Appendix 1. Technical Appendix

This appendix provides additional technical information regarding the choices of methods that are used in SQUEAC and SLEAC.

Active and Adaptive Case-Finding

The active and adaptive case-finding method (see Box 3, page 65) was originally developed for use in CSAS coverage surveys. Active and adaptive case-finding was tested during the development of the CSAS coverage survey method using capture-recapture studies. It was also common practice to test active and adaptive case-finding procedures using capture-recapture studies when the CSAS coverage survey method was used to assess program coverage in a particular area for the first time. These tests found that, when well done, the active and adaptive case-finding method:

- Finds all or nearly all SAM cases in sampled communities
- Consistently performed better than central-location screening and house-to-house screening
- Performed as well as house-to-house screening with local informants and verbal household censuses at finding sick or sleeping children that may be 'hidden' to avoid them being disturbed by the survey team
- Is more efficient than other case-finding methods

It should be noted that these findings apply only to using the active and adaptive case-finding method to find cases of SAM.

The active and adaptive case-finding method has been tested for finding cases of moderate acute malnutrition (MAM) and has been found to perform considerably worse than house-to-house screening. If SQUEAC or SLEAC are used to assess the coverage of, for example, an SFP, then house-to-house screening should be used and extra time allocated to survey activities.

The active and adaptive case-finding method has been known to fail in some urban and camp settings. Some examples of this may be found in the case studies section of this document. In such settings, it is advisable to use house-to-house screening and allow additional time for survey activities.

When finding cases of SAM using house-to-house screening, it is advisable to use local informants and to take a (verbal) household census before asking to measure children. This avoids the problem of sick or sleeping children being 'hidden' to avoid them being disturbed by the survey team.

Calculating the Required Likelihood Survey Sample Size

The formula for calculating the sample size for a likelihood survey is:

$$n_{Likelihood} = \left[\frac{mode \times (1 - mode)}{(precision \div 1.96)^2} - (\alpha_{Prior} + \beta_{Prior} - 2) \right]$$

where *mode* is the mode of the prior, α_{Prior} and β_{Prior} are the shape parameters of the prior, and *precision* is the precision (i.e., the approximate half-width of the 95% credible interval) required for the posterior estimate.

The first part of the formula:

$$\frac{mode \times (1 - mode)}{(precision \div 1.96)^2}$$

is the standard formula for calculating the required sample size for a survey with a non-informative prior. The value *1.96* closely approximates the standard normal deviate for the 97.5^{th} percentile point of the standard normal distribution and is used to specify a 95% credible interval. Other values (e.g., 2.58 for a 99% credible interval) may be used.

The second part of the formula:

$$(\alpha_{Prior} + \beta_{Prior} - 2)$$

arises from a $Beta(\alpha_{Prior}, \beta_{Prior})$ prior being equivalent to finding:

$$\alpha_{Prior} - 1$$

successes in:

$$\alpha_{Prior} + \beta_{Prior} - 2$$

trials. A prior specified as $Beta(\alpha_{Prior}, \beta_{Prior})$ provides the same information as a survey with a sample size of:

$$n = \alpha_{Prior} + \beta_{Prior} - 2$$

The required sample size for the likelihood survey is the sample size required when using a non-informative prior minus the sample size that is represented by the information summarized by an informative prior.

The Simulation Approach to Calculating the Required Likelihood Survey Sample Size

The simulation approach to calculating the required likelihood survey sample size can be thought of as a directed search for the required sample size. The general approach can be illustrated using the example of a likelihood survey and a non-informative prior (i.e., a *Beta(1, 1)* prior). If we have an expected coverage (p) of 50% and we want the survey to estimate this with a 95% credible interval of ± 10 percentage points, then we might start with a test sample size ($n_{Likelihood}$) of 60. If we performed a survey with $p = 0.5$ (i.e., 50%) and $n_{Likelihood} = 60$, we would have an approximate 95% credible interval with width:

$$\pm 1.96 \times \sqrt{\frac{p \times (1 - p)}{n_{Likelihood}}} = \pm 1.96 \times \sqrt{\frac{0.5 \times (1 - 0.5)}{60}} = \pm 0.1265$$

This is wider than desired (i.e., 12.65% is greater than 10%). The test sample size is, therefore, smaller than that required to deliver the desired precision.

Simulating surveys with different test sample sizes selected using a simple directed search strategy can quickly find the required sample size. This is illustrated in **Table A1-1**.

Table A1-1. Finding a sample size for a desired precision using simulation and a simple directed search strategy

Test Sample Size	Precision	Interpretation	Action
60	$\pm 1.96 \times \sqrt{\frac{0.5 \times (1-0.5)}{60}} = \pm 0.1265$	The test sample size is **too small** to deliver the desired precision.	Double the test sample size and simulate.
120	$\pm 1.96 \times \sqrt{\frac{0.5 \times (1-0.5)}{120}} = \pm 0.0895$	The test sample size is **more than adequate** to deliver the desired precision.	Choose a new test sample size that is half-way between 60 and 120 and simulate.
90	$\pm 1.96 \times \sqrt{\frac{0.5 \times (1-0.5)}{90}} = \pm 0.1033$	The test sample size **almost delivers** the desired precision.	Choose a new test sample size that is half-way between 90 and 120 and simulate.
105	$\pm 1.96 \times \sqrt{\frac{0.5 \times (1-0.5)}{105}} = \pm 0.0956$	The test sample size is **more than adequate** to deliver the desired precision.	Choose new a test sample size that is half-way between 90 and 105 and simulate.
97	$\pm 1.96 \times \sqrt{\frac{0.5 \times (1-0.5)}{97}} = \pm 0.0995$	The test sample size may be a **little larger** than is required to deliver the desired precision.	Choose a new test sample size that is very slightly smaller than 97 and simulate.
96	$\pm 1.96 \times \sqrt{\frac{0.5 \times (1-0.5)}{96}} = \pm 0.1000$	This is the **required** sample size.	Plan the survey to achieve a minimum sample size of 96.

It is possible to use a similar approach with an informative prior. In this case, the precision of a simulated survey is calculated as:

$$\pm 1.96 \times \sqrt{\frac{p \times (1-p)}{test\ sample\ size + (\alpha_{Prior} + \beta_{Prior} - 2)}}$$

When doing calculations by hand, it is much easier to use the sample size formula:

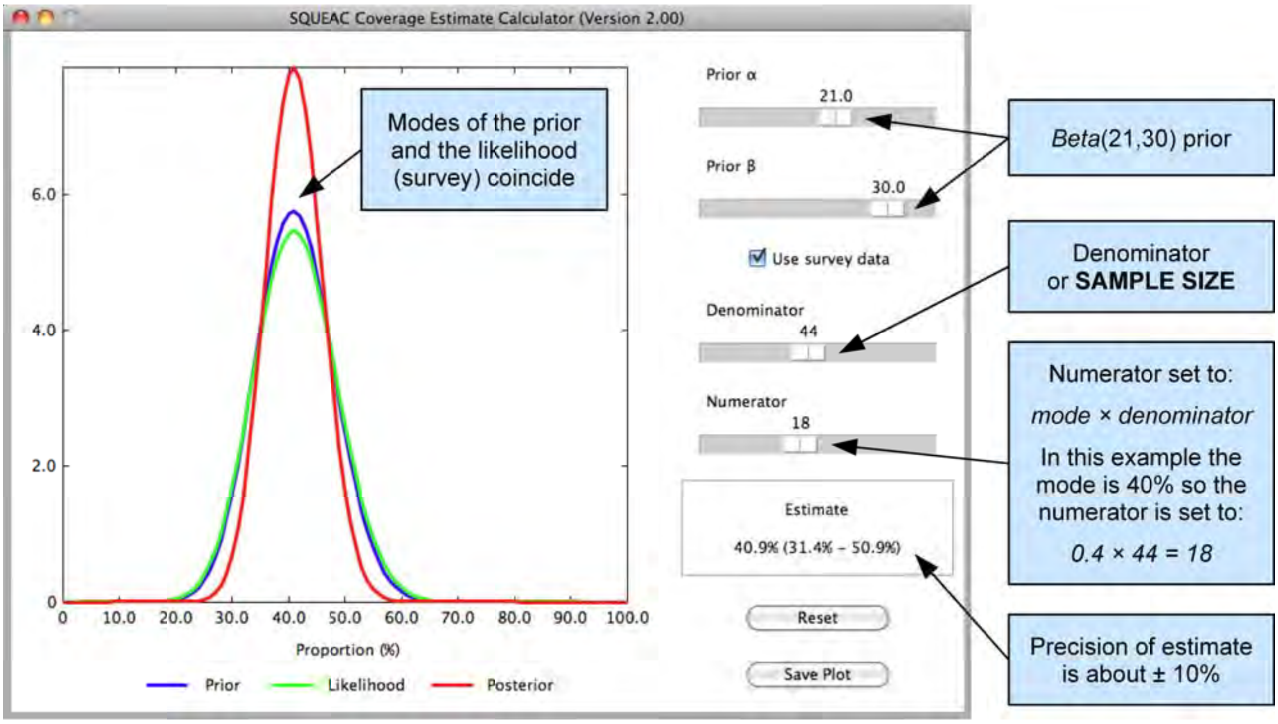
$$n_{Likelihood} = \left\lceil \frac{mode \times (1 - mode)}{(precision \div 1.96)^2} - (\alpha_{Prior} + \beta_{Prior} - 2) \right\rceil$$

It is easier to use the simulation approach if you have software, such as **BayesSQUEAC**, that reduces the overhead of having to perform repeated hand calculations. Spreadsheet software could also be used.

The simulation approach will return the same (or very similar) results as when the sample size formula is used. For example, **Figure A1-1** shows the likelihood survey sample size required to estimate coverage of 40% with a 95% credible interval of about ± 10 percentage points with a *Beta(21, 30)* prior found by simulation using **BayesSQUEAC**. The sample size found by the simulation shown in Figure A1-1 ($n_{Likelihood}$ = 44) is the same as that found using the sample size formula.

$$n_{Likelihood} = \left\lceil \frac{0.4 \times (1 - 0.4)}{(0.1 \div 1.96)^2} - (21 + 30 - 2) \right\rceil = 44$$

Figure A1-1. Sample size calculation by simulation using **BayesSQUEAC**



Additional Sample Size Guidelines

Early SQUEAC use studies found a tendency amongst program staff to specify a prior that was too strong/too narrow based on the available prior information. The method was subsequently adjusted to overcome these tendencies by:

- Specifying a sample size that ensures the likelihood is as at least as strong as the prior
- Placing an upper limit of 35 on both α_{Prior} and β_{Prior} shape parameter values
- Stressing in documentation that uncertainty about the position of the prior mode should seldom be specified as better than about ± 20 percentage points

A $Beta(\alpha_{Prior}, \beta_{Prior})$ prior provides the same information as a survey with a sample size of:

$$n = \alpha_{Prior} + \beta_{Prior} - 2$$

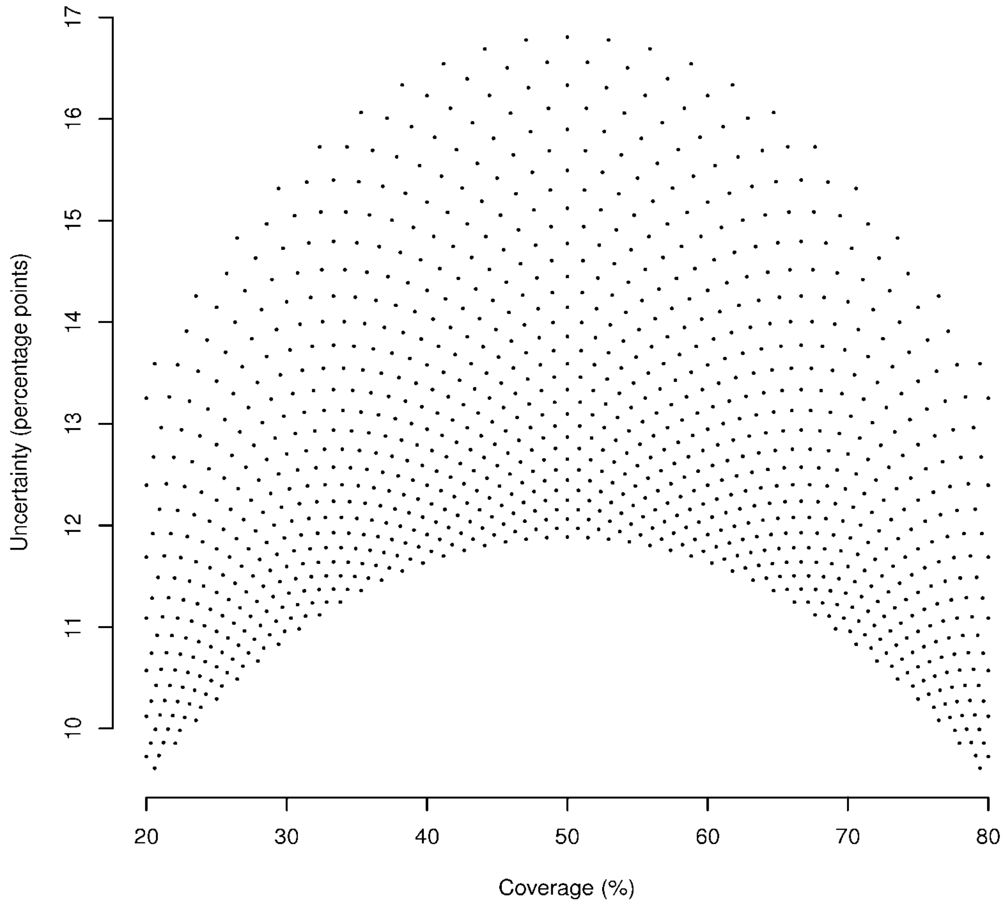
The sample size used for the likelihood survey should be sufficiently large to ensure that the likelihood data are able to correct a poorly specified prior. The minimum sample size guideline:

$$n_{min} = \alpha_{Prior} + \beta_{Prior} - 2$$

ensures that the likelihood is at least as strong as the prior.

The limit of 35 on both α_{Prior} and β_{Prior} is also based on considerations of the equivalent sample size associated with the prior, and limits this to between $n = 34$ and $n = 68$. This places limits on the uncertainty about the position of the prior mode. **Figure A1-2** shows the ranges of uncertainty associated with sample sizes between 34 and 68 over the typical range of coverages achieved by CMAM programs.

