

# A field test of three LQAS designs to assess the prevalence of acute malnutrition

Megan Deitchler,<sup>1\*</sup> Joseph J Valadez,<sup>2</sup> Kari Egge,<sup>3</sup> Soledad Fernandez<sup>4</sup> and Mary Hennigan<sup>5</sup>

**Accepted** 3 April 2007

**Background** The conventional method for assessing the prevalence of Global Acute Malnutrition (GAM) in emergency settings is the 30 × 30 cluster-survey. This study describes alternative approaches: three Lot Quality Assurance Sampling (LQAS) designs to assess GAM. The LQAS designs were field-tested and their results compared with those from a 30 × 30 cluster-survey.

**Methods** Computer simulations confirmed that small clusters instead of a simple random sample could be used for LQAS assessments of GAM. Three LQAS designs were developed (33 × 6, 67 × 3, Sequential design) to assess GAM thresholds of 10, 15 and 20%. The designs were field-tested simultaneously with a 30 × 30 cluster-survey in Siraro, Ethiopia during June 2003. Using a nested study design, anthropometric, morbidity and vaccination data were collected on all children 6–59 months in sampled households. Hypothesis tests about GAM thresholds were conducted for each LQAS design. Point estimates were obtained for the 30 × 30 cluster-survey and the 33 × 6 and 67 × 3 LQAS designs.

**Results** Hypothesis tests showed GAM as <10% for the 33 × 6 design and GAM as ≥10% for the 67 × 3 and Sequential designs. Point estimates for the 33 × 6 and 67 × 3 designs were similar to those of the 30 × 30 cluster-survey for GAM (6.7%, CI = 3.2–10.2%; 8.2%, CI = 4.3–12.1%, 7.4%, CI = 4.8–9.9%) and all other indicators. The CIs for the LQAS designs were only slightly wider than the CIs for the 30 × 30 cluster-survey; yet the LQAS designs required substantially less time to administer.

**Conclusions** The LQAS designs provide statistically appropriate alternatives to the more time-consuming 30 × 30 cluster-survey. However, additional field-testing is needed using independent samples rather than a nested study design.

**Keywords** Acute malnutrition, assessment, emergency, Ethiopia, lot quality assurance sampling, LQAS, wasting

## Introduction

In emergency settings, humanitarian workers must rapidly detect areas where a population's health and survival are

at risk. Due to the close association between acute malnutrition prevalence and crude mortality rates,<sup>1</sup> the prevalence of acute malnutrition among children 6–59 months can be used as an indicator of elevated risk. Governments, donors and relief agencies often use a predetermined threshold of acute malnutrition to judge whether a response is justified and to guide the level of aid provided. WHO's classification scale defines the severity of malnutrition in a community based on the prevalence of wasting [Weight-for-Height-Z-(WHZ)-score <−2 Standard Deviations (SDs)]: Acceptable: <5%; Poor: 5–9%; Serious: 10–14%; Critical: ≥15%. These thresholds can be used to guide the implementation of selective and general feeding programmes.<sup>2,3</sup> Global Acute Malnutrition (GAM) is defined as WHZ-score <−2 SDs and/or bipedal oedema.<sup>4–6</sup>

<sup>1</sup> Academy for Educational Development, FANTA Project, Washington DC, USA.

<sup>2</sup> The World Bank, Washington DC, USA.

<sup>3</sup> Independent Consultant, Formerly with Catholic Relief Services, Emergency Response Team, Nairobi, Kenya.

<sup>4</sup> Ohio State University, Center for Biostatistics, Columbus, Ohio, USA.

<sup>5</sup> Catholic Relief Services, Baltimore, Maryland, USA.

\* Corresponding author. Academy for Educational Development, Food and Nutrition Technical Assistance (FANTA) Project, 1825 Connecticut Avenue NW, Washington DC, 20009. E-mail: mdeitchl@aed.org

In emergency settings, governments and international agencies often apply the WHO thresholds to GAM in lieu of wasting. In these situations, the primary concern is to detect areas with a GAM prevalence 10% or above in order to determine where food aid is needed.

