VOLUME 20 NO 12 PP 1756–1770 DECEMBER 2015

Comparing two survey methods of measuring health-related indicators: Lot Quality Assurance Sampling and Demographic Health Surveys

Sarah C. Anoke[1], Paul Mwai[1], Caroline Jeffery[2], Joseph J. Valadez[2] and Marcello Pagano[1]

1 *Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA*
2 *Department of International Public Health, Liverpool School of Tropical Medicine, Liverpool, UK*

Abstract

OBJECTIVES Two common methods used to measure indicators for health programme monitoring and evaluation are the demographic and health surveys (DHS) and lot quality assurance sampling (LQAS); each one has different strengths. We report on both methods when utilised in comparable situations.

METHODS We compared 24 indicators in south-west Uganda, where data for prevalence estimations were collected independently for the two methods in 2011 (LQAS: n = 8876; DHS: n = 1200). Data were stratified (e.g. gender and age) resulting in 37 comparisons. We used a two-sample two-sided Z-test of proportions to compare both methods.

RESULTS The average difference between LQAS and DHS for 37 estimates was 0.062 (SD = 0.093; median = 0.039). The average difference among the 21 failures to reject equality of proportions was 0.010 (SD = 0.041; median = 0.009); among the 16 rejections, it was 0.130 (SD = 0.010, median = 0.118). Seven of the 16 rejections exhibited absolute differences of <0.10, which are clinically (or managerially) not significant; 5 had differences >0.10 and <0.20 (mean = 0.137, SD = 0.031) and four differences were >0.20 (mean = 0.261, SD = 0.083).

CONCLUSION There is 75.7% agreement across the two surveys. Both methods yield regional results, but only LQAS provides information at less granular levels (e.g. the district level) where managerial action is taken. The cost advantage and localisation make LQAS feasible to conduct more frequently, and provides the possibility for real-time health outcomes monitoring.

keywords monitoring and evaluation, stratified sampling, cluster sampling, lot quality assurance sampling, demographic and health survey, Uganda

Introduction

The importance of monitoring and evaluation (M&E) to assess interventional programmes, inform allocation of resources and improve evidence-based policy has been commented on by several authors [1–3]. Two common sampling and survey methodologies used to track health programme indicators for M&E are the demographic and health surveys (DHS) [4] and lot quality assurance sampling (LQAS) [5].

DHS and LQAS differ in structure because they serve different purposes: DHS for international comparisons and benchmarking, LQAS for intranational comparisons, benchmarking and health system management. A unique benefit of LQAS is the 'locality' of the methodology. LQAS gives local (e.g. subdistrict, county or subcounty) information, which, if need be, can subsequently be further aggregated into district and regional information. The disaggregation helps overcome the ecological fallacy problem, the assumption that all subregions perform at the regional mean. Additionally, LQAS gives more distributive information about how the subregional estimates vary across the region, which allows for identification of geographical disparities.

Further, LQAS surveys are shorter, cheaper to implement, and the data obtained are readily available. With regard to this last point, LQAS data are hand tabulated within a week of data collection to permit district managers to classify subdistrict units according to predetermined coverage targets; also, more formal reports with districts and regional prevalence measures can be produced within 6 weeks of data collection. Thus, the surveys can be done more frequently, perhaps within the three- to five-year interim between DHS implementations. This increased frequency of measurement allows LQAS data to be used for health system management whereas

DHS data, because of the need for international consistency, take several months after collection to process and several additional months to compile into a final report. The increased frequency of LQAS surveys also positively impacts the building of local capacity, because local district teams incorporate LQAS data collection into their regular health system responsibilities, whereas a DHS may temporarily employ individuals every few years.

An LQAS survey also is flexible and can be adapted to obtain information most useful for programme management; survey items relevant to the region of implementation are easily added or removed, and these modifications do not hinder either the data collection process or the data analysis. Comparatively, a DHS is a large and expensive undertaking, making it difficult to modify the data collection and analysis process. This inertia, combined with the DHS' occasional reference as a 'gold standard', underscores the importance of identifying the best use of a specific survey tool, rather than assuming it serves all informational purposes.