Figure A1-2. Uncertainty/precision associated with sample sizes between 34 and 68 over the typical range of coverages achieved by CMAM programs



The stress that uncertainty about the position of the prior mode should seldom be specified as better than about ± 20 percentage points means that a sample size for the likelihood survey larger than about:

$$n_{likelihood} = \left\lceil \frac{7}{9}\alpha_{Prior} + \frac{7}{9}\beta_{Prior} - \frac{14}{9} \right\rceil$$

is required to estimate coverage with a precision of ± 15 percentage points or better and a sample size for the likelihood survey larger than about:

$$n_{likelihood} = \lceil 3\alpha_{Prior} + 3\beta_{Prior} - 6 \rceil$$

is required to estimate coverage with a precision of ± 10 percentage points or better.

The guidelines related to sample size help to ensure a reasonably large sample size for the likelihood survey and help prevent posterior estimates being dominated by an overly strong prior.

Additional sample size guidelines apply when using small values of α_{Prior} *and* β_{Prior} and analysing data by hand. These are discussed in the following section on 'Formula for the 95% Credible Interval on the Posterior Mode'.

Formula for the 95% Credible Interval on the Posterior Mode

The formula given to calculate the 95% credible interval on the posterior mode:

$$95\%\,CI = mode \pm 1.96 \times \sqrt{\frac{\alpha_{Posterior} \times \beta_{Posterior}}{(\alpha_{Posterior} + \beta_{Posterior})^2 \times (\alpha_{Posterior} + \beta_{posterior} + 1)}}$$

returns an approximate *equal-tailed* 95% credible interval on the posterior mode rather than an exact *highest posterior density* (HPD) 95% credible interval on the posterior mode. The term:

$$\frac{\alpha_{Posterior} \times \beta_{Posterior}}{(\alpha_{Posterior} + \beta_{Posterior})^2 \times (\alpha_{Posterior} + \beta_{posterior} + 1)}$$

is the variance of $Beta(\alpha_{Posterior}, \beta_{Posterior})$, and *1.96* closely approximates the standard normal deviate for the 97.5th percentile point of the standard normal distribution. This approach is analogous to using the normal approximation when calculating an approximate 95% confidence interval on a binomial proportion. This is considered safe when:

$$np > 5 \quad \text{and} \quad n - np > 5$$

where:

$$n = \alpha_{Posterior} + \beta_{Posterior} - 2$$

$$p = posterior\ mode$$

Problems occur at extremely low and extremely high values of *p* or when very small effective sample sizes (*n*) are used. The purpose of the guideline:

$$\alpha_{Posterior} + \beta_{Posterior} - 2 \geq 30$$

is to ensure that a safe effective sample size (*n*) is used for values of *p* between about 17% and 83%. Most CMAM programs achieve levels of coverage within this range.

A normal approximation will only be safe when the $Beta(\alpha_{Posterior}, \beta_{Posterior})$ distribution can be approximated by a normal distribution. A $Beta(\alpha, \beta)$ distribution can be approximated by a normal distribution when α and β are sufficiently large such that:

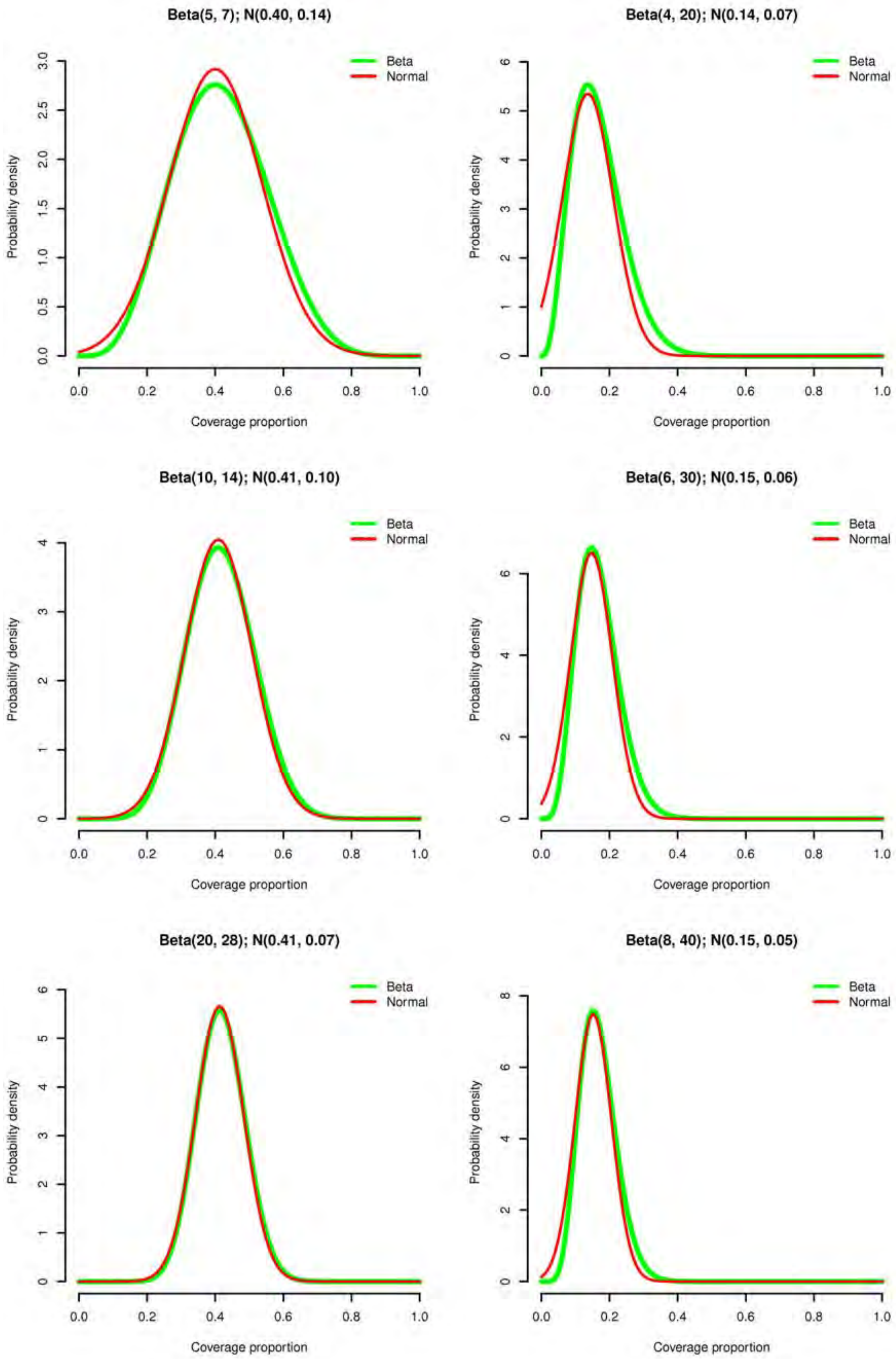
$$\frac{\alpha + 1}{\alpha - 1} \approx 1 \quad \text{and} \quad \frac{\beta + 1}{\beta - 1} \approx 1$$

A common rule-of-thumb is that a normal approximation is probably safe to use when both α and β are greater than or equal to 6 and that the normal approximation is safe to use when both α and β are greater than or equal to 10. For example, **Figure A1-3** shows how well the normal approximation fits $Beta(\alpha, \beta)$ distributions with two different modes and different values of α and β. The purpose of the guideline:

$$\lfloor \alpha_{Prior} + mode \times n_{Likelihood} \rfloor \geq 10 \quad \textbf{and} \quad \lfloor \beta_{prior} + n_{Likelihood} - mode \times n_{Likelihood} \rfloor \geq 10$$

is to ensure that the normal approximation returns reasonably accurate results.

Figure A1-3. Normal approximation to *Beta(α, β)* distributions with two different modes and different values of α and β



The 95% Credible Interval on the Posterior Mode in BayesSQUEAC

BayesSQUEAC estimates the mode of the posterior from α_{Prior} and β_{Prior} using the standard formula:

$$mode = \frac{\alpha_{Posterior} - 1}{\alpha_{Posterior} + \beta_{Posterior} - 2}$$

The 95% credible interval is found by a bootstrap aggregation ('bagging') method:

1. 100 sets of 100 replicates drawn at random from the $Beta(\alpha_{Posterior}, \beta_{Posterior})$ distribution are generated.

2. The 0.025 and 0.975 quantiles of each of the 100 sets of replicates are found.

3. The means of the sets of 0.025 and 0.975 quantiles from Step 2 are calculated.

The bagging algorithm is faster and more accurate than a simple bootstrap algorithm. Using 10,000 (i.e., 100 sets of 100) replicates the bagging algorithm provides an 'approximately exact' equal-tailed 95% credible interval.

BayesSQUEAC does **not** use a normal approximation to $Beta(\alpha_{Posterior}, \beta_{Posterior})$. This means that the guidelines:

$$\alpha_{Posterior} + \beta_{Posterior} - 2 \geq 30$$

and:

$$\lfloor \alpha_{Prior} + mode \times n_{Likelihood} \rfloor \geq 10 \ \ \textbf{and} \ \ \lfloor \beta_{prior} + n_{Likelihood} - mode \times n_{Likelihood} \rfloor \geq 10$$

need **not** be applied if data are to be analysed using **BayesSQUEAC**.

Formulas for Finding Suitable Values for α_{Prior} and β_{Prior}

The formulas used for finding suitable values for α_{Prior} and β_{Prior} are:

$$\mu = \frac{minimum + 4 \times mode + maximum}{6}$$

$$\sigma = \frac{maximum - minimum}{6}$$

$$\alpha_{Prior} = \mu \times \left(\frac{\mu \times (1 - \mu)}{\sigma^2} - 1 \right)$$

$$\beta_{Prior} = (1 - \mu) \times \left(\frac{\mu \times (1 - \mu)}{\sigma^2} - 1 \right)$$

The formulas for μ and σ are taken from the *three-point estimation* approach to task-duration modelling used in the Project Evaluation and Review Technique (PERT) and Critical Path Method (CPM) project management techniques.

Distributions that can be defined by the three points of mode, minimum, and maximum values are useful in PERT and CPM because task completion times are commonly specified as the most likely case (mode), the worst case (maximum), and the best case (minimum). The reason that the three-point estimation approach is used in PERT and CPM is because it is quick, easy, accurate, reliable, and easy to teach.

The approach used in PERT and CPM is:

1. Make informed guesses of the mode, minimum, and maximum.

2. Estimate the mean (μ) and the standard deviation (σ) of a *beta* distribution using the mode, minimum and maximum:

$$\mu = \frac{minimum + 4 \times mode + maximum}{6}$$

$$\sigma = \frac{maximum - minimum}{6}$$

3. Find suitable values for the α and β shape parameters of the required *beta* distribution using the standard formulas:

$$\alpha_{Prior} = \mu \times \left(\frac{\mu \times (1 - \mu)}{\sigma^2} - 1 \right)$$

$$\beta_{Prior} = (1 - \mu) \times \left(\frac{\mu \times (1 - \mu)}{\sigma^2} - 1 \right)$$

The same process is used in SQUEAC. The method is *approximate*. This is acceptable in Bayesian analysis because a prior is never exact.

Simplified LQAS and the Rule-of-Thumb Formula

The simplified LQAS method was developed to meet the need for SQUEAC and SLEAC assessments to be conducted without using computers. The method is 'simplified' because it does not bother the user with such matters as the selection of appropriate probability distributions, specification of lower and upper triage thresholds, or the specification of provider and consumer errors. Unlike conventional sequential methods, the simplified LQAS method **starts** with a given sample size and works from that. The rule-of-thumb formula provides a 'rough and ready' way of creating an LQAS sampling plan.

Applying the rule-of-thumb formula to 50% standard (p) with a sample size of 11 gives:

$$d = \left\lfloor \frac{11}{2} \right\rfloor = \lfloor 5.5 \rfloor = 5 \quad \text{which is the same as} \quad d = \left\lfloor 11 \times \frac{p}{100} \right\rfloor = \left\lfloor 11 \times \frac{50}{100} \right\rfloor = \lfloor 5.5 \rfloor = 5$$

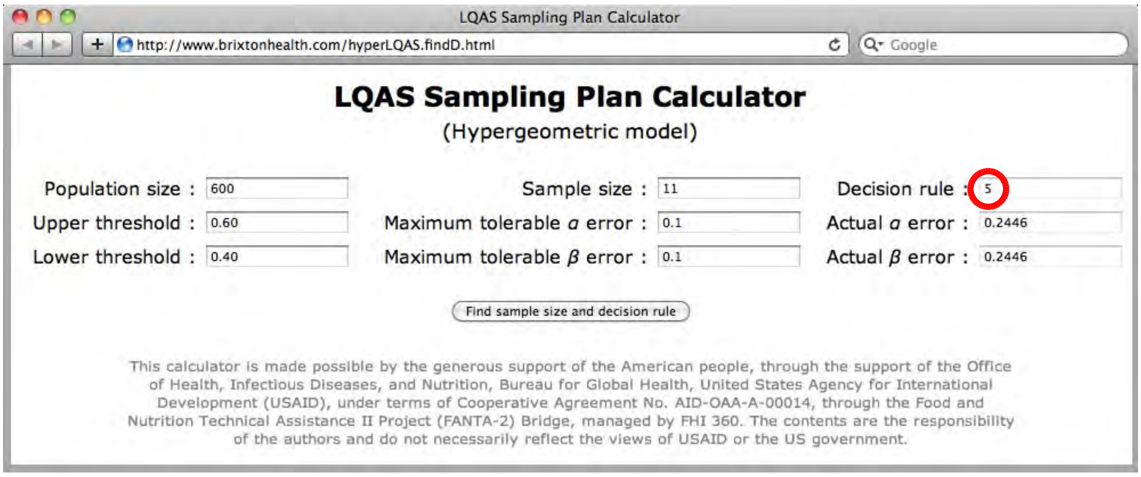
Conventional LQAS would use the *hypergeometric* distribution for this problem because the sample is drawn **without** replacement from a (usually) small population of SAM cases. Using a reverse hypergeometric LQAS calculator designed to find the best sampling plan for a given sample size:

`http://www.brixtonhealth.com/hyperLQAS.findD.html`

and specifying:

Upper value :	0.60	These specify a 50% swing point
Lower value :	0.40	
Sample size :	11	
Other parameters :	Use default values	

the calculator returns:



The calculator returns $d = 5$, which is the same value as returned by the rule-of-thumb formula.

The rule-of-thumb formula will return the same or very similar values for a given sample size and swing point as conventional LQAS using the binomial or hypergeometric distribution.