In emergency settings, the prevalence of GAM needs to be assessed quickly, accurately, and often times repeatedly, to determine when and where to start and stop humanitarian aid. The most common approach for assessing GAM is a two-stage 30 × 30 cluster-survey<sup>7</sup>—which can be time-consuming and expensive. Alternative, less time- and resource-intensive sampling approaches are needed to rapidly identify areas requiring humanitarian assistance, thereby saving lives and resources. One of the most frequently used quality control statistical methods in international health is Lot Quality Assurance Sampling (LQAS).<sup>8</sup>

This article examines three adaptations of LQAS to assess GAM and other child-level indicators in food-insecure settings. Earlier sample-size and precision limitations of LQAS for GAM assessment are addressed.<sup>9</sup> A field-test of the LQAS designs was conducted in the Siraro woreda of Ethiopia in June 2003 during a period of heightened food insecurity, at a time when the government used GAM thresholds to help prioritize woredas for food aid. The LQAS designs were administered simultaneously alongside a conventional 30 × 30 cluster-survey, using a nested sampling approach. This article describes the LQAS designs and compares the results of those designs with the 30 × 30 cluster-survey—in terms of point estimates and 95% CIs for GAM and other child-level indicators, classification results for GAM thresholds of 10, 15 and 20%, and the amount of time required for data collection.

## Methods

### LQAS principles

Cumulative binomial probabilities are used in LQAS analyses to detect if a critical threshold has been reached for an indicator. To design an LQA sampling plan, the threshold of interest for an indicator (e.g. GAM prevalence), and tolerable statistical error (alpha and beta) are defined a priori. LQAS is an hypothesis test to determine whether the outcome is  $\geq$  or  $<$  the defined threshold. The alpha error is the probability of incorrectly classifying an area as not being at risk when the true GAM prevalence is  $\geq$  the threshold of interest. The beta error is the probability of incorrectly classifying an area as being at risk when the true GAM prevalence is  $<$  the threshold of interest.

LQAS uses two thresholds, an upper and a lower threshold, to define the alpha and beta errors. The upper threshold is a GAM prevalence at which an area is at risk (e.g.  $\geq 15\%$ ). This is the threshold the data are tested against. The lower threshold is the GAM prevalence at which an area would not be considered a priority for an intervention (e.g.  $\leq 10\%$ ). The alpha error is calculated for the upper threshold and the beta error for the lower threshold.

The null hypothesis assumes the GAM prevalence is  $\geq$  the upper threshold. To classify GAM prevalence as  $\geq$  or  $<$  the upper threshold, the number of children with GAM is counted

and then compared against a Decision Rule (DR) determined using binomial probabilities.<sup>8,10</sup> We judge GAM prevalence as  $\geq$  the upper threshold if the number of children with GAM in the sample is  $>$  the DR and judge GAM prevalence as  $<$  the upper threshold if the number of children with GAM is  $\leq$  the DR.

To illustrate how LQAS is applied in international health settings, assume a design with a sample size of 198, a 15% upper threshold, 10% lower threshold, alpha error  $\leq 0.10$  and beta error  $\leq 0.20$ . If more than 23 (DR) children are malnourished, the area is judged to be at risk (GAM  $\geq 15\%$ ). If 23 or less children are malnourished, we conclude the area is not at risk (GAM  $< 15\%$ ).

### GAM sampling design and protocol

GAM prevalence is estimated using a specific sampling protocol. Traditional LQAS requires a simple random sample (SRS) of households and one randomly sampled interviewee per household.<sup>11</sup> In contrast, the sampling protocol to obtain an unbiased estimate of GAM in emergency settings requires that every child 6–59 months in a sampled household be measured.<sup>5,6,12</sup> A two-stage 30 × 30 cluster-survey is most commonly used.<sup>7</sup> The standard method involves selecting 30 clusters by Probability Proportional to Size (PPS) and the first household within each cluster by a random method such as spin-the-bottle. Subsequent households to be sampled within a cluster are typically selected by proximity to the first random household selected.<sup>5,6,12</sup> Data collected by the 30 × 30 cluster-method are clusters of children within households as well as clusters of households within clusters. In short, this means the children selected to obtain an unbiased estimate of GAM are not a SRS—they are clusters. Therefore, one cannot assume that LQAS can be used for analysing GAM.

