Large Country-Lot Quality Assurance Sampling:

A New Method for Rapid Monitoring and Evaluation of Health, Nutrition and Population Programs at Sub-National Levels

Bethany L. Hedt, Casey Olives, Marcello Pagano, Joseph J. Valadez

May 2008

HNP
HEALTH, NUTRITION, AND POPULATION
THE WORLD BANK

Large Country-Lot Quality Assurance Sampling:

A New Method for
Rapid Monitoring and Evaluation of
Health, Nutrition and Population Programs
at Sub-National Levels

Bethany L. Hedt; PhD, MS
Casey Olives; MS
Marcello Pagano; PhD
Joseph J. Valadez; PhD, SD, MPH

May 2008

Health, Nutrition and Population (HNP) Discussion Paper

This series is produced by the Health, Nutrition, and Population Family (HNP) of the World Bank's Human Development Network. The papers in this series aim to provide a vehicle for publishing preliminary and unpolished results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character. For free copies of papers in this series please contact the individual author(s) whose name appears on the paper.

Enquiries about the series and submissions should be made directly to the Editor, Homira Nassery (HNassery@worldbank.org). Submissions should have been previously reviewed and cleared by the sponsoring department, which will bear the cost of publication. No additional reviews will be undertaken after submission. The sponsoring department and author(s) bear full responsibility for the quality of the technical contents and presentation of material in the series.

Since the material will be published as presented, authors should submit an electronic copy in a predefined format (available at www.worldbank.org/hnppublications on the Guide for Authors page). Drafts that do not meet minimum presentational standards may be returned to authors for more work before being accepted.

For information regarding this and other World Bank publications, please contact the HNP Advisory Services at healthpop@worldbank.org (email), 202-473-2256 (telephone), or 202-522-3234 (fax).

Health, Nutrition and Population (HNP) Discussion Paper

Large Country-Lot Quality Assurance Sampling:
A New Method for Rapid Monitoring and Evaluation of Health, Nutrition and Population Programs at Sub-National Levels

Bethany L. Hedt; PhD, MS[a]   Casey Olives; MS[a]
Marcello Pagano; PhD[a]   Joseph J. Valadez; PhD, SD, MPH[b]

[a] Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA
[b] Malaria Implementation Resource Team, AFTHD, The World Bank, Washington, DC, USA

Abstract: Sampling theory facilitates development of economical, effective and rapid measurement of a population. While national policy makers value survey results measuring indicators representative of a large area (a country, state or province), measurement in smaller areas produces information useful for managers at the local level. It is often not possible to disaggregate a national survey to obtain local information if that was not the intent of the original survey design. Cluster sampling is typically used for national or large area surveys because sampling in clusters lowers the cost of a survey. Lot Quality Assurance Sampling (LQAS) is used to measure results at a local level, since it requires small random samples and produces results useful to local managers. However, current LQAS methodology requires all local areas (strata) be included in the survey in order to be aggregated to produce point estimates for the nation or state. In large countries it is not feasible to sample all strata for logistical and financial reasons. This paper resolves this problem by presenting Large Country-Lot Quality Assurance Sampling (LC-LQAS), a method with two concurrent objectives: (1) provide local managers with accurate local information to enable data driven decisions, and (2) provide central policy makers with the aggregate information they require. These are achieved by integrating cluster sampling with LQAS methodologies. Two examples of the implementation of LC-LQAS are provided, in an HIV/AIDS program in Kenya and a Malaria Booster Project in Nigeria. Classifications of local health units into performance categories and aggregate estimates of coverage, with associated confidence intervals, are provided for select indicators in order to demonstrate its use, analysis, and costs. This paper is written as a manual to support the use of LC-LQAS by others.

Disclaimer: The findings, interpretations and conclusions expressed in the paper are entirely those of the authors, and do not represent the views of the World Bank, its Executive Directors, or the countries they represent.

Correspondence Details: Professor Marcello Pagano, Department of Biostatistics, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA, Telephone: 1-617-432-4911, Pagano@biostat.harvard.edu

Table of Contents

v

ACKNOWLEDGEMENTS

ACRONYMS

AIDS	Acquired Immunodeficiency Syndrome
CA	Catchment Area
CI	Confidence Interval
DE	Design Effect
DHS	Demographic and Health Survey
HIV	Human Immunodeficiency Virus
ICC	Intraclass Correlation
ITN	Insecticide Treated Net
LC-LQAS	Large Country-Lot Quality Assurance Sampling
LGA	Local Government Authority
LQAS	Lot Quality Assurance Sampling
M&E	Monitoring and Evaluation
MAP	Multi-Country AIDS Program
NACC	National AIDS Control Council
NMCP	Nigerian National Malaria Control Program
PPS	Probability Proportionate to Size
SA	Supervision Area
SRS	Simple Random Sampling
WHO	World Health Organization

LARGE COUNTRY-LOT QUALITY ASSURANCE SAMPLING:

A NEW METHOD FOR RAPID MONITORING AND EVALUATION OF HEALTH, NUTRITION AND POPULATION PROGRAMS AT SUB-NATIONAL LEVELS

INTRODUCTION

Effective management of decentralized health systems requires up-to-date information at the local level where programs are implemented. Information at this level allows program managers to know which health systems are meeting particular targets, and which ones are not. To help in acquiring this information we turn to lot quality assurance sampling (LQAS). Interest in applying LQAS to health assessments has been growing since the mid-1980s (Khoromana, Campbell et al. 1986; Lwanga and Abiprojo 1987; Lemeshow and Stroh 1989; Smith Gordon 1989; Wolff and Black 1989). The LQAS method was originally developed in the early part of the 20th century as a quality control technique for goods produced in factories, and it got a great boost in popularity during the second world war when it was used as a method for improving the quality of war materials. But the sampling concepts of the LQAS method have universal applicability, and in 1991, a World Health Organization (WHO) consultation on epidemiological and statistical methods for rapid health assessment concluded that LQAS was one of the more practical methods available, and encouraged its further development to monitor health programs (Anker 1991; Lanata and Black 1991; Lemeshow and Taber 1991). Since that time there have been advances in the LQAS methodology: Statistical reference tables, a text book, and software have been produced that have improved the understanding among public health professionals of what LQAS is, and facilitated its use by eliminating the need for carrying out tedious calculations (Valadez 1991). Also, training manuals have been developed to facilitate usage of LQAS in field settings by public health practitioners (Valadez, Weiss et al. 2003; Valadez, Weiss et al. 2003). This advancement empowers local managers by giving them a rapid and easy to use method for monitoring and evaluating programs, as well as for making program relevant management decisions.

The growing interest in using the LQAS method was captured in a 1997 WHO review of 34 LQAS applications assessing immunization coverage, antenatal care, use of oral rehydration therapy, growth monitoring, family planning, disease incidence, and the technical skills and knowledge of health workers (Robertson, Anker et al. 1997). Subsequently, a 2006 review by the WHO and the World Bank (Robertson and Valadez 2006) included more than 800 LQAS applications, evidence of a huge increase in interest in a short time frame. The LQAS method has also been used to assess the accuracy of health records, outreach of community health workers, and health worker training programs (Valadez 1991; Valadez, Brown et al. 1996; Valadez, Transgrud et al. 1997). In Armenia, Malawi, and Nicaragua, networks of nongovernmental organizations used LQAS to track national disaster relief and reproductive

health programs (Valadez, Leburg et al. 2001; Campos, Valadez et al. 2002; Valadez, Mobley et al. 2003). In Uganda, it has been used to assess the performance of HIV/AIDS control programs at the district and sub-district levels as well as at the national level. There are now several examples applying LQAS for ongoing supervision, and the most well documented example is from a maternal and child health project in rural Nepal (Valadez and Devkota 2002).

In the World Bank, LQAS has been used in support of several health projects including HIV/AIDS projects in Kenya, Uganda, Eritrea, and the Dominican Republic; malaria projects in Nigeria, Ghana and Eritrea, and nutrition projects in the Dominican Republic, Uzbekistan and Ghana. The prime interest LQAS has held for country level managers and those in the Bank is that the resulting information is produced in a timely manner and can be used for results based management.

This paper represents an advance in the development of the LQAS method for carrying out national level assessments of coverage indicators. As described below, LQAS in its most typical current application takes small random samples from all program areas (or strata). As a result when these data are pooled, they form a stratified random sample of the program area. While this application has been feasible for most applications of LQAS, a technical problem has arisen as we enter an era in which countries want to use this methodology on a national level. In some cases, the number of strata can be so large that the costs would be too high, and logistics too complex, to measure them all simultaneously. In such instances it is more sensible to go to scale gradually by introducing LQAS into a small sample of areas, and then build national capacity so that all areas can be eventually covered. However, prior to reaching the ability to have total coverage, it would be propitious to pool the few areas that had had LQAS assessments and still have an accurate point estimate for an entire region or country.

The implication is that, at the initial stages, not all areas within the program catchment area would be represented in LQAS applications; only a subset of the areas would be selected. This being the case, the pooled data would no longer be a stratified random sample, but instead it would more appropriately be called a cluster sample, if the areas had been chosen at random. This paper presents a protocol to integrate LQAS with cluster sampling. It presents a method for combining a small number of LQAS-clusters, drawn as a random sample from all strata in a greater catchment area, and calculating an aggregated result to obtain a point-estimator, with associated confidence interval, of an indicator for the whole catchment area.

General Principles of the LQAS Method

During the 1980s, health system evaluators explored applications of industrial quality control methods to assess health worker performance (Stroh 1983b; Reinke 1988; Valadez, Vargas et al. 1988). LQAS was originally developed in the 1920s to control the quality of industrially produced goods (Dodge and Roming 1959). The principle is that a supervisor inspects a small random sample of a recently manufactured batch or lot of goods from a production unit, such as an assembly line or machine. If the number of defective goods in the sample exceeds a predetermined allowable number, then the batch or lot is rejected; otherwise it is accepted as being of reasonable quality. The number of allowable defective goods is based on a production standard and a predetermined sample size (Dodge and Roming 1959). The sample size is set so that a manager has a high probability of accepting lots in which a predetermined proportion of the goods are of reasonable quality, and a high probability of rejecting lots that fail to reach a production quality standard.

In health systems, an example of a production standard is a predetermined population coverage target for an intervention such as immunization, contraceptive use, or pregnant women requesting to be tested for human immunodeficiency virus (HIV) infection. Health coverage targets are set by health system managers at the national or district level. In health systems, a lot most often consists of a supervision area, e.g., a community or a health facility catchment area. The production unit is the set of health workers working under the supervisor who manages the supervision area. In this setting, there are two primary reasons for using LQAS: first, to determine, within given levels of confidence, whether a specific supervision area has reached a predetermined coverage target, and second, to prioritize allocation of resources based on the outcomes of different supervision areas.

To use LQAS, health system managers need to identify two thresholds. The upper threshold or coverage target (e.g., 80%), which is the proportion of the community that health workers wish to reach during a predetermined period, such as one year. The lower threshold is an unacceptably low level of coverage (e.g., 50%) that should provoke managers to identify the problem causing the failed service delivery and to resolve it with a focused investment of time and resources. A predetermined decision rule is selected so that supervision areas reaching the coverage target have a high probability of being so labeled. The decision rule is also selected so that supervisions areas at or below the lower threshold are detected with low amounts of error. LQAS was developed to make these classifications, and it does it well. However, there can be supervision areas with coverage targets between these two thresholds. In practice, supervision areas with coverage closer to the upper threshold are more likely to be classified as reaching the target, while supervision areas with coverage closer to the lower threshold are more likely to be prioritized as substandard. In either instance, the error is manageable. For example, if a supervision area is slightly below the upper threshold and judged as having reached it, then this error is not pernicious since it is more important for the health system to focus on supervision areas where larger proportions of the population are at risk. Correspondingly, a supervision area that is slightly above the lower threshold and judged as below the coverage target is not worrisome since the supervision area has not

reached the coverage target and would have to be dealt with sooner or later.

Several characteristics have made LQAS attractive to health system evaluators. First, only a small sample is needed to judge whether a supervision area has reached the predetermined coverage target; the further apart the two thresholds are, the smaller the required sample for fixed levels of uncertainty. With a small sample, data collection does not seriously compete with time for provision of health services. Second, the LQAS sampling procedures and analyses are relatively simple and as a result it is easy to build up local capacity. This simplicity is also welcome to overworked supervisors and health workers, who need management tools that can easily be understood and applied. These two characteristics have made LQAS valuable as a practical management tool for monitoring and evaluation of community health services, both for use in a single application and as an ongoing tool for monitoring and surveillance.

Another attractive feature of LQAS designs is that the data from individual supervision areas or strata can be pooled into an estimate of coverage for an entire program area. In typical applications, when all strata are pooled the result from each supervision area is weighted by the size of its population (Valadez, Weiss et al. 2003). To date, the data aggregation across multiple supervision areas assumes that data has been collected in all supervision areas. The following sections show how LQAS can now be used when only a sampling of, rather than all, supervision areas are included in an assessment.

Background for Developing a Large Country-LQAS

In 2001, Kenya National AIDS Control Council (NACC) received $50 million from the World Bank to fund HIV prevention, care and support programs through the Multi-Country AIDS Program (MAP). Over half of the money was channeled to the Provincial level, additional amounts went to the District level, and then finally to the Constituency level, the next administrative unit after the District. Most NACC community development programs are implemented and managed at the Constituency level.

As NACC completed its funding cycle, it faced the challenge of assessing its achievements. NACC used this opportunity to measure the current status of key HIV/AIDS indicators — especially at the Constituency level. While the DHS and other Behavioral Surveillance Surveys have been used to measure key indicators at the Provincial level or national level, they are not intended to be used at decentralized levels, and are too costly to implement frequently. For this reason, NACC selected LQAS as the preferred M&E method, as it was developed for decentralized monitoring of programs. Additionally, using LQAS met another priority for NACC by engaging program managers within Constituencies, rather than delegating management tasks to outsiders.

