the survey design, and use the svy prefix before standard commands like prop, mean etc. for the analysis.
	+ R: svydesign in the survey package (<https://cran.r-project.org/web/packages/survey/survey.pdf>) for instructing the survey design, and use specific package functions like svymean, svytotal, etc. for the analysis.
* As a result, the software calculates the design effect and uses it to produce the confidence intervals. The DEFF is a measure of the inter-cluster heterogeneity and is expressed as the ratio of the actual variance to the theoretical expected variance as if the survey was carried out on a simple random sample of the population. The DEFF could be considered, therefore, as the excess of variance attributable to the design of the survey.
In this perspective, you may consider that:
	+ A DEFF of 1 means that the actual variance is the same of the theoretical expected variance
	+ A DEFF greater than 1 means that the actual variance is greater than the theoretical expected variance. This by far the most frequent effect.
	+ A DEFF smaller than 1 means that the actual variance is smaller than the theoretical expected variance. This is rare, but it might happen.
* As the DEFF is imputed to calculate the confidence intervals, a DEFF greater than 1 will enlarge the confidence intervals limits.
* If the actual DEFF is greater than the anticipated DEFF used to calculate the sample size, there will be ultimately a loss of desired precision in the estimates.
* In stratified sampling design you must use a weight, when the proportion of the sample in each stratum is not identical. The weight to be used in the inverse of the proportion.
Consider, for example, a survey with two strata of (A) 10,000 and (B) 50,000 population from which a sample of 200 individuals in each stratum were sampled. The weight will be 10,000/200 for the stratum A and 50,000/200 for the stratum B.
The weight can be set by the same instructions svyset for STATA and svydesign for R.
* You may consider using a weight also in a cluster sampling design when the anticipated equal size of clusters resulted to be unequal. In this case the weight would be the ratio of the expected cluster size to the actual cluster size. As for the stratified sampling design, the weight can be set by the same instructions svyset for STATA and svydesign for R.

# Ethical considerations

## Ethical references

The survey will be conducted in accordance with the Declaration of Helsinki (6), the International Ethical Guidelines for Epidemiological Studies (7) and the International Ethical Guidelines for Health-related Research Involving Humans (8).

These are the most updated references at the time of editing this template. Check whether these reference versions are still the most recent ones when writing the protocol.

The survey responds as well to the [research guidelines/…] of [Country] published on [date] (indicate here the reference of the document).

The Principal Investigator is responsible for ethical compliance of the survey.

## Ethical approval

The MSF Ethics Review Board approved the template that was used to edit this survey protocol. The MSF-[OCX] Medical Director determined that this protocol met the MSF Ethics Review Board’s criteria exempting it from further review by the MSF ERB.

This protocol will also be submitted to the [Name of relevant national or MoH ethics committee or institutional review board]of [Country where the survey will be carried out] for approval.

## Administrative endorsement

In some circumstances, where local administrative authorities are in place, you may need to add the following sentence. [The ethics approved copy of this protocol, together with an overall explanation of how the survey will be carried out, will be presented to the local administrative authorities.] In addition to the sentence above, local administrative authorities may require their written or verbal endorsement to carry out the survey. If this is the case, also add the following sentence. [and the survey will be carried out after their [written] [verbal] chose one endorsement.

The community key persons in the survey area such as village heads, religious leaders and opinion makers, will be informed about the purpose of the survey and their agreement will be sought. Add here how this will be done i.e., by organising a community meeting to present the survey. The awareness of the survey in the community should help improve survey content relevance and enhance security for both survey staff and participants.

## Verbal consent

A dedicated information notice (Appendix 1) translated into the local language will be read to and a verbal consent will be sought from every head of the selected households.

The head of the household may choose to delegate answering the questionnaire to another member of the household who is more informed regarding the questions that will be asked.

All participants included in the surveys will have the investigations explained to them in a language with which they are familiar. Everyone will be offered the opportunity to refuse participation in the survey at any time without penalty and no incentives or inducements will be provided to any respondents. Everyone approached for the survey is completely free to participate or not.

A copy of the information notice should always remain within the household after conducting the interviews.