Finally, another advantage of LQAS is that the data are almost real time in that the data collectors see the immediate and local impact of the data they collect, as opposed to a detached central 'black box' repository and its distant possible impact on health policy. This may favourably affect the quality of the data, and it certainly influences the cost of providing national, or aggregated summaries, as the inputs to such summaries are the data that were gathered to provide local information, an aim that presumably justifies the cost of obtaining the data. Thus, the marginal cost of aggregation is minimal compared to the cost of acquiring the data.

The goal of this study is to provide substantive evidence to support the above claims about LQAS' relative utility, by conducting a formal statistical comparison of indicators common between the two surveys. These indicators cover several aspects of Ugandan public health, such as HIV prevention, malaria treatment and prevention, family planning and reproductive health, sanitation, maternal, newborn and child health, and nutrition.

Methods

Selection of region and indicators for comparison

We selected Uganda for this comparison because data exist from both DHS and LQAS surveys collected around the same time: between July and August 2011 for the LQAS, and between June and December 2011 for the DHS.

DHS is a national survey; the sample collected represents all 112 districts in Uganda. Seventy-eight of these districts are engaged in USAID-funded projects that use LQAS for their monitoring. The best geographic overlap between the two surveys is in the DHS-defined south-west region, where LQAS surveys were conducted in each of this region's constituent districts. In this study, we compare indicators calculated for the south-west region.

The choice of indicators to compare started with a 'core set' of 59 national indicators created to track social service performance in Uganda. This list was created by a Technical Working Group of the USAID-funded STAR-E LQAS project comprising representatives from several Ugandan institutions, projects and programmes. Twenty-five LQAS indicators had definitions comparable to those contained in the DHS Final Report; we report on 24 of these comparisons. We replicated all but one DHS result using the DHS data set supplied by Inner City Fund (ICF) International. The indicator for which we could not reproduce the reported DHS estimate and the 33 LQAS indicators we did not find within the DHS Final Report were omitted.

Sampling schemes and data collection

The DHS Programme is implemented by ICF International under contract from the U.S. Agency for International Development (USAID) [4]. The Programme administers several surveys internationally, including the eponymous demographic and health survey (DHS). Although there is a general structure, each survey is tailored to the needs of the specific country. Here, we discuss the structure of the 2011 Ugandan DHS (UDHS), which was implemented jointly with the Uganda Bureau of Statistics (UBOS).

As discussed in the 2011 UDHS Final Report, the sample for the 2011 UDHS was designed 'to provide population and health indicator estimates for the country as a whole and for urban and rural areas separately' as well as for 10 regions, whose boundaries are administratively defined by the DHS Programme [6]. This two-stage stratified cluster sample was selected by sampling households in each of 405 clusters, where stratification was by urban/rural status and region. The sampling frame for the selection of the clusters was the 2002 Population Census provided by UBOS. A three-month household listing operation was conducted in the 405 selected clusters, starting in April 2011. Data collection took place over a six-month period, from the end of June 2011 to early December 2011. Women aged 15–49 years in all households and men aged 15–54 in one-third of households were eligible for interview.

In the first stage of sampling within the south-west region, 40 clusters were selected from a total of 8369 with 7983 being rural and 386 urban. The 40 selected clusters comprise five urban and 35 rural areas. In the second stage of sampling, the DHS sampled 1200 of 685 695 households; 150 were urban and 1050 rural. The expected number of completed interviews for the region was 1097 (96.3% completed) for women 15–49 years and 477 (92.6% completed) for men 15–54 years. We report the actual sample sizes with the results. The national DHS first stage of sampling comprised 405 clusters selected from 48 715 clusters (42 675 rural, 6040 urban), and included 119 urban and 286 rural areas. The second stage comprised 12 150 households (8580 rural, 3570 urban) of 5 076 534 households. The expected number of completed interviews was 9885 for women 15–49 and 3628 for men 15–54.

The three subsurveys of interest are the household survey, the women's survey (asked in all households), and the men's survey (asked in approximately every third household). All three subsurveys were conducted within the same household.