In theory LQAS could be used by management teams in each Constituency. However, as Kenya has more than 200 Constituencies, it was not feasible to implement LQAS simultaneously in all Constituencies at this initial stage, as there would be too many data collection teams to train (one team for each Constituency) and supervise. Also, the costs of training, data collection and analysis could be too high. To carry out the decentralized monitoring, while addressing the logistical constraint of implementing LQAS in 200 constituencies, we adapted Lot Quality Assurance Sampling (LQAS) to evaluate performance of a sample of Constituencies while at the same time aggregating the Constituencies' data using cluster sampling theory to estimate overall Provincial effects. The remainder of this paper summarizes the following key tasks carried out to develop this new procedure, henceforth called Large Country-Lot Quality Assurance Sampling (LC-LQAS):

1. The theory: Integrate cluster sample theory with LQAS to establish a feasible Monitoring and Evaluation (M&E) system.

2. The tool: Develop and apply LC-LQAS as a program M&E tool.

3. Next steps: Delineate additional developments needed for LC-LQAS.

While this report details the process designed for the NACC in Kenya, it is written so that the procedure can be applied in other settings. In a later section, we discuss an application in Nigeria as a second example.

Integrating Cluster Sample Theory with LQAS

Both cluster sampling and LQAS are useful tools for M&E. In this section we first describe the standard uses and limitations of each procedure and then discuss the usefulness of combining the two methodologies in order to meet the objectives of NACC.

Cluster Sampling to Obtain Provincial Estimates of Indicators

Cluster sampling is a technique for obtaining estimates when simple random sampling is neither feasible nor the preferred approach, due to the size of the area or for financial reasons. With single-stage cluster sampling, an area is divided into non-overlapping units, called clusters. A limited number of clusters are randomly sampled for inclusion into the study, and then all units within the selected clusters are sampled. Two-stage cluster sampling uses the same techniques as single-stage cluster sampling for selecting the clusters, but subsequently, two-stage cluster sampling uses simple random sampling to sample subjects within the cluster, instead of including all of the units in the cluster.

Cluster sampling is less expensive to implement than simple random sampling because it geographically restricts the sampling area, and thus it reduces the time and costs of traveling when compared to the resources required to travel throughout an entire region. If it is expensive or difficult to obtain a list of all units eligible for sampling in the entire area (i.e., to develop a sampling frame), cluster sampling is preferable because it only requires sampling frames for the clusters that are sampled.

Conventional cluster sampling generally focuses on obtaining an estimate for a large area. In practice, no inferences are typically made about individual cluster estimates or about a program's performance within the cluster. Further, the design may not guarantee a sufficient sample size at the cluster level to guarantee reliable cluster level decisions and therefore would not be appropriate for evaluation of SA performance. Another drawback of cluster sampling is that the sampling design usually inflates the variance of an estimate when compared to simple random sampling of the same size. Thus cluster sampling requires a larger sample than simple random sampling to obtain estimates with the same precision.

Lot Quality Assurance Sampling in Large Countries

In monitoring and evaluation of public health programmes, as in the industrial setting, LQAS is used to provide a classification of the programme, based on a binary indicator, into an "acceptable" or "unacceptable" class. This is its primary function. It is not intended for accurately measuring a point estimate of the prevalence of indicators. However, this does not preclude the use of the gathered information to provide a point estimate for the prevalence of these indicators. If some indicators are properly aggregated over a number of subregions,

the resultant estimator could well be based on sufficiently large samples to provide a very accurate regional estimate.

One reason LQAS is less expensive than many other study designs is it uses local people as samplers and interviewers — people who are often already employed by the local program and thus require no additional salary.

Combining LQAS with Cluster Sample Theory

Due to the decentralized nature of program implementation and management, the NACC wanted an M&E system to assess relevant indicators at both the Constituency and Provincial levels. Because of its prior success as a community level M&E tool in Uganda (Mukaire, Kaweesa et al. 2004; Valadez and Nsubuga 2004), NACC decided to use LQAS to evaluate Constituency efforts. If all Constituencies implemented LQAS, then the data from the Constituencies could then be aggregated to obtain Provincial estimates using stratified sampling methodology (Hoshaw-Woodard 2001). However, it was not practical to obtain measures in each of the more than 200 Constituencies in Kenya due to budgetary limits, limited human resources and time, especially since there was no local M&E system to build on.

However, NACC could, with very limited resources, obtain Provincial estimates while evaluating the program at the Constituency level by integrating LQAS and cluster sampling theory. In the process of developing these methods for NACC, we:

1. combine LQAS and cluster sampling theory with specified constraints to develop formulae for calculating Provincial point estimates, 95% confidence intervals and sample sizes;

2. use the new formulae to identify the minimum number of Constituencies to sample in order to establish a Constituency level M&E system, and still result in an accurate Provincial measure of the indicators; and

3. implement LC-LQAS in order to judge performance at the Constituency level and calculate Provincial estimates with 95% confidence intervals.

The following section discusses each of these issues in more detail. In order to synthesize terminology of the two methodologies, clusters are henceforth referred to as supervision areas (SA) and the larger area, or collection of supervision areas, is referred to as the catchment area (CA). These terms are used in typical applications of LQAS, and use of these terms reinforces the idea that with LC-LQAS, we are interested in both the analyses of each cluster as well as with the aggregated Provincial estimates.

7

Development and Application of LC-LQAS as a Program M&E Tool

Constraints

The LC-LQAS procedure is shaped by four methodological decisions and constraints imposed by NACC. These constraints are not unique to the Kenyan NACC and are experienced by many programs globally. Therefore, the technical solutions provided here have general application.

1. **Define Constituencies as the SA** — NACC wanted managers to assess the impact of programs via LQAS at the level of implementation. The Constituency is therefore the supervision area (SA) for LC-LQAS.

2. **Restrict the length of the 95% confidence intervals for Provincial indicators** — When applying LQAS in a public health setting, the conclusion that an area is performing "acceptably" or "unacceptably" on a key indicator is based on the behavior of the subjects in the SA. Aggregating the data across supervision areas in a Province provides a regional estimate of the coverage proportion, the proportion of the population that embraces the behavior of interest measured by the indicator. The formulae below ultimately provide point estimates and associated variances for these coverage proportions. NACC requested that the SA sample size formula be developed with the goal of restricting the length of the 95% confidence interval for Provincial coverage proportions to ±10%. In order to allow flexibility, we discuss the sample size formulae with respect to any maximum length, ℓ_{max}. In the case of Kenya, we set $\ell_{max} = 0.20$.

3. **Fix the within Constituency sample size for each Province** — In order to immediately use previously developed LQAS training materials and to simplify training, NACC requested that the LQAS sample size in all Constituencies remain constant in a Province. As a result, one set of LQAS decision rule tables can be used everywhere in a Province, thereby simplifying the training.

4. **Select Constituencies using SRS without replacement** — We recommended NACC to select Constituencies for inclusion using simple random sampling (SRS) without replacement rather than probability proportional to size sampling because this guaranteed that any given Constituency would only be sampled once. This choice also maximizes the number of Constituencies in the sample and again keeps the recommended sample size for each Constituency constant within a Province. Further, this allows all Constituencies, independent of size, an equal opportunity for local level program evaluation.

Principle 1: Use Simple Random Sampling to select the Supervision Areas rather than Probability Proportional to Size sampling. This will guarantee that a given SA will be sampled at most once and assigns all SA equal probability of inclusion, independent of their size.

LC-LQAS Sample Size Formulae

Accommodating the above constraints, the number of SAs included in the sample, were determined by the LC-LQAS sample size formula (derived in Appendix B). The number of SAs to be sampled, n, is dependent on six parameters: (1) the number of samples collected in each SA, m; (2) the total number of SAs in a catchment area, N; (3) the total population in the catchment area (usually based on a national census), N_{cen}^* ; (4) the average of the square of the populations in each SAs, $\overline{M^2}$; (5) an estimate of the intraclass correlation, $\hat{\rho}$; and (6) the maximum desired length for the confidence interval, ℓ_{max}, which in this case, has a value of 0.2.

$$ n = N(1 + (m-1)\widehat{\rho}) \left\{ \left[\left(\frac{\ell_{max} N_{cen}^*}{1.96} \right)^2 \left(\frac{(m-1)(1-\widehat{\rho})}{N\overline{M^2}} \right) + m\widehat{\rho} \right] \right\}^{-1} $$

Three of the parameters, namely, the number of SAs, the total population size, and the average of the square of the populations in each SA, are obtained directly from the census of the population. The size of the sample collected in each SA is determined by the minimum sample required to apply LQAS decision rules with acceptable error. It is typically set to either 19 or 20 for two reasons: (1) previously developed and field tested training materials can be used immediately (Valadez, Weiss et al. 2003) to carry out an LQAS analysis of each SA, and (2) these sample sizes have been used successfully in many applications globally. Thus, an estimate of the intraclass correlation, $\hat{\rho}$, is the only unknown quantity and we discuss methods of obtaining this estimate in the following section.

In practice, the form presented in Appendix C facilitates using the above formula for calculating the number of supervision areas required for sampling. We show how to use this form with an example from Kenya later in this section.

Principle 2: Use the LC-LQAS sample size formula to calculate the minimum number of Supervision Areas / Districts to sample from the universe of all SA / Districts.

Estimating the Intraclass Correlation Coefficient

An estimate of the intraclass correlation coefficient (ICC), $\hat{\rho}$, is required to use the above formula to calculate the number of SAs to include in the sampling frame. Overestimating the intraclass correlation leads to larger samples than necessary to meet the imposed constraints, and unnecessarily inflates the costs of the survey. Using too small a value of $\hat{\rho}$ results in failing to constrain the length of the 95% confidence interval to be 20%.

One possibility is to use intraclass correlation coefficient estimates from identically designed surveys looking at similar indicators, such as an LC-LQAS survey previously implemented in the same area. When these estimates are not available, one can turn to ICC estimates based on design effects from other multistage surveys in the area. Fenn and her team describe methods for estimating the intraclass correlation from the Demographic Health Surveys (DHS) (Fenn, Morrisa et al. 2004). The DHS is generally a stratified two-stage cluster survey, that collects data on numerous health indicators in many countries. The intraclass correlation is related to the design effect of this survey, using the relationship $DE = 1 + (\overline{m}_{DHS} - 1)\hat{\rho}$, where $\overline{m}_{DHS}$ is the average cluster sample size for the DHS. By solving for ICC, or $\hat{\rho}$, the following equation results and can be used for estimation, $\hat{\rho} = (DE - 1)/(\overline{m}_{DHS} - 1)$.

This DHS estimate may result in multiple recommendations for the ICC — one for each indicator, region and stratum. The median of these ICC serves as a first recommendation for the sample size formula. The 75th percentile would be preferable if more precision is required and the necessary resources are available. However, if there are not adequate resources for the survey then the 25th percentile is recommended. Nevertheless, the median should be used in most circumstance, as both precision and cost are both important issues.

If the DHS is not available for a particular country, then these same methods can be applied to other surveys to estimate ICC. Once LC-LQAS has been implemented in a region, then it is advantageous to use the intraclass correlation estimates from this survey directly to determine the new sampling frame, as we discuss below.

Principle 3: Use existing surveys to obtain an estimate of the intraclass correlation coefficient. Identify the design effect for key indicators/regions and calculate the corresponding ICC using the relationship $\hat{\rho} = (DE - 1)/(\overline{m} - 1)$. Organize the resulting ICCs into quartiles. Use the median value if both precision and cost are priority design issues. If greater precision is required and funds are available, then use a higher estimate of ICC such as by the 75th percentile. If fewer funds are available, use the 25th percentile.

Case Example 1.1: Determining a Sampling Frame for Nyanza Province

In 2004, the Kenyan NACC used the LC-LQAS methodology to monitor programs in Nyanza Province and Western Province. The process to determine the sampling frame for Nyanza Province, Kenya, is outlined here. The same process was also used in Western Province. NACC set the constituency sample size in Nyanza to 19. A full discussion of setting sample sizes within an SA (or lot) can be found in several publications and is not repeated here (see: Khoromana, Campbell et al. 1986; Lwanga and Abiprojo 1987; Smith Gordon 1989; Wolff and Black 1989; Lemeshow and Taber 1991; Valadez 1991; Valadez, Weiss et al. 2003). The number of supervision areas to sample was dependent on the following information for Nyanza:

- There were 32 constituencies in Nyanza, all of which have NACC funded programs ($N = 32$);

- The total population in the Province was 4,392,196 people ($N^*_{cen} = 4,392,196$); and

- The average square population size of a Constituency was 19,512,141,396 people ($\overline{M^2} = 19,512,141,396$).

We used the 2003 Kenyan Demographic and Health Surveys (DHS) to estimate the intraclass correlation. Four indicators in the 2003 Kenya DHS were relevant to the LC-LQAS survey planned for Kenya during 2004 as they relate to reproductive health or poverty. In addition to reporting national values, DHS reported a design effect for each of the eight provinces in Kenya. Table 1 presents national design effects and the design effects for Nyanza and Western Province for these four indicators by gender (CBS, 2003). Table 1 also presents the associated ICCs calculated using the previously presented formula.

