

Guidance notes on sampling methodology for the MSM

Once the immediate needs and interventions have been identified, it will be necessary to gather more in depth information to inform the future design of the programme and to provide a baseline for monitoring and evaluation. The 'baseline'¹ survey provides a detailed assessment of sanitation and hygiene practices. It should draw on both qualitative data (obtained from focus group discussions, three pile sorting exercises etc.) and quantitative data obtained from a random sample of the population, usually using a questionnaire. The data and analysis obtained from the rapid assessment should feed into the baseline survey.

Qualitative data collection uses a form of sampling called **purposive sampling**. Participants in focus groups or other discussion groups are selected **intentionally** because they meet certain criteria e.g. women with children under five, male heads of household, families with latrines.

Quantitative data can be obtained by asking closed questions, by noting observations of sanitary conditions and also by taking fixed measurements e.g. the depth of latrine pits. It is important to base the questionnaire content on the outcome of the initial assessment and to keep it as simple as possible, using not more than 15-20 questions/observation or measurement criteria. You will also need to decide who you will want to question e.g. women with children under five or male heads of household.

A **random sample** of the population is used in order to ensure that the sample is as representative as possible. The degree of reliability of the sample and the extent to which the results can be generalised to the rest of the population is known as the confidence interval. The **sample size of 100 for any population over 1000 people** is reliable to 90% and this is considered as a 'good enough' confidence interval when conducting a sanitation and hygiene survey in an emergency context.

Ensuring that the sample of people who will answer questions is as random as possible is vital in ensuring the reliability of the data. The following information shows you how to work out the number of interviews you need to do in each location if you are working in camps or settlements of unequal size.

Identifying Interview Locations

- Draw up a **table of communities/villages/camps/settlements** and record the population figures for each (if no population figures are available estimate the number of households and multiply by the average number of children)
- Identify the **number of interviews to be held in each location** by using the following formula:

Number of people in community/camp/location X number of the total sample required / total number of population being studied.

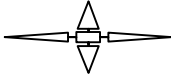
- The finished table is known as a **sampling frame** as shown below:

Sampling Frame for Identifying Interview Locations for Questionnaire Survey

Name of community/camp/location	Total population	Formula to determine weighting	Number of interviews
Kabila	978	$978 \times 100 / 19440 = 5.03$	5
Umani	1360	$1360 \times 100 / 19440 = 6.99$	7
Bendela	90	$90 \times 100 / 19440 = 0.46$	0
Abrama	500	$500 \times 100 / 19440 = 2.57$	3

ETC. ETC.....

¹ In an emergency situation some interventions need to start immediately and therefore the survey will not always reflect the original baseline conditions. A baseline survey can also be taken to mean a much broader assessment of the context but these guidance notes emphasise the use of the baseline for monitoring and evaluation of the MSM intervention.



Bemba	939	$939 \times 100 / 19440 = 4.85$	5
Kirunga	3879	$3879 \times 100 / 19440 = 19.95$	20
Lungi	1790	$1790 \times 100 / 19440 = 9.21$	9
TOTAL	19,440		100

NB In some communities/locations where the population size is relatively small e.g. Bendela, there may not be anyone selected for interview. It is still important, however, that each location is included in the sampling frame.

Once you know how many interviews you will do in each location, you will need to **randomly select the actual households** where you will carry out the interview.

Selecting households in each location

- If a list of all households is available then use the numbered list to choose households at random using the random number table (see instructions below).
- If the number of households is ≤ 30 – draw a map and number the households
- If there are >30 households subdivide the area and select one area at random – draw a map and number the households
- Select random numbers using the random number table to identify the household(s) where the interviews will take place
- If no one that meets the interviewee criteria is available e.g. women with children under five years or female refugee then go to the next house

Spin a Bottle Method

This method is less reliable and easily misused. It is also not good for large areas. It should be possible, however, to divide a large area into more manageable areas as above.

- Go to the centre of the location
- Place the bottle on the ground so that it rotates
- Walk in a straight line in the direction of the bottle
- Count the number of houses along this line until you come to the boundary e.g. 10
- Choose a random number from one to ten (or however many houses are along the line) to allow you to select the household where you will conduct the interview
- If no one that meets the interviewee criteria is available (e.g. women with children under five years or female refugees) then go to the next house

Using a Random Number Table

- Use the random number table in the annex
- Close your eyes and allow a pen to pick out a number on the table
- Determine how many digits you need to refer to e.g. you want to pick a random number between 1 and 1000 so you will use only four of the digits on the selected row of the table
- You can read from right to left or left to right to select your four digits and therefore the random number

Example of Activity Chart/Gantt Chart for conducting the questionnaire survey*

Activity	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13
Prepare questionnaire	■												
Translate (if necessary)		■	■										
Prepare sampling frame		■	■										
Train survey team				■	■	■							
Pre-test questionnaire					■								
Amend questionnaire						■							
Data collection							■	■	■	■			
Data collation										■	■		
Data analysis												■	
Report writing													■

Use of findings

- To inform indicators/ set targets within the logframe (i.e. this is an intrinsic part of the monitoring system)
- Once you have completed the analysis, it is good practice to try to discuss the results with the community. Feedback to the community should help to identify subsequent community and agency actions

* The preparation and execution of the questionnaire survey will need to be carried out simultaneously with other hygiene promotion activities and should not require the cessation of these other activities for the full thirteen days.