The LQAS methodology is a health science derivative of Statistical Quality Control, a set of tools developed by Dodge and Romig, and Shewhart [7]. The data are sampled from a local administrative unit called a supervision area (SA; e.g. county, subcounty or parish within a district), which is classified as 'acceptable' or 'unacceptable' according to a coverage target. Although the goal is classification, it is also possible to aggregate SA-level data to construct prevalence estimates for the respective districts and regions; here, the classification decisions do not in any way impact the estimation of indicators [5, 7].

LQAS in Uganda during 2011 included more than 11 400 interviewees; the sample of 8676 households in the south-west was also selected using a stratified two-stage process. Districts in south-west region were divided into SAs based on how the district managed health services. Within each SA, a sample of 19 or 24 villages was selected with probability of selection into the sample proportional to the village population size (PPS). To maintain an approximate minimum district sample size of 96, districts with only 4 SAs required an SA sample size of 24 (4 SAs × 24). In each selected village, the interviewer constructed a map of the village with the help of a chief or other local leader, and divided the map into equivalent segments based on visible landmarks and the number of households in each segment. One segment was selected randomly. The interviewer then enumerated the households in the selected segment and selected one randomly. If the selected segment had 30 or more households, it was further segmented and a subsegment selected randomly; all households in the final segment were enumerated and one chosen randomly. To accommodate the fact that there could be a nearby household with zero probability of selection (e.g. it was omitted from the map because it was hidden behind vegetation), the next house with the closest door was selected for the first interview. Thereafter, the household with the next closest door was selected for each subsequent subpopulation. Only one individual from each subpopulation was interviewed in the sampled village.

The five subsurveys of interest correspond to particular subpopulations: mothers of children 0–11 months, mothers of children 12–23 months, women 15–49 years, men 15–54 years and youth 15–24 years. All five subsurveys were conducted in different households, comparatively different from what was employed by DHS. To accomplish this, from a randomly selected house, an interviewee is selected who is either a woman aged 15–49 years, a man aged 15–54 years, the mother of an infant aged 0–11 months, the mother of an infant aged 12–23 months or a youth aged 15–24 years. Subsequent households were selected to find interviewees from the remaining populations, taking care not to select two interviewees from the same household.

Weighting

Within both the UDHS and LQAS data sets, individuals had different probabilities of being sampled. To construct valid, representative estimates from these data, we calculated sampling weights based on each sampling design.

DHS. In the 2011 UDHS, sampling weights were calculated based on the two-stage stratified cluster design used to sample households (See Appendix A.4 of [6] for details). These weights are provided within the 2011 UDHS data set.

LQAS. In the LQAS data, we calculated weights based on the two-stage stratified design used to sample households. Within each SA, a fixed number (either 19 or 24) was sampled irrespective of the SA population size. To adjust for differences in SA sample sizes, individual observations are weighted by the number of individuals a response represents. For example, if an observation is one of 19 sampled from an SA with a population of 2000, then each observation is weighted by 2000/19. In another SA, if an observation is one of 24 sampled from an SA with a population of 4500, then each observation is weighted by 4500/24. We use these weights to construct a representative district point estimate, and a representative regional point estimate (Figure 2).

Sampling errors

DHS. The DHS Programme provides a formula in Appendix B of the 2011 UDHS Final Report [6] for calculating sampling errors based on the two-stage stratified cluster design used to sample individuals. For indicators considered to be of 'primary interest' by the DHS Programme, sampling errors are provided in the report. Where possible, we use these sampling errors. For indicators where sampling errors are not provided, we calculated them using the formulae provided.

LQAS. The survey data software within Stata® 13 was used to calculate standard errors at both the district and regional levels [8]. For details on the formulae used, refer to the Stata Survey Data Reference Manual [9]. At the regional level, we used the Wilson score interval to construct confidence intervals [10].