Table 1: Design Effects for Kenya, Nationally and for Nyanza and Western Provinces, for Four Health Indicators ($\overline{m}_{DHS,women}$=20.5, $\overline{m}_{DHS,men}$=8.9)

Indicator		National		Nyanza		Western	
		Women	Men	Women	Men	Women	Men
Knows at least one contraceptive method	DE	5.24	1.01	0.64	N/A	2.54	N/A
	ICC	0.217	0.001	-0.018		0.079	
Knows at least one modern contraceptive method	DE	5.56	1.39	0.64	2.05	2.54	N/A
	ICC	0.234	0.049	-0.018	0.133	0.079	
Has no formal education	DE	5.85	2.81	2.57	2.01	2.69	1.36
	ICC	0.249	0.229	0.081	0.128	0.087	0.046
Has secondary education or higher	DE	3.78	2.37	3.13	2.88	2.82	2.16
	ICC	0.143	0.173	0.109	0.238	0.093	0.147

The relationship between the design effect and ICC, $\hat{\rho} = (DE - 1)/(\overline{m}_{DHS} - 1)$, is dependent on the average cluster size. The 2003 Kenya DHS Survey included 8,195 women and 3,578 men from 400 clusters, corresponding to average sample sizes of 20.5 and 8.9 for women and men respectively. The median ICC estimate is 0.087 (25th percentile: 0.079; 75th percentile: 0.128) for the two provinces. We used the 50th percentile of the intraclass correlation coefficients from the Nyanza Province and Western Province since both accuracy and cost were priorities. Therefore, the ICC estimate used for the sample size calculations was 0.087. See the following Case Example 1.1 to review how this ICC estimate was used to determine that 12 constituencies was the total number of supervision areas to include in the sample for Nyanza Province. Ultimately, the NACC decided to include 16 SA in the sample. They decided on this number thinking that LC-LQAS could be scaled up faster in Kenya if more staff were trained in the method. Later, NACC realized that doubling the sample size increased the logistical complexity enormously, and that it would have been better to have embraced the recommended sample size of 12.

Case Example 1.1: Calculating the Number of Supervision Areas to Sample in Nyanza Province

Total number of Supervision Areas	N	**32**
Sample Size per SA	m	19
Total Population	N_{cen}^*	^{A} **4,392,196**
Average Square SA Population	$\overline{M^2}$	^{B} **19,512,141,396**
Estimate of Design Effect	DE	*varies*
Estimate of Intraclass Correlation	$\widehat{\rho} = (DE - 1)/(m^\dagger - 1)$	0.087

12

Supervision Area	Population M_i	Population Squared M_i^2
Homa Bay, CACC 1	144,270	20,813,832,900
Homa Bay, CACC 2	144,270	20,813,832,900
Kisii Central, CACC 1	122,946	15,115,718,916
Kisii Central, CACC 2	122,946	15,115,718,916
Kisii Central, CACC 3	122,947	15,115,964,809
Kisii Central, CACC 4	122,947	15,115,964,809
Kisumu, CACC 1	168,119	28,263,998,161
Kisumu, CACC 2	168,120	28,264,334,400
Kisumu, CACC 3	168,120	28,263,334,400
Kuria, CACC 1	151,887	23,069,660,769
Migoria, CACC 1	128,725	16,570,125,625
Migoria, CACC 2	128,724	16,569,868,176
Migoria, CACC 3	128,724	16,569,868,176
Migoria, CACC 4	128,724	16,569,868,176
N. Kisii, CACC 1	166,034	27,567,289,156
N. Kisii, CACC 2	166,034	27,567,289,156
N. Kisii, CACC 3	166,034	27,567,289,156
Rachuonyo, CACC 1	153,563	23,581,594,969
Rachuonyo, CACC 2	153,563	23,581,594,969
Siaya, CACC 1	160,062	25,619,843,844
Siaya, CACC 2	160,061	25,619,523,721
Siaya, CACC 3	160,061	25,619,523,721
Suba, CACC 1	77,833	6,057,975,889
Suba, CACC 2	77,833	6,057,975,889
Bondo, CACC 1	119,390	14,253,872,100
Bondo, CACC 2	119,390	14,253,872,100
Nyando, CACC 1	99,977	9,995,400,529
Nyando, CACC 2	99,977	9,995,400,529
Nyando, CACC 3	99,976	9,995,200,576
Gucha, CACC 1	153,647	23,607,400,609
Gucha, CACC 2	153,646	23,607,093,316
Gucha, CACC 3	153,646	23,607,093,316
	Total Population $N^* = \sum_{i=1}^{N} M_i$ $=^A$ 4,392,196	Average Population Squared $\overline{M^2} = \frac{1}{N} \sum_{i=1}^{N} M_i^2$ $=^B$ 19,512,141,396

$$n = N(1 + (m-1)\widehat{\rho}) \left\{ \left[\left(\frac{\ell_{max} N^*_{cen}}{1.96} \right)^2 \left(\frac{(m-1)(1-\widehat{\rho})}{N \overline{M^2}} \right) + m\widehat{\rho} \right] \right\}^{-1}$$

$$= 32(1 + (19-1)0.087) \left[\left(\frac{0.2(4,392,196)}{1.96} \right)^2 \left(\frac{(19-1)(1-0.087)}{32(19,512,141,396)} \right) + 19(0.087) \right]^{-1}$$

$$= 11.83$$

Number of Supervision Areas to Sample = _12_ (smallest integer bigger than n).

Selecting the Final Sample of Supervision Areas and Individuals

Three steps are used to sample the SAs from the universe of all SAs in Nyanza Province and individuals within each selected SA. First, Simple Random Sampling (SRS) is used to sample supervision areas, which in the case of Nyanza Province is 16 SAs. The most important component of the sampling procedure is that SAs be chosen randomly without replacement and without being influenced by the demographics or characteristics of the SA (such as accessibility). Selecting the SAs in a nonrandom way could lead to a biased estimate of the coverage proportion. One simple way to implement SRS for SA selection in a field setting, is to list each name of the N supervision areas on a sheet of paper, and randomly draw n of them.

Once the supervision areas are selected, LQAS requires a random sample of size m within the SA. This is achieved in two steps. First, the m individuals to be sampled are divided between the villages using probability proportionate to size. Then the number assigned to a particular village are sampled using SRS. This two step process is described in more technical detail elsewhere (Module Three of Assessing Community Health Programs (Valadez, Weiss et al. 2003)).

Principle 4: Selecting the final sampling frame requires three steps: 1) select n of the N SAs using SRS, 2) select villages within the SAs using probability proportionate to size, and 3) randomly select individuals in the villages.

Analysis of LC-LQAS Data

The next critical steps for implementing LC-LQAS take place after data collection and concern the analysis of performance of supervision areas, and the estimation of catchment area coverage proportions with 95% confidence intervals. Module Six in <u>Assessing Community Health Programs</u> (Valadez, Weiss et al. 2003) describes the procedures for evaluating an SA using LQAS. LC-LQAS uses the same system at the supervision area level. Because the LQAS analysis is well covered in the literature, no further discussion of this method is presented here. Rather, discussion focuses exclusively on the new components related to LC-LQAS.

> **Principle 5:** Use the standard LQAS analysis methods to classify SAs as having reached a standard or target.

With LC-LQAS, we may not treat the data as a stratified random sample (which is what we would do in a typical LQAS design which pools the SA data). Instead, we must account for the multistage cluster design in the analysis. We derive the estimators for this analysis in Appendix A. It is important to note that the N_{sam}^* used in this estimator is the total population for the catchment area as estimated from the survey data by multiplying the population sizes in the surveyed supervision areas by the inverse of the fraction of supervision areas sampled ($N_{sam}^* = (N/n) \sum_{i=1}^{n} M_i$). In Appendix D, we present the form we use to organize the information needed to calculate the indicator point estimates by hand, and provide an illustration of this tool in Case Example 1.2.

While it is possible to calculate the coverage proportion with corresponding 95% confidence interval for the catchment area by hand, this task is cumbersome. To make the calculations more manageable, we advise the use of a statistical package to assist in these calculations. Most statistical packages, including Stata, SAS and SPSS, offer the necessary tools to calculate these estimates incorporating the two-stage cluster sampling design (either as an add-on or as a standard component of the software package). Similarly, we can use a spreadsheet to facilitate the computation of point estimates, variance, confidence intervals and intraclass correlation for multiple indicators.

> **Principle 6:** Use the LC-LQAS formulae to pool the SA data to calculate a point estimate for an indicator and corresponding 95% Confidence Interval. The point estimate is calculated with the following formula:
> $$\frac{N}{n(N_{sam}^*)} \sum_{i=1}^{n} M_i \widehat{p}_i.$$

Case Example 1.2: Analysis for Male Respondents on the Indicator "Know ways to prevent sexual transmission of HIV infection" in Nyanza Province, Kenya

For the purpose of this example, we restrict our discussion to one indicator for men in Nyanza province: "Knows ways to prevent sexual transmission of HIV infection". Suppose that the target coverage for this indicator is that 80% of men should know ways to prevent sexual transmission of HIV. Given the in-cluster sample size of 19, this corresponds to a decision rule of 13 (Module One of Assessing Community Health Programs (Valadez, Weiss et al. 2003)). In other words, if 13 or more of the 19 sampled men in a supervision area can correctly name the ways to prevent sexual transmission of HIV, then we classify that SA as having reached the 80% coverage target. If less than 13 men are able to name the ways to prevent sexual transmission of HIV, then we classify the SA as *not* meeting the 80% performance target.

We report the constituency level results in Table 2. Using the decision rule of 13 correct responses, we classify 11 of the 16 areas as having reached the 80% coverage target. However, five constituencies are performing below standard; namely, Rarieda, Kasipul Kabondo, Kisumu Town East, Kisumu Town West, and Nyatike. These areas should be specifically targeted for future adult education activities and their current education programs need to be examined and possibly redesigned.

Table 2: Summary of Results of Supervision Area for Men in Nyanza Province: "Knows ways to prevent sexual transmission of HIV infection"

Supervision Area	Number of Correct Responses	Supervision Area	Number of Correct Responses
West Mugirango	15	**Kisumu Town East**	**11**
North Mugirango	18	**Kisumu Town West**	**11**
Bondo	19	Kitutu Chache	17
Rarieda	**12**	Nyaribari Chache	16
Gem	13	Ndhiwa	19
Ugenya	13	Rangwe	18
Gwasi	16	Nyando	17
Kasipul Kabondo	**8**	**Nyatike**	**9**

We use the forms in Appendix D to calculate the coverage proportion, corresponding 95% Confidence Interval, and Intraclass Correlation for men in Nyanza Province for this indicator and display it in the Case Example 1.2 text box which follows. The analysis, done by hand here, can be repeated in Stata using the commands summarized in Appendix E. We estimate

that 73.5% of men in Nyanza province have heard of voluntary testing and counseling, with corresponding 95% Confidence Interval of (65.5%, 81.4%). The intraclass correlation estimate for this indicator in Nyanza province is 0.151.

Catchment Area Coverage Proportions

INDICATOR: "Know ways to prevent sexual transmission of HIV infection"

Sampled Supervision Areas	Population M_i	Number with Positive Result t_i	SA Proportion $\hat{p}_i = \frac{t_i}{19}$	Population * Proportion $M_i(\hat{p}_i)$
West Mugirango	99,910	15	0.789474	78,876
North Mugirango	92,149	18	0.947368	87,299
Bondo	54,687	19	1.000000	54,687
Rarieda	74,856	12	0.631579	47,277
Gem	77,363	13	0.684211	52,933
Ugenya	42,506	13	0.684211	29,083
Gwasi	32,955	16	0.842105	27,752
Kasipul Kabondo	120,970	8	0.421053	50,935
Kisumu Town East	85,115	11	0.578947	49,277
Kisumu Town West	101,778	11	0.578947	58,924
Kitutu Chache	43,914	17	0.894737	39,291
Nyaribari Chache	43,206	16	0.842105	36,384
Ndhiwa	56,781	19	1.000000	56,781
Rangwe	55,916	18	0.947368	52,973
Nyando	44,632	17	0.894737	39,934
Nyatike	31,524	9	0.473684	14,932
	Total $= \sum_{i=1}^{n} M_i$ 1,058,262			Total $= \sum_{i=1}^{n} M_i(\hat{p}_i)$ $=^C$ 777,338

Total number of Supervision Areas	N	32
Number of sampled SAs	n	16
Total Population	$N^*_{sam} = \frac{N}{n} \sum_{i=1}^{n} M_i$	2,116,524
Sum of SA pop times SA proportion	$\sum_{i=1}^{n} M_i(\hat{p}_i)$	^{C} 777,338

$$\hat{P} = \frac{N}{n(N^*_{sam})} \sum_{i=1}^{n} M_i(\hat{p}_i)$$

$$= \frac{32}{16(2,116,524)}777,338$$

$$= 0.735$$

Estimation of Confidence Interval

INDICATOR: "Know ways to prevent sexual transmission of HIV infection"

Sampled Supervision Area	Population M_i	Square of Population M_i^2	SA Proportion $\hat{p}_i$	SA Sample Variation $s_{p_i}^2 = \frac{\hat{p}_i(1-\hat{p}_i)}{18}$	Pop Squared * Variation $M_i^2(s_{p_i}^2)$	Square Error $(\hat{p}_i - \hat{P})^2$
West Mugirango	99,910	9,982,008,100	0.789474	0.009234	92,169,973	0.003017
North Mugirango	92,149	8,491,438,201	0.947368	0.002770	23,521,989	0.045295
Bondo	54,687	2,990,667,969	1.000000	0.000000	0	0.070467
Rarieda	74,856	5,603,420,736	0.631579	0.012927	72,435,725	0.010602
Gem	77,363	5,985,033,769	0.684211	0.012004	71,842,511	0.002533
Ugenya	42,506	1,806,760,036	0.684211	0.012004	21,687,794	0.002533
Gwasi	32,955	1,086,032,025	0.842105	0.007387	8,022,397	0.011570
Kasipul Kabondo	120,970	14,633,740,900	0.421053	0.013543	198,179,317	0.098276
Kisumu Town East	85,115	7,244,563,225	0.578947	0.013543	98,110,428	0.024210
Kisumu Town West	101,778	10,358,761,284	0.578947	0.013543	140,284,856	0.024210
Kitutu Chache	43,914	1,928,439,396	0.894737	0.005232	10,090,326	0.025662
Nyaribari Chache	43,206	1,866,758,436	0.842105	0.007387	13,789,536	0.011570
Ndhiwa	56,781	3,224,081,961	1.000000	0.000000	0	0.070467
Rangwe	55,916	3,126,599,056	0.947368	0.002770	8,660,939	0.045295
Nyando	44,632	1,992,015,424	0.894737	0.005232	10,422,980	0.025662
Nyatike	31,524	993,762,576	0.473684	0.013850	13,764,025	0.068047
		Total = $\sum_{i=1}^{n} M_i^2$ $=^D$ 81,314,083,094		SSE = $\frac{m}{n} \sum_{i=1}^{n} s_{p_i}^2$ ^{E} 0.156067	Total = $\sum_{i=1}^{n} M_i^2(s_{p_i}^2)$ $=^F$ 782,982,796	Between SA Variance $s_B^2 = \frac{\sum_{i=1}^{n}(\hat{p}_i - \hat{P})^2}{n-1}$ $=^G$ 0.035961