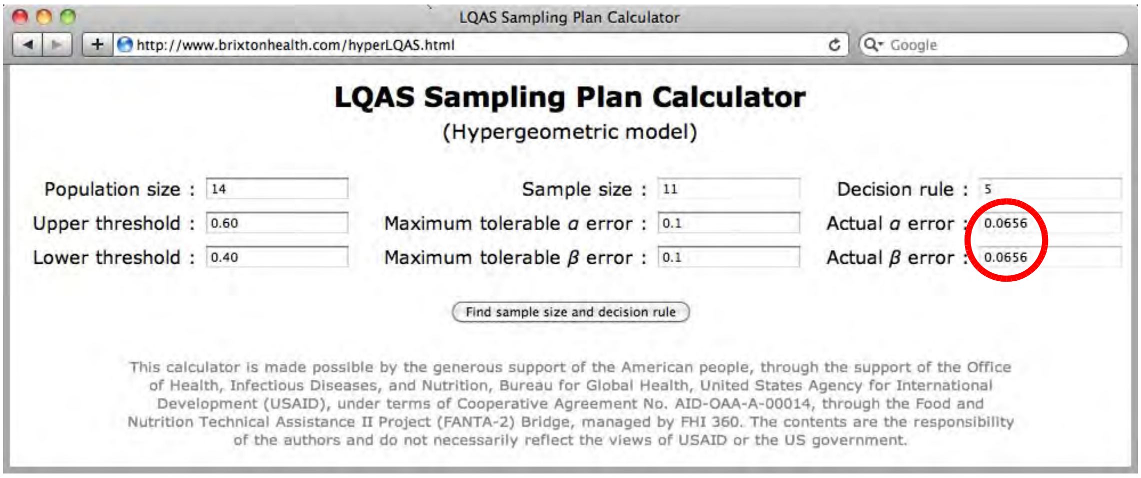
The simplified LQAS method does not quantify error, but there are good reasons to believe that errors will be small. The population size is small and the active and adaptive case-finding method (if applied correctly) yield a large sampling proportion. Assuming the case-finding method has an 80% exhaustivity (i.e., the method finds 80% of all SAM cases in the area sampled) and continuing with the $p = 50\%$, $n = 11$ example, then the population size would be:

$$\textit{Population size} = \left\lceil 11 \times \frac{1}{0.8} \right\rceil = 14$$

Using the reverse hypergeometric LQAS calculator and specifying:

Upper value : 0.60
Lower value : 0.40 — These specify a 50% swing point
Sample size : 11
Population size : 14
Other parameters : Use default values

the calculator returns:



The errors associated with the decision rule found using the rule-of-thumb formula are reasonably small. It is important to realize that error in small-area surveys is more complicated than classifying coverage **without** a prior hypothesis about a proportion. SQUEAC uses a two-stage model that aims to confirm or deny a prior hypothesis regarding coverage. In such circumstances, the practical meaning of errors (i.e., the positive and negative predictive values) are very different from the naive LQAS case. This is analogous to a two-stage screening-test model.

Although error is not specified in the simplified LQAS method, it is possible to favour one error over another by altering the rounding rule used in the rule-of-thumb formula:

Rounding rule		
Round down (*floor*)	**Round to nearest whole number**	**Round up (*ceiling*)**
$d = \left\lfloor n \times \frac{p}{100} \right\rfloor$	$d = \left\Vert n \times \frac{p}{100} \right\Vert$	$d = \left\lceil n \times \frac{p}{100} \right\rceil$

Different rounding rules favour different errors:

- The rule-of-thumb formula used in the SQUEAC documentation employs the rounding down (floor) rule. For measures of success (e.g., coverage), this favours the provider. For measures of failure (e.g., defaulting), this favours the consumer.

- Rounding to the nearest whole number favours neither consumer nor the provider.

- Rounding up favours the consumer for measures of success (e.g., coverage) and favours the provider for measures of failure (e.g., defaulting).

To avoid confusion, the SQUEAC documentation presents a single rounding rule. The rounding down (floor) rule was chosen for simplicity. More advanced users of the simplified LQAS method might like to use rounding up for measures of success (e.g., coverage) and rounding down for measures of failure (e.g., defaulting). Such a strategy will always favours the consumer.

Modelling Coverage Using the Binomial Distribution

The use of the beta-binomial conjugate analysis assumes that it is safe to model likelihood survey data using the binomial distribution. The binomial model assumes that the likelihood survey samples *n* observations from a population of size *N* **with** replacement. The likelihood survey violates this assumption by sampling **without** replacement. This violation is not usually considered to be problematic if the sampling proportion:

$$\text{Sampling proportion} = \frac{n}{N}$$

remains small. The most frequently used rule-of-thumb is that it is safe to use the binomial model when:

$$n \leq N \times 0.1$$

At this sampling proportion (10%), under-dispersion (the ‘finite population correction’) is just below 0.95. With a sampling proportion of 20%, it is just below 0.90.

The likelihood survey estimates coverage in the program area using a sample of current and recovering SAM cases in the program area found in a sample of villages in the program area. Experience with CSAS, SQUEAC, and SLEAC indicates that it is rare for more than about 10% or 15% of villages to be sampled to achieve the required sample size. The proportion of villages sampled may be used as a proxy for the sampling proportion and will usually overestimate the true sampling proportion because villages are not selected using population proportional sampling.

Data arising from a sample drawn from a small population **without** replacement should be modelled using the hypergeometric distribution. Working with the hypergeometric distribution requires good estimates of the population size (i.e., the total number of current and recovering SAM cases in the program area). It is uncommon to have good estimates of the population size with which to calculate exact hypergeometric probabilities. The binomial model does not require this information. SQUEAC likelihood surveys, therefore, use the binomial approximation for the hypergeometric. The nature of the error associated with the use of the binomial distribution is related to the relative under-dispersion of the hypergeometric compared to the binomial. The estimate of mode (i.e., the point estimate of coverage) is **not** affected. Use of the binomial model results in the use of a slightly larger sample size and produces slightly wider credible intervals than if the hypergeometric model were used.

It is important to note that the beta-binomial conjugate analysis is convenient in the sense that the method of analysis is relatively simple. This informed the choice of method because of the need to provide methods that could be performed without using a computer. It is not feasible to use the hypergeometric distribution without using a computer. Software to perform the required analysis (e.g., **`winBUGS`**, **`openBUGS`**, **`JAGS`**) is designed to be used by professional statisticians and is probably too complicated for the majority of the SQUEAC user group.

The Use of n = 40 as the Standard Sample Size for SLEAC

The use of n = 40 as the standard sample size for SLEAC surveys was informed by the need to provide accurate and reliable classification of coverage as being above or below Sphere minimum standards using small sample sizes. This sample size was selected after computer-based simulation of the two-class and three-class simplified LQAS methods. Simulation parameters were:

<table>
<tr><th colspan="2">Parameter</th><th>Values</th></tr>
<tr><td colspan="2">Sample sizes</td><td>20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80</td></tr>
<tr><td colspan="2">Coverage proportion</td><td>0%, 1%, 2%, …, 98%, 99%, 100%</td></tr>
<tr><td colspan="2">Simulated surveys</td><td>1,000 simulated surveys at each coverage proportion</td></tr>
<tr><td colspan="2">Population size</td><td>400</td></tr>
<tr><td rowspan="4">Thresholds</td><td rowspan="2">Two classes</td><td>50% (Sphere rural)</td></tr>
<tr><td>70% (Sphere urban/camp)</td></tr>
<tr><td rowspan="2">Three classes</td><td>20%/50%</td></tr>
<tr><td>30%/70%</td></tr>
</table>

The population size of 400 is derived from the following assumptions:

- The largest service delivery unit for which coverage will be classified is a health district and these are seldom larger than 100,000 total population.
- Approximately 20% of the population are aged between 6 and 59 months.
- A SAM prevalence of 2%.

These parameters yielded a population size of:

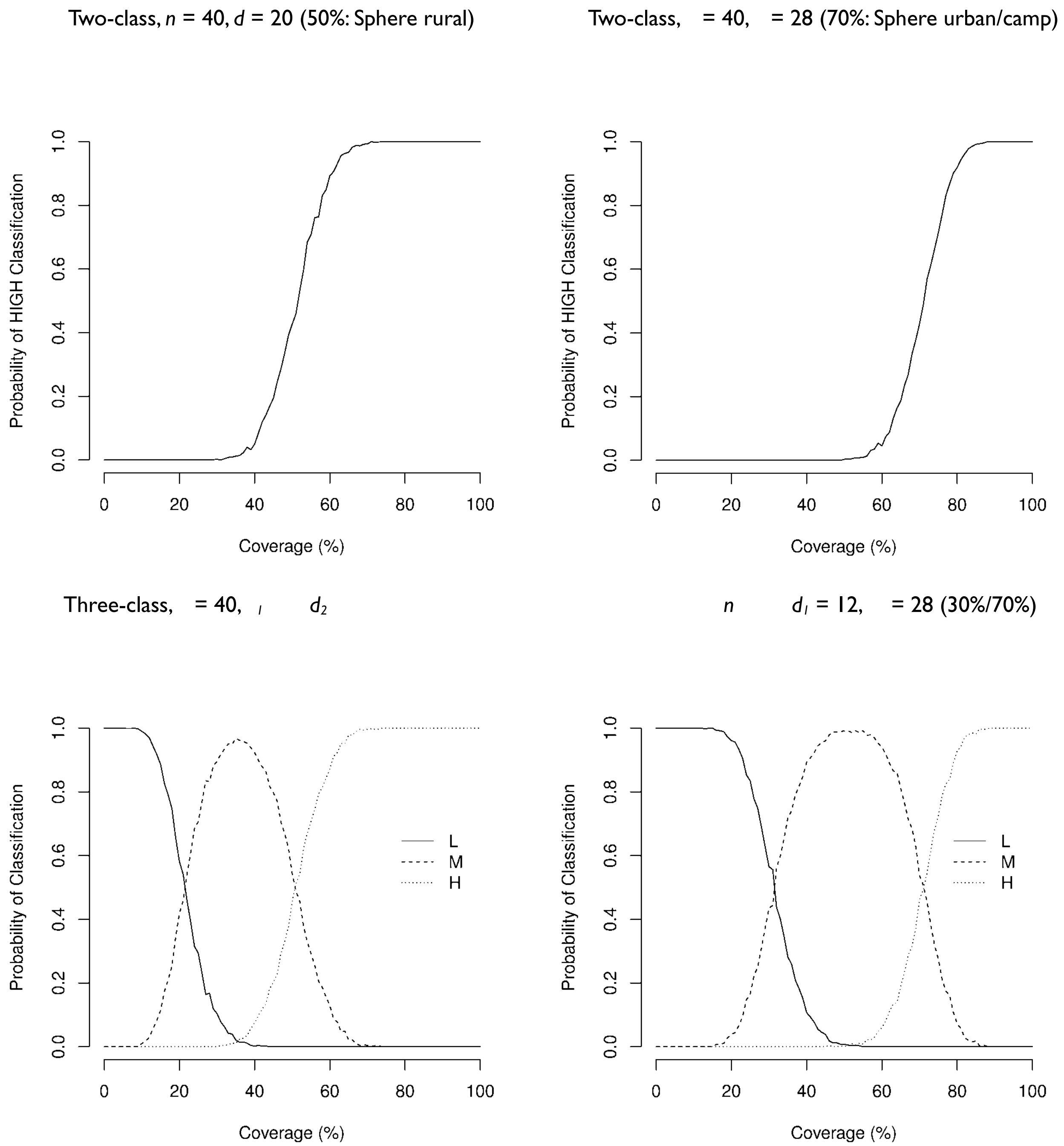
$$\textit{Population Size} = 100{,}000 \times \frac{20}{100} \times \frac{2}{100} = 400$$

High parameter values were selected to maximise the population size.

The performance of the method was assessed by examination of operating characteristic curves and probability of classification plots.

A minimum sample size of n = 40 was selected as appropriate for **SLEAC** surveys. **Figure A1-4** shows the operating characteristic plots and probability of classification plots found for n = 40. The curves are steep and there are no gross misclassifications. At lower prevalences or in smaller populations, the method will perform better than indicated by the operating characteristic curves and probability of classification plots shown in Figure A1-4 as the sampling proportion increases.

Figure A1-4. Operating characteristic and probability of classification plots found from simulations of two-class and three-class SLEAC methods with n = 40



Smoothing Time-Series Data Using Moving Averages

SQUEAC requires you to analyse a variety of time-series, such as admissions and exits over time. This type of data usually requires smoothing before being plotted using line charts.

You can smooth data using the charting functions in spreadsheet applications. If you do this, make sure that you use a smoothing function that is suited to time-series data. These will usually be called something like ‘moving average’ or ‘running average’.

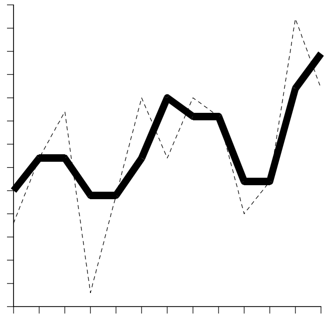
Do not worry if your spreadsheet package does not provide moving average functions. It is easy to program these functions yourself. **Figure A1-5** shows how to program a spreadsheet for three different types of moving average and the effect that these have on a time-series.