The verbal consent will be documented in the [paper/electronic] questionnaire by an interviewer declaration that the verbal consent was obtained from the head of the household (Appendix 2).

When using this protocol template, verbal consent will usually be appropriate because the survey presents minimal risk of harm to participants and involves no procedures for which written consent is normally required. However, a written consent may be required by the country ethics committee. In this case, the text should be modified accordingly.

Moreover, although the consent is given verbally, the interviewers are requested to declare in paper or electronically that the head of the household, as representative of the interviewee, understood the information sheet and agreed to participate in the survey.

## Risks and benefits for participants and contingency plans

The retrospective mortality survey does not cause any physical harm to participants. Nevertheless, asking for details of recent deaths in the household may be upsetting and/or intrusive.

When undertaking and designing surveys consider the impact that such questions may have on participants. Train your interviewers on how to detect these issues and on existing referral pathways for mental health support. Identify any mental health support that may be available prior to starting the survey and facilitate referrals when operationally possible. If the mental health impact of conducting a survey is big and no mitigation strategies can be put in place you should reconsider the feasibility of the survey.

There is also the risk to communities of breach of confidentiality and/or stigmatisation at community level.

There will be no direct benefit for the participants, however, over the course of the survey, emergency cases identified during interviews will be referred to the nearest healthcare facility, and the fees will be paid by MSF.

In addition, the community may benefit from the survey results:

* A better understanding of the rates and causes of death in the area will facilitate more appropriate program development and a more efficient use of resources.
* Specific data regarding mortality and estimates of the causes of mortality are crucial for appeals at the national and international levels.
* The data collected will enable a better understanding of the results of the MSF interventions.

No financial or in-kind compensation will be provided in exchange for participation in the survey.

Benefits can be seen both at the survey participant level and at the community level. A better understanding of the rates and causes of mortality in the area will allow better tailored programming and more efficient use of resources. Accurate data on mortality and estimates regarding causes of mortality are of tremendous importance for advocacy on national and international level. Nevertheless, investigators must be aware that, if a high mortality rate was found, the community may be in the spotlight, and even stigmatised. All efforts are meant to ensure that the dissemination of the results favour the wellbeing of the community.

[An external threat could be…] insert here any external threat that could influence the field part of the survey and therefore the results.

Consider an additional review of confidentiality, safeguarding procedures and referral paths according to OCA recommendations [https://msfintl-my.sharepoint.com/:w:/g/personal/isidro\_carrion-martin\_london\_msf\_org/EU2S6hj0uklDocg7UsnlofIBDoDnXplZ9\_GiBvfvb9wzzg?e=PwSRe7](https://msfintl-my.sharepoint.com/%3Aw%3A/g/personal/isidro_carrion-martin_london_msf_org/EU2S6hj0uklDocg7UsnlofIBDoDnXplZ9_GiBvfvb9wzzg?e=PwSRe7).

If MSF (or other organisation) has already in place a mental health support service, consider to refer the interviewee or any person in the household to the mental health service.

## Confidentiality

Participants’ privacy will be respected during the interview process. Personnel will be trained in how to determine appropriate conditions for maintaining confidentiality during the interview process, including choosing the best location when the context makes it difficult to have privacy (for example, a single room dwelling).

Participant names will not be recorded on questionnaire, and individual person records will be linked only to a household number throughout the data entry and analysis process.

The confidentiality of the data collected from the participants will be ensured both during and after the survey. Individual responses will not be associated with a household number at any time during data entry and analysis. The electronic database will be password-protected. Any data that could be combined with other data sources to potentially make individuals’ responses identifiable will not be distributed outside of the survey site and will not appear in any report or publication.

Remove the paragraph below if a spatial sampling procedure was not used.

[The house will be identified by geographical coordinates, but the names of the households that participate will not be recorded in the questionnaire. All geographical coordinates and village names will be used for sampling purposes only. They will then be delated once the data have been collected, the database created, and the analysis completed. This will reduce the risk of participants being identified once the survey has been completed.]

# Collaboration

This survey will be carried out in collaboration between Epicentre, MSF-[OCX] and [local collaborators, such as MoH of country where the survey will be carried out, who may also be co-investigators]. The collaborating partners will try their best to involve the community in the preparation and implementation of the survey and the dissemination of the results.