Total number of Supervision Areas	N	32
Number of sampled SAs	n	16
Total Population	N^*_{sam}	2,116,524
Total of Squared Supervision Area Population	$\sum_{i=1}^{n} M_i^2$	^{D} 81,314,083,094
Sum of SA pop square times SA variance	$\sum_{i=1}^{n} M_i^2(s_{p_i}^2)$	^{F} 782,982,796
Between SA variance	s_B^2	^{G} 0.035961
MSE		^{E} 0.156067
MSC	$m(s_B^2)$	^{H} 0.683261

Variance

$$\widehat{var}(\hat{P}) = \frac{1}{(N^*_{sam})^2}\left[\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 s_B^2 + \frac{N}{n}\sum_{i=1}^{n}M_i^2 s_{p_i}^2\right]$$

$$= \frac{1}{(2,116,524)^2}\left[\frac{(32)^2}{(16)^2}(1-\frac{16}{32})81,314,083,094(0.035961) + \frac{32}{16}782,982,796\right]$$

$$= 0.001655084$$

95% Confidence Interval

The 95% Confidence Interval is formed by:

$$\left[\hat{P} - 1.96\sqrt{\widehat{var}\hat{P}}, \hat{P} + 1.96\sqrt{\widehat{var}\hat{P}}\right]$$

$$\left[0.735 - 1.96\sqrt{0.001655084}, 0.735 + 1.96\sqrt{0.001655084}\right]$$

$$\left[0.735 - 1.96(0.0214), 0.735 + 1.96(0.0214)\right]$$

$$\left[0.655, 0.814\right]$$

Estimated Intraclass Correlation

The intraclass correlation is estimated by:

$$\widehat{\rho} = \frac{msc - mse}{msc + (m-1)mse}$$

$$= \frac{0.683261 - 0.156067}{0.683261 + (19-1)0.156067}$$

$$= 0.151$$

Improving the Estimate of ICCs

For the first implementation of LC-LQAS in a country, the intraclass correlation estimate is based on previous sources of information. In the case of Kenya, the ICC estimate came from results of the 2003 Kenya DHS. However, once the LC-LQAS data have been collected, the ICC estimate can then be recalculated yielding a more tailored recommendation for the number of SA to be sampled in the future in this locale and setting. The formula and an explanation presented in Appendix D is used to calculate the intraclass correlation directly, namely, $\widehat{\rho} = (msc - mse)/(msc + (m-1)mse)$. Alternatively, the ICC can be indirectly estimated by using the relationship between design effect and ICC, namely, $\widehat{\rho} = (DE-1)/(m-1)$. If the survey is complex then using a statistical package to estimate design effect (for example, Stata programming code is shown in Appendix E) and calculate ICC indirectly may be a more practical solution.

The strategy used in Kenya for obtaining an observed value of ICC was to implement LC-LQAS in two catchment areas with extremely different conditions before using it throughout the entire country. One CA had little expected variation between SAs and the other had large expected variation. By choosing these two extreme cases, the average variation can then be used to estimate ICC for use across all SA. NACC selected Nyanza and Western as the initial Provinces. Nyanza had high HIV prevalence of about 15.1% and prevention programs which had been ongoing for several years. Western, on the other hand, had HIV prevalence of about 4.9% and newer prevention programs. Nyanza was the province where we expected high variation due to the maturity of the HIV/AIDS programs. In other words, we expected some constituencies to be more successful than others in carrying out the planned work. Similarly, we expected Western Province, which has nascent programs, to have homogeneous constituencies. Post implementation of the first LC-LQAS in Kenya, we updated the ICC estimates, using the data from both Provinces, to provide more appropriate recommendations for sample sizes for future surveys.

Table 3: ICC Estimates for Select Indicators and Subpopulations, Kenya		
Indicator	Nyanza ICC	Western ICC
Knows ways to prevent sexual transmission (Men)	0.151	0.027
Knows ways to prevent sexual transmission (Women)	0.101	0.202
Knows ways to prevent sexual transmission (Mothers)	0.202	0.019
Knows ways to prevent sexual transmission (Youth)	0.182	0.012
Knows HIV can be transmitted from mother to child (Men)	0.023	0.008
Knows HIV can be transmitted from mother to child (Mothers)	0.003	0.015
Median ICC	**0.126**	**0.017**

Table 3 shows the estimated intraclass correlation for six key indicators in Nyanza and Western Province using the LC-LQAS survey data. The 25th, 50th and 75th percentiles of the ICCs for these indicators, when combining both Provinces, are 0.014, 0.025, and 0.159, respectively. The choice of which of these values to use in future sample size estimates continues to be a trade off between accuracy and cost. Using the lower quartile will reduce the required sample, but at the risk of accuracy being lower than desired, resulting in wider confidence intervals than originally intended. The upper quartile will improve the accuracy of the estimates, but with the disadvantage of increased cost. Nevertheless, the observed ICC estimate improves the sample size recommendations for the remaining provinces in Kenya, since it is based on this specific survey and indicators. We continue to use the median value of these ICCs since it was used in the initial sample size calculation. When using the median ICC (0.025), the recommended number of SAs for sampling in Nyanza then drops from 12 to 8. This reduction represents nearly a 33% savings in subsequent applications of LC-LQAS in Kenya. One expects that this reduction in constituency sample size would translate into a corresponding reduction in the budget needed to implement LC-LQAS subsequently, because of the corresponding reduction in the number of data collectors, supervisors, vehicles, petrol, questionnaires, data entry and the like. Note, however, in some cases, more supervision areas may be recommended for sampling, suggesting that a larger sample is required to meet the imposed constraints.

In reality, the process for calculating the new value of ICC for Kenya is slightly more complex. All ICC estimates for each indicator and region in the assessment must be considered, which can be computationally intensive if calculating by hand. If using a statistical package to support data analysis, then it is less intensive to have the package automatically estimate

the design effect, and then indirectly calculate the intraclass correlation estimate using the relationship $\widehat{\rho} = (DE - 1)/(m - 1)$.

Principle 7: Once the data are collected, pool the design effects for key indicators, and take the median value to use in subsequent surveys. If greater precision is required, then use the 75% percentile value.

Replicating LC-LQAS in Other Countries

While this report specifically describes the development of LC-LQAS for Kenya, all of the above procedures can be adapted to other countries. LC-LQAS is particularly useful for countries that have a large number of supervision areas and/or limited resources for M&E. By using the forms in the Appendices and in this report, as well as the LQAS Training Manuals (Valadez, Weiss et al. 2003), other countries can:

- Estimate ICC using the median design effects of several relevant indicators from other surveys, and use this estimate to determine the minimum number of supervision areas to sample,

- Sample supervision areas using simple random sampling without replacement, and sample subjects in them using a randomizing process,

- Collect the data,

- Analyze the data at the supervision area level using LQAS principles, and pool the supervision areas sampled in the catchment area using the LC-LQAS formulae presented here, and

- Calculate an observed value of ICC that can be used in future application of LC-LQAS in the country.

By following these steps, countries can monitor and evaluate indicators using population based sampling at community and provincial levels. The results provide efficient information, leading to better program management and implementation that is focused directly on the management units responsible for program implementation at the local level.

Case Example 2.1: An Application of LC-LQAS Methodology in the National Malaria Control Project in Nigeria

We adapted the LC-LQAS protocol developed for Kenya to aid the Nigerian National Malaria Control Program (NMCP) in establishing a decentralized M&E system in seven Nigerian states (Kano, Jagawa, Bauchi, Gombe, Rivers, Akwa Ibom, and Anambra). Each state represents a distinct catchment area for the World Bank funded malaria project. States are subdivided into administrative areas called Local Government Authorities (LGAs). Therefore, the LGA serves as the supervision area (SA) and the state serves as the catchment area (CA). Although several indicators were measured in the application, we focus on one indicator in this example to demonstrate a second application of LC-LQAS. The indicator we use is the "percent of children 0–59 months of age that slept under an insecticide treated bednet (ITN) last night".

We required an a priori estimate of the intraclass correlation to calculate the number of LGAs to sample within each CA. Since LC-LQAS had not been previously conducted in Nigeria, we used the design effects from the 2003 Nigeria DHS to estimate the intraclass correlation, ICC (NPC, 2004). However, as design effects were not available for the specific indicator of interest, we considered five related indicators: (i) *infant mortality for the last 10 years,* (ii) *infant mortality for the last 5 years,* (iii) *child mortality for the last 10 years,* (iv) *child mortality for the last 5 years,* and (v) *sick child taken to a healthcare provider.* Table 4 gives the regional and national estimates of these design effects.

Table 4: Nigeria 2003 DHS Design Effects for Five Health Indicators

Indicator	National	North Central	North East	North West	South East	South South	South West
Infant Mortality (last 10 Years)	NA	1.53	1.44	1.56	5.64	2.50	1.79
Infant Mortality (last 5 Years)	2.19	NA	NA	NA	NA	NA	NA
Child Mortality (last 10 Years)	NA	2.25	1.62	2.06	2.73	0.58	1.68
Child Mortality (last 5 Years)	1.97	NA	NA	NA	NA	NA	NA
Sick child taken to Healthcare Provider	3.67	0.72	1.68	4.28	2.79	0.95	1.27

Table 5: Intraclass Correlation Estimates Calculated from Nigeria 2003 DHS Design Effects for Five Health Indicators, $\overline{m}_{DHS} = 22.6$.

Indicator	National	North Central	North East	North West	South East	South South	South West
Infant Mortality (last 10 Years)	NA	0.025	0.020	0.026	0.215	0.069	0.036
Infant Mortality (last 5 Years)	0.055	NA	NA	NA	NA	NA	NA
Child Mortality (last 10 Years)	NA	0.058	0.029	0.049	0.080	-0.019	0.031
Child Mortality (last 5 Years)	0.045	NA	NA	NA	NA	NA	NA
Sick child taken to Healthcare Provider	0.124	-0.013	0.031	0.152	0.083	-0.002	0.013

DHS does not provide state level indicator estimates for Nigeria; instead regional measures which comprise several states are given. The regional design effects provide the most insight into the estimate of ICC for the states. The range of design effects across all regions and indicators is 0.58–5.64 (Table 4). Assuming that on average states have 22.6 LGAs, ICCs were estimated with a range of values of -0.019 to 0.215 (Table 5) (25th percentile = 0.025, median = 0.036, and 75th percentile = 0.069). These calculations were made with the equation included in Principle 3, presented earlier in this report.

In order to conserve resources, and because the ICC in states may be lower than in regions, we used the 25th percentile estimate for ICC to calculate the number of LGAs to sample in each of the seven states. The total number of LGAs per state (N), total population per state (N^*), average square LGA population for each state ($\overline{M^2}$), and the population of each sampled LGA were needed to calculate the LGA sample sizes for each state. The population data were taken from the 1991 National Census, which is the latest available census. We present the LGA sample size calculations for Kano state in Appendix F and summarize the recommendations for all seven states in Table 6.

Table 6: Total number of LGAs per State and the Number Sampled from Each State.		
State	**Total Number of LGAs (N)**	n
Northern States		
Bauchi	20	9
Gombe	11	8
Jagawa	27	9
Kano	44	9
Southern States		
Akwa Ibom	31	9
Anambra	21	9
Rivers	23	10

The next step was to sample n LGAs from each state using SRS without replacement. Within each sampled LGA, we then randomly sampled m=19 individuals analyzed according to traditional LQAS principles (Valadez, Weiss et al. 2003). Using the formulae previously given, we estimated the proportion of children 0–59 months of age who slept under an ITN last night and a 95% confidence interval for each state. The calculations for Kano are shown in Appendix F and all seven states are summarized in Table 7. In some cases, the 95% Wald confidence interval reports negative lower limits. In these cases, we report the lower limit as zero.

Table 7: Point Estimates ($\widehat{P}$) , 95% Confidence Interval, and Intraclass Correlations for Seven Nigerian States.

State	$\widehat{P}$	95% CI	$\widehat{\rho}$
Bauchi	0.015	(0.000,0.038)*	0.0150
Gombe	0.035	(0.000,0.074)*	0.0461
Jagawa	0.068	(0.019,0.116)	0.0507
Kano	0.053	(0.018,0.089)	0.0001
Akwa Ibom	0.017	(0.000,0.041)*	0.0286
Anambra	0.036	(0.000,0.081)*	0.0876
Rivers	0.026	(0.000,0.060)*	0.0259

* Lower limit of the 95% CI < 0. In these cases we report the lower limit as zero

The prevalence of ITN use among children 0–59 months of age was low in all seven project states. This was to be expected, as the seven states were selected for the project because of their need for technical assistance and commodities. It is also important to note that this malaria project is confined to the seven intervention states and two control states (not presented here) and therefore, the estimates in these states are not meant to be representative of Nigeria as a whole.

