Introduction

The Simple Spatial Survey Method (S3M) was developed from the CSAS coverage survey method as a response to the widespread adoption of community management of acute malnutrition (CMAM) by ministries of health. Large-scale programs need a large-scale survey method and S3M was developed to meet that need.

S3M was designed to:

- Be simple enough for MoH, NGO, and UNO personnel to perform.
- Be able to survey areas up to ten times larger than the CSAS method at approximately twice the cost whilst maintaining the spatial resolution of CSAS surveys.
- Provide a general survey method. S3M can be used to survey and map the coverage of selective entry programs such as CMAM and universal programs such as EPI or GMP, the values of indicators for (e.g.) IYCF and WASH, and prevalence over wide areas.

This document is intended as a short guide for those undertaking S3M surveys.

This document is currently in draft. Please address comments and suggestions to:

mark@brixtonhealth.com
Step 1: Find a map

The first step in a 3SM survey is to find a map of the survey area.

Try to find a map showing the locations of all towns and villages in the survey area.

If you are surveying a very large area then you will find it useful to have:

- A small scale map of the entire survey area. This map does not need to show the location of all towns and villages in survey area.
- A collection of larger scale maps showing the locations of all towns and villages. This collection of maps should cover the entire survey area.

The small scale map will be useful for identifying initial sampling locations.

The large scale maps will be useful for identifying the precise location of sampling points and for selecting the communities to be sampled.
Step 2: Decide the area to be represented by each sampling point

The easiest way of thinking about this as a function of the intended maximum distance \((d)\) of any community from the nearest sampling point.

If you are surveying the coverage of a program treating a rare condition (e.g. SAM) then it is a good idea to select a small value for \(d\). This helps to ensure that large areas are not represented by very few distant sampling points just because some sampling points yield no cases. The alternative to choosing a small value for \(d\) is to sample more intensively around each sampling point (i.e. taking a bigger sample from each community or sampling in more communities).

A value for \(d\) of 10km will probably be small enough in most circumstances.
Step 3 : Draw a grid over the map

The next step is to draw a grid over the map.

The size of the grid is determined by the distance \((d)\) that you decided in Step 2.

The grid is rectangular. This means that the width of the grid in the east-west \((x)\) direction is different from the height of the grid in the north-south \((y)\) direction.

The width of the grid in the east-west \((x)\) direction is calculated using:

\[
x = \frac{3d}{2}
\]

The height of the grid in the north-south \((y)\) direction can be calculated using:

\[
y = \frac{\sqrt{3}d}{2}
\]
Step 3: Draw a grid over the map

For example, if $d = 10$ km then:

\[ x = \frac{3d}{2} = \frac{3 \times 10}{2} = \frac{30}{2} = 15\text{ km} \]

and:

\[ y = \frac{\sqrt{3}d}{2} \approx \frac{1.73 \times 10}{2} \approx \frac{17.3}{2} \approx 8.7\text{ km} \]

The table below shows the grid sizes for different values of $d$:

<table>
<thead>
<tr>
<th>$d$</th>
<th>$x$</th>
<th>$y$</th>
<th>$d$</th>
<th>$x$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>7.5</td>
<td>4.3</td>
<td>13</td>
<td>19.5</td>
<td>11.3</td>
</tr>
<tr>
<td>6</td>
<td>9.0</td>
<td>5.2</td>
<td>14</td>
<td>21.0</td>
<td>12.1</td>
</tr>
<tr>
<td>7</td>
<td>10.5</td>
<td>6.1</td>
<td>15</td>
<td>22.5</td>
<td>13.0</td>
</tr>
<tr>
<td>8</td>
<td>12.0</td>
<td>6.9</td>
<td>16</td>
<td>24.0</td>
<td>13.9</td>
</tr>
<tr>
<td>9</td>
<td>13.5</td>
<td>7.8</td>
<td>17</td>
<td>25.5</td>
<td>14.7</td>
</tr>
<tr>
<td>10</td>
<td>15.0</td>
<td>8.7</td>
<td>18</td>
<td>27.0</td>
<td>15.6</td>
</tr>
<tr>
<td>11</td>
<td>16.5</td>
<td>9.5</td>
<td>19</td>
<td>28.5</td>
<td>16.5</td>
</tr>
<tr>
<td>12</td>
<td>18.0</td>
<td>10.4</td>
<td>20</td>
<td>30.0</td>
<td>17.3</td>
</tr>
</tbody>
</table>

When drawing the grid make sure that it covers the entire survey area.