Statistical comparison of indicators

A two-sample two-sided *Z*-test of proportions was used to test whether the proportions as estimated from DHS data and LQAS data were statistically equivalent. Standard errors for test statistics were calculated by taking the square root of the sum of the squared standard errors from the two estimated proportions. In two cases (Table A4), it was necessary to calculate a weighted average and accompanying standard error of two LQAS subpopulation estimates for comparison to a single DHS measure. The weights used were the proportion of the aggregated sample that belonged to a particular subpopulation. For example, for an aggregated sample consisting of members from two subpopulations with 1353 and 752 members, respectively, the corresponding weights are 1353/(1353+752) and 752/(1353+752).

Results

Regional comparisons

The 24 selected indicators cover several aspects of Ugandan public health; including HIV knowledge, counselling, and behaviour (8 indicators), malaria treatment and prevention (3), family planning & reproductive health (4), child health (3), nutrition (4) and sanitation (2). The results of the 37 comparisons are summarised as a forest plot (Figure 1). Point estimates, confidence intervals and the results of statistical comparisons are shown in the Appendix (Tables A1–A8). In Tables A9 and A10 (also in the Appendix), we summarise our comparisons. For 6 indicators (Table A1 and TableA3), we refine the comparison by making subpopulation comparisons (e.g. men, women, male youths, female youths) resulting in additional comparisons. In total, we assessed 38 comparisons; 1 comparison using a cohort of male youths (Table A3) was eliminated due to the UDHS having insufficient comparable data, thereby reducing the number of comparisons to 37. We did not reject equality of the proportions in 21 of 37 (56.8%). The average difference between LQAS and DHS estimates for the 37 comparisons was 0.062 (SD = 0.093; median = 0.039). The average difference among the 21 failures to reject equality of proportions was 0.010 (SD = 0.041; median = 0.009); among the 16 rejections, it was 0.130 (SD = 0.010, median = 0.118). As the large standard deviation, and lower median value compared to the mean indicate considerable variation among these rejections, we examined the variation further. Seven of the 16 rejections exhibited differences of <0.10, which are clinically (or managerially) not significant; five more had differences >0.10 and

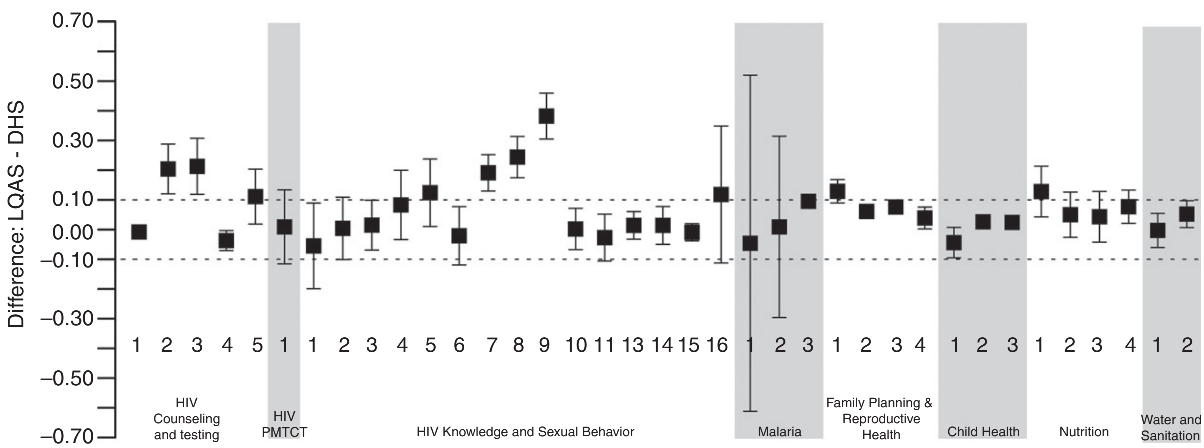

Figure 1 A forest plot of 37 comparisons of DHS and LQAS data collected in south-west Uganda during 2011.

<0.20 (mean = 0.137, SD = 0.031) and 4 differences were >0.20 (mean = 0.261, SD = 0.083). We consider the more interesting of the 16 rejections in the Discussion below.