The observed intraclass correlation for four variables are presented in Table 8: (a) percent of children 0–59 months of age who slept under an ITN, (b) percent of pregnant women who took two or more doses of SP/fansidar taken during pregnancy (i.e., for intermittent preventive treatment of malaria, IPT2), (c) percent of households owning one or more ITNs, and (d) percent of pregnant women who had one or more antenatal care visit. These observed ICCs can be used to estimate ICC for the design of future LC-LQAS. The median ICC estimate varies widely across each indicator. The indicators associated with malaria programming (ITN ownership and use, and IPT), which is a new program, exhibit homogeneity across the states, and has a low intraclass correlation. However, the indicator concerning ANC visits, which is an established activity, results in high heterogeneity across the states. Therefore, the intraclass correlation estimates is also high for the ANC indicator. The overall median ICC is 0.0304 (25th percentile= 0.0142, 75th percentile = 0.1146).

Table 8: Estimated ICC Using LC-LQAS Data

State	0–59 m.o. child slept under ITN	IPT	HH with ITN	ANC visit
Bauchi	0.0150	0.0118	-0.0132	0.0917
Gombe	0.0461	0.0290	0.0087	0.1875
Jagawa	0.0507	0.1718	0.0295	0.1082
Kano	0.0001	0.1224	0.0198	0.2984
Akwa Ibom	0.0286	0.0164	-0.0045	0.1120
Anambra	0.0876	0.0312	-0.0066	0.3426
Rivers	0.0259	0.2419	-0.0077	0.1242
Median	**0.0286**	**0.0312**	**-0.0045**	**0.1242**

As mentioned above, larger estimates of ICC lead to larger overall sample sizes. As an illustration of this, we recalculated the minimal number of LGAs required in the sample (n) for each state using updated, observed values of ICC. Specifically we consider the first, second, and third quartiles of the observed values of ICC from the data (Table 9). As expected, the larger values of ICC result in larger sample sizes. However, using the median estimate of the intraclass correlation results in only an additional two LGAs sampled in all seven states.

Table 9: Comparing Number of LGAs to Be Sampled per State (n) Using the Original Versus Updated ICC Values.

States	Original Value ICC=0.0250	25th percentile ICC = 0.0142	Median ICC = 0.0304	75th percentile ICC=0.1146
Bauchi	9	8	9	13
Gombe	8	7	8	9
Jagawa	9	8	9	14
Kano	9	8	10	17
Akwa Ibom	9	8	10	15
Anambra	9	8	9	13
Rivers	10	9	10	15
Total	**63**	**56**	**65**	**96**

Although the 25th percentile was selected for the original calculation, this choice was based on the fact that DHS design effects were available at the regional level only. Therefore, we expected larger variation in this estimate since each region reflected multiple states. Once the LC-LQAS values are available on a state basis we can revert to using the median value as more representative of variation across the states.

Additional Considerations for LC-LQAS

Using LC-LQAS to Obtain National Estimates

The LC-LQAS protocol describes methods for obtaining information to classify the performance of local supervision areas while aggregating the SA data to obtain catchment area estimates — in the case of Kenya this was at the Provincial level while in Nigeria it was at the state level. It is often valuable to also estimate the coverage proportion nationally or on a program-wide basis. Catchment level data can also be aggregated to calculate national estimates using survey sampling theory. National level calculations can be made either when all of the catchment areas are sampled (equivalent to a stratified design) or when only some of the catchment areas are included in the LC-LQAS activity, and those are selected randomly as a cluster design. This is not the case in Kenya, where two very different provinces were specifically sampled, nor was it the case in Nigeria where only the seven states with World Bank programs were included of the 36 states comprising Nigeria. While we do not explore these methods further here, it is important for countries to consider the feasibility and priority of national estimates when designing an LC-LQAS survey. National level LQAS have already been carried out in Costa Rica and the Dominican Republic, and are planned for Malawi, Benin, and Uzbekistan (see Robertson and Valadez 2006 for more examples); to date, LC-LQAS with national estimates has also been carried out in one country, Eritrea.

Using Multiple ICC Estimates for Sample Size Calculations

For simplicity, we only used one ICC estimate for all sample size estimates for both Kenya and Nigeria. This was not unreasonable, since the programs in both countries were relatively new and the ICC was estimated to be consistent across all provinces and states. However, if a condition arises in which one assumes a priori that there is large variability in the ICC across catchment areas, then it is reasonable to use a different ICC for each catchment area. This allows for increased sample sizes in areas with a suspected higher intraclass correlation, while keeping sample sizes small in areas believed to be more homogeneous.

Conclusion

This paper presents a new protocol for establishing a decentralized M&E system (e.g., provincial, region or district level) in countries in which it is feasible to start data collection and evaluation in some but not all supervision areas. In addition to outlining the statistical logic for merging LQAS and cluster sampling methodologies, this paper outlines the steps to take in order to implement LC-LQAS. Two examples are provided — one in Kenya, where LC-LQAS was implemented in two provinces, and one in Nigeria, where LC-LQAS was implemented in seven states.

Like other cluster sampling methods, LC-LQAS reduces the amount of resources required to obtain catchment area estimates by geographically restricting the areas visited. The real power of this methodology is in producing measurement results for the supervision areas (i.e., the clusters) that were sampled to include in the survey, while simultaneously calculating point estimates for the catchment area. This dual feature of LC-LQAS allows program managers to direct resources where they are needed at the local level to improve current programs, and to decide where new activities could be planned.

Having demonstrated the versatility of LC-LQAS in two countries implementing different programs, we conclude that the LC-LQAS is ready for application in other program settings. This protocol provides the necessary detail and forms to support this procedure.

APPENDIX A: DERIVATION OF LC-LQAS ESTIMATORS

NOTATION

y_{ij} A variable to indicate whether or not individual j person in SA i has successful outcome

t_i The total number of successful outcomes observed in a sample from SA i

p_i The probability of a successful outcome in subregion i

$s_{p_i}^2$ The within subregion i variance

N The total number of subregions in a region

n The number of subregions sampled in a region

M_i The total population of subregion i

m The number of individuals sampled in subregion i

N^* The total population of the region

P The probability of success in the region

T The total number of individuals with a successful outcome in a region

s_B^2 The between subregion variance

S_k The kth sample of subregions from the region

ESTIMATION IN THE SUBREGIONS

We outline the estimation of the coverage proportion of an indicator for one subregion. This value will not be used for conclusions about the subregions's performance but will be used for calculating the coverage proportion in the entire region.

Suppose that the proportion of people performing acceptable on an indicator in a subregion is p_i. Let y_{ij} be an indicator variable that is 1 if an individual performs acceptably or 0 if an individual does not.

$$y_{ij} = \begin{cases} 1 & \text{if performs at an acceptable level} \\ 0 & \text{if does not perform at an acceptable level} \end{cases}$$

$$\begin{aligned} P(y_{ij} = 1) &= p_i \\ P(y_{ij} = 0) &= 1 - p_i \end{aligned}$$

Let t_i equal the total number of people in a sample of size m to perform at an acceptable level on an indicator; so,

$$t_i = \sum_{j=1}^{m} y_{ij}$$

and $t_i \sim Binomial(m, p_i)$, if the individuals are chosen at random.

The binomial distribution requires that p_i be constant for each individual. When sampling without replacement from a finite population this assumption is compromised, and the hypergeometric is a more precise distribution for t_i. However, because here the sample size is small relative to the total region population, i.e. $m << M_i$, the binomial provides a close approximation to the hypergeometric distribution. The binomial distribution also assumes that each individual observation is independent of the other observations.

The Maximum Likelihood Estimator for the Binomial is $\widehat{p}_i = t_i/m$, which is unbiased, and is the estimator with minimum variance. An unbiased estimator for the variance of $\widehat{p}_i$ is

$$s_{p_i}^2 = (1 - \frac{m}{M_i}) \frac{\widehat{p}_i(1 - \widehat{p}_i)}{m - 1}.$$

Since the sample size is very small compared to the region population, the sampling fraction, m/M_i, is negligible. Thus the estimator for the variance can simply be expressed as:

$$s^2_{p_i} = \frac{\widehat{p}_i(1 - \widehat{p}_i)}{m - 1}.$$

ESTIMATION IN REGIONS

An estimator of the overall proportion of the region population that performs at an acceptable level can be obtained using the sampled subregions. These subregions are randomly sampled with equal probability, so a regional estimate obtained by simply averaging the sampled subregions proportions and standard binomial variance estimators is not optimal because:

- The estimator is biased because the subregion sizes are different but selected with equal probability (i.e. people in smaller subregions have a higher overall probability of being selected than people in larger subregions).

- The variance is underestimated because people within clusters are expected to be more similar (have similar knowledge levels) than people from across the entire region.

For these reasons, the estimated proportions from sampled subregions must be weighted to provide a more accurate estimator for the regional proportion and the variance estimator should accommodate within cluster and between cluster variability.

Derivation of the Weights

Each individual observation is weighted by the inverse of the probability that that individual is selected (see Lohr, p226).

$$w_{ij} = \frac{1}{P(\text{ith subregion Selected}) * P(\text{jth individual selected} \mid \text{ith subregion selected})}$$

$$= \frac{1}{\frac{n}{N}\left(\frac{m}{M_i}\right)}$$

$$= \frac{NM_i}{nm}.$$

Estimator of Proportion

The total proportion of the population in the region that performs acceptably, P, is the ratio of the total number of people in the region that perform acceptably, T, to the total population living in region, N^*.

$$N^* = \sum_{i=1}^{N} M_i$$

$$P = \frac{T}{N^*}.$$

In order to obtain a reliable estimator for P in the region, we derive an unbiased estimator for $T = \sum_{i=1}^{N} M_i p_i$ by using the weighted observations from the sampled subregions.

$$\widehat{T} = \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}(y_{ij})$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{NM_i}{nm}(y_{ij})$$

$$= \sum_{i=1}^{n} \frac{NM_i}{n}\left(\frac{\sum_{j=1}^{m} y_{ij}}{m}\right)$$

$$= \sum_{i=1}^{n} \frac{NM_i}{n}\left(\frac{t_i}{m}\right)$$

$$= \sum_{i=1}^{n} \frac{NM_i}{n}(\widehat{p}_i).$$

Using this estimator of T, we can obtain an estimator for P:

$$\widehat{P} = \frac{\widehat{T}}{N^*}$$

$$= \frac{1}{N^*}\left(\frac{N}{n}\sum_{i=1}^{n}M_i(\widehat{p}_i)\right).$$

Noting that $\widehat{T}$ is an unbiased estimator for T (see following proof), it follows that $\widehat{P} = \widehat{T}/N^*$ is an unbiased estimator for P. The variance for $\widehat{P}$ will have terms for both the within cluster and between cluster variability. As shown in later in this appendix, the best estimator for the variance of $\widehat{P}$ is

$$\widehat{var}(\widehat{P}) = \frac{1}{(N^*)^2}\left[\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 s_B^2 + \frac{N}{n}\sum_{i=1}^{n}M_i^2 s_{p_i}^2\right]$$

with

$$s_B^2 = \frac{\sum_{i=1}^{n}(\widehat{p}_i - \widehat{P})^2}{n-1}$$

$$s_{p_i}^2 = \frac{\widehat{p}_i(1-\widehat{p}_i)}{m-1}.$$

Wald-like Confidence Intervals for P can be constructed using the normal approximation to the binomial (see Casella and Berger, p492-493, 497).

$$= \widehat{P} \pm 1.96\sqrt{\widehat{Var}[\widehat{P}]}$$

$$= \widehat{P} \pm \frac{1.96}{N^*}\sqrt{\left[\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 s_B^2 + \frac{N}{n}\sum_{i=1}^{n}M_i^2 s_{p_i}^2\right]}.$$

Proof that $\widehat{T}$ is an Unbiased Estimator of T

Let S be the set of the n subregions that are sampled, and denote expectation by E[]:

$$
\begin{aligned}
E[\widehat{T}] &= E[E[\widehat{T}|i \in S]] \\
&= E[E[\frac{N}{n} \sum_{i=1}^{n} M_i \widehat{p}_i | i \in S]] \\
&= E[\frac{N}{n} \sum_{i=1}^{n} M_i p_i].
\end{aligned}
$$

There are $\binom{N}{n}$ ways to select a sample of n from N woredas, and because of random selection, each has a probability of $1/\binom{N}{n}$ of being selected. Thus, continuing from above:

$$
\begin{aligned}
E[\widehat{T}] &= \left(\frac{N}{n} \sum_{i \in S_1} M_i p_i \right) \frac{1}{\binom{N}{n}} + \left(\frac{N}{n} \sum_{i \in S_2} M_i p_i \right) \frac{1}{\binom{N}{n}} + \cdots + \left(\frac{N}{n} \sum_{i \in S_{\binom{N}{n}}} M_i p_i \right) \frac{1}{\binom{N}{n}} \\
&= \left(\frac{N}{n} \right) \frac{1}{\binom{N}{n}} \left[\binom{N-1}{n-1} M_1 p_1 + \binom{N-1}{n-1} M_2 p_2 + \cdots + \binom{N-1}{n-1} M_N p_N \right] \\
&= \frac{N(N-1)!}{n(n-1)!(N-1-(n-1))!} \frac{1}{\binom{N}{n}} \left[\sum_{i=1}^{N} M_i p_i \right] \\
&= \binom{N}{n} \frac{1}{\binom{N}{n}} \left[\sum_{i=1}^{N} M_i p_i \right] \\
&= \sum_{i=1}^{N} M_i p_i \\
&= \sum_{i=1}^{N} T_i = T
\end{aligned}
$$

which proves the assertion that $\widehat{T}$ is an unbiased estimator of T.

