LQAS FREQUENTLY ASKED QUESTIONS

2008

The LOT Quality Assurance Sampling (LQAS)Frequently Asked Questions document was produced by the CORE Group M&E working group as a result of a TAG meeting held in August 2008,in which practitioners developed new guidelines for parallel sampling in order to improve the quality of monitoring and evaluation efforts.

CORE Group

*Working Together in Health for
Mothers, Children, and Communities*

1. **Q: When should/shouldn't I use LQAS?**

A: If your project is <u>ONLY</u> interested in obtaining the frequency (mean) and confidence interval for the various indicators of interest at the level of the project area as a whole, then 30 cluster methodology may be preferable. 30-cluster sampling requires interviews in 30 sites compared to about 95 for LQAS. It is likely to be less expensive and less complex logistically---if, again, your only goal is to obtain mean values and confidence intervals on project indicators at the project area level. However, if the project is also interested having data about performance of indicators at the sub-project level, and not only at baseline and final but throughout the life of the project (in order to make adjustments along the way), then LQAS is the recommended methodology. It will be less expensive than carrying out several 30-cluster surveys within the project area.

The following table offers some advantages and disadvantages of 30 cluster and LQAS methodologies:

	30 Cluster Methodology	**LQAS Methodology**
USES	Baseline, midterm, and final evaluations	Baseline, midterm, final evaluations, and regular monitoring of project interventions
RESULTS	The frequency and confidence interval of all indicators chosen, for the entire project area as a whole (assuming 30 or more sites are included in the sample for each indicator). For example, 30.26% (22.76%, 37.6%) of mothers of children aged 0-23 months in the project area received at least 2 tetanus toxoid immunizations before the birth of the youngest child.	▪ Whether or not the performance of indicators is adequate or not in each of the supervision areas (SA) <u>within</u> the project catchment area. ▪ The frequency and confidence interval of all indicators chosen, for the entire project area as a whole (assuming 95 or more sites are included in the sample for each indicator).
ADVANTAGES	▪ Allows the calculation of frequencies (means) and confidence intervals for the entire project area (but these results should be adjusted for design effects of sampling in clusters). ▪ Often less costly in baseline and final evaluations due to the requirement to visit only 30 sites---assuming survey sites are not to be	▪ Allows for the calculation of frequencies (means) and confidence intervals for the entire project area (but these results should be weighted for SA population sizes) ▪ Also allows for the decentralization of data management to the field level, which means that managers have information to judge performance of indicators in individual supervision areas within an entire project area. ▪ Also allows for routine monitoring of project interventions for management purposes. ▪ Simplified data analysis allows for hand calculation of indicators at the supervision area level

	30 Cluster Methodology	LQAS Methodology
	visited during routine visits to communities but as a one-off effort to collect the data quickly.	▪ In routine monitoring as well as at baseline and final, data collection may be carried out by project staff, while they are in the sampled communities during performance of their regularly scheduled duties, thus lowering data collection costs and staff time required ▪ Allows for testing a hypothesis regarding a particular sub-group's relative performance in an intervention area
DISADVANTAGES	▪ Only allows projects to calculate means (coverage/prevalence) and confidence intervals for the entire project area. ▪ Analysts must use computer programs allowing for complex sampling designs to adjust for design effects when calculating confidence intervals (e.g., Epi Info).	▪ Requires visits to more sites than a 30-cluster survey requires. ▪ Data analysis at the level of the project as a whole is relatively more complex as it requires users to analyze data at two levels: supervision area and project area. Supervision area level analysis, however, can be done manually; computers (spreadsheets at minimum) are very helpful for project level analysis.
WHEN TO USE	▪ When the ONLY result desired is a mean (coverage/prevalence) and confidence interval for each indicator for the entire project area ▪ When use in monitoring is not desired	▪ When a project is also interested in having data about performance of indicators at the sub-project level, not just at the level of the project as a whole. ▪ Whenever monitoring of the project is also desired, because it combines receiving means and confidence intervals at the project area level for Baseline and Final Evaluations with the ability to monitor all project indicators on a routine basis. ▪ When the project is sub-divided into management areas or supervision areas with persons assigned to manage/supervise the performance of the project in these sub-project areas. LQAS will provide these manager/supervisors information to help them prioritize which indicators require more attention. ▪ LQAS may also be used at baseline and final even if the project does not plan to use it for routine monitoring. For example, it may be used at baseline, MTE, FE to let projects understand what is happening at the SA level so that project managers can differentiate their program interventions where I see variation that may even appear at baseline. Therefore, this is useful even without a plan to collect data annually with LQAS.