Distribution of prevalences across districts

The limit of inference when using UDHS data is at the regional level; however, district health system managers cannot use such results without making the strong assumption that the districts within the region perform similarly, with the regional estimate reflective of the overall mean. This assumption is unnecessary, and indeed, becomes a testable hypothesis, when making inferences using LQAS data, because we are able to provide information at both the regional and subregional (i.e. district) levels. This information includes identification of highly and poorly performing districts (and highly and poorly performing SAs within the district), and a measure of the geographic variability of the regional estimator.

To illustrate this point, in Figures 2–4 are maps of south-west region displaying the 14 constituent districts with population sizes, and prevalence estimates calculated using LQAS data from that district for two indicators (contraceptive prevalence, and fully vaccinated children 12–23 months of age). Each smaller filled circle represents one of the 40 DHS clusters sampled from this region; note that the DHS prevalence is estimated such that the comparative map would contain a single colour covering the whole region. In the lower portion of each of these maps is the overall regional prevalence from both surveys.

Discussion

Discrepancies between prevalence estimates

When comparing two indicators, we first need to ensure that the indicators are measuring the same phenomenon. This is often difficult to ensure when the two are defined in different surveys by different individuals. Our choice

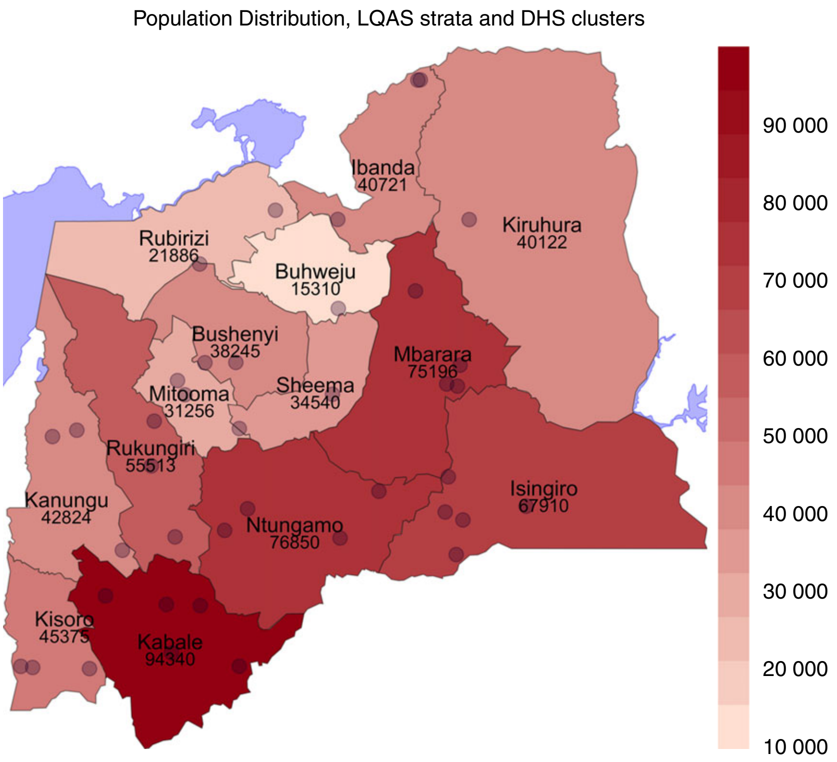

Figure 2 Population distribution by district, across the 14 districts of the south-west region.§
§These population counts were calculated from LQAS sampling frames created during sampling of the data used in this writing, and were used to calculate the weighted regional prevalence estimate for each LQAS indicator. The 40 clusters that were sampled by the DHS Programme for inclusion in their survey are denoted by translucent circles. The LQAS population counts are congruous with the distribution of DHS clusters, which were selected based on a distribution proportional to the population density.