DERIVATION OF THE VARIANCE FOR $\widehat{P}$

Let S be the set of the n subregions that are sampled. The variance of $\widehat{P}$ is denoted as:

$$Var(\widehat{P}) = \frac{Var(\widehat{T})}{(N^*)^2}$$

$$= \frac{E[Var(\widehat{T}|i \in S)] + Var[E(\widehat{T}|i \in S)]}{(N^*)^2}.$$

Looking at each component separately,

$$E(\widehat{T}|i \in S) = E[\frac{N}{n} \sum_{i=1}^{n} M_i \widehat{p}_i | i \in S]$$

$$= \frac{N}{n} \sum_{i=1}^{n} M_i p_i.$$

$$Var[E(\widehat{T}|i \in S)] = Var[\frac{N}{n} \sum_{i=1}^{n} M_i p_i]$$

$$= \frac{N^2}{n^2}(1 - \frac{n}{N}) \sum_{i=1}^{n} M_i^2 Var(p_i).$$

Note that the variance of p_i is not the variance that arises from the binomial, $p_i(1-p_i)$. This term, $Var(p_i)$, denotes the variance of the estimates of the proportions between the clusters, S_B^2. Thus,

$$Var[E(\widehat{T}|i \in S)] = \frac{N^2}{n^2}(1 - \frac{n}{N}) \sum_{i=1}^{n} M_i^2 S_B^2.$$

Focusing on the second component, we get

$$Var(\widehat{T}|i \in S) = Var[\frac{N}{n}\sum_{i=1}^{n} M_i\widehat{p}_i|i \in S]$$

$$= \frac{N^2}{n^2}\sum_{i=1}^{n} M_i^2 S_{p_i}^2$$

with $S_{p_i}^2$ representing the variance with cluster i. Also,

$$E[Var(\widehat{T}|i \in S)] = E[\frac{N^2}{n^2}\sum_{i=1}^{n} M_i^2 S_{p_i}^2]$$

$$= \frac{N^2}{n^2}E[\sum_{i=1}^{n} M_i^2 S_{p_i}^2].$$

Again, there are $\binom{N}{n}$ ways to select a sample, each of which has equal probability of selection, and so

$$E[Var(\widehat{T}|i \in S)] = \frac{N^2}{n^2}\left[\left(\sum_{i\in S_1} M_i^2 S_{p_i}^2\right)\frac{1}{\binom{N}{n}} + \left(\sum_{i\in S_2} M_i^2 S_{p_i}^2\right)\frac{1}{\binom{N}{n}} + \cdots + \left(\sum_{i\in S_{\binom{N}{n}}} M_i^2 S_{p_i}^2\right)\frac{1}{\binom{N}{n}}\right]$$

$$= \frac{N^2}{n^2}\frac{1}{\binom{N}{n}}\left[\binom{N-1}{n-1} M_1^2 S_{p_1}^2 + \binom{N-1}{n-1} M_2^2 S_{p_2}^2 + \cdots + \binom{N-1}{n-1} M_N^2 S_{p_N}^2\right]$$

$$= \frac{N}{n}\sum_{i=1}^{N} M_i^2 S_{p_i}^2.$$

Thus, the variance of $\widehat{P}$ can be expressed as:

$$Var(\widehat{P}) = \frac{1}{(N^*)^2}\left[\frac{N^2}{n^2}(1 - \frac{n}{N})\sum_{i=1}^{n} M_i^2 S_B^2 + \frac{N}{n}\sum_{i=1}^{N} M_i^2 S_{p_i}^2\right],$$

which in turn can be estimated unbiasedly by (see Lohr, p210):

$$\widehat{var}(\widehat{P}) = \frac{1}{(N^*)^2}\left[\frac{N^2}{n^2}(1 - \frac{n}{N})\sum_{i=1}^{n} M_i^2 s_B^2 + \frac{N}{n}\sum_{i=1}^{n} M_i^2 s_{p_i}^2\right].$$

37

APPENDIX B: DERIVATION OF LC-LQAS SAMPLE SIZE FORMULA

Based on the work of Fleiss (1981) and Ridout et al (1999), we have an ANOVA estimate of the intraclass correlation (ICC) for binary data, $\widehat{\rho} = (msc - mse)/(msc + (m-1)mse)$, where

$$
\begin{aligned}
msc &= \sum_{i=1}^{n} \frac{m(\widehat{p}_i - \widehat{P})^2}{n-1} \\
&= ms_B^2
\end{aligned}
$$

$$
\begin{aligned}
mse &= m \sum_{i=1}^{n} \frac{\widehat{p}_i(1 - \widehat{p}_i)}{n(m-1)} \\
&= \frac{m}{n} \sum_{i=1}^{n} s_{p_i}^2.
\end{aligned}
$$

Using these relationships, we can represent the between cluster variance as a function of the within cluster variance:

$$
\begin{aligned}
s_B^2 &= \frac{1 + (m-1)\widehat{\rho}}{n(1 - \widehat{\rho})} \sum_{i=1}^{n} s_{p_i}^2 \\
&= K_{\widehat{\rho}} \sum_{i=1}^{n} s_{p_i}^2.
\end{aligned}
$$

From Appendix A, the length of the 95% Wald-like Confidence Interval is:

$$
\begin{aligned}
\text{length} &= 2 * 1.96 \sqrt{\widehat{Var}[\widehat{P}]} \\
&= \frac{2 * 1.96}{N^*} \sqrt{\left[\frac{N^2}{n^2} (1 - \frac{n}{N}) \sum_{i=1}^{n} M_i^2 s_B^2 + \frac{N}{n} \sum_{i=1}^{n} M_i^2 s_{p_i}^2 \right]}.
\end{aligned}
$$

Substituting the relationship of the between to within cluster variation based on the ICC, the maximum length of the confidence interval can be expressed as:

$$
\begin{aligned}
max[\text{length}] \;=\; & \frac{2*1.96}{N^*}\sqrt{max\left[\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 K_{\widehat{\rho}}\sum_{i=1}^{n}s_{p_i}^2 + \frac{N}{n}\sum_{i=1}^{n}M_i^2 s_{p_i}^2\right]} \\[2mm]
=\; & \frac{2*1.96}{N^*}\sqrt{\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 K_{\widehat{\rho}}\sum_{i=1}^{n}max(s_{p_i}^2) + \frac{N}{n}\sum_{i=1}^{n}M_i^2 max(s_{p_i}^2)} \\[2mm]
=\; & \frac{2*1.96}{N^*}\sqrt{\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 K_{\widehat{\rho}}\sum_{i=1}^{n}\frac{max(\widehat{p}_i(1-\widehat{p}_i))}{m-1} + \frac{N}{n}\sum_{i=1}^{n}M_i^2\frac{max(\widehat{p}_i(1-\widehat{p}_i))}{m-1}}.
\end{aligned}
$$

This value is maximized at $\widehat{p}_i = 1/2$. Thus we obtain:

$$
\begin{aligned}
max[\text{length}] = \ell_{max} \;=\; & \frac{2*1.96}{N^*}\sqrt{\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 K_{\widehat{\rho}}\sum_{i=1}^{n}\frac{(0.5)^2}{m-1} + \frac{N}{n}\sum_{i=1}^{n}M_i^2\frac{(0.5)^2}{m-1}} \\[2mm]
=\; & \frac{2*1.96}{N^*}\sqrt{N^2(1-\frac{n}{N})\left(\frac{\sum_{i=1}^{n}M_i^2}{n}\right)K_{\widehat{\rho}}\frac{(0.5)^2}{m-1} + N\left(\frac{\sum_{i=1}^{n}M_i^2}{n}\right)\frac{(0.5)^2}{m-1}}.
\end{aligned}
$$

This has the advantage of expressing the problem in terms of a single parameter, ρ. Also note that the quantity $(\sum_{i=1}^{n}M_i^2)/(n)$ is specific to our sample but can be best estimated with the expression $\overline{M^2} = (\sum_{i=1}^{N}M_i^2)/(N)$.

$$\ell_{max} = \frac{2*1.96}{N^*}\sqrt{N^2(1-\frac{n}{N})\overline{M^2}K_{\widehat{\rho}}\frac{(0.5)^2}{m-1} + N\overline{M^2}\frac{(0.5)^2}{m-1}}$$

$$= \frac{2*1.96}{N^*}\sqrt{N^2(1-\frac{n}{N})\overline{M^2}K_{\widehat{\rho}}\frac{(0.5)^2}{m-1} + N\overline{M^2}\frac{(0.5)^2}{m-1}}$$

$$= \frac{2*1.96}{N^*}\sqrt{N\overline{M^2}\frac{(0.5)^2}{m-1}\left[N(1-\frac{n}{N})K_{\widehat{\rho}}+1\right]}$$

$$= \frac{2*1.96}{N^*}\sqrt{N\overline{M^2}\frac{(0.5)^2}{m-1}\left[\frac{N}{n}\left(\frac{1+(m-1)\widehat{\rho}}{1-\widehat{\rho}}\right) - \frac{m\widehat{\rho}}{1-\widehat{\rho}}\right]}$$

$$\Rightarrow \left(\frac{\ell_{max}N^*}{2*1.96*0.5}\right)^2\left(\frac{m-1}{N\overline{M^2}}\right) + \frac{m\widehat{\rho}}{1-\widehat{\rho}} = \frac{N}{n}\left(\frac{1+(m-1)\widehat{\rho}}{1-\widehat{\rho}}\right)$$

$$\Rightarrow \quad n = N\frac{(1+(m-1)\widehat{\rho})}{\left[\left(\frac{\ell_{max}N^*}{1.96}\right)^2\left(\frac{(m-1)(1-\widehat{\rho})}{N\overline{M^2}}\right) + m\widehat{\rho}\right]}$$

Comments on the sample size formula

The sample size formula above is bounded below by n=0, corresponding to the case of when $\widehat{\rho} = -1/(m-1)$. Given the relationship between the ICC and design effect, this occurs at the unrealistic scenario when the design effect is zero. When $\widehat{\rho} = 1$, corresponding to the case when the design effect is at a maximum, $DE = m$, then we have $n = N$. In other words, the maximum number of clusters that can be recommended to include in the frame is N, the total number of clusters available for sampling.

Additionally, when $DE = 1$ corresponding to $\widehat{\rho} = 0$, we have

$$n = \frac{N^2\overline{M^2}1.96^2}{\ell_{max}^2(m-1)(N^*)^2}.$$

Note that when all of the subregions have the same population size, then the recommended sample size is $(1.96^2)/(\ell_{max}^2*(m-1))$ which reduces to $96/(m-1)$ when restricting the length to $\pm10\%$. This results in approximately the total sample size of 96 which is required to meet the same restriction under simple random sampling. Also, as shown below, the multiplier of

this term, $N^2\overline{M^2}/(N^*)^2$ is bounded below by 1. This implies that the recommended number clusters is at least what is expected if there is no design effect, and is only substantially bigger when the clusters' population sizes vary widely (corresponding with higher variability in the weighted estimator of regional proportions).

Showing that $N^2\overline{M^2}/(N^*)^2$ is bounded below by 1, we have:

$$\frac{N^2\overline{M^2}}{(N^*)^2} = \frac{N\sum_{i=1}^{N}M_i^2}{(\sum_{i=1}^{N}M_i)^2}$$

$$= N\sum_{i=1}^{N}X_i^2,$$

where $X_i = (M_i)/(\sum_{i=1}^{N}M_i)^2)$. Note that

$$0 \leq \sum_{i=1}^{N}(X_i - \bar{X})^2 = \sum_{i=1}^{N}X_i^2 - 2\bar{X}\sum_{i=1}^{N} + N(\bar{X})^2$$

$$\leq \sum_{i=1}^{N}X_i^2 - 2N(\bar{X})^2 + N\bar{X} = \sum_{i=1}^{N}X_i^2 - N(\bar{X})^2 = \sum_{i=1}^{N}X_i^2 - 1/N$$

$$\Rightarrow \frac{1}{N} \leq \sum_{i=1}^{N}X_i^2 = \frac{\sum_{i=1}^{N}M_i^2}{(\sum_{i=1}^{N}M_i)^2}$$

$$\Rightarrow 1 \leq \frac{N\sum_{i=1}^{N}M_i^2}{(\sum_{i=1}^{N}M_i)^2} = \frac{N^2\overline{M^2}}{(N^*)^2}$$

APPENDIX C: SAMPLE SIZE CALCULATION FORM

Total number of Supervision Areas	N	
Sample Size per SA	m	19
Total Population†	N^*_{cen}	A
Average Square SA Population	$\overline{M^2}$	B
Estimate of Design Effect	DE	
Estimate of Intraclass Correlation	$\widehat{\rho} = (DE - 1)/(m^{\dagger\dagger} - 1)$	

† For estimating the sample size, use total population estimates based on the most recent census.
††Note that the m used for the ICC estimate is the average cluster size for the survey from which the design effect was drawn.

Supervision Area	Population M_i	Population Squared M_i^2
	Total Population $N^*_{cen} = \sum_{i=1}^{N} M_i$ $= {}^A$	Average Population Squared $\overline{M^2} = \frac{1}{N} \sum_{i=1}^{N} M_i^2$ $= {}^B$

$$n = N(1 + (m-1)\widehat{\rho}) \left\{ \left[\left(\frac{\ell_{max} N^*_{cen}}{1.96} \right)^2 \left(\frac{(m-1)(1-\widehat{\rho})}{N\overline{M^2}} \right) + m\widehat{\rho} \right] \right\}^{-1}$$

$$= \underline{}^N (1 + (19-1)\underline{}^{\widehat{\rho}}) \left\{ \left[\left(\frac{0.2\underline{}^{N^*_{cen}}}{1.96} \right)^2 \left(\frac{(19-1)(1-\underline{}^{\widehat{\rho}})}{\underline{}^N \underline{}^{\overline{M^2}}} \right) + 19\underline{}^{\widehat{\rho}} \right] \right\}^{-1}$$

$$=$$

Number of Supervision Areas to Sample = __ (smallest integer bigger than n).