2. **Q: If I just want a coverage estimate (mean and confidence interval) at the level of the project area, and am not worried about performance of indicators within sub-project areas (e.g., supervision areas), is LQAS for me?**

A: If the desire is to have means with confidence intervals for indicators for the entire project area ONLY, then the 30 cluster methodology would usually be preferable because it would usually be less expensive to carry out and easier logistically. LQAS is more desirable when data at both the project level and sub-project level are desired for either baseline or final and/or for routine monitoring purposes.

3. **Q: What criteria should guide the selection of Supervision Areas?**

A: Supervision Areas should be chosen that make programmatic sense for the project, not "invented" in order to perform LQAS. Each SA should be a performance area that will be evaluated utilizing LQAS in order to make management and programmatic decisions, like health center catchment areas or areas covered by different supervisors. In addition, performance of indicators at the project level can also be calculated when the sample size of all supervision areas combined is about 95 or greater. This ensures that the confidence interval will be +/- 10% or less. Note that the calculation will require that the results (means and confidence intervals) are weighted by the relative population size of each SA. Five to six supervision areas (each with a sample of 19) is ideal. If the project only has four or fewer supervision areas, then the sample size in each area should be increased so that the combined number is about 95 or greater. For example, a project with four supervision areas should consider including a sample size of 24 in each supervision area for a combined sample size of 96.

4. **Q: May LQAS results be stratified by ethnic group, gender, and other variables of interest?**

A: Theories regarding different strata of a variable may be tested by utilizing the Decision rule in certain cases. The process of determining whether LQAS is proper for this type of analysis should be based upon what level of precision is desired:
 i. If you have a subset of the data upon which you would like more information, this subset can be analyzed using LQAS decision rules to determine if performance meets certain targets or benchmarks.
 ii. If, however, you would like a mean and confidence interval for this sub-group, then one must only calculate the mean and confidence interval at the project level with the data from all supervision areas combined. The confidence interval is likely to be greater than desirable, however (greater than +/- 10%), if the combined sample size is less than 95 for the indicator of interest. [1]

---

[1] Both 30-cluster sampling and LQAS allow for different types of stratification. After conducting a 30-cluster sample, project staff may want to stratify a particular indicator (which has been asked of all 300 mothers) by age, for example. This allows project staff to compare each age group to the whole, but not necessarily allow comparison of each group against each other, because the relative sample sizes will be too low and therefore the confidence intervals will be too large. [This may also be accomplished with LQAS data by calculating the weighted mean and CI separately for boys and girls or for independent age groups]. For example, using 30-cluster sampling, a community health project may have gathered information on weight-for-age for children under two years of age. In this example, perhaps 1.3% (confidence interval: 0.0, 7.1) of children age 0-5 months and 12.3% (6.1, 21.5) of children 6-12 months were moderately underweight. In this example, despite the appearance of a considerable difference between these two results, the confidence intervals of the age groups overlap and no statistically significant difference between the groups can be demonstrated. However, even if there is overlapping, the difference between 1.3% and 12.3% is large enough that project managers may use the data to make programmatic inferences and decide to focus extra attention on nutrition for children aged 6-12 months. For stratifying the differences between groups for programmatic (management) purposes, LQAS may be more appropriate. For example, if a project has a nutrition intervention, and project staff think that one group (children aged 6-12 months) are not doing as well as expected, this theory can be tested

iii.If you need to have a confidence interval less than +/- 10% or you want to determine if there is a statistically significant difference between this population subgroup and the other groups, then a separate survey for each of these groups needs to be carried out using sample size calculations to guide sample size determinations.

5. Q: What Criteria/factors should be considered when aggregating results obtained from different LQAS surveys?

A: The following items should be considered when considering aggregating results:

1. **Differences in Questionnaires:** If the questionnaires are different or ask similar questions but in a different format, then care should be taken to determine how these differences would bias the results. Aggregating indicators taken from dissimilar questionnaires will yield biased results.

2. **Geographic areas:** Careful consideration should be given to the survey areas and Supervision Areas chosen for each survey. If the surveys have overlapping areas, or would have gaps due to some areas between the different surveys not being included, then aggregating results will not yield reliable, unbiased results.

3. **Population Size:** It is important to weight the results of each SA according to its relative population size, in order to avoid giving too much influence to SAs with smaller population sizes and too little influence to SAs with larger population sizes.