By using computer simulations, we investigated whether the DRs derived from binomial probabilities were accurate when GAM was assessed using data collected in small clusters. Different population sizes as well as different intra-cluster correlations (0, 0.25 and 0.5) were used in the simulations. These intra-cluster correlation levels represent the GAM correlations that could occur from sampling multiple children in a household as well as from sampling multiple households within a cluster. Results from the simulations revealed that for an intra-cluster correlation of 0.25, the alpha error ranged from 0.045–0.057 and the beta error ranged from 0.258–0.287, when data were collected in households having one to six children<sup>13</sup> (Table 1). In other words, small cluster-sizes of size two to six had errors similar to a SRS of the same size for GAM thresholds of 10, 15, and 20%.

Three LQAS designs were developed from those simulation results: (1) 33 × 6 (33 clusters, six children in each cluster); (2) 67 × 3; and (3) a 67 × 3 sequential design. The designs use the smallest sample-sizes possible to detect GAM upper thresholds of 10 and 15% (with lower thresholds of 5 and 10%, respectively), while maintaining alpha errors  $\leq 0.10$  and beta errors  $\leq 0.20$ . The designs can also detect a GAM upper threshold of 20% (with lower threshold of 15%) with slightly larger errors (alpha  $\leq 0.13$ , beta  $\leq 0.24$ ).<sup>13</sup> The latter threshold was relevant to Ethiopia, given the emergency classification scale used by the Government.<sup>14</sup> In addition,

**Table 1** Alpha and Beta errors associated with clusters of size 1–6 in computer simulations for  $\geq 15\%$  GAM<sup>a</sup>, using varying levels of intra-cluster correlation, where  $N=2000$ ,  $n=200$ , decision rule = 23 (13)

Intra-Cluster					
Correlation	Cluster- Size	Alpha Error	Beta Error	Total Error	
0	1	0.049	0.235	0.284	
0	2	0.050	0.231	0.281	
0	3	0.047	0.247	0.294	
0	4	0.054	0.238	0.292	
0	5	0.051	0.236	0.287	
0	6	0.063	0.221	0.284	
0.25	1	0.046	0.268	0.314	
0.25	2	0.045	0.273	0.318	
0.25	3	0.046	0.287	0.333	
0.25	4	0.049	0.277	0.326	
0.25	5	0.047	0.266	0.313	
0.25	6	0.057	0.258	0.315	
0.5	1	0.061	0.264	0.325	
0.5	2	0.066	0.260	0.326	
0.5	3	0.057	0.274	0.331	
0.5	4	0.060	0.263	0.323	
0.5	5	0.062	0.262	0.324	
0.5	6	0.064	0.246	0.310	

<sup>a</sup>Beta errors computed using a lower threshold of 10%.

the  $33 \times 6$  and  $67 \times 3$  designs provide point estimates and 95% CIs for GAM and other child-level indicators.

The sequential design uses a multi-stage sampling plan which allows GAM thresholds to be detected before the full sample has been collected. The design can be used when the GAM threshold is defined a priori and point estimates for GAM and other child-level indicators are not desired. The full sample plan is based on the  $67 \times 3$  design (67 clusters, three children in each cluster), however, the design allows for data analysis once a minimum sample-size of 25 is attained. To allow for multiple hypothesis testing, the sequential design uses lower errors ( $\alpha \leq 0.01$ ,  $\beta \leq 0.05$ ) for DRs at the smaller sample-sizes.<sup>13</sup> Once the clusters are selected by PPS, the clusters can be sampled in any order. Sampling stops when one can make a reliable classification about the GAM threshold of interest. If the number of malnourished children is neither sufficiently high nor sufficiently low, a classification cannot be made about the threshold and sampling continues. Up to four hypothesis tests can be conducted over the course of the sampling while still conforming to overall error limits of  $\alpha \leq 0.10$  and  $\beta \leq 0.20$ .