APPENDIX D: FORM FOR ESTIMATING CATCHMENT AREA COVERAGE PROPORTION WITH 95% CONFIDENCE INTERVALS

CATCHMENT AREA COVERAGE PROPORTIONS

INDICATOR:

Sampled Supervision Areas	Population M_i	Number with Positive Result t_i	SA Proportion $\hat{p}_i = \frac{t_i}{19}$	Population * Proportion $M_i(\hat{p}_i)$
	Total $= \sum\limits_{i=1}^{n} M_i$			Total $= \sum\limits_{i=1}^{n} M_i(\hat{p}_i)$ $= C$

Total number of Supervision Areas	N	
Number of sampled SAs	n	
Total Population[†]	$N^*_{sam} = \frac{N}{n} \sum\limits_{i=1}^{n} M_i$	
Sum of SA pop times SA proportion	$\sum\limits_{i=1}^{n} M_i(\hat{p}_i)$	C

[†] For data analysis, use total population estimates based on the sample, $N^*_{sam} = N \sum\limits_{i=1}^{n} M_i / n$.

$$\hat{P} = \frac{N}{n(N^*_{sam})} \sum\limits_{i=1}^{n} M_i(\hat{p}_i)$$

$$= \frac{\frac{N}{\underline{\quad}}}{\underline{\quad}n\Big(\underline{\qquad\quad}N^*_{sam}\Big)} \underline{\quad}\frac{C}{\quad}$$

$$=$$

95% Confidence Interval

INDICATOR:

Sampled Supervision Area	Population M_i	Square of Population M_i^2	SA Proportion $\hat{p}_i$	SA Sample Variation $s_{p_i}^2 = \frac{\hat{p}_i(1-\hat{p}_i)}{18}$	Pop Squared * Variation $M_i^2(s_{p_i}^2)$	Square Error $(\hat{p}_i - \hat{P})^2$
		Total = $\sum_{i=1}^{n} M_i^2$ $=^D$		SSE = $\frac{m}{n} \sum_{i=1}^{n} s_{p_i}^2$ $=^E$	Total = $\sum_{i=1}^{n} M_i^2(s_{p_i}^2)$ $=^F$	Between SA Variance $s_B^2 = \frac{\sum_{i=1}^{n}(\hat{p}_i - \hat{P})^2}{n-1}$ $=^G$

Total number of Supervision Areas	N	
Number of sampled SAs	n	
Total Population	N_{sam}^*	
Total of Squared Supervision Area Population	$\sum_{i=1}^{n} M_i^2$	D
Sum of SA pop square times SA variance	$\sum_{i=1}^{n} M_i^2(s_{p_i}^2)$	F
Between SA variance	s_B^2	G
MSE		E
MSC	$m(s_B^2)$	H

Variance

$$\widehat{var}(\hat{P}) \;=\; \frac{1}{(N^*_{sam})^2}\left[\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 s_B^2 + \frac{N}{n}\sum_{i=1}^{n}M_i^2 s_{p_i}^2\right]$$

$$=\; \frac{1}{(\underline{\qquad}N^*_{sam})^2}\left[\frac{(\underline{\quad}N)^2}{(\underline{\quad}n)^2}(1-\underline{\quad}\frac{n}{N})\underline{\qquad}\underset{D}{}\underline{\qquad}\underset{G}{} + \underline{\quad}\frac{N}{n}\underline{\qquad}\underset{F}{}\right]$$

$$=$$

95% Confidence Interval

The 95% Confidence Interval is formed by:

$$\left[\hat{P}-1.96\sqrt{\widehat{var}(\hat{P})},\,\hat{P}+1.96\sqrt{\widehat{var}(\hat{P})}\right]$$

$$\left[\underline{\quad}\hat{P}-1.96\sqrt{\underline{\qquad}\widehat{var}},\,\underline{\quad}\hat{P}+1.96\sqrt{\underline{\qquad}\widehat{var}}\right]$$

$$\left[\underline{\qquad},\underline{\qquad}\right]$$

INTRACLASS CORRELATION

The intraclass correlation is estimated by:

$$\widehat{\rho} \;=\; \frac{msc-mse}{msc+(m-1)mse}$$

$$=\; \frac{\underline{\qquad}^H - \underline{\qquad}^E}{\underline{\quad}^H + (19-1)\underline{\quad}^E}$$

$$=$$

46

APPENDIX E: STATA COMMANDS FOR COVERAGE PROPORTIONS WITH 95% CONFIDENCE INTERVALS

The example below provides the necessary commands to analyze this data in the statistical computing package, Stata. It is important to note that there are subtleties in how each package defines weights for each stage of the sample and carries out a finite population correction. The reader should look into more detail before the use of each package. In order for Stata to appropriately incorporate the survey design, the key features of the sampling design must be set using the **svyset** command. First, the variable that contains the cluster or supervision area number in our dataset is referred to as ***const***, and the variable that contains the identification number for each individual surveyed, ***q_num***. Next, the sampling weight, as described in Section A.2.1, must be defined. In our data set, the variable ***weight*** contains the population size for each sampled SA or Constituency. The following command creates a new variable with the appropriate sampling weight.

$$\text{generate float samp\_weight} = 32\text{*weight}/19/16$$

where 32 is the number of SA in Nyanza, 19 is the LQAS sample size in each SA sampled, and 16 is the number of SA of the 32 total that were sampled. Note that our previous calculations indicated that a minimum of 11 SA were needed. However, the NACC concluded that they could accommodate 5 additional Constituencies in the training and data collection. Therefore the number was increased to 16 for management reasons.

Finally, the finite population correction, in Stata defined as the fraction of eligible units sampled, must be defined and since 16 out of 32 districts were sampled in the Kenya survey, the following command results.

$$\text{generate float fpc1} = 16/32$$

Now, the following command sets the survey design into Stata memory.

$$\text{svyset const [pweight=samp\_weight], fpc(fpc1) || q\_num}$$

The variable ***KnowPrev*** has the "yes" and "no" responses to the question "Can you name at least two ways to successfully prevent sexual transmission of HIV?". The command below calculates the coverage proportion of the men in Nyanza province that responded yes to this question, incorporating the design effect into the calculation.

svy: proportion KnowPrev

Finally, this last command reports the estimated design effect for this proportion.

estat effects, deff

Running these commands in sequence leads to the following output from Stata.

```
. generate float samp_weight = 32*weight/19/16

. generate float fpc1 = 16/32

. svyset const [pweight=samp_weight], fpc(fpc1) || q_num

      pweight: samp_weight
          VCE: linearized
    Strata 1: <one>
        SU 1: const
       FPC 1: fpc1
    Strata 2: <one>
        SU 2: q_num
       FPC 2: <zero>

. svy: proportion KnowPrev (running proportion on estimation sample)

Survey: Proportion estimation

Number of strata =         1          Number of obs    =       304
Number of PSUs   =        16          Population size  = 2.1e+06
                                      Design df        =        15


------------------------------------------------------------------
           |                 Linearized       Binomial Wald
           | Proportion   Std. Err.     [95% Conf. Interval]
-----------+------------------------------------------------------
KnowPrev   |
         0 |    .2654571   .0429534      .1739042      .35701
         1 |    .7345429   .0429534       .64299     .8260958
------------------------------------------------------------------
```

```
. estat effects, deff


--------------------------------------------------
            |               Linearized
            | Proportion    Std. Err.        Deff
------------+-------------------------------------
KnowPrev    |
         0  |   .2654571     .0429534       2.8674
         1  |   .7345429     .0429534       2.8674
--------------------------------------------------
```
Note: Weights must represent population totals for deff to be correct
    when using an FPC; however, deft is invariant to the scale of weights.

APPENDIX F: SAMPLING FRAME AND ANALYSIS FOR LC-LQAS IN KANO STATE, NIGERIA

SAMPLE SIZE CALCULATION

Total number of Supervision Areas	N	44
Sample Size per SA	m	19
Total Population[†]	N^*_{cen}	[A] 5,792,097
Average Square SA Population	$\overline{M^2}$	[B] 20,820,689,926
Estimate of Design Effect	DE	1.54
Estimate of Intraclass Correlation	$\widehat{\rho} = (DE - 1)/(m^{\dagger\dagger} - 1)$	0.025

[†] For estimating the sample size, use total population estimates based on the most recent census.
[††]Note that the m used for the ICC estimate is the average cluster size for the survey from which the design effect was drawn.

Supervision Area	Population M_i	Population Squared M_i^2	Supervision Area	Population M_i	Population Squared M_i^2
Rogo	140,447	19,725,359,809	Ajingi	104,212	10,860,140,944
Kiru	156,584	24,518,549,056	Warawa	81,666	6,669,335,556
Bebeji	118,833	14,121,281,889	Tofa	64,796	4,198,521,616
Rano	88,886	7,900,720,996	Bichi	182,674	33,369,790,276
Tundu Wada	141,288	19,962,298,944	Dawakin Tofa	156,443	24,474,412,249
Sumaila	164,242	26,975,434,564	Minjibir	139,750	19,530,062,500
Kiblya	75,277	5,666,626,729	Kumbotso	166,558	27,741,567,364
Garko	125,001	15,625,250,001	Gezawa	154,629	23,910,127,641
Takai	130,007	16,901,820,049	Gabasawa	152,899	23,378,104,201
Bunkure	122,856	15,093,596,736	Ungogo	168,373	28,349,467,129
Karaye	89,711	8,048,063,521	Dambatta	144,268	20,813,255,824
Gwarzo	118,778	14,108,213,284	Tsanyawa	83,942	7,046,259,364
Shanono	84,861	7,201,389,321	Makoda	69,954	4,893,562,116
Kabo	90,158	8,128,464,964	Doguwa	83,365	6,949,723,225
Kura	89,342	7,981,992,964	Kunchi	75,277	5,666,626,729
Madobi	78,924	6,228,997,776	Dala	316,137	99,942,602,769
Rimingado	60,622	3,675,026,884	Fagge	156,342	24,442,820,964
Bagwai	106,645	11,373,156,025	Gwale	177,437	31,483,888,969
Dawakin Kudu	163,668	26,787,214,224	Kano Municipal	270,764	73,313,143,696
Wudil	108,441	11,759,450,481	Nassarawa	335,729	112,713,961,441
Albasu	116,603	13,596,259,609	Tarauni	135,846	18,454,135,716
Gaya	135,689	18,411,504,721	Garum Mallam	64,173	4,118,173,929
				Total Population $N^*_{cen} = \sum\limits_{i=1}^{N} M_i$ $= 5,792,097$ [A]	Average Population Squared $\overline{M^2} = \frac{1}{N} \sum\limits_{i=1}^{N} M_i^2$ $= 20,820,689,926$ [B]

$$n = N(1 + (m-1)\widehat{\rho}) \left\{ \left[\left(\frac{\ell_{max} N_{cen}^*}{1.96} \right)^2 \left(\frac{(m-1)(1-\widehat{\rho})}{N\overline{M}^2} \right) + m\widehat{\rho} \right] \right\}^{-1}$$

$$= 44(1 + (19-1)0.025) \left[\left(\frac{0.2(5,792,097)}{1.96} \right)^2 \left(\frac{(19-1)(1-0.025)}{44(20,820,689,926)} \right) + 19(0.025) \right]^{-1}$$

$$= 8.90$$

Number of Supervision Areas to Sample $= \underline{9}$ (smallest integer bigger than n).

COVERAGE PROPORTIONS

INDICATOR: Child Used an ITN.

Sampled Supervision Areas	Population M_i	Number with Positive Result t_i	SA Proportion $\hat{p}_i = \frac{t_i}{19}$	Population * Proportion $M_i(\hat{p}_i)$
Ajingi	104,212	0	0.000	0
Bagwai	106,645	0	0.000	0
Dala	316,137	2	0.105	33,278
Dambatta	144,268	1	0.053	7,593
Garko	125,001	1	0.053	6,579
Minjibir	139,750	0	0.000	0
Shanono	84,861	1	0.053	4,466
Tarauni	135,846	2	0.105	14,300
Rano	88,886	0	0.000	0
	Total $= \sum\limits_{i=1}^{n} M_i$ 1,245,606			Total $= \sum\limits_{i=1}^{n} M_i(\hat{p}_i)$ $= 66,216$ [C]

Total number of Supervision Areas	N	44
Number of sampled SAs	n	9
Total Population[†]	$N_{sam}^* = \frac{N}{n} \sum\limits_{i=1}^{n} M_i$	6,089,629
Sum of SA pop times SA proportion	$\sum\limits_{i=1}^{n} M_i(\hat{p}_i)$	66,216 [C]

[†] For data analysis, use total population estimates based on the sample, $N_{sam}^* = N \sum\limits_{i=1}^{n} M_i/n$.

$$
\begin{aligned}
\hat{P} &= \frac{N}{n(N_{sam}^*)} \sum_{i=1}^{n} M_i(\hat{p}_i) \\
&= \frac{44}{9(6,089,629)} 66,216 \\
&= 0.0532
\end{aligned}
$$

CONFIDENCE INTERVAL

INDICATOR: Child Used an ITN.