4. **Date of Surveys:** If the different surveys were performed greater than 12 months apart, it is not recommended that they be aggregated. Too many confounding factors could influence the results over this time frame, including results of programmatic interventions carried out during this time.

5. **Overriding Difference in Areas:** Other differences in the sampling areas should be considered as well, such as whether both areas are rural or urban, etc. These differences could also bias the cumulative results.

6. Q: Is it advisable to use LQAS in a heterogeneous (diverse ethnically, culturally, geographically, etc.) community?

A: Due to the sampling strategy of LQAS, at least 95 survey sites (clusters) will be chosen to collect the data, versus 30 for 30 cluster sampling methodology. Therefore, more of the project area's diversity is captured using LQAS and as a result it is more representative of the entire community.

7. Q: What is the layman's explanation of upper and lower thresholds, as well as Alpha and Beta errors?

A: The driving assumption of LQAS is that it will identify whether or not a minimal proportion of responses to questions regarding an indicator are "correct", showing that the SA is performing adequately with regard to that indicator. This performance adequacy is a function of the performance benchmark or target, and is called the upper threshold. The purpose of the LQAS study is to identify indicators that perform below the upper threshold (the benchmark or target). Using a small sample size, we understand that some indicators that are performing below the

---

by doing parallel sampling to make sure that 19 of this group are included, and then use the decision rule to see if the group of children aged 6-12 months differs significantly from the average or threshold. In this example, the 6-12 month old group would not meet the decision rule, and therefore would be deemed not of adequate weight for age, and it could then be stated that it is statistically and significantly below the average.

target/benchmark will be classified as acceptable---especially indicators that are performing just below the target. The lower threshold is set as an unacceptably low level of coverage that should not go undetected. For management purposes, we typically are setting a lower threshold at a value that is 30% below the target. It should provoke managers to focus on determining the causes of the failure and resolve it.

To better understand thresholds, it is necessary to also understand Alpha and Beta errors. 1-Alpha is equivalent to specificity, which is the probability of correctly identifying SAs that reach their performance benchmarks (upper threshold). This means that alpha error is the probability that projects will identify a SA as having performed inadequately with regard to an indicator that has actually reached the performance standard. 1-Beta is equivalent to sensitivity, which is the probability of correctly identifying SAs that do not reach the determined performance standard (lower threshold). This means that B error measures the probability that a SA will be identified as having reaching the performance standard with regard to an indicator when in fact it has not. In LQAS, both Alpha and Beta errors for sample sizes of 19 are less than 10%. The links below will also provide more detailed information on thresholds.

8. Q: Why is a sample size of 19 sufficient? When is a sample size of 19 too small?

A: When looking at the LQAS Table for Decision Rules, you will notice that 19 is the smallest sample size that allows at least 90% sensitivity and specificity for all benchmarks or targets of 10%-95%. This means that when the decision rule is <u>not</u> met you have evidence that the target <u>has not</u> been achieved (this will be true 90% or more of the time). Note that if the decision rule is met, you cannot say you have evidence that the target has been achieved; this is a limitation of a small sample. What you can say is that there is "no evidence that it <u>has not</u> been achieved." What you do have evidence for is that the percent achieved is at minimum a value 30% below the target (this will be true 90% or more of the time)---the achievement is likely to be higher than this but statistical evidence is limited to a value 30% below the target. A sample size of 19 is too small when you want more precision: that when the decision rule is met you have evidence that the percent actually achieved is more than the value of the target % - 30%. For more in depth information regarding sample size determinations please see the following references:

- *Assessing Community Health Programs Using LQAS for Baseline Surveys and Regular Monitoring*[2]
- The *KPC Trainer of Survey Trainers curriculum*[3]
- Session summary from the CORE Spring meeting 2003[4]

9. Q: How do you select a decision rule when your target/threshold is very low?

---

[2] *Assessing Community Health Programs Using LQAS for Baseline Surveys and Regular Monitoring (Participants manual and workbook and the trainer's guide)* by Joseph J. Valadez, William Weiss, Corey Leburg, and Robb Davis http://www.coregroup.org/working_groups/lqas_train.html and http://www.coregroup.org/working_groups/LQAS_Participant_Manual_L.pdf
[3] http://www.coregroup.org/working_groups/kpc_training/welcome.html
[4] http://www.coregroup.org/resources/meetings/april03/thur03_LQAS_vs_Cluster.pdf

A: If the threshold is very low (10% or below), LQAS may not be the appropriate survey methodology to use. However, with a target of 15%, the sample size may be increased from 19 to 25 in order to obtain a Decision Rule of 1.