Figure A1-5. Programming a spreadsheet for three different moving averages with a span of three successive data points

Running medians-of-three (M3)

	A	B
1	Raw	M3
2	98	=MEDIAN(A2:A3)
3	112	=MEDIAN(A2:A4)
4	122	=MEDIAN(A3:A5)
5	83	=MEDIAN(A4:A6)
6	104	=MEDIAN(A5:A7)
7	125	=MEDIAN(A6:A8)
8	112	=MEDIAN(A7:A9)
9	125	=MEDIAN(A8:A10)
10	121	=MEDIAN(A9:A11)
11	100	=MEDIAN(A10:A12)
12	107	=MEDIAN(A11:A13)
13	147	=MEDIAN(A12:A14)
14	127	=MEDIAN(A13:A14)

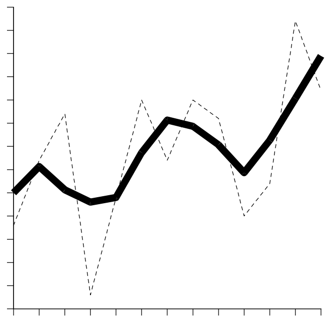
	A	B
1	Raw	M3
2	98	105
3	112	112
4	122	112
5	83	104
6	104	104
7	125	112
8	112	125
9	125	121
10	121	121
11	100	107
12	107	107
13	147	127
14	127	135



Running averages-of-three (A3)

	A	B
1	Raw	A3
2	98	=AVERAGE(A2:A3)
3	112	=AVERAGE(A2:A4)
4	122	=AVERAGE(A3:A5)
5	83	=AVERAGE(A4:A6)
6	104	=AVERAGE(A5:A7)
7	125	=AVERAGE(A6:A8)
8	112	=AVERAGE(A7:A9)
9	125	=AVERAGE(A8:A10)
10	121	=AVERAGE(A9:A11)
11	100	=AVERAGE(A10:A12)
12	107	=AVERAGE(A11:A13)
13	147	=AVERAGE(A12:A14)
14	127	=AVERAGE(A13:A14)

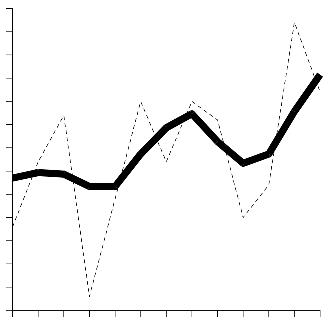
	A	B
1	Raw	A3
2	98	105
3	112	111
4	122	106
5	83	103
6	104	104
7	125	114
8	112	121
9	125	119
10	121	115
11	100	109
12	107	116
13	147	125
14	127	135



Running medians-of-three followed by running averages-of-three (M3A3)

	A	B	C
1	Raw	M3	A3
2	98	=MEDIAN(A2:A3)	=AVERAGE(B2:B3)
3	112	=MEDIAN(A2:A4)	=AVERAGE(B2:B4)
4	122	=MEDIAN(A3:A5)	=AVERAGE(B3:B5)
5	83	=MEDIAN(A4:A6)	=AVERAGE(B4:B6)
6	104	=MEDIAN(A5:A7)	=AVERAGE(B5:B7)
7	125	=MEDIAN(A6:A8)	=AVERAGE(B6:B8)
8	112	=MEDIAN(A7:A9)	=AVERAGE(B7:B9)
9	125	=MEDIAN(A8:A10)	=AVERAGE(B8:B10)
10	121	=MEDIAN(A9:A11)	=AVERAGE(B9:B11)
11	100	=MEDIAN(A10:A12)	=AVERAGE(B10:B12)
12	107	=MEDIAN(A11:A13)	=AVERAGE(B11:B13)
13	147	=MEDIAN(A12:A14)	=AVERAGE(B12:B14)
14	127	=MEDIAN(A13:A14)	=AVERAGE(B13:B14)

	A	B	C
1	Raw	M3	A3
2	98	105	109
3	112	112	110
4	122	112	109
5	83	104	107
6	104	104	107
7	125	112	114
8	112	125	119
9	125	121	122
10	121	121	116
11	100	107	112
12	107	107	114
13	147	127	123
14	127	135	131



Moving averages can be applied several times. This involves applying a smoothing method to previously smoothed data, as is done with the M3A3 smoother shown in Figure A1-5. With the M3A3 smoother, the data are smoothed by taking the medians of sets of three successive data points (M3). The results are then smoothed by taking the arithmetic means of sets of three successive smoothed data points (A3). The more times you apply a moving average, the more smoothing is applied to the data.

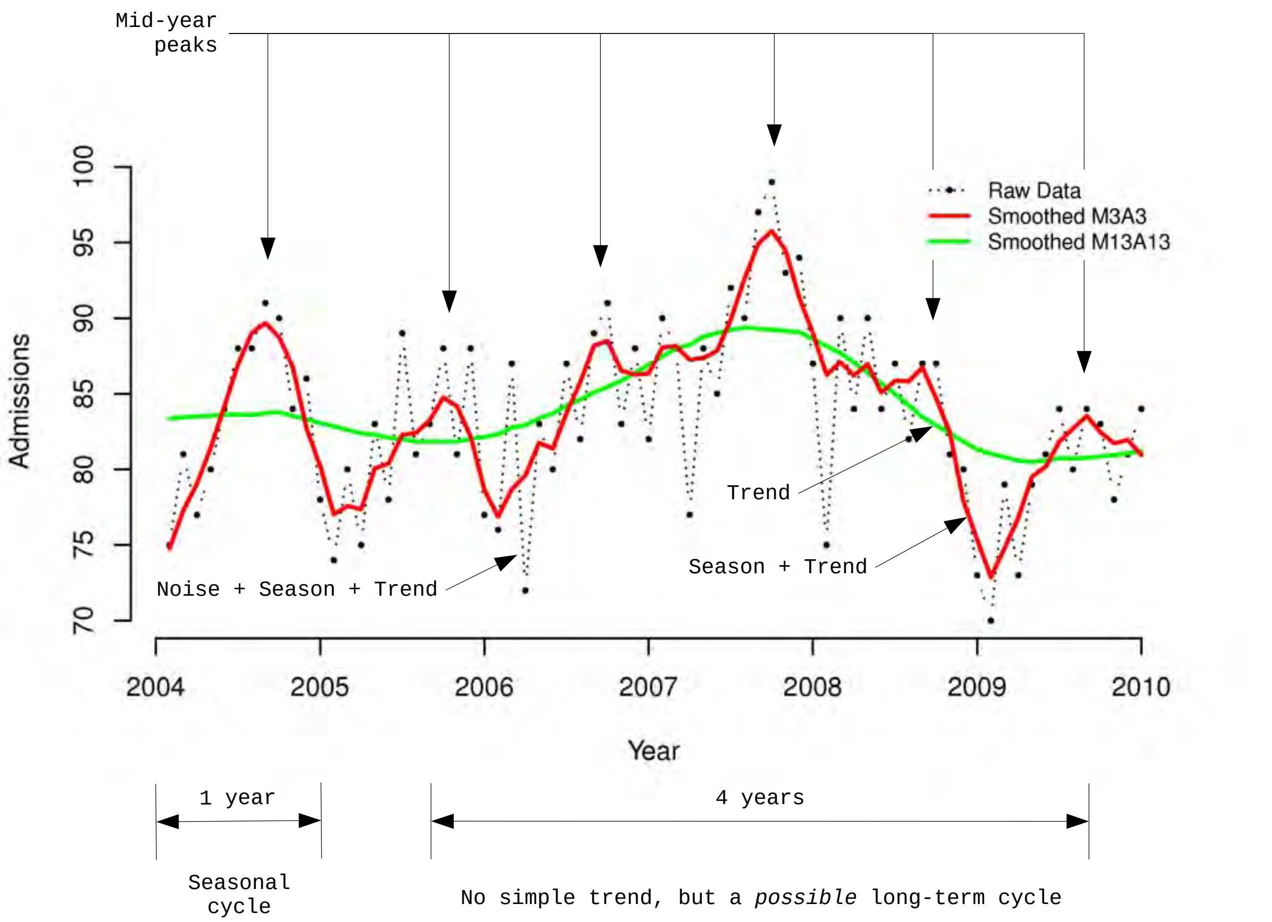
A time-series can be thought of as a combination of random (irregular or 'noise'), seasonal, and trend components. Judicious application of smoothing techniques, such as moving averages, hides some of these components and helps uncover other components of the time-series:

- Smoothing using moving averages of short spans (i.e., of just a few successive data points) will tend to hide the random 'noise' component and help reveal the seasonal and trend components of the time-series.

- Smoothing using moving averages of longer spans (i.e., of enough data points to cover an entire seasonal cycle) will tend to hide both the random 'noise' and seasonal components and help reveal the trend component of the time-series.

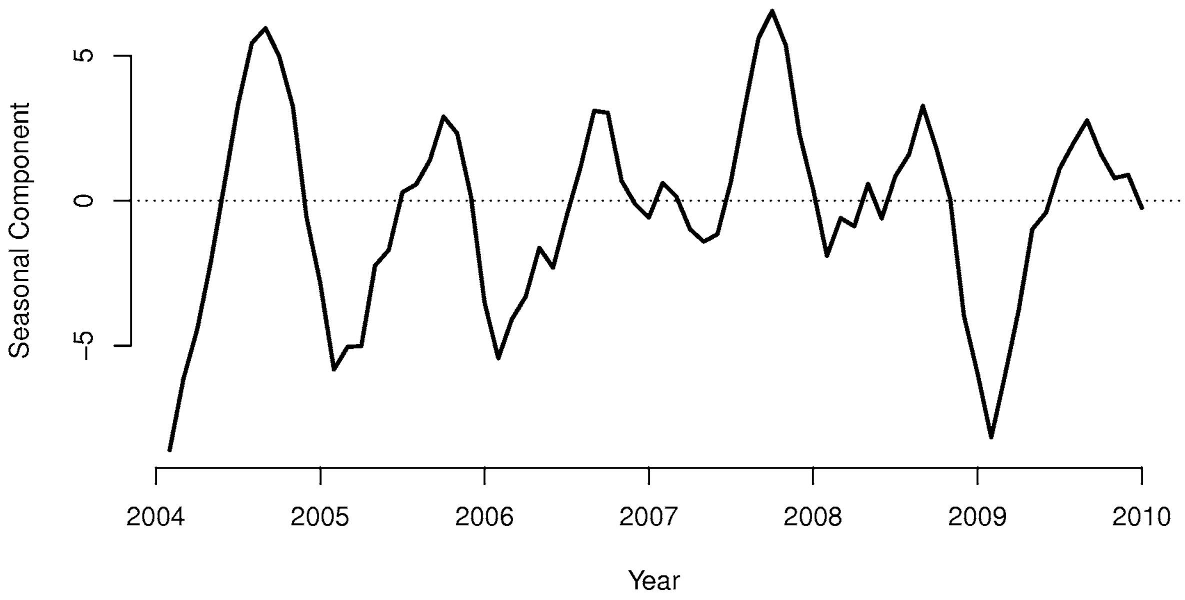
Figure A1-6 shows the effect of smoothing 6 years of monthly data using smoother spans of 3 to hide the random (irregular or 'noise') component and smoother spans of 13 to hide both the random (irregular or 'noise') and seasonal components.

Figure A1-6. Effects of moving average smoothers of spans 3 (M3A3) and 13 (M13A13) on 6 years of monthly admissions data



It is also possible to find the seasonal component alone by subtracting the trend component (i.e., the data smoothed using M13A13) from the combined season and trend components (i.e., the data smoothed using M3A3). This is shown in **Figure A1-7**.

Figure A1-7. Seasonal component of 6 years of monthly admissions data found by subtracting data smoothed using M13A13 from data smoothed using M3A3



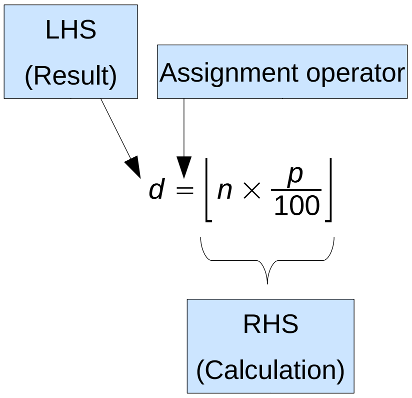
Appendix 2. Working with Formulas

Formulas are devices that precisely describe calculations. A number of mathematical formulas are presented in this document. This appendix contains some basic information to help you work with these formulas.

Parts of Formulas

Formulas are made up from a number of basic parts. These are:

Sides. A formula usually has two sides. One side, usually the left-hand side (LHS), is reserved for the result of the calculation. The other side, usually the right-hand side (RHS), precisely describes the data and calculations required to find the result. The two sides are separated by an assignment (=) operator. For example:



Constants. Constants are numbers that are the same for all contexts. They are shown in formulas as numbers. In this formula:

$$d = \left\lfloor n \times \frac{p}{100} \right\rfloor$$

the number 100 is a constant. The value 100 is used whenever this formula is used. It is a constant because it does not change.

Variables. Variables are placeholders for values that change (i.e., are different in different contexts). You have to provide the appropriate values for each of the variables when you use a formula. In this formula:

$$d = \left\lfloor n \times \frac{p}{100} \right\rfloor$$

the letters n and p are variables. The letters or names given to variables frequently indicate the type of data that is required to perform the calculation and the type of data that is produced as the result of the calculation. In the example formula, d is a number that is used to make a decision (i.e., d for 'decision'), n is used to indicate a sample size (i.e., n for 'number') and p is used to indicate a percentage (i.e., p for 'percentage'). The use of variables allows formulas to present calculations and methods in a way that enables the same formula to be used in different contexts with different data.

Operators. Operators specify the operations that are performed on the constants and variables in a formula. The operators used in SQUEAC and SLEAC formulas are shown in **Table A2-1**. This formula:

$$d = \left\lfloor n \times \frac{p}{100} \right\rfloor$$

uses the 'floor' (round down), multiply, and divide operators. The operators are applied to the constants and variables in a specific order called the *order of precedence*.