Keep the following sentence below if Epicentre is the implementing agent.

Epicentre is the implementing agent and is in charge of carrying out in collaboration with MSF the entire survey. Epicentre is also in charge of the field activity supervision, the data analysis and of the report writing.

MSF-[OCX] is the survey sponsor and is responsible for the funding. Keep the following sentence if MSF is the implementing agent MSF is also in charge of the field part of the survey, the analysis and report writing. Permission for publication must be obtained from MSF-[OCX] and the MoH.

If necessary, a Data Sharing Agreement should be signed between MSF and the collaboration partners. This is necessary if the collaboration partners will have access to the individual participant data. Both MSF and collaborating partners are to be considered custodians of the data.

Survey results will belong to MSF-[OCX] and the MoH of [Country where the survey will be carried out], the country where the survey will be conducted.

If there are any further collaborating partners, they must be listed in this chapter and their tasks and responsibilities in the survey explained.

# Dissemination of results

MSF-[OCX] commits to sharing survey results with the community surveyed. The local community will be involved and informed through[complete with the dissemination strategy of the results, such as to distribute posters that show the survey results in the health clinics]. The MSF medical team will decide about the best venues to display the results.

Describe how findings will be disseminated: including translation of research into booklets or other advocacy materials as appropriate. Here below the list of recipients you shall consider:

* MSF – project, mission, headquarters
* Community and community representatives
* In country partners (including MoH)
* International dissemination (including WHO and other agencies, scientific publication)

Questions to consider when planning dissemination include:

* What is the most appropriate way to share findings with communities? For example, community meetings, posters, booklets, drama etc. Ideally these should involve the opportunity for two-way discussions, communities to share feedback and validations
* How might findings impact communities? This may include foreseeing mental health support for communities/specific groups to process death related to trauma.
* How do communities perceive/interpret findings?
* How do communities think findings would be shared further – with who – for what purpose?

# Field implementation

## Selection and tasks of the interviewer teams

Dedicated teams of two interviewers will be to collect the necessary data for the survey. We estimated that [number of teams] teams will be able to complete the data collection in [number of days] days (see chapter 7.5).

To better estimate the number of team it is advised to estimate the average time to complete one questionnaire and to consider the availability of transport in the field including security constraints.

If you opted for a cluster design survey, you may consider to plan the survey so that one team can complete one cluster per day. In any case do not let teams work more than 6 days in a row. Consider also having a back-up team to manage unexpected events.

Job application advertisement and distribution will be done under the national work regulations. Interviewers will be selected according to the following criteria:

* Able to read and write in [MSF local working language, such as English or French], *and*
* Fluent in the local languages [insert all languages spoken in the surveyed area], *and*
* Available for the entire time of the survey (training and interview days), *and*
* Motivated to participate in the survey, *and*
* Have no conflict of interest, *and*
* Experience with interviews in difficult settings and survey populations would be an advantage, *and*
* Expected to be accepted by the population

Make all efforts to ensure that selection of interviewers is independent, particularly when the context is politically sensitive or when there is a risk of manipulation, e.g., interviewers appointed or imposed by local leaders or officials may be problematic.

## Supervision

The principal investigator is the overall responsible for the final version of the protocol, overall quality of the survey and data analysis, and the final report.

The principal investigator, the study Coordinator and/or a senior investigator chose depending on the OC will ensure that the following tasks are performed:

These tasks should be carried out either through delegation or by the principal investigator him/herself (as s/he is the final responsible for them).

* Preparation of all necessary documents (protocol, [paper and/or electronic] questionnaire, informed consent forms)
* Secure the necessary ethics approvals (the exemption from the MSF ERB review from the MSF-[OCX] Medical Director and the approval of the country ethics committee)
* Preparation of the field component of the survey (training of the survey teams, logistics, materials) together with the MSF team in the field
* Follow-up of the field component of the survey
* Validation of data entry
* Data analysis
* Report writing
* Ensuring ethical compliance during implementation of the survey through supervision and training.