### Assessment site

Siraro woreda, located in the Oromiya Region of Ethiopia, comprises 62 geographically defined Peasants Associations (PAs) and has a population of about 200 000. The main source of livelihood is subsistence agriculture. During 2002–2003, the Famine Early Warning System Network (FEWS NET) estimated that approximately 4 million in the Oromiya region of Ethiopia were affected by drought-related

effects of food insecurity.<sup>15</sup> At the time of this study, 22 of the 62 PAs in Siraro had lost 60–70% of their crops, 23 had lost 40–60% and 17 had lost 25–40%.<sup>16</sup> Agencies working in the area suspected elevated GAM prevalence and that a priority situation existed.

### Sampling methodology

To assess the situation in Siraro, Catholic Relief Services (CRS), a relief and development agency implementing programs in Ethiopia, in collaboration with the FANTA Project managed by the Academy for Educational Development (AED), administered a conventional  $30 \times 30$  cluster-survey in accordance with the Government of Ethiopia's emergency nutrition survey guidelines.<sup>14</sup> The  $33 \times 6$  and  $67 \times 3$  designs were administered simultaneously using the same questionnaires and a nested design. The data collected for the  $67 \times 3$  design were also used for the sequential design.

The clusters (PAs) in the  $30 \times 30$  cluster-survey and the  $67 \times 3$  design were selected independently from the same sampling frame of PAs using PPS sampling. The first 30 clusters of the  $33 \times 6$  design were the 30 clusters selected for the  $30 \times 30$  cluster-survey. The last three clusters were selected using PPS from among the clusters sampled for the  $67 \times 3$  design. During data collection, two of the PAs selected for data collection were inaccessible. Replacement clusters were selected from PAs on the basis of close proximity and similar population size.

All children 6–59 months in a household were included in the study and measured using standard anthropometric methods.<sup>14,17</sup> To randomly identify the first household to be sampled in each cluster, the spin-the-bottle method was used.<sup>12,14</sup> Subsequent households in a cluster were selected by geographic proximity. This sequence of household selection continued until the minimum prescribed sample-size for the cluster of each design was attained.

If a PA was selected as a primary sampling unit by more than one survey design, the data were shared across survey designs, until the sample-size required for the cluster of a given design was attained.

### Data management

To collect data for this study, 15 interviewers (five interview teams) were hired. The teams received 35 h of training. During data collection (June 4–11 2003), three technical advisors provided supervision and technical support.

For the time study, the teams used an odometer to record the distance travelled from the base camp to the first cluster for each work-day and used a stopwatch to record the time required to complete each survey. Additional data collected by one technical advisor included the amount of time required to: (i) arrive at the first randomly selected household in each cluster, (ii) walk the distance between households, and (iii) complete each segment of travel. To calculate the time required for data collection, average measures of time expenditure were applied to the cluster and sample-size specifications of each design.

Data were checked for out-of-range and missing-values and entered (June 5–13 2003) using SPSS V7.0. Identical data cleaning and transformation procedures were applied to

**Table 2** GAM point estimate results with assumption of simple random sample (i.e. design effect = 1.000) compared with GAM point estimate results by CSample analysis (i.e. accounting for design effect)

Design type	Analysis method			
	SRS analysis method		CSample analysis method, weighting inversely proportional to achieved cluster-size	
	Point estimate, 95% CI	Design effect	Point estimate, 95% CI	Design effect
30 × 30 Cluster-survey	NA	NA	7.4 (4.8, 9.9)	2.175
33 × 6 LQAS design	6.7 (3.2, 10.2)	1.000	6.7 (2.9, 10.4)	1.167
67 × 3 LQAS design	8.2 (4.3, 12.1)	1.000	8.7 (4.9, 12.5)	0.931

**Table 3** Point estimates with 95% confidence intervals for 12 Key nutrition and health indicators, derived from three sampling designs