Table A2-1. Operators used in SQUEAC and SLEAC formulas

Symbol	Meaning	Example	Notes
$=$	Assignment	$d = \left\lfloor n \times \frac{p}{100} \right\rfloor$	Separates the result and the calculations in a formula.
$+$	Add	$4 + 2 = 6$	Simple addition
$-$	Subtract	$4 - 2 = 2$	Also used to indicated negation as in: $-2 = 0 - 2$ or 'minus 2' or 'negative 2'
$\times$	Multiply	$4 \times 2 = 8$	A 'dot-product' operator may be used in formulas. The operator may also be missing in formulas. For example: $a \times b = a \cdot b = ab$
$\div$	Divide	$4 \div 2 = 2$ or $\frac{4}{2} = 2$	The 'over' operator is useful for grouping expressions. For example: $\frac{2+2}{3-1} = (2+2) \div (3-1) = 4 \div 2 = 2$
$\pm$	Add and subtract	$6 \pm 3 = \{ 3, 9 \}$	This operator return two values. It is used when calculating credible limits. For example: $95\%\,CI = mode \pm 1.96 \times \sqrt{\frac{\alpha \times \beta}{(\alpha + \beta)^2 \times (\alpha + \beta + 1)}}$
$\sum$	Sum a series of numbers	$\sum \{ 1, 2, 3 \} = 1 + 2 + 3 = 6$	Used when calculating a mean (average) value. If, for example: $x = \{ 6, 9, 7 \}$ Then the mean value is: $Mean = \frac{\sum x}{3} = \frac{6 + 9 + 7}{3} = \frac{22}{3} = 7.33$
$\lfloor x \rfloor$	Round x towards zero (*floor*)	$\lfloor 2.7 \rfloor = 2$	Used when finding classification thresholds for simplified LQAS sampling plans. For example: $d = \left\lfloor 11 \times \frac{70}{100} \right\rfloor = \lfloor 11 \times 0.7 \rfloor = \lfloor 7.7 \rfloor = 7$ Rounding operators act like brackets grouping expressions and altering the normal order of precedence.
$\lceil x \rceil$	Round x away from zero (*ceiling*)	$\lceil 2.7 \rceil = 3$	Use in sample size calculations. For example: $n = \left\lceil \frac{41}{600 \times \frac{20}{100} \times \frac{1}{100}} \right\rceil = \lceil 34.17 \rceil = 35$ Rounding operators act like brackets grouping expressions and altering the normal order of precedence.
$\Vert x \Vert$	Round to nearest whole number	$\Vert 2.3 \Vert = 2$ and $\Vert 2.7 \Vert = 3$	May be used when finding classification thresholds for simplified LQAS sampling plans as explained in Appendix 1. Rounding operators act like brackets grouping expressions and altering the normal order of precedence.
(x)	Group expression	$3 \times (2 + 1) = 3 \times 3 = 9$	Brackets alter the order of precedence. The calculations in brackets are completed before all other calculations.
$[x]$	Group expression	$3 \times [2 + 1] = 3 \times 3 = 9$	As above. Square brackets are used to help avoid confusion in long or complicated calculations. For example: $Coverage = \sum \left[\left(\frac{N}{\sum N} \right) \times \frac{c}{n} \right]$
x^y	Power or exponentiation	$3^2 = 3 \times 3 = 9$	Exponentiation and roots are the same underlying operation as with, for example: $\sqrt{x} = x^{0.5}$
$\sqrt{x}$	Square root	$\sqrt{9} = 3$	

Order of Precedence

Formulas describe calculations precisely. For this to be possible, there are rules defining the order in which operations should be performed. These rules are called the order of precedence. The order in which operations should be performed is:

Order	Operations	Examples
1	Operations grouped together by brackets, rounding functions, or above or below the 'over' function Work from the inside out with nested brackets	Complete operations in brackets first: $3 \times (1 + 2) = 3 \times 3 = 9$ Work from the inside out with nested brackets: $4 + [3 \times (4 - 1)]^2$ $= 4 + [3 \times 3]^2$ $= 4 + 9^2$ $= 4 + 81$ $= 85$
2	Exponentiation and roots	Perform exponentiation first: $2 + 3^2 = 2 + 9 = 11$ Roots are a form of exponentiation: $11 - \sqrt{9} = 11 - 9^{0.5} = 11 - 3 = 8$
3	Division and multiplication evaluated from left to right	Work from left to right: $15 \div 3 \times 4 = 5 \times 4 = 20$ ✓ **not** $15 \div 3 \times 4 = 15 \div 12 = 1.25$ ✗
4	Addition and subtraction evaluated from left to right	Work from left to right: $10 - 3 + 2 = 7 + 2 = 9$ ✓ **not** $10 - 3 + 2 = 10 - 5 = 5$ ✗

One way to remember the order of precedence is **BEDMAS,** which stands for:

B	Brackets	
E	Exponentiation and roots	
D	Division	Equal precedence, evaluated left to right
M	Multiplication	
A	Addition	Equal precedence, evaluated left to right
S	Subtraction	

If you are not sure how to carry out the calculation specified in a formula, work through the examples in the text and check your results against the given results.

Most problems with using formulas are caused by failure to follow the order of precedence. For example:

$$15 \div 3 \times 4 = 5 \times 4 = 20 \quad ✓ \quad \textbf{\underline{not}} \quad 15 \div 3 \times 4 = 15 \div 12 = 1.25 \quad ✗$$

Percentages and Proportions

Many formulas in this document require you to use proportions or return a result that is expressed as a proportion. Conversions between proportions and percentages are straightforward:

$$Proportion = \frac{Percentage}{100} \quad \text{and} \quad Percentage = Proportion \times 100$$

Chains of Formulas

Some formulas are chained together. This is done to simplify calculations, with the results of one formula being used in subsequent formulas. For example:

$$\mu = \frac{minimum + 4 \times mode + maximum}{6}$$

$$\sigma = \frac{maximum - minimum}{6}$$

$$\alpha_{Prior} = \mu \times \left(\frac{\mu \times (1 - \mu)}{\sigma^2} - 1 \right)$$

$$\beta_{Prior} = (1 - \mu) \times \left(\frac{\mu \times (1 - \mu)}{\sigma^2} - 1 \right)$$

In this set of formulas, the results that we are interested in are α_{Prior} and β_{Prior}, which are calculated using μ and σ, which are calculated from the variables called *minimum*, *mode*, and *maximum*.

It is possible, for example, to calculate β_{Prior} directly using *minimum*, *mode*, and *maximum* as:

$$\beta_{Prior} = \left(1 - \frac{minimum + 4 \times mode + maximum}{6}\right) \times \left[\frac{\frac{minimum + 4 \times mode + maximum}{6} \times \left(1 - \frac{minimum + 4 \times mode + maximum}{6}\right)}{\left(\frac{maximum - minimum}{6}\right)^2} - 1 \right]$$

but the calculation would be complicated, repetitious, and prone to error.

If in Doubt …

If you are not sure how to carry out the calculation specified in a formula, work through the examples given in the text and check your results against the results given in the text. Most problems are caused by failure to follow the order of precedence. Another common problem is using percentages when proportions are required. Most formulas in the text require you to use proportions and return proportions.

Appendix 3. Glossary of terms

Accuracy. The degree of closeness of a measurement of a quantity to the measured quantity's true value. See *precision*.

Active and adaptive case-finding. A type of sampling used in SQUEAC small-area surveys, SQUEAC likelihood surveys, and SLEAC surveys. This type of sampling searches actively for cases of SAM with the intention of finding all (or nearly all) cases of SAM in sampled communities. This type of sampling is also known as 'snowball sampling', 'optimally biased sampling', or 'chain-referral sampling'. See *sample, sampling*.

Active case. See *current case*.

Acute malnutrition. A form of undernutrition caused by infection and/or a decrease in food intake or uptake resulting in rapid weight loss (wasting) or bilateral pitting oedema. Acute malnutrition is defined by the presence of bilateral pitting oedema or wasting (low MUAC or low weight-for-height). See *bilateral pitting oedema*, *global acute malnutrition*, *kwashiorkor, mid-upper arm circumference*, *moderate acute malnutrition*, *severe acute malnutrition*, *visible severe wasting*.

Admission criteria. Rules describing individuals that are eligible for admission to a program. Also known as the program's 'case definition'. In CMAM programs, the admission criteria is usually 'MUAC < 115 mm or bilateral pitting oedema or visible severe wasting in a child aged between 6 months and 5 years'. Some CMAM programs may also use weight-for-height for admission. See *bilateral pitting oedema*, *case definition*, *discharge criteria*, *visible severe wasting*.

Aetiology. The cause, origin, or reason for something (usually of a disease or abnormal condition).

Anthropometric criteria. Admission or discharge criteria using anthropometry (usually MUAC). See *admission criteria*, *anthropometry*, *discharge criteria*, *mid-upper arm circumference*.

Anthropometry. Measurement of the proportion, size, or weight of the human body or a human body part. Anthropometric measurements are used to assess the nutritional status of individuals and population groups, and as admission and discharge criteria for nutrition support programs. CMAM programs use MUAC (for screening/case-finding, admission, monitoring response to treatment, discharge) and weight (for monitoring response to treatment and discharge). Some CMAM programs may also use weight-for-height for admission. See *case-finding*, *mid-upper arm circumference*, *screening*.

Areal. A description of one of more areas. In SLEAC and SQUEAC, for example, quadrats may be described as 'areal sampling units'. See *centric systematic area sampling*, *quadrat*, *systematic area sampling*.

ARI. Acronym for 'acute respiratory infection'. The acronym ARTI (for 'acute respiratory tract infection') is also in common use.

Attack phase. A phase in the program cycle during which coverage is increased. The term is usually applied to the first few months of program activity, but may be used to describe the period following program reforms designed to improve coverage.

Bar chart. A chart drawn using rectangular bars with lengths proportional to the values that they represent. The bars may be drawn vertically or horizontally. See *Pareto chart*.

Barrier. Anything that restrains, obstructs, or delays access to a program or restrains coverage. See *booster.*

Bayesian. The interpretation of probability as a measure of confidence (or belief) that something is true. In Bayesian inference, belief is modified as fresh evidence is observed. At each step, the initial belief is called the 'prior', the fresh evidence is called the 'likelihood', and the modified belief is called the 'posterior'. See *frequentist, likelihood, posterior, prior.*

Beneficiary record card. A card recording beneficiary data, including (but not limited to) identifying, locating, demographic, clinical, and anthropometric data. The card is used to record all relevant information about a beneficiary and the treatment episode. Beneficiary record cards usually follow the design developed by Valid International for use in CTC programs.

Best practice. A method or technique that has consistently shown results superior to those achieved by other means. In addition, a 'best' practice can evolve to become better as improvements are discovered. SQUEAC uses the clinical audit approach to evolve best practice (i.e., the practices that maximise program coverage). See *clinical audit*.

Beta-binomial conjugate analysis. In Bayesian inference, a type of conjugate analysis in which the prior and posterior are modelled using the beta distribution and the likelihood is modelled using the binomial distribution. See *beta distribution*, *binomial distribution*, *conjugate analysis, likelihood, posterior, prior.*

Beta distribution. A family of probability distributions defined on the interval [0, 1] parameterised by two positive shape parameters denoted α and β. The beta distribution is suited to the statistical modelling of proportions. SQUEAC uses the beta distribution to model the coverage proportion.

Bilateral pitting oedema. A sign of SAM caused by an accumulation of fluid in the interstitial tissue spaces (the areas surrounding the body's cells and blood vessels). See *acute malnutrition, kwashiorkor, severe acute malnutrition.*

Binary variable. A variable that can take one of two complementary values. In SQUEAC**,** this is usually the coverage status of a SAM case (i.e., the SAM case is either 'covered' or 'not covered'). See *variable*.

Binomial distribution. A probability distribution suited to the statistical modelling of proportions of a binary variable. SQUEAC uses the binomial distribution to model coverage in likelihood surveys. See *binary variable*, *likelihood*, *variable*.

Booster. Anything that encourages or enables access to a program or leads to an increase in coverage. See *barrier.*

Branching hierarchy. A way of organising findings in a mind-map. SQUEAC investigations tend to use mind-maps organised using a Central Theme → Data Source/Method → Individual Findings hierarchy. See *mind-map*.

Case, active. See *current case*.

Case, current. See *current case*.

Case, recovering. See *recovering case*.

Case definition. The method by which cases of SAM are defined for purposes of admission to a program. The term may apply to other situations, such as defining defaulters, defining failure to respond to treatment, or defining when a beneficiary has recovered and may be discharged from a program. See *admission criteria*, *discharge criteria*.

Case-finding. Activities aimed at finding and recruiting current cases. Effective CTC/CMAM programs usually employ many case-finding strategies, including screening children attending health centres, screening children in the community by program staff or CHWs, screening children attending vaccination sites, screening children attending growth monitoring programs, screening children by CBVs, and referring cases to the program by carers of program beneficiaries. The use of MUAC facilitates the use of diverse case-finding strategies. See *community-based volunteer, community health worker, current case*, *mid-upper arm circumference*, *screening*.

Case history. A detailed account of the facts affecting the development or condition of a person (or group) under treatment or study.

Catchment area. The area served by a service delivery unit such as a health centre or health post. See *service delivery unit*.

CBV. See *community-based volunteer.*

Census. The procedure of systematically acquiring and recording information about all members of a given population or household. See *census sample*.

Census sample. A sample that is designed to include all (or nearly all) members of a given population. SQUEAC uses census samples in small-area surveys and in likelihood surveys through the use of active and adaptive case-finding and house-to-house/door-to-door sampling. See *active and adaptive case-finding*, *sample*, *sampling*.

Centric systematic area sampling (CSAS). A way of taking a spatially stratified/spatially representative sample that involves drawing a grid of equally sized squares (‘quadrats’) over the area to be sampled and sampling the community or communities located closest to the centre of each square. See *sample*, *sampling*, *spatial*, *systematic area sampling*, *systematic sampling*.

CHW. See *community health worker.*

Clinic workload returns. Routine statistics on clinic activities usually including (but not limited to) counts of cases of specific diseases seen at a clinic in a reporting period.

Clinical audit. A quality improvement and monitoring method that seeks to improve service delivery through systematic review against specific criteria and standards and the implementation of change. SQUEAC uses clinical audit to evolve best practice. See *best practice*.

Clinical trial. A carefully controlled study conducted to test the effectiveness and safety of new drugs, medical products, protocols, or techniques

CMAM. See *Community-Based Management of Acute Malnutrition.*

Community-Based Management of Acute Malnutrition (CMAM). Refers to a program delivering therapeutic feeding to the majority of cases of severe wasting as outpatients. Effective CMAM programs include community outreach, mobilisation, and sensitisation activities to help ensure early detection, referral, and recruitment of cases and the follow-up of cases in the community. See *Integrated Community-Based Management of Acute Malnutrition*, *Community-Based Therapeutic Care*, *mobilisation*, *outpatient*, *outreach*, *sensitisation*.

Community-Based Therapeutic Care (CTC). Usually refers to NGO-run programs delivering therapeutic feeding to the majority of cases of severe wasting as outpatients in natural and civil emergencies. Effective CTC programs include community outreach, mobilisation, and sensitisation activities to help ensure early detection, referral, and recruitment of cases and the follow-up of cases in the community. See *Community-Based Management of Acute Malnutrition, Integrated Community-Based Management of Acute Malnutrition, mobilisation, outpatient, outreach, sensitisation*.

Community-based volunteer (CBV). A member of the community who assists with program activities (usually community mobilisation and case-finding and referral). CBVs typically do unpaid work on a 'little and often' basis. A CBV can receive an incentive (e.g., for attending training or per-referral), but no regular remuneration. Effective CTC and CMAM programs usually recruit and involve very many CBVs to assist with program activities. See *case-finding, community health worker, mobilisation*.

Community health worker (CHW). A member of a community who provides basic health and medical care to the community. CHWs may deliver CMAM services, such as community-based screening and referral, facility-based screening and diagnosis, and treatment. See *community-based volunteer.*

Community mobilisation. See *mobilisation*.

Community sensitisation. See *sensitisation*.