## MSF field support

* Administrative support for survey preparation at the field level and during field part, such as [support that will be required from the administration team on field level, such as presentation of the survey protocol to the relevant national or MoH ethics committee or institutional review board, payment of survey teams]
* Human resources support, such as [support that will be required from the HR team on field level, such as hiring survey team/interviewers, a translator for the principal investigator if needed]
* Logistic support for survey preparation at the field level and during field part, such as [support that will be required from the logistic team on field level, such as organizing sufficient cars including drivers for the field part of the survey, providing communication tools and MSF ID (e. g. aprons, vests or arm bands) to the survey teams, stationary, printing the questionnaire and consent forms].

## Training and pilot phase

A three four-day training will be given to all interviewers to familiarise them with the background of the survey, the questionnaire, the information sheet, and the informed consent procedures. The training will be given in [MSF local working language, with translation if needed] by the principal investigator. It consists of an intensive review of the questionnaire and the information sheet including role-plays. As the interviews will be held in the national language, the principal investigator should ensure that all interviewers are using the same and correct wording for providing information to the households and for the interviews.

The questionnaire will be tested in the field, in a village that is not selected for the survey, during a pilot day, and will be improved if needed. The selected village will be made aware of the survey the day before the teams arrive. Approximately 15 households will be tested by each team. Verbal consent for participation in the pilot will be sought in the same way as for the survey. The data collected during the pilot phase will not be included in the survey data and will be destroyed immediately following the day of the test.

A preparation meeting will be held with associated staff such as logisticians, community liaisons, drivers, and data clerks to explain the overall survey and their roles and expectations.

Carefully consider the number of days needed to for the training. And, most importantly, do not rush during the training, even if you are asked to carry out the survey as quick as possible. If you are in doubt between three and four days, chose four days. A one additional training day might help to avoid troubles and misunderstandings that will be more difficult to solve when the survey is on its way.

During the training, investigators must ensure that the interviewers go through the following points:

* Identification of the sampled households
	+ Training on the tool and standardising procedures to identify the household according to the sampling method
	+ What to do if the interviewers were not able to identify the household
* Introduction to the household and to the head of the household
	+ What to do if there is more than one head of household (consider there might be two different households, discuss how is culturally defined the head of the household)
	+ How to proceed with the verbal consent; to highlight that the interviewers shall let the head of the household to freely decide whether to participate or not; also reassure interviewers that the refusals to participate is not a marker of poor performance
* Identify a suitable location for the interview
	+ Suggest the head of the household that the interview might include sensitive question and that a quiet private place is suitable
* Interview
	+ Make sure that questions are posed in a standardised way. Although different interviewers might have different approaches, make sure that the ultimately the question is well understood by the interviewees
	+ Make sure that sensitive issues like deaths and violence are dealt with empathy and compassion; to explain that some interviewees might need more time to express themselves and might have different ways to do so.
* Farewell
	+ Make the interviewers understand that the farewell of the household, after finishing with the interview, is also very important. To thanks and praise their collaboration

The location for the pilot day should be easily accessible by all interviewers by foot. This allows the principal investigator to physically supervise all teams at least once during that day. Paper printing of the survey questionnaire should be done after the pilot day, as useful suggestions to modify or to better address the questions might be raised during the training and the pilot day. In addition, the information sheet/consent forms should be pre-tested to identify any problems with understanding or translation.

A training agenda template is provided in Appendix 5 of this document.

## Safety and security of interviewers

During the training, interviewers will be briefed on security, and will follow all along the field activity the MSF security rules that are implemented in the project.

The investigators will check with the MSF referent person, the security status every morning, before sending the interviewers to the field.

A communication system will be put in place so that interviewers will be reachable, while in the field, and can rapidly return to the MSF premises, if necessary.

## Timeframe in the field

The timeframe is always very context related. Consider the example below as a frame that surely needs to be adapted.