Annex 1

Example Random Number Table

87172	43062	39719	10020	32722	86545	86985	04962	54546	23138	62135	55870	97083	67875
28900	50851	30543	89185	16747	95104	49852	26467	58869	79053	06894	23975	34902	23587
86248	71156	55044	13045	33161	95604	57876	23367	10768	78193	60477	70307	06498	48793
10531	51391	41884	69759	32741	70072	01902	96656	90584	59263	49995	27235	40055	20917
02481	90230	81978	39127	93335	74259	25856	52838	49847	69042	85964	78159	40374	49658
23988	13019	78830	17069	58267	69796	94329	34050	25622	55349	10403	93790	77631	74261
37137	47689	82466	24243	10756	54009	44053	74870	28352	66389	38729	80349	50509	56465
38230	82039	34158	90149	82948	60686	27962	39306	53826	47852	76144	38812	76939	03119
98745	08288	19108	84791	58470	59415	45456	44839	86274	25091	42809	56707	47169	95273
44653	58412	91751	14954	87949	81399	51105	29718	82780	11262	23712	99782	42829	26308
88386	66621	16648	19217	52375	05417	26136	05952	71958	25744	52021	20225	01377	47012
50660	58138	01695	69351	25445	20797	74079	60851	47634	36633	93999	96345	58484	12506
36732	74234	84240	46924	62744	39238	78397	60869	26426	55588	56963	59506	17293	45096
34187	78277	83678	34754	46616	45250	25291	04999	19717	60324	66915	03473	98329	82447
26095	98131	79362	39530	53870	87445	26277	90551	28604	39865	40686	05435	74511	69866
00067	74289	20706	74076	28206	36960	09231	82988	57062	35331	08212	68111	52199	05065
42104	26434	30953	15259	76676	63339	75664	23993	63538	34968	47655	44553	61982	13296
82580	46580	87292	23226	21865	60338	04115	33807	38395	98484	40387	69877	24910	13317
89266	14764	17681	68663	66030	12931	17372	35601	63805	55739	42705	30549	31697	33478
47100	92329	89435	69974	40783	52649	93444	41317	02749	19052	34647	92814	88046	34020
59566	26527	44706	85670	96223	36275	82013	82673	60955	62617	90214	24589	59715	57612
10946	24676	66513	56743	96911	89042	08263	70753	89045	39189	04306	06090	94515	17772
34013	69250	27977	84597	55192	65088	55739	35953	18533	39339	78037	32827	68269	69218
21606	11751	30073	71431	53569	27865	90215	34772	21779	11734	64313	49764	30816	56852
56620	92612	77157	90231	90144	29781	01683	52503	60080	73703	70080	80686	47379	33279
49238	90475	84356	87159	21222	40106	02671	52684	38514	68434	16407	58164	13341	48142
50738	21999	73539	51802	78179	27872	57937	29696	67783	29373	96563	74619	77099	17190
58761	21571	71692	19723	25088	10483	71430	47068	78378	80237	32113	09381	62931	29243
55335	71937	22025	33538	04648	74232	57839	62431	61835	04784	06732	34202	93497	72070
26515	31143	83795	78445	32869	31489	81587	90354	97672	70106	35008	37899	36246	97805
32625	36806	00082	26902	26250	28919	38054	49027	22209	42696	46980	17065	61288	30208
20311	96089	20141	30362	04980	32703	04202	91080	28660	89691	84660	73433	70169	11273
10941	73003	87930	85620	06956	38719	88711	61454	64076	13316	02203	54437	54306	78229
56982	46636	34070	30803	39095	80387	08971	25067	07377	70704	13629	68474	99229	05535
14661	10670	15811	00454	81124	46977	89983	48836	48182	17054	06344	24267	16686	21401
52760	78118	23277	29760	00099	97325	54762	43117	73199	19621	24599	11030	64809	35088
48874	20831	02286	73635	93771	54264	49801	22653	01524	84621	91023	64028	29278	15987
44817	77408	48447	25934	22912	43086	68126	92970	91833	26418	72454	97636	94593	07880
17896	79375	70883	70135	21589	51181	71969	32951	35036	17219	27357	96517	55307	84470
27166	22347	92146	92189	16301	15747	72837	59174	75024	39459	54910	95335	95013	47068
13665	30490	63583	73098	19976	03001	94645	40476	43617	85698	66512	42759	20973	98759
58644	73840	08103	97926	57340	63077	08114	10031	35668	21740	33787	44756	20527	65367
72570	36278	06602	56406	85679	85529	08576	50874	59706	01019	29980	56742	05356	04810
92041	68829	02163	59918	83041	71241	90678	79835	86324	13075	29913	99831	25688	53648
71240	74119	53090	23693	14007	90107	68804	54927	68964	26535	28184	21630	12362	67990