Compliance. The degree of constancy and accuracy with which a prescribed regimen (treatment protocol) is followed. The term is usually applied in the negative sense (i.e., non-compliance or poor compliance). The term may be applied to beneficiaries when, for example, RUTF is shared within the household, clinic visits are missed, or drugs are not given/taken. The term may be applied to a program or to program staff when, for example, the full CMAM protocol is not delivered.

Concept map. A diagram showing the relationships between findings. A graphical tool for organising and representing knowledge. Findings are represented as boxes or circles and connected using labeled arrows. The relationship between findings may be expressed using such phrases as 'gives rise to', 'contributes to', 'results in', and 'is required by'. See *mind-map*.

Confidence interval. In frequentist inference, an interval used to indicate the reliability (precision) of an estimate. See *credible interval, frequentist*.

Confidence limits. Upper and lower end-points of a confidence interval. See *confidence interval*.

Conjugate analysis. In Bayesian inference, a prior can be used that produces a posterior distribution of the same form as the prior distribution. Such a prior is called a 'conjugate prior'. When a conjugate prior is used, the prior to posterior Bayesian analysis is called a 'conjugate analysis'. SQUEAC uses beta-binomial conjugate analysis in which a beta-distributed prior is modified by a binomially distributed likelihood, resulting in a beta-distributed posterior. See *beta-binomial conjugate analysis, beta distribution, binomial distribution, likelihood, posterior, prior.*

Consumer. A synonym for beneficiary. A term used in LQAS. See *consumer probability of error, Lot Quality Assurance Sampling, provider, provider probability of error.*

Consumer probability of error (CPE). The risk that an investigation will conclude that coverage is high when it is (in fact) low. Also known as 'consumer risk'. See *Lot Quality Assurance Sampling, provider probability of error.*

Coverage. The proportion of all people needing or eligible to receive a service that actually receive that service. Also known as ‘treatment coverage’.

Coverage failure. An event or circumstance that results in people that need a service or are eligible for a service failing to receive that service. Examples of coverage failures are defaulters, DNA referrals, late admissions, and lack of proximity of services to the beneficiary population.

CPE. See *consumer probability of error.*

CPM. See *Critical Path Method, Program Evaluation and Review Technique*.

Credible interval. In Bayesian inference, a way of summarising the posterior distribution that gives an interval within which most (usually 95%) of the posterior distribution lies. A credible interval predicts that the true value of a parameter has a particular probability (usually 95%) of being in the credible interval given the observed data. The credible interval may be seen as the Bayesian equivalent of the frequentist confidence interval. See *Bayesian, confidence interval, frequentist, posterior.*

Credible limits. Upper and lower end-points of a credible interval. See *credible interval*.

Credible value. In Bayesian inference, a value for a parameter that is consistent with the available data about that parameter. See *Bayesian*.

Critical incident. An event in which there has been a significant or extreme occurrence (usually, but not necessarily, involving an undesirable outcome for the beneficiary) that is analysed in a systematic and detailed way to ascertain what can be learned about the overall quality of care and to indicate changes that might lead to future improvements. Also known as a ‘significant event’.

Critical Path Method (CPM). See *Program Evaluation and Review Technique*.

CSAS. See *centric systematic area sampling*.

CTC. See *Community-Based Therapeutic Care.*

Current case. A child meeting the program’s admission criteria. See *admission criteria*.

Cyclical process. A process that is characterised by moving in cycles or by happening at regular intervals. SQUEAC uses the cyclical process of clinical audit to evolve best practice. See *best practice, clinical audit.*

Defaulter. A beneficiary who was admitted to a program but who left the program without being formally discharged. Note that some beneficiaries may leave the program because they have moved away from the program area or have died. If they can be identified, such cases should be classified as having moved or died rather than as defaulters.

Denominator. The number or expression written below the line in a fraction (e.g., a coverage estimator). See *estimator, numerator.*

Did not attend (DNA). Used to indicate a case who was referred to a program but did not attend the program. A DNA case is a direct coverage failure. See *coverage failure, direct coverage failure*.

Direct coverage estimate. An estimate of coverage made by finding cases and ascertaining whether or not they are in a suitable treatment program. CSAS, S3M, SLEAC, and SQUEAC are direct methods. See *centric systematic area sampling*, *indirect coverage estimate*, *simple spatial survey method*.

Direct coverage failure. An event or circumstance that has a direct and immediate negative effect on coverage. Examples of direct coverage failures are DNA cases, defaulters, and late admissions. See *coverage failure, defaulter, did not attend, indirect coverage failure*.

Discharge criteria. Rules describing the circumstances in which beneficiaries may be discharged from a program. Discharge criteria vary between programs and will depend on whether beneficiaries are discharged to the community or to a less intense nutritional support program (e.g., an SFP). Discharge criteria will typically include loss of oedema, consistent weight gain, being clinically well, MUAC above a given threshold, or proportional weight gain above a given threshold. Discharge criteria may also include rules for transferring patients to more intensive nutritional support (e.g., inpatient therapeutic feeding) and for discharging cases as not responding to treatment. See *admission criteria*, *inpatient*, *supplementary feeding program*.

Distribution. The arrangement of the values of a variable. Usually represented using histograms and summary statistics (continuous variables, probability distributions) or using bar charts, Pareto charts, and tables (categorical variables). See *bar chart*, *histogram*, *Pareto chart*, *variable*.

DNA. See *did not attend.*

Equal-tailed credible interval. In Bayesian inference, a credible interval in which the probability that the parameter's value is below the lower end of the credible interval is equal to the probability that it is above the upper end of the credible interval. Also known as a 'central interval'. An equal-tailed credible interval is usually slightly wider than the equivalent HPD credible interval. The hand calculation methods given in SQUEAC documentation and the **`BayesSQUEAC`** software both produce equal-tailed 95% credible intervals. See *credible interval*, *highest posterior density (HPD) credible interval*.

Estimator. A function applied to a sample of a population used to estimate a parameter (e.g., the coverage proportion) of the whole population.

Evaluation. A management information process that measures how well a program's activities have met expected objectives and/or the extent to which changes in outcomes can be attributed to program activities. See *management information, monitoring*.

Food security. The availability of safe and nutritious food and access to it. A household is considered 'food secure' when its occupants do not live in hunger or in fear of hunger. A population is 'food secure' when all people at all times have physical, social, and economic access to sufficient, safe, and nutritious food that meets their dietary needs and food preferences for an active and healthy life.

Frequency distribution. See *distribution*.

Frequentist. The interpretation of probability that defines the probability of an event as the limit of the relative frequency of the event in a large number of trials. See *Bayesian*.

Fuzzy number. An extension of a regular number in the sense that it does not refer to one single value but rather to a connected set of possible values. The set is usually a range of possible values attached to a descriptive term. In SQUEAC, fuzzy numbers are used to link times or distances to the descriptive terms 'very near', 'near', 'not far', 'not near', 'far', and 'very far' when investigating distance as a barrier to accessing CMAM services. See *barrier.*

GAM. See *global acute malnutrition.*

Geographical coverage. The availability of CMAM through the decentralisation and scale-up of CMAM services. Geographical coverage can be defined as the ratio of primary healthcare facilities in a program area that deliver CMAM services to the total number of primary healthcare facilities in the program area. This indicator should **not** be confused with treatment coverage and can be biased by the use of different numerators and denominators. Geographic coverage is an indirect or proxy estimator of treatment coverage and should not be confused with spatial coverage. See *coverage, denominator, direct coverage estimate, indirect coverage estimate, numerator, spatial coverage, treatment coverage*.

Geographical information system (GIS). A system (usually computerised) designed to capture, store, manipulate, analyse, manage, and present geographically referenced data. GIS merges cartography (mapping), statistical analysis, and database management. SLEAC and SQUEAC were designed to be used without computerised GIS, but can make use of computerised GIS if available.

GIS. See *geographical information system.*

Global acute malnutrition (GAM). Usually defined as MUAC < 125 mm or bilateral pitting oedema. Some programs and surveys may also use a weight-for-height case definition. GAM is the sum of MAM and SAM. See *acute malnutrition, moderate acute malnutrition, severe acute malnutrition*.

Global Positioning System (GPS). A space-based global navigation satellite system that provides accurate and precise location information (latitude, longitude, altitude, and time) anywhere on the Earth. The system is maintained by the United States government and is freely accessible by anyone with a GPS receiver. Other satellite navigation systems are available, but not in common use.

GPS. See *global positioning system.*

Half-distance. An approximate method for finding distances that carers are willing or able to walk to access services, namely, half of the average distance between villages and towns with markets in the program area.

Headline coverage estimate. See *overall coverage estimate*.

Highest posterior density (HPD) credible interval. In Bayesian inference, the narrowest interval within which most (usually 95%) of the posterior distribution lies. An HPD credible interval is usually slightly narrower than the equivalent equal-tailed credible interval. See *credible interval, equal-tailed credible interval*.

Histogram. A graphical representation of the distribution of data in which tabulated frequencies are presented using adjacent rectangles with areas proportional to frequency in non-overlapping intervals. Histograms can be used to show an estimate of a probability distribution. Histograms are used in SQUEAC to describe and summarise prior belief about coverage. See *histogram prior, prior.*

Histogram prior. In Bayesian inference, a graphical tool used to describe and summarise prior belief.

Household census. A list of all individuals belonging to a single household.

Hypergeometric distribution. A probability distribution suited to the statistical modelling of proportions of a binary variable when data are sampled without replacement from a small population. SQUEAC does not use the hypergeometric distribution because of uncertainty over population sizes. See *binary variable*, *binomial distribution*, *sampling without replacement*, *variable*.

Hypothesis. A tentative theory about the world that is not yet verified but that if true would help explain certain facts or phenomena.

IDP. See *internally displaced person(s).*

IGD. See *informal group discussion*.

IMCI. See *Integrated Management of Childhood Illness.*

IMNCI. See *Integrated Management of Neonatal and Childhood Illness.*

Incidence. The number of new cases of a condition occurring in a population over a given time. See *prevalence*.

Indirect coverage estimate. An estimate of coverage made using data collected for other purposes ('secondary data') or proxy measures of coverage. For example, coverage of a CMAM program may be estimated by comparing program numbers with numbers predicted from the estimate of the prevalence of SAM found in a nutritional anthropometry survey multiplied by an estimate of the population in the program area and adjusted (using informed guesses) for incidence, spontaneous recovery, and death. Indirect estimates are usually inaccurate and imprecise. See *direct coverage estimate*, *geographical coverage*, *incidence*, *prevalence*.

Indirect coverage failure. An event or circumstance that has an indirect, delayed, and probable long- term negative effect on coverage. For example, late admission is an indirect coverage failure because it is associated with the need for inpatient care, longer treatment, defaulting, and poor treatment outcomes (e.g., death). These can all lead to the circulation of poor opinions of a program in the host population which can, in turn, lead to more late presentations and admissions and a cycle of negative feedback may develop. See *coverage failure*, *direct coverage failure*.

Informal group discussion (IGD). A data collection technique base on group discussion in which the discussion is informal and conversational and informants are encouraged to express themselves in their own terms rather than those dictated by the interviewer.

Informant. A person able to provide (useful) information.

Informative prior. In Bayesian inference, a prior that contains information about the value of a quantity. See *non-informative prior, prior.*

Inpatient. A patient who stays in a hospital while under treatment. See *outpatient*.

Integrated Community-Based Management of Acute Malnutrition. A CMAM program delivered by a local MOH or by an NGO in partnership with the local MOH as a component of an IMCI strategy. Note that 'integrated' usually indicates more than delivery of a vertical program at MOH facilities. See *Community-Based Management of Acute Malnutrition*, *Integrated Management of Childhood Illness*.

Integrated Management of Childhood Illness (IMCI). A systematic approach to children's health, including curative care and prevention of disease. The approach was developed by the World Health Organisation (WHO) and UNICEF. The management of SAM may be integrated into both facility-based or community-based IMCI. See *Community-Based Management of Acute Malnutrition, severe acute malnutrition*.

Integrated Management of Neonatal and Childhood Illness (IMNCI). May be used instead of IMCI. See *Integrated Management of Childhood Illness*.

Interface. Communication between different health facilities or programs for the purpose of transferring patients between facilities or programs.

Internally displaced person(s) (IDP). IDPs are people that are forced to flee their home but that remain within the borders of their home country. IDPs are often referred to as 'refugees', although they do not fall within the recognised legal definition of a refugee. See *refugee*.

Interview. A conversation between two people (the interviewer and the interviewee/informant/respondent) where questions are asked by the interviewer in order to obtain (useful) information from the interviewee/informant/respondent.

Iterative process. A method by which progress is made in a stepwise fashion with new depth and detail of information added and incorporated at each step. SQUEAC may be described as an iterative method.

Key informant. A person able to provide collective and important viewpoints and opinions. Key informants usually have a special role in the community (e.g., religious leaders, teachers, traditional healers, TBAs, village chiefs). See *lay informant*.

Kwashiorkor. A clinical term for a form of SAM. Bilateral pitting oedema is always present in kwashiorkor. Other clinical signs or symptoms of kwashiorkor include irritability; poor appetite; (pigmentary) dermatosis; depigmentation of hair; and hair that is sparse, loose, or unusually straight. See *bilateral pitting oedema*.

Late admission. An admission that is late in the course of a disease. In the management of SAM, for example, a MUAC < 105 mm or severe (+++) oedema is a late admission. Late admissions are both direct and indirect coverage failures. Very late admissions (e.g., MUAC < 95 mm) should be treated as critical incidents. See *coverage failure*, *critical incidents*, *direct coverage failure*, *indirect coverage failure*.

Lay informant. A person able to provide individual viewpoints and opinions. Lay informants usually have no special role in the community (e.g., carers of SAM cases). See *key informant*.

Leading question. A question in an interview that prompts a respondent to answer in a particular way. Leading questions are undesirable because they result in false or biased information. See *interview, open-ended question*.

Likelihood. In Bayesian inference, the information provided by new evidence. The likelihood is used to modify the prior to arrive at the posterior. In SQUEAC, this is the information provided by a survey (the likelihood survey). See *beta-binomial conjugate analysis*, *conjugate analysis*, *posterior, prior.*

Likelihood survey. In SQUEAC, a survey designed to provide evidence to modify prior belief about coverage using a beta-binomial conjugate analysis. See *beta-binomial conjugate analysis*, *conjugate analysis, posterior, prior.*

Line chart. A type of graph that displays information as a series of data points connected by lines. Line charts are often used to plot data over intervals of time (a time-series). SQUEAC uses line charts to plot time-series, such as admissions and exits over time. See *time-series*.