Sampled Supervision Area	Population M_i	Square of Population M_i^2	SA Proportion $\hat{p}_i$	SA Sample Variation $s_{p_i}^2 = \frac{\hat{p}_i(1-\hat{p}_i)}{18}$	Pop Squared * Variation $M_i^2(s_{p_i}^2)$	Square Error $(\hat{p}_i - \hat{P})^2$
Ajingi	104,212	10,860,140,944	0.000	0.000000	0	0.002826
Bagwai	106,645	11,373,156,025	0.000	0.000000	0	0.002826
Dala	316,137	99,942,602,769	0.105	0.005232	522,937,595	0.002715
Dambatta	144,268	20,813,255,824	0.053	0.002770	57,654,448	0.000000
Garko	125,001	15,625,250,001	0.053	0.002770	43,283,241	0.000000
Minjibir	139,750	19,530,062,500	0.000	0.000000	0	0.002826
Shanono	84,861	7,201,389,321	0.053	0.002770	19,948,447	0.000000
Tarauni	135,846	18,454,135,716	0.105	0.005232	96,559,036	0.002715
Rano	88,886	7,900,720,996	0.000	0.000000	0	0.002826
		Total = $\sum\limits_{i=1}^{n} M_i^2$ $= 211{,}700{,}714{,}096$ [D]		SSE = $\frac{m}{n}\sum\limits_{i=1}^{n} s_{p_i}^2$ $= 0.039636$ [E]	Total = $\sum\limits_{i=1}^{n} M_i^2(s_{p_i}^2)$ $= 740{,}382{,}767$ [F]	Between SA Variance $s_B^2 = \frac{\sum\limits_{i=1}^{n}(\hat{p}_i - \hat{P})^2}{n-1}$ $= 0.002092$ [G]

Total number of Supervision Areas	N	44
Number of sampled SAs	n	9
Total Population	N_{sam}^*	6,089,629
Total of Squared Supervision Area Population	$\sum\limits_{i=1}^{n} M_i^2$	211,700,714,096 [D]
Sum of SA pop square times SA variance	$\sum\limits_{i=1}^{n} M_i^2(s_{p_i}^2)$	740,382,767 [F]
Between SA variance	s_B^2	0.002092 [G]
MSE		0.039636 [E]
MSC	$m(s_B^2)$	0.039744 [H]

Variance

$$\widehat{var}(\hat{P}) = \frac{1}{(N^*_{sam})^2}\left[\frac{N^2}{n^2}(1-\frac{n}{N})\sum_{i=1}^{n}M_i^2 s_B^2 + \frac{N}{n}\sum_{i=1}^{n}M_i^2 s_{p_i}^2\right]$$

$$= \frac{1}{(6,089,629)^2}\left[\frac{(44)^2}{(9)^2}(1-\frac{9}{44})211,700,714,096(0.002092) + \frac{44}{9}(740,382,767)\right]$$

$$= 0.00032467$$

95% Confidence Interval

The 95% Confidence Interval is formed by:

$$\left[\hat{P} - 1.96\sqrt{\widehat{var}(\hat{P})}, \hat{P} + 1.96\sqrt{\widehat{var}(\hat{P})}\right]$$
$$\left[0.0532 - 1.96\sqrt{0.00032467}, 0.0532 + 1.96\sqrt{0.00032467}\right]$$
$$[0.0178434, 0.08847594] \rightarrow [0.0178, 0.0885]$$

INTRACLASS CORRELATION

The intraclass correlation is estimated by:

$$\hat{\rho} = \frac{msc - mse}{msc + (m-1)mse}$$

$$= \frac{0.039744 - 0.039636}{0.039744 + (19-1)0.039636}$$

$$= 0.0001$$

APPENDIX G: SUMMARY OF PARAMETERS AND RESULTS FOR 7 NIGERIAN STATES

Summary of Parameters and Results for 7 States for the indicator "Percentage of Children 0-59 Months of Age Who Slept Under and ITN/LLIN the Night Preceding the Survey"

I. Establishing SA Sample Size

	Total Number of Supervision Areas	Sample Size per SA	Total Population: Census	Average Square of SA Population	Estimate of Design Effect	Estimate of Intraclass Correlation	Number of SA to Sample
Kano	44	19	5,792,097	20,820,689,926	1.54	0.025	9
Jagawa	27	19	2,875,525	13,555,490,675	1.54	0.025	9
Gombe	11	19	1,489,120	21,238,046,074	1.54	0.025	8
Bauchi	20	19	2,861,887	25,399,210,182	1.54	0.025	9
Rivers	23	19	3,187,864	26,737,333,186	1.54	0.025	10
Akwa Ibom	31	19	2,409,314	7,239,849,229	1.54	0.025	9
Anambra	21	19	2,796,475	21,570,376,364	1.54	0.025	9

Summary of Parameters and Results for 7 States for the indicator "Percentage of Children 0-59 Months of Age Who Slept Under and ITN/LLIN the Night Preceding the Survey"

II. Calculating Indicator Proportion

	Total Population Using the Sampled SA Population to Estimate	Sum of SA Population times SA Proportion	Coverage Proportion
Kano	6,089,629	66,216	0.0532
Jagawa	2,936,856	66,148	0.0676
Gombe	1,350,924	29,783	0.0346
Bauchi	2,409,933	16,179	0.0149
Rivers	2,634,588	30,276	0.0264
Akwa Ibom	2,888,435	13,979	0.0167
Anambra	3,022,367	47,064	0.0363

Summary of Parameters and Results for 7 States for the indicator "Percentage of Children 0-59 Months of Age Who Slept Under and ITN/LLIN the Night Preceding the Survey"

III. Calculating Variance, 95% Confidence Intervals and Intraclass Correlation Using LC-LQAS Data								
	Total of Squared SA Population	Sum of SA Population Square times SA Variance	Between SA Variance	MSC	Variance	95% Confidence Interval	MSE	Updated Estimate of Intraclass Correlation
Kano	211,700,714,096	740,382,767	0.002092	0.039744	0.000325	[0.018, 0.089]	0.039636	0.0001
Jagawa	128,130,441,212	461,589,896	0.005099	0.096876	0.000615	[0.019, 0.116]	0.04808	0.0507
Gombe	127,209,981,789	155,579,804	0.004217	0.080123	0.000398	[0.000, 0.074]	0.041771	0.0461
Bauchi	159,308,923,676	64,519,488	0.001543	0.029312	0.0001396	[0.000, 0.038]	0.022742	0.0150
Rivers	193,523,402,093	89,923,430	0.003150	0.059852	0.000292	[0.000, 0.060]	0.039766	0.0259
Akwa Ibom	86,838,905,462	69,711,086	0.001386	0.026332	0.000150	[0.000, 0.041]	0.016894	0.0286
Anambra	229,850,628,060	338,622,685	0.00541	0.102792	0.00051	[0.000, 0.081]	0.036387	0.0876

REFERENCES

[1] Anker, M. (1991). "Epidemiological and statistical methods for rapid health assessment: introduction." World Health Statistics Quarterly **44**: 94-97.

[2] Campos, F., J. Valadez, et al. (2002). The Nyasaland Network: restoring community health activities in Nicaragua after Hurricane Mitch, a final evaluation (December 1999-October 2001). Washington, DC, NGO Networks for Health.

[3] Central Bureau of Statistics (CBS) [Kenya], Ministry of Health (MOH) [Kenya], and ORC Macro. 2004. *Kenya Demographic and Health Survey 2003*. Calverton, Maryland: CBS, MOH, and ORC Macro.

[4] Dodge, H. F. and H. G. Roming (1959). Sampling Inspection Tables. Single and Double Sampling, 2nd ed. New York, John Wiley.

[5] Fenn, B., S. S. Morrisa, et al. (2004). "Do childhood growth indicators in developing countries cluster? Implications for intervention strategies." Public Health Nutrition **7**: 829-834.

[6] Fleiss, J. L. (1981). Statistical Methods for Rates and Proportions. Wiley-Interscience.

[7] Hoshaw-Woodard, S. (2001). Description and comparison of the methods of cluster sampling and lot quality assurance sampling to assess immunization coverage. Geneva, World Health Organization.

[8] Khoromana, C. O., C. C. Campbell, et al. (1986). "In vivo efficacy of chloroquine treatment for Plasmodium falciparum in Malawian children under five years of age." Journal of Tropical Medicine and Hygiene **35**(3): 463-471.

[9] Lanata, C. F. and R. E. Black (1991). "Lot quality assurance sampling techniques in health surveys in developing countries: advantages and current constrains." World Health Statistics Quarterly **44**: 133-139.

[10] Lemeshow, S. and G. Stroh (1989). "Quality assurance sampling for evaluating health parameters in developing countries." Survey Methodology Statistics Canada **15**(1): 71-81.

[11] Lemeshow, S. and S. Taber (1991). "Lot quality assurance sampling: single- and double-sampling plans." World Health Statistics Quarterly **44**: 115-132.

[12] Lohr, S. L (1999). Sampling: Design and Analysis. Duxbury Press.

[13] Lwanga, S. K. and N. Abiprojo (1987). "Immunization coverage surveys: methodological studies in Indonesia." Bulletin of the World Health Organization **65**(6): 847-853.

[14] Mukaire, J., D. Kaweesa, et al. (2004). Assessment of HIV/AIDS Related Knowledge, Practices and Coverage In 19 Districts Supported by the Uganda AIDS Control Project. Kampala, Uganda AIDS Control Project.

[15] National Population Commission (NPC) [Nigeria] and ORC Macro. 2004. *Nigeria Demographic and Health Survey 2003.* Calverton, Maryland: National Population Commission and ORC Macro.

[16] Reinke, W. A. (1988). Industrial sampling plans: prospects for public health applications. Johns Hopkins Institute for International Programs Occasional Paper No. 2. Baltimore, Johns Hopkins University.

[17] Ridout, M. S., Demtrio, C. G. B., et al. (1999). "Estimating Intraclass Correlation for Binary Data." Biometrics **55**(1), 137-148.

[18] Robertson, S. E., M. Anker, et al. (1997). "The lot quality technique: a global review of applications in the assessment of health services and disease." World Health Statistics Quarterly **50**: 199-209.

[19] Robertson, S. E. and J. J. Valadez (2006). "Global review of health care surveys using lot quality assurance sampling (LQAS), 1984-2004." Social Science and Medicine **63**: 1648-1660.

[20] Robertson, S. E. and J. J. Valadez (2006). "Global review of health care surveys using lot quality assurance sampling (LQAS), 1984-2004." Social Science and Medicine **63**: 1648-1660.

[21] Smith Gordon, S. (1989). "Development of rapid epidemiological assessment methods to evaluate health status and delivery of health services." International Journal of Epidemiology **18**(4): S2-S15.

[22] Stroh, G. (1983b). Rationale for the development and trial of Q.C. sampling designs in program evaluation. Geneva, World Health Organization.

[23] Valadez, J., C. Leburg, et al. (2001). Network for Health in Armenia: baseline survey for three partner organizations (ADRA, CARE, and Save the Children). Washington, DC, NGO Networks for Health.

[24] Valadez, J., W. Weiss, et al. (2003). Assessing community health programs: using LQAS for baseline surveys and regular monitoring. A participant's manual and workbook. St Albans, Teaching-aids at Low Cost (TALC).

[25] Valadez, J., W. Weiss, et al. (2003). Assessing community health programs: using LQAS for baseline surveys and regular monitoring. A trainer's guide. St Albans, Teaching-aids at Low Cost (TALC).

[26] Valadez, J. J. (1991). Assessing child survival programs in developing countries: testing lot quality assurance sampling. Boston, Harvard University Press.

[27] Valadez, J. J., L. D. Brown, et al. (1996). "Using lot quality assurance sampling to assess measurement in growth monitoring in a developing country's primary health care system." International journal of Epidemiology **25**(2): 381-387.

[28] Valadez, J. J. and B. R. Devkota (2002). Decentralized supervision of community health programs: using LQAS in two districts of Southern Nepal. Chapter 9. Community-Based Health Care: Lessons from Bangladesh to Boston. J. Rhode and J. Wyon. Boston, MSH: Chapter 9.

[29] Valadez, J. J., J. Mobley, et al. (2003). Evaluation survey results for 8 partner organizations: ADRA, Adventist Health Services, Word Alive Ministries International, DAPP, Ekwendeni Hospital, MACRO, Malamulo Hospital, and PAPHAM. Washington DC, NGO Networks For Health Project: 71.

[30] Valadez, J. J. and P. Nsubuga (2004). "Learning by Doing: Uganda's AIDS Control Project Empowers Local Managers." Development Outreach **6**(2): 18:22.

[31] Valadez, J. J., R. Transgrud, et al. (1997). "Assessing family planning service-delivery skills in Kenya." Studies in Family Planning **28**(2): 143-150.

[32] Valadez, J. J., W. V. Vargas, et al. (1988). Final report: quality control of primary health care in Costa Rica, Ministry of Health of Costa Rica, Harvard Institute for International Development, Pan American Health Organization, and PRICOR.

[33] Valadez, J. J., B. Weiss, et al. (2003). Assessing Community Health Programs: Using LQAS for Baseline Surveys and Regular Monitoring. London, Teaching-aids at Low Cost.

[34] Wolff, M. C. and R. E. Black (1989). Manual for conducting lot quality assessments in oral rehydration therapy clinics. Baltimore, Johns Hopkins University School of Hygiene and Public Health, Institute for International Programs.

HEALTH, NUTRITION,
AND POPULATION

HNP
HEALTH, NUTRITION, AND POPULATION
THE WORLD BANK

HUMAN DEVELOPMENT NETWORK

THE WORLD BANK

About this series...

This series is produced by the Health, Nutrition, and Population Family (HNP) of the World Bank's Human Development Network. The papers in this series aim to provide a vehicle for publishing preliminary and unpolished results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character. For free copies of papers in this series please contact the individual authors whose name appears on the paper.

Enquiries about the series and submissions should be made directly to the Editor Homira Nassery (hnassery@worldbank.org) or HNP Advisory Service (healthpop@worldbank.org, tel 202 473-2256, fax 202 522-3234). For more information, see also www.worldbank.org/hnppublications.

THE WORLD BANK

1818 H Street, NW
Washington, DC USA 20433
Telephone: 202 473 1000
Facsimile: 202 477 6391
Internet: www.worldbank.org
E-mail: feedback@worldbank.org