Indicator	30 × 30 Cluster-Survey	33 × 6 LQAS Design	67 × 3 LQAS Design
Global acute malnutrition (GAM)	7.4 (4.8, 9.9)	6.7 (3.2–10.2)	8.2 (4.3–12.1)
Oedema	1.5 (0.4, 2.6)	1.6 (0.0, 3.4)	4.2 (1.4, 7.0)
Wasting	6.3 (3.8, 8.7)	5.6 (1.9, 9.3)	5.0 (1.9, 8.2)
Severe acute malnutrition (SAM)	2.3 (1.0, 3.6)	2.1 (0.1, 4.1)	5.7 (2.3, 9.1)
Underweight	40.3 (36.2, 44.5)	41.8 (33.4, 50.1)	49.2 (40.9, 57.6)
Stunting	51.5 (47.2, 55.9)	52.7 (44.0, 61.4)	57.3 (49.2, 65.4)
Had vaccination card	10.3 (6.2, 14.3)	12.1 (5.3, 19.0)	11.2 (5.8, 16.5)
Received BCG vaccination	36.4 (24.9, 47.9)	40.3 (28.1, 52.5)	41.7 (32.7, 50.8)
Received measles vaccination	26.1 (16.0, 36.1)	27.5 (17.0, 38.0)	27.0 (18.9, 35.1)
Received vitamin A capsule in last 6 months	31.5 (16.9, 46.0)	35.4 (20.8, 49.9)	26.9 (16.7, 37.0)
Diarrhea in last 2 weeks	24.4 (17.3, 31.5)	31.2 (22.8, 39.5)	32.5 (23.9, 41.1)
Fever in last 2 weeks	24.8 (18.8, 30.8)	30.7 (21.5, 39.9)	37.5 (28.4, 46.6)

the 30 × 30, 33 × 6 and 67 × 3 data sets. Anthropometric Z-scores were calculated using the 1978 NCHS growth references and the EpiNut sub-routine of Epi-Info 6.04. Children with out-of-range anthropometric values flagged by Epi-Info were excluded from analysis, unless bipedal oedema was present.<sup>18</sup> The data cleaning resulted in some clusters having less than the prescribed sample-size. Also, certain clusters were larger than the planned design since the sampling protocol called for data to be collected on all children 6–59 months in a sampled household. Oversized clusters were treated by randomly eliminating the excess number of cases. The final sample-size of the 30 × 30, 33 × 6 and 67 × 3 designs was 871, 195 and 194, respectively. To treat unequal cluster-sizes within designs, data were weighted inversely proportional to the achieved sample-size during point estimate calculations.

The 30 × 30 cluster-survey data were analysed using the CSample sub-routine of Epi-Info 6.04 which calculated the design effect for GAM (DE = 2.175) and other child-level indicators. Because computer simulations revealed the errors associated with GAM to approximate those of a SRS for the 33 × 6 and 67 × 3 designs,<sup>13</sup> accounting for design effects for GAM was not essential. The primary analysis of GAM prevalence for the 33 × 6 and 67 × 3 data sets was therefore conducted using SPSS V12.0. For this analysis, the data were not weighted. To perform a test of the SRS assumption, a secondary analysis of GAM prevalence was also conducted on the 33 × 6 and 67 × 3 data, using CSample-the same analysis

method applied to the 30 × 30 cluster-data. All other indicators for the 33 × 6 and 67 × 3 designs were analysed using the Epi-Info CSample sub-routine.

## Results

### Point estimate analysis

The prevalence of GAM from the 33 × 6 and 67 × 3 designs found estimates of 6.7% (CI = 3.2–10.2%) and 8.2% (CI = 4.3–12.1%), respectively. The results are similar to the 30 × 30 cluster-design which estimated a GAM prevalence of 7.4% (CI = 4.8–9.9%). The CSample analysis of GAM prevalence for the 33 × 6 and 67 × 3 data produced results similar to those which assumed a SRS (Table 2). As evidenced by the overlapping CIs, indicators related to child morbidity and vaccination status also yielded similar results across all designs (Table 3).

### GAM threshold analysis

Table 4 shows the DRs used to test the 33 × 6 and 67 × 3 designs for the 10, 15 and 20% GAM thresholds. GAM was found in 13 children in the 33 × 6 design ( $n = 195$ ) and in 16 children in the 67 × 3 design ( $n = 194$ ). With DR = 23, both designs indicate the GAM prevalence is <15% (Table 4). The assessment of whether the 10% threshold had been reached (DR = 13) yielded mixed results across designs.