Lot Quality Assurance Sampling (LQAS). A classification and hypothesis testing technique used in both SLEAC and SQUEAC surveys. SLEAC and SQUEAC use a simplified approach to LQAS that uses simple rule-of-thumb formulas to create sampling plans. See *sampling plan*.

LQAS. See *Lot Quality Assurance Sampling.*

MAM. See *moderate acute malnutrition.*

Management information. Information that is needed and used to manage a program efficiently and effectively.

Maximum probable value. The maximum value that the coverage proportion is likely to be (given all available evidence).

Mean. The arithmetic mean or 'standard' average (i.e., the sum of all values divided by the number of values summed). See *median*.

Median. The value separating the upper half of a sample (or population or distribution) from the lower half. The median of a list of numbers is found by arranging the numbers from lowest value to highest value and picking the middle one. If there is an even number of observations, then the median is the mean of the two middle values. See *mean*.

Migration. The movement of persons from one country or locality to another.

Mid-upper arm circumference (MUAC). The circumference of the upper arm measured at the mid-point between the tip of the shoulder and the tip of the elbow. MUAC is the best available and practical indicator of mortality risk associated with acute malnutrition. See *acute malnutrition, admission criteria, anthropometric criteria, anthropometry, case-finding, discharge criteria, late admission*.

Mind-map. A graphical way of storing and organising data and ideas. A mind-map organises findings using tree structures organised around a central theme. See *branching hierarchy, concept map*.

Minimum probable value. The minimum value that the coverage proportion is likely to be (given all available evidence).

Mobilisation. Activities designed to foster the participation of the host population in key program activities, such as sensitisation, case-finding/referral, and the follow-up of cases. See *case-finding, sensitisation*.

Mode. The mode is the value that occurs most frequently in a data set or a probability distribution. In SQUEAC, the mode is used to summarise belief about coverage in both the prior and the posterior. See *posterior, prior.*

Moderate acute malnutrition (MAM). MAM or moderate wasting is defined as MUAC between 115 mm and 125 mm **without** bilateral pitting oedema in children between 6 and 59 months old. Some programs and survey reports may also use a weight-for-height case definition. See *acute malnutrition, bilateral pitting oedema, global acute malnutrition, severe acute malnutrition*.

Monitoring. A management information process that focusses on implementation and the progress made toward the achievement of program objectives. See *evaluation*, *management information*.

Morbidity. A diseased state, disability, or poor health due to any cause.

Mortality. Death.

Moving average. In statistics, a method of smoothing a set of data points (usually a time-series) by creating a series of averages of ordered subsets of the full data set. A moving average is obtained by taking the average of the first subset. The fixed subset size (the 'span') is then shifted forward by a single unit in time, creating a new subset of data to be averaged. This process is repeated over the entire dataset. A moving average smooths away short-term fluctuations to highlight longer-term trends or cycles in the data. The threshold between short term and long term depends on the application, and the span of the moving average is set accordingly. Different types of average (e.g., mean, median) may be used. Moving averages can be applied several times (i.e., the method can be applied to previously smoothed data). SQUEAC uses moving averages (medians with a span of 3 months followed by means with a span of 3 months) to reveal seasonality and trend in time-series of admissions and exits. SQUEAC investigations may also use moving averages (medians with a span of 13 months followed by means with a span of 13 months) to reveal longer-term trends in, for example, time-series of admissions and exits. See *smoothing*, *time-series*.

MUAC. See *mid-upper arm circumference.*

NGO. See *non-governmental organisation.*

Nomogram. A graphical device designed to allow the approximate (graphical) computation of a formula or function. A nomogram is the graphical equivalent of a look-up table. Nomograms can be used to calculate the threshold value (d) for a given sample size (n) and standard (p) for use in a simplified LQAS sampling plan. See *Lot Quality Assurance Sampling*, *sampling plan*.

Non-governmental organisation (NGO). Used to refer to an organisation that does not form part of the government and is not a conventional for-profit business.

Non-informative prior. In Bayesian inference, a prior that contains no information, reflecting a state of total ignorance about the value of a quantity. In SQUEAC, a non-informative prior is defined as *Beta(1, 1)*. See *informative prior, prior.*

Normal approximation. The practice of using the normal distribution as an approximation to, for example, the binomial distribution to simplify calculations. SQUEAC uses this approach for the hand calculation of 95% credible intervals.

Normal distribution. A theoretical frequency distribution for a set of data usually represented by a bell-shaped curve that is symmetrical about the mean and is defined by the mean and standard deviation. See *mean*, *standard deviation*, *standard normal distribution*.

Numerator. The number or expression written above the line in a fraction (e.g., a coverage estimator). See *denominator, estimator.*

Observational study. In SQUEAC, a study that relies on observing behaviours or processes (also called 'naturalistic observation'). Note that this is a different meaning from the epidemiological term, namely, a study that draws inferences about the effect of a treatment on subjects where the assignment of subjects into treatment and control groups is outside the control of the investigator (also known as a 'natural experiment').

Observer effect. In SQUEAC, this is the short-term boost in coverage caused by a SQUEAC investigation that is independent of program reform. It is due to, among other things, case-finding and referral, defaulter tracing, DNA tracing, and the mobilising effect of observation. SQUEAC investigations that are repeated too frequently are likely to observe these short-term improvements in program coverage and spatial reach and, mistakenly, attribute such improvements to the remedial actions implemented as a result of the assessment.

Open-ended question. An investigative tool designed to encourage a full and meaningful answer using the subject's own knowledge. It is the opposite of a closed-ended question, which encourages short (usually single-word) answers. Open-ended questions (or discussions) tend to be less 'leading' than closed-ended questions. See *leading question*.

OTP. See *outpatient therapeutic program.*

Outpatient. A patient who receives medical treatment without being admitted to a hospital. See *inpatient*.

Outpatient therapeutic program (OTP). A program treating cases of uncomplicated SAM as outpatients. OTP is the central CTC/CMAM program service. See *outpatient*.

Outreach. Program activities that connect a program to the efforts of other organisations, groups, specific audiences, or the general public. Outreach implies active engagement and mobilisation rather than just dissemination of program messages. Effective CTC and CMAM programs usually have specific outreach activities for community mobilisation, case-finding, and referral.

Outreach worker. A member of program staff undertaking outreach work. See *outreach*.

Overall coverage estimate. A summary coverage estimate for an entire program area. Only useful when coverage is similar across the entire program area.

Pareto chart. A type of bar chart in which the bars are organised in order of ascending or descending size. Pareto charts may also use a line to represent cumulative totals. SLEAC and SQUEAC use Pareto charts to analyse and present findings, such as the relative importance of barriers to service access and uptake. See *bar chart*.

Participatory rural appraisal (PRA). See *rapid rural appraisal*.

Patchy. Adjective describing spatially uneven coverage (also called 'spatial heterogeneity').

Peri-urban. Relating to or characteristic of areas immediately adjoining an urban area (i.e., between the suburbs and the countryside). See *rural, urban*.

Period coverage. Coverage estimated using both current and recovering cases. The rationale for using recovering cases is that they are children that should be in the program because they have not yet met program discharge criteria. See *coverage*, *point coverage*.

PERT. See *Program Evaluation and Review Technique.*

Pie chart. A circular chart divided into sectors intended to illustrate (relative) proportion. The arc length, central angle, and area of each sector is proportional in size to the quantity it represents. SQUEAC does **not** use pie charts to analyse or present findings because many people find it difficult to compare the sizes of items in a chart when area is used instead of length and when different items have different shapes. SQUEAC uses bar charts and Pareto charts to investigate and illustrate proportion. See *bar chart, Pareto chart*.

Planned discharges. Cases discharged as cured or as not responding to treatment (i.e., all discharges excluding transfers, deaths, and defaulters). This group of beneficiaries are used to investigate the duration of treatment episodes.

Plot. A graphical device used for analysing and presenting data. SQUEAC uses a broad variety of plots to present and analyse data.

Point coverage. Coverage estimated using current cases only. See *coverage, period coverage*.

Posterior. In Bayesian inference, the posterior is the result of modifying prior belief using new evidence. See *beta-binomial conjugate analysis*, *beta distribution*, *conjugate analysis, likelihood*, *prior.*

Powers of 2. The result of exponentiation with the number 2 as the base and any non-negative whole number (including 0) as the exponent. SQUEAC uses powers of 2 to generate random numbers from coin tosses.

PPE. See *provider probability of error.*

PRA. See *rapid rural appraisal*.

Precision. The degree to which repeated measurements under unchanged conditions show the same results. See *accuracy.*

Prevalence. The proportion of a population with a given condition at a given time. See *incidence*.

Prior. In Bayesian inference, the prior is a probabilistic representation of available knowledge about a quantity. In SQUEAC, the prior is a probabilistic representation of knowledge relating to program coverage. SQUEAC uses a beta-distributed prior. See *beta-binomial conjugate analysis, beta distribution*, *conjugate analysis*, *likelihood*, *posterior.*

Probability density. A function that describes the relative likelihood for a particular value of a variable. See *variable*.

Probability distribution. A function that describes the relative likelihood for a particular value of a variable. See *variable*.

Producer. See *provider.*

Program Evaluation and Review Technique (PERT). A statistical tool used in project management that is designed to analyse and represent the tasks involved in completing a given project. It is commonly used in conjunction with the Critical Path Method (CPM), which is a systematic approach to scheduling a set of project activities. SQUEAC uses the three-point estimation technique (used in PERT and CPM for task-duration modelling) to find appropriate shape parameters for the prior. See *prior, task duration modelling*, *three-point estimation approach*.

Proof-of-cure. A period of time (usually 2 weeks) during which a beneficiary is retained in a TFP after having met the program's discharge criteria that is intended to ensure that the beneficiary has been cured and is unlikely to relapse. The beneficiary is discharged at the 'proof-of-cure visit'. See *relapse*.

Protocol. A plan for a course of medical treatment.

Provider. A provider of CTC/CMAM services.

Provider probability of error (PPE). The risk that an investigation will conclude that coverage is low when it is in fact high. Also known as 'provider risk'. See *consumer probability of error*, *Lot Quality Assurance Sampling*.

Proximity. Nearness (usually in space).

Quadrat. A square area defined for sampling purposes. SQUEAC uses quadrats to locate sampling points in some survey activities. See *centric systematic area sampling*.

Qualitative research. A method of inquiry employed in the social sciences. Qualitative research aims to provide an in-depth understanding of human behaviours and the reasons that govern such behaviours. SQUEAC uses both qualitative and quantitative research methods. See *quantitative research*.

Quantitative research. The systematic and empirical investigation of social phenomena using statistical, mathematical, or computational techniques. SQUEAC uses both qualitative and quantitative research methods. See *qualitative research*.

Range. The extent to which (or the limits between which) variation is possible. The set of all possible values of a variable. The lowest and highest values of a variable. The difference between the lowest and highest values of a variable. See *variable*.

Rapid rural appraisal (RRA). Also known as 'participatory rural appraisal' (PRA). An approach used by NGOs and other agencies involved in community development. The RRA approach aims to incorporate the knowledge and opinions of rural people in the planning and management of development projects and programs. SQUEAC uses some elements of the RRA approach.

Ready-to-use therapeutic food (RUTF). A prepared and packaged nutrient- and energy-dense therapeutic food designed to require no preparation by the end-user and to be shelf-stable for extended periods. RUTF has a similar nutritional profile to Formula-100 therapeutic milk. A common RUTF is a spread made of ground peanuts with milk powder, sugar, oil, minerals, and vitamins.

Recovering case. A child who recently met a program's admission criteria (i.e., was recently a SAM case but is no longer a SAM case), but does not yet meet the program's discharge criteria. Usually applied to children in the program. See *admission criteria*, *discharge criteria*, *severe acute malnutrition*.

Refugee. Someone who has been forced to flee his or her home country. Refugee is a legal definition under Article (1)(A) of the United Nations Convention Relating to the Status of Refugees (1951) and the Convention's 1967 Protocol. See *internally displaced person(s)*.

Relapse. To suffer deterioration after a period of improvement (of someone suffering from a disease). This term is also used, particularly when reporting routine program data, to mean a new episode of SAM in a patient who was previously discharged as cured. See *routine program data*.

Retention. The ability of a program to keep beneficiaries in the program until they are formally discharged. Retention is achieved by minimising defaulting. See *defaulter.*

Rounding rule. A rule regarding the expression of a rational or real number as a whole number. Rounding rules include 'always round down' ('floor'), 'always round up' ('ceiling'), and 'round to nearest integer' ('round'). SQUEAC and SLEAC use rounding rules to decide LQAS sampling plans and when calculating sampling intervals when using systematic sampling to select communities to be sampled. See *Lot Quality Assurance Sampling*, *systematic sampling*.

Routine program data. Data that should be collected and reported by all TFPs, including (but not limited to) admissions, proportion of exits cured, proportion of exits defaulting or lost to follow-up, proportion of exits died, proportion of exits transferred to another facility or program, and proportion of exits discharged as not responding to treatment.

RRA. See *rapid rural appraisal*.

Rule-of-thumb. A simple method of wide applicability not intended to be strictly accurate or reliable in every possible situation. SQUEAC and SLEAC use rule-of-thumb formulas to define LQAS sampling plans. See *Lot Quality Assurance Sampling*.

Rural. Relating to or characteristic of the countryside rather than towns or cities. See *peri-urban*, *urban*.

RUTF. See *ready-to-use therapeutic food*.

S3M. See *simple spatial survey method.*

SAM. See *severe acute malnutrition*.

Sample. A subset of a population. Samples are collected in surveys and studies, and statistics are calculated from the samples to make inferences about the population from which the samples are collected. See *survey.*

Sample size. The number of observations in a sample.

Sampling. The process of collecting a sample. See *sample*.

Sampling frame. The source from which a sample is drawn. See *sample*.

Sampling plan. In LQAS, a set of rules that may include minimum and maximum sample sizes and decision rules that when applied to survey data are used to make classifications (or test hypotheses) about the level of an indicator. See *Lot Quality Assurance Sampling*.

Sampling to redundancy. A social science technique in which data are collected until no new information comes to light. This technique is often combined with triangulation. SQUEAC makes extensive use of both triangulation and sampling to redundancy. See *triangulation*.

Sampling with replacement. A sampling method that allows members of a population to be chosen more than once. This is **not** the usual survey process. See *sampling*.

Sampling without replacement. A sampling method that deliberately avoids choosing any member of a population more than once. This is the usual survey process. See *sampling*.

Satellite imagery. Images of the Earth made by means of sensors (e.g., cameras, radar) carried by artificial satellites. SLEAC and SQUEAC may use satellite imagery when useful maps are not available.