If you are drawing the grid directly onto the map then use a soft pencil (e.g. a “2B” or “Number 1” pencil). A soft pencil will not damage the surface of the map and is easy to erase should you make a mistake or need to draw a different grid.
Step 4: Create an even spread of sampling points

Sampling points are located at the intersections of the rectangular grid in a staggered fashion. Alternate intersections of the grid are used:

Make sure that your sample points go right to the edge (or even over the edge) of the survey area.

If you identify fewer than about thirty sampling points then you should go back to Step 3 and choose a smaller value for \( d \).
Step 5: Select the communities to sample

Select the community (or communities) closest to the sampling points identified in Step 4.

If you are investigating the coverage of a selective program for a rare condition (e.g. SAM) then you will need to select several communities in order to be confident of finding cases.

If you are investigating the coverage of a universal program with a limited age-range (e.g. EPI or GMP programs) or are investigating the prevalence of a condition or a behaviour then you should select a sufficient number of communities to give you a sample size of 21 or more from each sampling point.

The position of the sampling point is moved to the position of the selected community. This is shown in the diagram above.

If more than one community is selected then the position of the sampling point is moved to the middle of the selected communities. This is shown in the diagram below.

You may drop sampling points if you find that many sampling points are clustered closely together. You may add sampling points if you find that there are populated areas that do not contain sampling points. The aim is to create an even spread of sampling points over the majority of the survey area. You need to have a minimum of about 30 sampling points.
Step 5: Select the communities to sample

Select the community (or communities) closest to the sampling points identified in Step 4.

If you are investigating the coverage of a selective program for a rare condition (e.g. SAM) then you will need to select several communities in order to be confident of finding cases.

If you are investigating the coverage of a universal program with a limited age-range (e.g. EPI or GMP programs) or are investigating the prevalence of a condition or a behaviour then you should select a sufficient number of communities to give you a sample size of 21 or more from each sampling point.

The position of the sampling point is moved to the position of the selected community.

If more than one community is selected then the position of the sampling point is moved to the middle of the selected communities. This is shown in the diagram above in which three communities are selected for each sampling point.

You may drop sampling points if you find that many sampling points are clustered closely together. You may add sampling points if you find that there are populated areas that do not contain sampling points. The aim is to create an even spread of sampling points over the majority of the survey area. You need to have a minimum of about 30 sampling points.
Step 6 : Label each sampling point

Give each sampling point a unique identifying label:

- The label may be a number or a name.
- The label must be unique.
- The label is used to identify which community belongs to which sampling point.

The label is used when collecting, organising, and analysing data.
Step 7: Plan data collection

Now you have identified the location and number of sampling points you can draw up a project plan and budget. When drawing up the timetable and budget it is common to assume that one team will sample one sampling point in one day.

---

### Project timetable

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
</tr>
<tr>
<td>g</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
</tr>
<tr>
<td>i</td>
<td></td>
</tr>
<tr>
<td>j</td>
<td></td>
</tr>
</tbody>
</table>

### Budget

<table>
<thead>
<tr>
<th>Activity</th>
<th>Budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>750.00</td>
</tr>
<tr>
<td>b</td>
<td>400.00</td>
</tr>
<tr>
<td>c</td>
<td>800.00</td>
</tr>
<tr>
<td>d</td>
<td>1,750.00</td>
</tr>
<tr>
<td>e</td>
<td>350.00</td>
</tr>
<tr>
<td>f</td>
<td>2,225.00</td>
</tr>
<tr>
<td>g</td>
<td>1,500.00</td>
</tr>
<tr>
<td>h</td>
<td>300.00</td>
</tr>
<tr>
<td>i</td>
<td>8,500.00</td>
</tr>
<tr>
<td>j</td>
<td>1,500.00</td>
</tr>
<tr>
<td>Total</td>
<td>18,075.00</td>
</tr>
</tbody>
</table>
Step 8: Within-community sampling

The sampling process that you use to select a sample from a community will depend on what the survey is investigating.

If you are investigating multiple indicators which apply to different groups of individuals then you may find it easier to use different sampling methods for different indicators. You can think of this as having different surveys for different indicators sampled from the same set of communities at the same time.