* 1-3 days for the principal investigator to reach the field
* 2-3 days for final preparation of the survey in the field, such as:
	+ defining the final survey area
	+ finalising and verify the feasibility of the sampling procedures
	+ finalising the frame of the recall period and the events calendar with experienced people who know the local context
	+ finalising the plan the survey in terms of human resources and logistics, and calculated the number of days to completed the field activity. The number of team
	+ to plan vehicle movements, to check materials for the survey, to organise photocopies of questionnaire and further required information, to define working conditions of the selected interviewers, such as working hours, per-diem (which usually should cover food and water during the time in the villages), payment
* 3 days training including 1 day as pilot day (a two-day training agenda template is provided in the Appendix 5)
* X number of days of data collection + one or two buffer days
* 1-3 travel days needed for the principal investigator to return from the field.

As an example of a two-stage cluster sampling survey one team is able to complete one cluster of maximum 30 households in one day.

For a survey of 30 clusters x 30 households with 5 interviewer teams it will require a minimum of 6 days to complete the data collection. It is advised to forecast 1-2 buffer days for unforeseen events including a debriefing at the field level to give a preliminary idea of the outcomes of the survey.

The transport of the teams is always a critical aspect of the logistics. The number of cars and the movement plan of the cars are strictly dependent on the existing safety rules. Where there is no security constraint, 2 or even 3 survey teams may be allowed to move with the same car. The teams will be dropped in their respective villages in the morning and then collected in the evening. In contexts with security issues, however, a car may need to stay constantly in the proximity of the team during the whole working day.

The tighter the security rules the more cars needed for the survey teams. This issue should be discussed in close collaboration with the MSF person responsible for security at the field level.

Supervision of the interviewers’ teams is crucial. As for the pilot day, a close supervision during the first days of the survey is also essential, so that mistakes or misunderstandings can be corrected as soon as possible. For this purpose, interviewers’ teams should have a supervisor (a supervisor can also supervise two or three teams depending on the context). For the first days of the survey is recommended that, the teams carry out the survey in clusters that are closer to the base, so the supervisors can easily move to the different team's locations and assure a close supervision of all teams.

A preliminary plan of the field part of the survey is presented in the table below.

|  |
| --- |
| Table 2.  Preliminary plan of the field part of the mortality survey, [Name of the MSF project, region, country, year] |
| **Date [dd/mm/20XX]** | **Nb of days** | **To do** |
| … | … |  |
|  |  | Travel days for arrival |
|  |  | Final preparation of the survey |
|  |  | Training including the pilot survey |
|  |  | Field part |
|  |  | Buffer days / debriefing |
|  |  | Travel days to return |
| … | … | … |
|  | **Total: XX days** |

# Budget

Add here a budget detailed with at least major expenditure categories:

* staff salaries or fees (local and international personnel)
* logistics (cars, etc.)
* other operating costs (stationary, administration etc.)
* external fees (i.e., some country ethics committees require paying a fee to examine the protocol)

etc.

# References

1. Dean AG, Sullivan KM SM. OpenEpi: Open Source Epidemiologic Statistics for Public Health [Internet]. Version. www.OpenEpi.com, updated 2013/04/06. 2013 [cited 2021 Mar 4]. Available from: http://www.openepi.com/Menu/OE\_Menu.htm

2. ENA for SMART [Internet]. 2011 [cited 2021 Mar 4]. Available from: http://www.nutrisurvey.de/ena2011/main.htm

3. Binkin N, Sullivan K, Staehling N, Nieburg P. Rapid Nutrition Surveys: How Many Clusters are Enough? Disasters. 1992 Jun;16(2):97–103.

4. Checchi F, Roberts L. Interpreting and using mortality data in humanitarian emergencies: A primer for non-epidemiologists. Humanit Pract Netw. 2005;

5. Kaiser R, Woodruff BA, Bilukha O, Spiegel PB, Salama P. Using design effects from previous cluster surveys to guide sample size calculation in emergency settings. Disasters. 2006 Jun;30(2):199–211.

6. WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects – WMA – The World Medical Association [Internet]. 2013 [cited 2021 Feb 23]. Available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

7. International Ethical Guidelines for Epidemiological Studies - CIOMS. 2009.

8. World Health Organization, Council for International Organizations of Medical Sciences. International ethical guidelines for health-related research involving humans [Internet]. 2017. Available from: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf

# Appendixes

## Appendix 1: Information sheet

A copy of the information sheet shall remain within the household after conducting the interviews. If necessary, the information sheet should be translated in the local language.