Scaling. In SQUEAC, a method of adjusting weights or scores associated with individual findings so that the mode of the prior is constrained to lie between 0% and 100%. See *prior, weight*, *weighting*.

Screening. A strategy used in a population to detect disease in individuals usually without overt symptoms of that disease. The intention of screening is to identify cases early to provide early intervention and reduce morbidity and mortality. See *case-finding*, *morbidity*, *mortality*, *two-stage screening test model*.

Seasonal calendar. An ordered collection of events and activities usually related to changes in the weather and how they tend to affect health, food availability, food prices, terms of trade, farming, labour demand, migration, etc.

Secondary data. Data collected by someone other than the current user or for purposes other than the current purpose.

Selective feeding program. A feeding program that admits individuals based on anthropometric, clinical, or social criteria. Programs such as CMAM (for SAM) and targeted SFPs (for MAM) are selective feeding programs. General food distributions and blanket SFPs are **not** selective feeding programs.

Self-referral. A patient who arrives at a health facility without being referred by program outreach staff, CHWs, CBVs, other clinical staff, or other program staff.

Semi-structured interview. An interview technique using a set of clear instructions comprising a list of questions that should be asked and topics that should be covered (the ‘interview guide’). The exact order and wording of questions may differ from informant to informant and change as data collection proceeds. See *interview.*

Sensitisation. Activities that promote understanding of program objectives and methods. Sensitisation activities include holding information sessions with community leaders and training sessions with CHWs and CBVs, announcing schedules of program activities (e.g., clinic days) to local people, and describing the target population based on local understanding of acute malnutrition and using local terminology to describe it. See *mobilisation*.

Sensitivity. The ability of a (screening) test to identify correctly those that have the disease being screened for. See *screening*, *specificity.*

Service delivery unit. The administrative unit or health facility responsible for delivering CMAM services. In the case of a national or regional program delivering CMAM services through health districts the service delivery unit is the health district. In the case of a district program delivering CMAM services through primary healthcare centres the service delivery unit is the primary healthcare centre.

Service provider. See *provider.*

Severe acute malnutrition (SAM). Usually defined as MUAC < 115 mm and/or bilateral pitting oedema in children between 6 and 59 months old. Some programs and survey reports may also use a weight-for-height case definition. See *acute malnutrition*, *bilateral pitting oedema*, *global acute malnutrition*, *moderate acute malnutrition*.

SFP. See *supplementary feeding program*.

Shape parameter. A parameter of a probability distribution that affects the shape of the distribution rather than simply shifting it (as a location parameter does) or stretching/shrinking it (as a scale parameter does). See *beta distribution*, *probability distribution*.

Simple spatial survey method (S3M). S3M is a development of the CSAS method that makes more effective use of survey data. S3M is a general survey method and can be used for purposes other than coverage assessment. See *centric systematic area sampling*.

Simple structured interview. An interview technique that exposes every informant to the same stimulus by asking the same questions in the same order.

Simulation. The imitation of a state of a system or a process. In SQUEAC, suitable sample sizes for likelihood surveys may be found by simulating surveys with different sample sizes using the **`BayesSQUEAC`** software.

Small area. An area smaller (usually much smaller) than the entire program area. See *small-area survey.*

Small-area survey. A survey investigating coverage in a small area. See *small area*.

Small study. A short, usually semi-quantitative piece of work focussing on testing a single hypothesis. The hypothesis being tested usually relates to processes that affect coverage rather than to coverage directly. Sampling and study design are directed by the hypothesis being tested. See *sampling*.

Small survey. A small-sample survey undertaken in population groups that are hypothesised to have high or low coverage.

SMART. See *Standardised Monitoring and Assessment of Relief and Transitions.*

Smoothing. In statistics, to create an approximating function that removes noise and fine-scale or rapidly changing phenomena from data (usually time-series or image data) in order to reveal patterns (trends) in the data. SQUEAC uses the ‘moving average’ algorithm to smooth data from time-series, such as admissions and exits over time. See *moving average*, *time-series*.

Spatial. Methods or findings regarding the relationship between coverage and location. SQUEAC uses maps (e.g., of outreach activities and the home locations of beneficiaries, defaulters, and DNA referrals), tables (e.g., of outreach activities and distance/time-to-travel), and plots (e.g., of time-to-travel) to analyse and present findings about coverage and location. SQUEAC also uses spatial sampling methods (e.g., CSAS, stratified spatial sampling) in wide-area surveys. See *centric systematic area sampling*, *defaulter*, *did not attend*, *stratified spatial sampling*.

Spatial coverage. The pattern of treatment coverage measured using a direct coverage method over the entire program area. Spatial coverage should not be confused with geographical coverage. See *coverage*, *denominator*, *direct coverage estimate*, *geographical coverage*, *indirect coverage estimate*, *numerator*, *treatment coverage*.

Spatial distribution. The pattern of an indicator over an entire program area. May also refer to the distribution of clinic sites over a program area.

Specificity. The ability of a (screening) test to correctly identify those that do not have the disease being screened for. See *screening*, *sensitivity.*

Sphere. A project launched by a group of humanitarian NGOs and the Red Cross and Red Crescent movements. The Sphere Project is an initiative to define and uphold the standards by which the global community responds to the plight of people affected by disasters, principally through a set of guidelines that are set out in the *Humanitarian Charter and Minimum Standards in Disaster Response* (commonly referred to as the 'Sphere Handbook'). SQUEAC assessments may use the Sphere minimum standards for TFPs for coverage, cure, and defaulting.

Standard. The level of an indicator, for example, that defines satisfactory performance. Standards may be set as minimum acceptable performance levels (e.g., as in Sphere) or as interim performance targets on the way to achieving best practice (e.g., as in clinical audit). See *best practice*, *clinical audit*, *Sphere*.

Standard deviation. A quantity calculated to indicate the extent of variation or 'dispersion' there is from the average (mean) of a variable. A low standard deviation indicates that the data points tend to be very close to the mean. A high standard deviation indicates that the data points are spread out over a large range of values. See *variable*.

Standard normal distribution. A normal distribution with mean of 0 and standard deviation of 1. All normal distributions are equivalent to this distribution when the unit of measurement is changed to measure standard deviations from the mean. This allows the standard normal distribution to be used to model any problem involving any normal distribution. See *normal distribution*, *z-score*.

Standard program indicator graph. A line chart showing the pattern of program exits over time, usually broken down into discharged as cured, discharged as not responding to treatment, died, defaulted, and transferred. Data are usually presented as proportions of the total number of exits at each time point.

Standardised Monitoring and Assessment of Relief and Transitions (SMART). A survey method for nutritional anthropometry, mortality, and household economy surveys. In SQUEAC documentation, SMART refers to the nutritional anthropometry surveys.

Stratification. The process of dividing members of a population into subgroups (strata) before sampling. SLEAC and SQUEAC frequently use spatial stratification. See *centric systematic area sampling*, *stratified spatial sampling*, *systematic area sampling*, *systematic sampling*.

Stratified spatial sampling. A systematic area sample. In SQUEAC, this refers to taking a systematic sample from lists of communities sorted by one or more areal (spatial) variables (e.g., district, chiefdom within district, village within chiefdom). See *centric systematic area sampling*, *systematic area sampling, systematic sampling, variable*.

Supplementary feeding program (SFP). A program intending to treat MAM or prevent MAM or SAM. See *moderate acute malnutrition, severe acute malnutrition.*

Survey. A research tool that uses a sample of individuals from a population to make (statistical) inferences about the population from which the sample is collected. See *sample*.

Swing point. In LQAS, a term applied to a threshold that defines different qualitative levels (e.g., high or low) of an indicator. See *Lot Quality Assurance Sampling.*

Symmetrical prior. A prior that is symmetrical about its mode. See *prior.*

Systematic area sampling. A sampling method that samples areal (spatial) units spread relatively evenly over the wider survey area. SQUEAC uses CSAS and systematic sampling from lists of communities sorted by area to take samples that are relatively evenly spread over wider survey areas. See *centric systematic area sampling, stratified spatial sampling*.

Systematic sampling. A sampling method involving the non-random selection of elements from an ordered sampling frame. See *sampling frame*.

Taboo. A strong social prohibition relating to any area of human activity or a behaviour that is forbidden based on moral judgement or religious belief.

Tabular analysis. A method of organising, analysing, and presenting data using tables.

Tabulation. The process of organising data using tables.

Tally plot. An integrated data collection, analysis, and presentation device. A tally sheet that draws a histogram as data are collected. SQUEAC uses tally plots for admission MUAC, time-to-travel, clinic visits made by defaulters, and the durations of treatment episodes. See *histogram*, *tally sheet*.

Tally sheet. An integrated data collection, analysis, and presentation device. SQUEAC uses tally sheets for data from surveys, such as counts of cases and counts of barriers. See *tally plot*.

Task-duration modelling. A method used in management and information systems for constructing an approximate probability distribution representing the duration of individual project activities. SQUEAC uses the three-point estimation technique (used in PERT and CPM for task-duration modelling) to find appropriate shape parameters for the prior. See *prior, Program Evaluation and Review Technique*, *three-point estimation approach*.

TBA. See *traditional birth attendant*.

Temporal. Relating to time.

Temporal coverage. The pattern of coverage over time.

TFC. See *therapeutic feeding centre*.

TFP. See *therapeutic feeding program*.

Therapeutic feeding centre (TFC). A facility treating all SAM cases as inpatients. CTC/CMAM programs usually treat cases of complicated SAM for short periods in inpatient facilities known as ‘stabilisation centres’. See *inpatient*.

Therapeutic feeding program (TFP). A program treating SAM. See *Community-Based Management of Acute Malnutrition*, *Community-Based Therapeutic Care*, *therapeutic feeding centre*.

Three-point estimation approach. In program management, an approach to task-duration modelling in which the duration of individual project activities is modelled using three parameters (best case = shortest time, worst case = longest time, and most likely case = mode) based on prior experience and informed guesses. SQUEAC borrows elements of this approach to find appropriate shape parameters for the prior. See *prior*, *Program Evaluation and Review Technique*, *task-duration modelling*.

Threshold value. A component of an LQAS sampling plan used to make a classification. See *Lot Quality Assurance Sampling, sampling plan*.

Time-series. A sequence of data points measured at successive times, usually at uniform time intervals. Time-series are frequently smoothed and plotted using line charts. SQUEAC investigations typically collect and analyse a number of time-series, such as admissions and exits over time. See *line chart*, *moving average*, *smoothing*.

Time-to-travel. A proxy for distance. In SQUEAC, methods and data used to investigate the relationship between location (or distance from CMAM sites) and coverage. See *spatial*, *spatial coverage*.

Traditional birth attendant (TBA). Also known as 'traditional midwife', 'community midwife', or 'lay midwife'. TBAs provide basic health care, support, and advice during and after pregnancy and childbirth, usually based on experience and knowledge acquired informally through the traditions and practices of the community in which they live. In SQUEAC, TBAs are an important type of key informant.

Traditional healer. A practitioner of traditional, indigenous, or folk medicine. Traditional healers are recognised by the community in which they live as competent to provide health care, using vegetable, animal, and mineral substances, as well as other methods derived from the knowledge, attitudes, and beliefs prevalent in a community. In some settings, traditional healers may also be religious leaders. In SQUEAC, traditional healers are an important type of key informant. Also known as 'indigenous healer' and 'traditional health practitioner' (abbreviated to THP).

Treatment coverage. See *coverage*.

Triangulation. A social science technique in which different methods and sources are used in an investigation to confirm findings. The rationale for triangulation is that the use of multiple methods and sources overcomes the weaknesses, intrinsic biases, and problems associated with using individual methods and sources. SQUEAC makes extensive use of triangulation. See *sampling to redundancy.*

Two-stage cluster sample. A two-stage sampling method that typically selects communities to sample using population proportional sampling (PPS) in the first stage and households to sample by proximity to a randomly selected household in the second stage. This type of sample is commonly used in nutritional anthropometry (e.g., SMART) surveys. SLEAC and SQUEAC do not use this type of sampling. See *Standardised Monitoring and Assessment of Relief and Transition*.

Two-stage screening test model. A method of screening that uses two tests. The first-stage test is typically of low cost with high sensitivity. The second-stage test is typically of higher cost and applied only to persons tested positive by the first-stage test; it also has high specificity. A combination of tests used in this way provides a low-cost screening method with low levels of error. SQUEAC uses a similar approach when, for example, using small-area surveys to identify areas of high and low coverage. The methods and data used to identify the areas to be surveyed can be seen as the first-stage test and the small-area survey as the second-stage test. The use of this model provides acceptably low levels of error with small sample sizes. See *screening*, *sensitivity*, *specificity.*

U.N. See *United Nations.*

United Nations (U.N.). An international organisation whose stated aims are facilitating cooperation in international law, international security, economic development, social progress, human rights, and achievement of world peace. The U.N. was founded in 1945 after World War II to replace the League of Nations, to stop wars between countries, and to provide a platform for dialogue. It contains multiple subsidiary organisations to carry out its missions.

Urban. Relating to or characteristic of towns or cities. See *peri-urban*, *rural*.

Validation. A process that ensures the soundness of findings. SQUEAC uses triangulation by source and method and sampling to redundancy to validate findings. See *sampling to redundancy*, *triangulation*.

Variable. A quantity or function that can assume any given value or set of values.

Visible severe wasting. A sign of SAM. Loss of muscles mass on the arms, thighs, and buttocks and/or sagging skin and buttocks ('baggy pants') evident from visible inspection of a child. See *acute malnutrition*.

Wasting. A form of acute malnutrition. It is defined by a MUAC < 125 mm (or a weight-for-height z-score < −2). See *acute malnutrition*.

Weight. In SQUEAC, the emphasis given to individual findings when deciding the prior. Also known as 'score'. In SLEAC, the emphasis given to the results of individual small-area surveys when they are combined to produce a wide-area estimate. See *prior.*

Weighting. The process of emphasising the contribution of some aspects of a set of data to a final result by giving them more 'weight' in the analysis. Some findings contribute more to the final result, rather than each finding contributing equally to the final result. SQUEAC uses weighting to help decide the prior. SLEAC uses weighting when estimating coverage over wide areas. See *prior.*

Wide area. An entire program area (usually a health district or larger). See *wide-area survey.*

Wide-area survey. A survey investigating coverage over an entire program area (usually a health district or larger). See *wide area*.

x axis. The horizontal axis of a chart. See *y axis*.

y axis. The vertical axis of a chart. See *x axis*.

z-score. In statistics, a z-score (or standard score) indicates how many standard deviations an observation is above or below the sample (or reference) mean.