10. **Q: At what point do you stop collecting data on an indicator?**

A: If you have inquired at every household in the randomly chosen community and cannot find a respondent that meets the requirements to be eligible to answer the questions, then it is suggested that the survey move to the next nearest community and to the next nearest doorway (from the last doorway visited in the previous community) in that community if possible. This is similar to the protocol in 30 cluster sampling.

11. **Q: How many times do you go back to a house?**

A: During the pre-implementation phase, particularly in the training sessions, use participatory principles to devise a protocol for all to follow (to ensure that it is standardized). For example, it may be decided that the surveyor should go back to the household two additional times during the course of that day, if they have been told that respondent is supposed to return home that day, or possibly within the next 2-3 days if logistics allow.

12. **Q: Why can't I express my LQAS results at the supervision area level as percentages, and why can't I calculate coverage/ prevalence rates for SAs?**

A: The Confidence Intervals would be too wide at this level, due to the small population size sampled, to make any meaningful inferences. The smallest sample size that provides a confidence interval +-10% or less for all rates between 0% - 100% is 95.

13. **Q: I used 30 cluster methodology for the project baseline. May I use LQAS for the final?**

A: Yes assuming that the questions one is trying to compare have not changed. There may be several reasons to do this. One example would be that a project decides, after the baseline, that it would like to do routine monitoring using LQAS, and would therefore like to use LQAS for the final evaluation. The change in method should be explained in the narrative of the results section of any subsequent reports comparing the final results to the baseline findings.

14. **Q: If there is more than one respondent in a household, should they be included in the survey?**

A: When the questions pertain to the respondents' practices or behaviors (% of mothers who wash their hands with soap at the appropriate times, for example), interviewing more than one respondent in the same household would introduce bias into the findings because respondents living in the same household tend to share the same practices or behaviors. In this case, only one of the respondents should be randomly chosen (by flipping a coin, etc.). When determining coverage rates (% of children 0-23 months who are fully immunized, for example) it is permissible to interview more than one respondent, but this still may introduce a similar bias in that respondents in the same household have

similar access to coverage. Therefore, while permissible because it is not a practice or behavior, potential bias should be avoided by choosing only one respondent per household if possible.

15. Q: What do you do if the project area is large and contains a large number of supervision areas? You want to add another indicator that has another target population, but you do not want to create another questionnaire.

A: Suppose you have a large project area with many supervision areas, for example 11, and want to make decisions about an additional indicator, such as child spacing[5], which has a different target population, but do not want to add another questionnaire. The target population for child spacing is a mother with two children, one of whom is less than 24 months old. One way to solve this problem is to add child spacing questions to the general questionnaire that all mothers can answer, if you can reasonably assume that you will have a total sample size of about 95 when the data from all supervision areas is pooled. You would not get 19 respondents from the child spacing target population for each supervision area, but you estimate that there will be enough respondents across the entire project area to calculate coverage at that level. For example, if there are 11 supervision areas, you might get 9 responses for each supervision area, which results in 99 responses for the entire project area. This number is more than 95, so coverage for the entire project can be calculated (with a confidence interval of at least +/- 10%). However, in a particular supervision area, you may or may not reach a sample size large enough to allow you to judge whether or not performance in that supervision area meets a benchmark or target.

16. What if for monitoring purposes, project managers are not interested in coverage for a specific indicator for the entire project area, but do want to know if the entire project area is meeting the target.

A: If the project area is not large, but you still want to incorporate another indictor with a different target area, you could add it to the general questionnaire. Although the number of respondents for this indicator from each supervision area may be too small to make decisions for each supervision area and the number may also be too small to calculate coverage for the entire project area with a confidence interval of +/- 10% or less, you will probably have at least 19 for the entire project area. Therefore, you would know if the entire project area was or was not meeting a target or benchmark using LQAS decision rules. This would be useful for project management. For example if you added a child spacing indicator, but you only had 5 supervision areas, you might get 9 responses for each supervision area for a total of 45 respondents for the entire project area. This is more than 19, so you could make decisions as to whether or not the entire project area met a target. However, you would not be able to calculate coverage for the entire project area, so this would not be adequate for Rapid CATCH information that must be sent to USAID.

---

[5] Percentage of children age 0-23 months who were born at least 24 months after the previous surviving child