**Information Sheet**

**Survey title: Retrospective mortality survey in the MSF catchment area [details where the survey takes place, i.e., region, district/province, country]**

**Introduction**

Dear Sir/Madam,

Thank you for taking the time to listen to our information about this survey.

My name is ............... and I work for Médecins Sans Frontières, known also as MSF.

As you might already know, MSF jointly with [insert name of collaborators, such as the Ministry of Health] is providing health care in [insert name of area].

Your family was selected to participate in a survey to evaluate the health conditions of the population living in [insert name of area]. The survey is carried out from [date of start] to [expected date of end of the survey].

Before you decide whether to participate, it is important that you understand the contents of this information notice, which explains how we are carrying out the survey.

You can read this notice, or have it read to you, and we will explain it to you. Please ask us to stop as we go through the information if you have any questions. We will answer any questions you may have now or later.

**What is the survey about?**

With this survey we aim to knowing better the health problems in your community, so that we can improve our work as a medical organisation especially concerning [add focus of the project, such as health for women and children, malaria, primary-health care]. To do so, we will ask questions related to recent deaths that occurred in your family, including deaths that occurred because of tragic events. Secondly, MSF might use the results of this survey to advocate on behalf of the population at regional, national or international level in order to raise awareness of the situation in your community.

Your participation to the survey will consist of answering to questions such as:

* The number of persons who has been living in household during a certain period, including member who are not there anymore, either because they left, or died
* The relation of each household member to the head of the household
* The age and sex of each household member
* For those members who died since [beginning of recall period], the reason of death (if known)
* Insert questions that were added to the protocol

**Why is my household being proposed to participate?**

Ideally, we would like to speak with every household in the community, but this would require a lot of time. So, we opted to select only a few households in each village in a ‘random’ way, which means that they are selected by chance or coincidentally. And your household is one of those selected.

**What happens if I agree to let my household participate in the survey?**

If you agree, we will ask to you – or to a person that you designate – to answer to the questions we mentioned above. Note that the surname of your family nor the name of any family members will not be recorded anywhere. However, a unique survey ID, which is a number, will be assign to the household. This number will be used exclusively to carry out the data analysis.

You or another member of your family may find it distressing to respond to deaths. We ensure we will give you time, and we will stop the interview if you feel to do so. [We will refer you or your family member to our mental health support service if you wish so. The service will be paid by MSF].

Add the sentence above if there is a mental health support service run by MSF or other actors. If a non-MSF entity is running the service, the investigators shall agree with the service the referral and the payment procedures.

**Do I have to let my household participate?**

The participation in this survey is voluntary. If you do not agree to participate, we will not ask you any question and we move on with the other households that have been selected. Moreover, even if you agree, you can refuse to answer to some or all questions that will be asked.

**What are the risks and benefits of participation?**

As with any collection of data, there is always a risk of loss of confidentiality of personal data. However, our staff received a training to ensure the security and confidentiality of all survey documents.

Neither you nor your family will receive any direct benefit such as food or payment as a reward for participating. However, if someone is ill in your household, he/she will be referred [insert only one statement, as relevant: one of our health centres / mental health support for free treatment / the nearest health facility, which may not be free].

**How will confidentiality be assured?**

Information collected will be kept confidential and used only for the purpose of the survey. This means we will not ask you for the surname your household, nor the name of your household members and we will not record the location of your house, so it will be impossible to identify your house. In addition, all documents (both paper and electronic) will be kept in a secure location accessible only to team members. Only the survey location of [insert name of overall survey area] may be used in the survey report for advocacy.

After the survey is completed, all paper documents will be archived in a secure room at MSF in [capital of the country] for 5 years and then destroyed. Electronic data will be stored on a secure MSF server in the European Union. After 5 years the survey ID used as identifier will be deleted from the electronic database.

**How will the results be used?**

Once we have the results of this survey, we will [add the dissemination strategy of the results, such as to distribute posters that show the survey results in the health clinics].

MSF might also use the results of this survey to advocate on behalf of the population at regional, national or international level in order to raise awareness of the situation in your community.

Do you have any questions? Please feel free to ask us, we are happy to answer.