If you are evaluating a CMAM program find cases using an active case-finding method. It is usually sufficient to ask community health workers, traditional birth attendants, traditional healers or other key informants to take you to see “children who are or have recently been sick, are thin, have swollen legs or feet, or go to the SAM program” and then ask mothers and neighbours of confirmed cases to help you find more cases. You could also use door-to-door screening. It is important that the case-finding method that you use finds all, or nearly all, cases in the sampled communities. Each case is confirmed by applying the program's entry criteria (e.g. MUAC < 115 mm and/or bilateral pitting oedema). When you find a confirmed case you should find out whether the child is in the program. Remember to follow-up children reported to be in a therapeutic feeding centre or at a distribution point on the day of the survey.

If you are evaluating a universal program with a narrow age-range (e.g. 0 to 24 months or 6 to 36 months) use the standard *EPI* sampling method to find eligible children:

**Selection of the first household:** Stand at a central point in the community and choose a direction at random (e.g. by spinning a bottle). Count the houses between the central point and the edge of the community in that direction. Select one of these houses at random using a random number table. If there is no eligible child in this house then select the nearest house with an eligible child. All eligible children in a selected household should be included in the sample.

**Selection of subsequent households:** Choose the house whose door is nearest to the door of the previously sampled household. If there is no eligible child in this house then select the nearest house with an eligible child. All eligible children in a selected household should be included in the sample.

**Stopping sampling:** Sampling stops when you have sampled the required number of children.

When you are surveying a group that makes up less than about 10% of the total population the *EPI* sampling method provides a sample that is representative of the sampled community.
Step 8 : Within-community sampling

If you are evaluating a universal program with a wider age-range or estimating prevalence with a wider age-range (e.g. prevalence of GAM in children aged 6 to 59 months) then you should use the *EPI5* sampling methods to find eligible children:

**Selection of the first household**: Stand at a central point in the community and choose a direction at random (e.g. by spinning a bottle). Count the houses between the central point and the edge of the community in that direction. Select one of these houses at random using a random number table. If there is no eligible child in this house then select the nearest house with an eligible child. All eligible children in a selected household should be included in the sample.

**Selection of subsequent households**: Choose a random direction (e.g. by spinning a bottle) and select the fifth (*EPI5*) nearest house in that direction. In small communities you may select the third house in that direction (*EPI3*). If there is no eligible child in this house then select the nearest house with an eligible child. All eligible children in a selected household should be included in the sample. If you reach the edge of the community then return to the central point and start again.

**Stopping sampling**: Sampling stops when you have sampled the required number of children.

If your are sampling in a large community then you should should use the *QTR* sampling method:

The community is divided into four quadrants. The appropriate sampling method (i.e. *EPI*, *EPI3*, or *EPI5*) is then used independently to select a quarter of the sample from each quadrant starting at a central point in each quadrant.

If you are investigating multiple indicators which apply to different groups of individuals then you may find it easier to use different sampling methods for different indicators. For example, a survey investigating coverage of CMAM and GMP programs. The table below shows example methods and sample sizes that might be used in a survey investigating the coverage of CMAM and GMP programs:

<table>
<thead>
<tr>
<th>Program</th>
<th>Eligible Children</th>
<th>Sampling method</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMAM</td>
<td>Children aged between 6 and 59 months with MUAC &lt; 115 mm or bilateral pitting oedema</td>
<td>Active / adaptive</td>
<td>Depends upon prevalence of SAM. The case-finding method should find all or nearly all cases in the sampled communities.</td>
</tr>
<tr>
<td>GMP</td>
<td>All children aged between 0 and 24 months.</td>
<td><em>EPI</em></td>
<td>Seven children from each community</td>
</tr>
</tbody>
</table>

Such a survey would probably use two teams to survey each community with each team collecting data for a single indicator.
Step 9 : Record the data

You will usually need to record very little data.

For a CMAM program you will need to record the number of cases you find, the number of these cases that are in the programs, and the number of non-cases that are in the program (i.e. recovering cases) that you find. For a GMP program you will need to record the number of children in the sample and the number of these children that had an up-to-date weight record on their “health passport” or “road to health” card. Data for these types of program can be recorded on simple tally sheet like the one shown above.

For programs with complicated definitions of coverage (e.g. EPI programs) you will probably need to record individual data and apply case-definitions after data-collection.

You will always need to record the unique sampling point identifier created in Step 6.