**Table 4** Decision rules<sup>a</sup> for 33 × 6 and 67 × 3 LQAS designs for various GAM thresholds

LQAS Design	<i>n</i>	# of Children with GAM	DR (# of Children with GAM)		
			≥10% GAM	≥15% GAM	≥20% GAM
33 × 6 LQAS Design	195	13	>13	>23	>33
67 × 3 LQAS Design	194	16	>13	>23	>33

<sup>a</sup> All DRs use a 5 percentage point difference between upper and lower GAM thresholds.

**Table 5** Decision rules<sup>a</sup> for 67 × 3 sequential design for various GAM thresholds, by cluster of data collection

Cluster #	Sample Size	# Children with GAM	DR ≥10% GAM	DR <10% GAM	DR ≥15% GAM	DR <15% GAM	DR ≥20% GAM	DR <20% GAM
10	30	2	>4	NA	>6	<1	>8	<1
11	33	2	>4	NA	>6	<1	>8	<2
12	36	2	>4	NA	>7	<1	>9	<2
13	39	2	>4	NA	>7	<1	>10	<3
14	42	2	>5	NA	>8	<2	>10	<3
15	45	3	>5	<1	>8	<2	>11	<3
36	107	8	>9	<4	>16	<8	>22	<12
37	110	8	>9	<4	>16	<8	>23	<13
38	113	8	>10	<5	>17	<9	>23	<13
39	115	9	>10	<5	>17	<9	>24	<14
40	118	9	>10	<5	>17	<9	>24	<14
59	171	13	>13	<9	>23	<15	>33	<23
60	174	13	>13	<9	>23	<16	>33	<23
61	177	14	>13	<9	>23	<16	>33	<24
62	180	14	>13	<9	>23	<16	>33	<24
63	183	14	>13	<9	>23	<17	>33	<24

<sup>a</sup>All DRs use a 5 percentage point difference between upper and lower GAM thresholds.

NA=Not Applicable—Sample-size is too small to make a decision about the threshold within the defined limits of errors.

The 33 × 6 design indicates the GAM prevalence is <10%, whereas the 67 × 3 design indicates the GAM prevalence is ≥10%.

The sequential design shows GAM classification results consistent with those of the 67 × 3 design. Our analysis indicates that judgments could be made before all the data were collected. Analysis by cluster reveals that sufficient information was available to judge GAM prevalence <20% after cluster 13 (*n*=39). After cluster 38 (*n*=113) GAM prevalence could be judged <15%. The judgment of GAM prevalence ≥10% was possible after cluster 61 when 14 of 177 children exhibited GAM (Table 5).

### Data collection time

Analysis of the time expenditure data showed, on average, two clusters in the 30 × 30 design could be completed by one interview team (three people) in a 12 h work day (30clusters/2 per day × 3 people = 45 person days). A 12 h work day was therefore assumed for all time estimation calculations. Our analysis shows the time required for the 30 × 30 cluster-survey and the three LQAS designs ranged from 5.22 to 45.00 person-days (Table 6). These time expenditure results are valid for this study only. Results would vary according to

the geography, infrastructure and prevalence of malnutrition in the area sampled.

## Discussion

This study compared three new LQAS designs with the 30 × 30 cluster-design commonly used for emergency nutrition surveys. Across designs the following results were compared: point estimates and 95% CIs for GAM and other child-level indicators, classification results for GAM thresholds of 10, 15 and 20%, and time required for data collection. Similar point estimates and comparable CIs for GAM and other child-level indicators were found from the 30 × 30, 33 × 6 and 67 × 3 designs. The CIs for the 33 × 6 and 67 × 3 designs are only slightly wider than those of the 30 × 30 design for most indicators tested; yet, the LQAS designs required substantially less time when compared with the 30 × 30 cluster-survey. The CSample analysis of GAM prevalence showed the 33 × 6 and 67 × 3 designs to conform generally to the SRS assumption. The slightly inflated design effect (>1.000) for the 33 × 6 design is likely related to the reduced sample-size. Although this study showed inconsistent classification results for the 33 × 6 and 67 × 3 designs at the 10% threshold, this is probably