***[After answering the questions]***

In case you have any questions after the interview has been completed, you are free to contact the survey supervisor at the below information:

* Survey Supervisor: [Name of Survey Supervisor]
* Phone number: [Phone number of survey supervisor]

You can also contact the representative of the Ethics Committee at the below information:

* Representative of the Ethics Committee of [Name of the Country]: [Name of the Representative]
* Phone number: [Phone number of the Representative]

Make sure this is a phone number that is also accessed by somebody who knows the local language

## Appendix 2: Interviewer consent declaration

A paper copy of the interviewer consent declaration, certifying that the head of the household, as representative of the interviewee, understood the information sheet and verbally agreed to participate in the survey.

This document does not need to be translated into the local language.

**Interviewer Consent Declaration**

**Survey title: Retrospective mortality survey in the MSF catchment area [details where the survey takes place, i.e., region, district/province, country]**

I undersigned, \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, declare having administered the information sheet to \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, as representative of a household that was selected to participate in this survey.

Furthermore, I declare that he/she understood the content of the information sheet, that I answered the questions he/she raised, and that he/she agreed his/her household to participate in the survey.

Date: ¦\_\_¦¦\_\_¦ / ¦\_\_¦¦\_\_¦ / ¦\_\_¦¦\_\_¦¦\_\_¦¦\_\_¦

Interviewer’s signature:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Appendix 3: Basic questionnaire: mandatory questions

You may want to consider starting the questionnaire with an open-ended question to build relationship with the household. An example can be “What are the main challenges the household faces in the last year (or a chosen period)?”

Responses can be summarised in a couple of key words and categorised into themes and subgroups during analysis

|  |
| --- |
| **CORE QUESTIONS** |
| **Variable** | **Values** | **Comments** |
| Team ID | Team ID number |  |
| Date of interview | dd/mm/yyyy |  |
| Cluster ID | Numeric | Unique for the survey |
| Household ID | Numeric | Unique in each cluster |
| Household member order | Numeric | Unique in each household |
| Verbal consent given by the head of the household | Y = YesN = No | If the answer is No, no further questions are to be asked.  |
| Gender | M = MaleF = Female |  |
| Age | Numeric (months/years) |  |
| Member of the household at the beginning start recall period | Y = YesN = No |  |
|  If No, reason | J = Joined the householdB = New-born |  |
|  If No, date of joining the household or date of birth | dd/mm/yyyy | Consider mm/yyyy if the day is unclear |
| Member of the household at the beginning start recall period | Y = YesN = No |  |
|  If No, reason | L = Left the householdD = Died |  |
|  If No, date of departure/death | dd/mm/yyyy | Consider mm/yyyy if the day is unclear |
|  If died, main cause of death | 1 = Diarrhoea2 = Respiratory infection3 = Malaria/Fever4 = During pregnancy5 = During/after delivery (<1 month)6 = Trauma/accident7 = Violence8 = Don’t know9 = Other (specify) | Most common causes of death. To be adapted to the context |
| Information source of the cause of death (optional) | 1 = Family perception 2 = verbal report3 = Written report |  |
|  If died, place of death | 1 = At home2 = At the health centre 3 = At the hospital 4 = Don’t know9 = Other (specify)*For displaced population can add:*5=home village6= during displacement7=in transit camp8= in current location | Most common places of death. To be adapted to the context. You may add ancillary variables for the name of the health centre and the hospital |
| **ADDITIONAL RECOMMENDED QUESTIONS RELATED TO HEALTH SEEKING** |
| Did s/he seek health care in the 2 weeks before dying? | 0 = no,1 = yes, 9 = don’t know |  |
| Place health care sought? | 1 = home,2 = primary health unit *(specify)*, 3 = primary health centre *(specify),*4 = hospital *(specify),* 5 = don’t know, 6 = other place *(specify)* | The possible answers need to be adapted according to the study area.All possibilities must be recorded.Check if place of death = home AND any health structure was visited. |
| What happened between leaving the last health structure and the time of death? |  | Open question, record as reported. |
| Main reason for not seeking care in a health structure? | 1 = immediate death,2 = no money/consultation too expensive3 = too sick for seeking care4 = not sick enough for seeking care5 = health facility too far away6 = went to traditional healer 7 = did not have time to go/ too busy to go8 = no confidence in health structure 9 = security problem10 = care was refused at health centre 11 = other reason *(specify)*12 = don’t know | The possible answers need to be adapted according to the study area. |