If you only have approximate location data then you may want survey teams to use GPS devices to find and record the location of the centres of the sampled communities (as in the tally sheet shown above).

If you want the survey to provide wide-area estimates then you should record the population of the sampled communities. This can be in terms of individuals or households but the units used must be consistent throughout the survey. Approximate population values are acceptable. Broad ranks such as 1 = small, 2 = medium, and 3 = large are also acceptable.

<table>
<thead>
<tr>
<th>Name of village</th>
<th>Population</th>
<th>Lat / Lon From GPS</th>
<th>Cases</th>
<th>Covered cases</th>
<th>Non-cases in program</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CMAM Coverage Survey : Yurugli Zone : November 2011

Sampling Point ID :          Lat / Lon (from map) :

Population

Covered cases

Non-cases in program

Case

Non-case

Recovered

Identify

Sampling

Point

ID

Lat / Lon

From

Map

Name of village

Population

Lat / Lon

From GPS

Cases

Covered cases

Non-cases in program

Population

Covered

cases

Non-cases in program
Step 10: Prepare the data for analysis and presentation

This requires a separate data table for each indicator with one row per sampled community containing the following variables:

- **id**: The unique sampling point identifier.
- **x**: The east-west position of the centre of the sampled community as a number. If you do not have this data then enter the east-west position of the sampling point as a number.
- **y**: The north-south position of the centre of the sampled community as a number. If you do not have this data then enter the north-south position of the sampling point as a number.
- **pop**: The population of the sampled community. This can be in terms of individuals or households but the units used must be consistent throughout the survey. Set this to 1 for all sampled communities if not collected or not known.
- **a**: The numerator of the estimator.
- **b**: The denominator of the estimator.

There can must be no missing data.

Location data (i.e. **x** and **y**) should be entered as numbers.

GPS devices can report position data in a number of formats. For example:

- **DDD MM SS**: Degrees, Minutes, and Seconds
- **DDD MM.MM**: Degree and Minutes

These should be converted into numbers using the following formulae:

\[
Number = DDD + \frac{MM}{60} + \frac{SS}{3600} \quad \text{or} \quad Number = DDD + \frac{MM.MM}{60}
\]

before data entry or data analysis.
Step 10: Prepare the data for analysis and presentation

Location (i.e. \( x \) and \( y \)) data may be taken from GPS devices or read from maps. The numbers used can be conventional latitudes and longitudes (converted to numbers - see above) or an arbitrary co-ordinate system such as the number of millimetres east and north measured from an arbitrary reference point on a map may be used:

![Diagram showing arbitrary reference point and sampled point](image)

Population (\( \text{pop} \)) is used when calculating wide area estimates. Approximate population values are acceptable. Broad ranks such as 1 = small, 2 = medium, and 3 = large are also acceptable. If this data is not collected or not known then you should enter 1 for all sampled communities.

The estimator has a numerator (\( a \)) and a denominator (\( b \)):

\[
\text{Estimate} = \frac{a}{b}
\]

For example, the period coverage estimator for a targeted feeding program is:

\[
\text{Coverage} = \frac{\text{number attending the program}}{\text{number of cases not attending the program} + \text{number attending the program}}
\]

For this estimator:

The numerator (\( a \)) is the number of respondents (current and recovering cases) attending the feeding program found in each sampled community.

The denominator (\( b \)) is the sum of the number of current cases not attending the feeding program and the number of respondents (current and recovering cases) attending the program found in each sampled community.
Step 10: Prepare the data for analysis and presentation

Data should be entered and saved in comma-separated-value (CSV) format using one line per sampled community with variable names recorded in the first line of the file.

You can prepare data using a text editor, a spreadsheet, or the data-entry page of the openS3M software.

Prepared data files should look like this:

```
id, x, y, pop, a, b
1,609,258,56,5,3
1,637,270,65,3,2
1,615,275,85,6,5
2,860,305,102,4,3
2,845,319,82,2,2
2,877,298,97,0,0
```

In this example there are three communities for each sampling point.

You will find it helpful to use a meaningful file name that identifies the survey and the indicator being investigated and ends with the extension .CSV or .csv. such as:

```
NorthWolloPointCoverageOTP.csv
```
Step 11 : Plot the data

Nothing here yet!

This section will be completed once the openS3M software has been developed.

Example map of period coverage from openS3M pilot code