**Table 6** Comparison of GAM point estimates, LQAS decisions and time expenditure for various sampling designs

Sampling design		Point estimate (%) (95% CI)	LQAS decision	Person-days required to collect data <sup>a</sup>
Cluster	30 × 30	7.4 (4.8–9.9)	NA	45.00
LQAS	33 × 6	6.7 (3.2–10.2)	<10% GAM threshold	16.95
LQAS	67 × 3	8.2 (4.3–12.1)	≥10% GAM threshold	25.68
Sequential LQAS	67 × 3	NA	≥10% GAM threshold	23.22
Sequential LQAS	67 × 3	NA	<15% GAM threshold	14.49
Sequential LQAS	67 × 3	NA	<20% GAM threshold	5.22

<sup>a</sup>Time estimation assumes a 12 h work-day.

due to the normal amount of error associated with surveys. Computer simulations have shown the statistical error associated with the 33 × 6 and 67 × 3 designs to be similar across designs at each GAM threshold.<sup>13</sup>

The most critical limitation of the study is the nested sample design. Ideally, data for each design should be sampled independently to allow for stricter comparison of results between designs. In addition, comprehensive time expenditure data would have been preferable. In this study, only one person (who rotated teams) collected time data related to driving, spin the bottle and walking household to household; with five teams working, one sample of time data may not represent all their efforts.

Despite these limitations, the LQAS designs described here can contribute to the methodological toolkit of humanitarian agencies. The 33 × 6 and 67 × 3 designs provide alternatives to the more time-consuming 30 × 30 cluster-survey. While alternative sampling methods may allow for a more precise CI for estimating GAM prevalence, the trade-off is a greater sample-size and more time for data collection than required by the 33 × 6 and 67 × 3 designs. The results from this study indicate the CIs for the LQAS designs are only slightly wider than the 30 × 30 design. However, because LQAS decision rules are based on binomial probabilities, in cases where the CI of a GAM point estimate overlaps a threshold for determining humanitarian action, the LQAS designs can be used for hypothesis testing, to make a decision about action. The LQAS decision rules minimize the alpha error so that the risk of not detecting a GAM threshold is small ( $\leq 0.10$ ), providing programme managers with a useful tool to decide whether resources

should be moved into an area. Furthermore, in comparison with a 30 × 30 cluster-survey, the LQAS designs are faster to implement, requiring fewer resources to obtain the information needed for action.

## Acknowledgments

The authors thank Stacy Hoshaw-Woodard for conducting the computer simulation work. The authors are also grateful to Cherinet Abuye for training the interview teams and supervising data collection, and to the country offices of CRS and Christian Children's Fund (CCF) in Ethiopia for the administrative and logistic support provided during the fieldwork. The authors also acknowledge Eunyong Chung, Bruce Cogill, Nadra Franklin, and Anne Swindale for their review of the manuscript. Special thanks are due to Gilles Bergeron who reviewed an earlier draft of the manuscript and provided valuable editorial comments. Financial support for this project was provided to CRS from the Andrew W Mellon Foundation and to the Food and Nutrition Technical Assistance (FANTA) Project by the Office of Health, Infectious Diseases and Nutrition of the Bureau for Global Health and the Ethiopia Mission at the US Agency for International Development (USAID), under terms of Cooperative Agreement No HRN-A-00-98-00046-00 awarded to the AED. The opinions expressed herein are those of the author(s) and do not necessarily reflect the views of the Andrew W Mellon Foundation or USAID.