## Appendix 4: Additional questions by topic

|  |  |  |
| --- | --- | --- |
| **Variable** | **Values** | **Comments** |
| Status | 1 = Local/host2 = Displaced3 = Refugee9 = Other |  |

**Education**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Values** | **Comments** |
| If ≥ 15 years: Able to read and write? | 0 = No1 = Yes |  |
| If yes:Highest education level | 1 = Primary2 = Secondary3 = Higher |  |

**Vaccination**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Values** | **Comments** |
| If aged 6 months to 5 years:Measles vaccination? | 0 = No1 = Yes |  |
| If yes:Vaccination card | 0 = No1 = Yes |  |

**Malaria**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Values** | **Comments** |
| Slept under a bed-net last night? | 0 = No1 = Yes |  |

**Pregnancy**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Values** | **Comments** |
| If female and 15-49 years: Pregnant? | 0 = No1 = Yes |  |
| If yes: Got malaria preventive treatment | 0 = No1 = Yes |  |
| If yes: Type of malaria treatment |  |  |

**Violence**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Values** | **Comments** |
| Experienced a violent episode | 0 = No1 = Yes |  |
| Violence was repeated more than once | 0 = No1 = Yes3 = Doesn’t want to respond4 = Not able to respond |  |
| Number of violent episodes experienced | Numeric |  |
| Date or period when violence occurred | From dd/mm/yyyyTo dd/mm/yyyy | Blank if unknown |
| Place where the violence occurred | String |  |
| Nature of violence | 1= Beaten2= Sexual3= Shot4= Detained/Kidnapped5= Doesn’t want to respond6= Other | Adapt the options to the survey area |
| Perpetrator was uniformed? | 0 = No1 = Yes |  |

## Appendix 5: Interviewers training program

The program needs to be adapted to the planned study in the field (i.e. training hours, time and program).

**First day**

|  |  |
| --- | --- |
| **Time** | **Program** |
| Morning | Introduction of investigators, supervisors, and interviewersIntroduction to MSF project and how the survey will help MSF project and ultimately the communityIntroduction to the survey protocol (reference population, sampling, recall period…) General introduction of the tools and procedures used in the surveySurvey ethics including the informed consent process and how to seek informed consent  |
| Afternoon | Local calendar of events during recall periodStructure of household Presentation of the questionnaireDiscussion of questionnaire questions one by one, including how to pose questions, and wording in the local language(s) |

**Second day**

|  |  |
| --- | --- |
| **Time** | **Program** |
| Morning | Recall of the topics presented in the previous dayRole-play on questionnaire including:* Introduction to the household
* Information sheet and consent form
* Interviews and questions
* Farewell from the household
 |
| Afternoon | Detailed presentation of tool and procedures on how to identify the sampled householdsPractical exercises on how to identify the sampled householdsPreparation of the pilot day |

**Third day**

|  |  |
| --- | --- |
| **Time** | **Program** |
| Morning | Pilot |
| Afternoon | Feed-back on the pilotPreparation of the survey day |

A meeting will be held with associated staff such as logisticians, community liaisons, drivers and data clerks to explain the overall study and their roles and expectations.

1. There is evidence that suggests that doing surveys on mortality only results in a more accurate mortality estimate than combining them with other survey items, malnutrition in particular. Combining anthropometric and mortality surveys was found to reduce the reporting of deaths in children under 5 by half. The inclusion of other components may increase a perception of judgement from the study team, and reluctance to answer or a desire to give more socially acceptable answers. For more detailed information see the two articles here <https://pubmed.ncbi.nlm.nih.gov/8307676/>, <https://pubmed.ncbi.nlm.nih.gov/8307670/>. [↑](#footnote-ref-2)
2. Adaptation of “The use of epidemiological tools in conflict-affected populations: open-access educational resources for policy-makers”, London School of Hygiene and Tropical Medicine, 2009 [↑](#footnote-ref-3)