**Conflict of interest:** None declared.

### KEY MESSAGES

- Three Lot Quality Assurance Sampling (LQAS) designs (33 × 6, 67 × 3, Sequential design) were developed to provide alternatives to implementing the 30 × 30 cluster-survey for assessment of acute malnutrition in emergency settings.
- Computer simulations confirmed that small-cluster sizes rather than a conventional simple random sample could be used for LQAS assessment of GAM thresholds of 10, 15 and 20%, a finding which has important implications for the sampling and analysis methods that can be applied for emergency nutrition surveys in developing countries.
- The results from this study show the 33 × 6 and 67 × 3 designs provide point estimates and 95% CIs similar to those of a 30 × 30 cluster-survey for child-level indicators, and that the LQAS designs can be administered in a fraction of the time required by the 30 × 30 cluster-survey.
- Humanitarian organizations and public health practitioners need no longer be confined to using the slower 30 × 30 cluster-survey for assessment of acute malnutrition. The 33 × 6, 67 × 3 and sequential designs are less time-intensive sampling approaches, allowing for rapid identification of areas requiring humanitarian assistance.

## References

- <sup>1</sup> Mason JB. Lessons on nutrition of displaced people. *J Nutr* 2002;**132**:2096S–2103S.
- <sup>2</sup> World Health Organization (WHO). *Management of Nutrition in Major Emergencies*. Geneva: World Health Organization, 2000.
- <sup>3</sup> World Health Organization (WHO). Physical status: the use and interpretation of anthropometry. *Report of a WHO expert committee*. Geneva: World Health Organization, 1995.
- <sup>4</sup> University of Nairobi. *Report of a Workshop on the Improvement of the Nutrition of Refugees and Displaced People in Africa*. Geneva: Standing Committee on Nutrition (SCN), 1995.
- <sup>5</sup> Standardized Monitoring and Assessment of Relief and Transitions (SMART). Measuring mortality, nutritional status and food security in crisis situations. Version 1, 2006. Available at: <http://www.smartindicators.org/>
- <sup>6</sup> Mediciens Sans Frontieres (MSF). *Nutrition Guidelines*. 1st edn. Paris: Mediciens Sans Frontieres, 1995.
- <sup>7</sup> Spiegel PB, Salama P, Maloney S, van der Veen A. Quality of malnutrition assessment surveys collected during famine in Ethiopia. *JAMA* 2004;**292**:613–18.
- <sup>8</sup> Valadez JJ. *Assessing child survival programs in developing countries: testing lot quality assurance sampling*. Cambridge: Harvard University Press, 1991.
- <sup>9</sup> Reinke W. Applicability of industrial sampling techniques to epidemiologic investigations: examination of an underutilized resource. *Am J Epidemiol* 1991;**132**:1222–32.
- <sup>10</sup> Valadez JJ, Weiss B, Leburg C, Davis R. *Assessing Community Health Programs: using LQAS for baseline surveys and regular monitoring*. London: Teaching-aids at Low Cost, 2003.
- <sup>11</sup> Roberston SE, Valadez JJ. Global review of health care surveys using lot quality assurance sampling (LQAS), 1984–2004. *Soc Sci Med* 2006;**63**:1648–60.
- <sup>12</sup> Save the Children (SC) UK. *Emergency Nutrition Assessment: Guidelines for Field Workers*. London: Save the Children, 2004.
- <sup>13</sup> Hoshaw-Woodard S. The use of LQAS sampling to assess malnutrition. Ohio State University. Version II, Final Draft, 2003.
- <sup>14</sup> Ethiopia Early Warning Department, Disaster Prevention and Preparedness Commission (DPPC). Ethiopia Early Warning System (EWS): Guideline on emergency nutrition assessment. Addis Ababa: Early Warning Department/Disaster Prevention and Preparedness Commission, 2002.
- <sup>15</sup> Famine Early Warning System Network (FEWS NET): Washington, DC: Ethiopia: unmistakable pre-famine conditions. Ethiopia monthly food security update, 2002 Dec 9. Available at: <http://www.fews.net>.
- <sup>16</sup> Ethiopia Disaster Prevention and Preparedness Commission. Siraro: Unpublished data, 2003 June.
- <sup>17</sup> Cogill B. *Anthropometric Indicators Measurement Guide*. Washington, DC: FANTA Project/Academy for Educational Development, 2003.
- <sup>18</sup> Dean A, Dean J, Coulombier D et al. *Epi Info: A Word Processing, Database, and Statistics Program for Public Health on IBM Compatible Microcomputers*. 6th edn. Atlanta: Centers for Disease Control and Prevention, 1996.