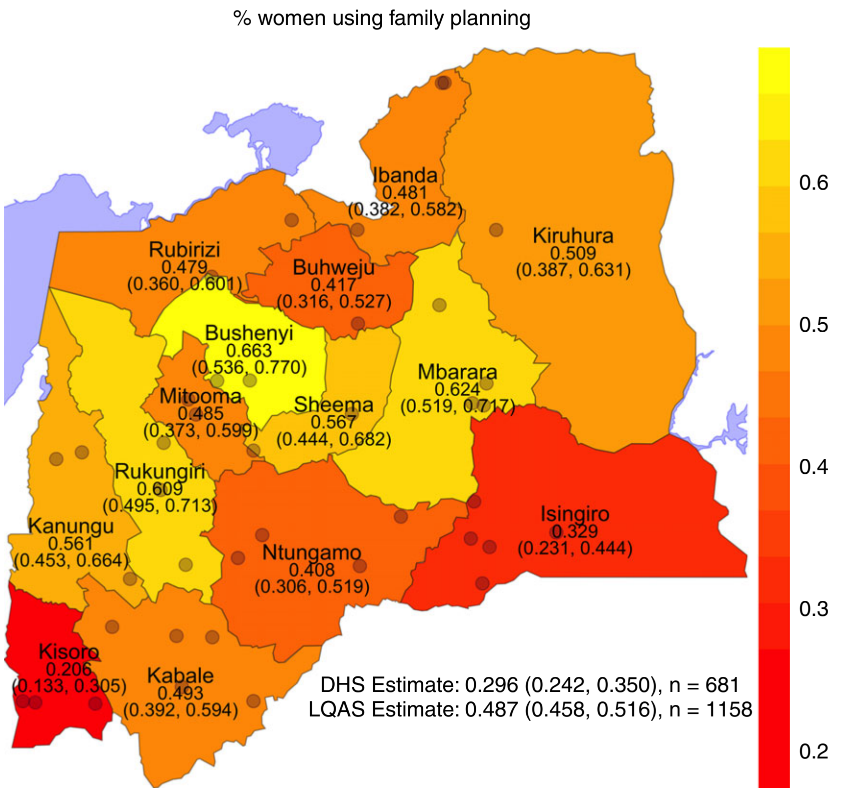

Figure 3 Distribution of the indicator '% of currently married women who are using any family planning method' across the 14 districts of the south-west region.§,¶ Test for homogeneity of prevalences: $P < 0.0005$.
§The 40 clusters that were sampled by the DHS Programme for inclusion in their survey are denoted by translucent circles.
¶Refer to Table A5.

of indicators to compare was influenced by how closely we could achieve comparability of indicators. Secondly, if two indicators are supposedly estimating the same quantity and the results differ, it is not possible, without importing extra information into the argument which we do not have available, to determine which indicator yields an answer that is closer to the 'truth'. With these *caveats,* we failed to find disagreement in 21 comparisons and another 7 show clinically insignificant difference (75.7%). However, there are discrepancies that reveal subtle differences between the UDHS and LQAS surveys. We discuss only a selection of extreme discrepancies to perhaps find explanation for these and other differences. For example, consider the 'HIV Counselling and Testing' indicators (Table A1), where, across all subpopulations, three of the five comparisons failed to disagree. While two indicators were found to be statistically different, their values are still reasonably close and clinically insignificant. For the five 'HIV Knowledge and Sexual Behaviour' indicators (Table A3), four failed to disagree for almost all subpopulation comparisons. For the indicator reporting the percentage of individuals who have had sexual intercourse with a non-marital or non-cohabiting sexual partner, the LQAS estimates were higher for all subpopulations. However, three of the four differences were clinically insignificant. In this example, the statistical difference masks the similarity of the prevalence estimates when considered from the point of view of the health system manager.

For the 'Prevention of Mother-to-Child Transmission' (PMTCT) indicator (Table A2), there was a significant difference. We believe this is attributable to the differing construction of the two indicators; the DHS asks several questions of respondents about receiving specific information related to PMTCT, while the LQAS survey asks a general question about whether the mother has received information about PMTCT.

Next, consider the indicator '% of mothers of children 0–11 months who received two of more doses of SP/Fansidar during their last pregnancy' (Table A4). From the way the corresponding DHS women's questionnaire item is structured (Item #425), respondents are asked to volun-

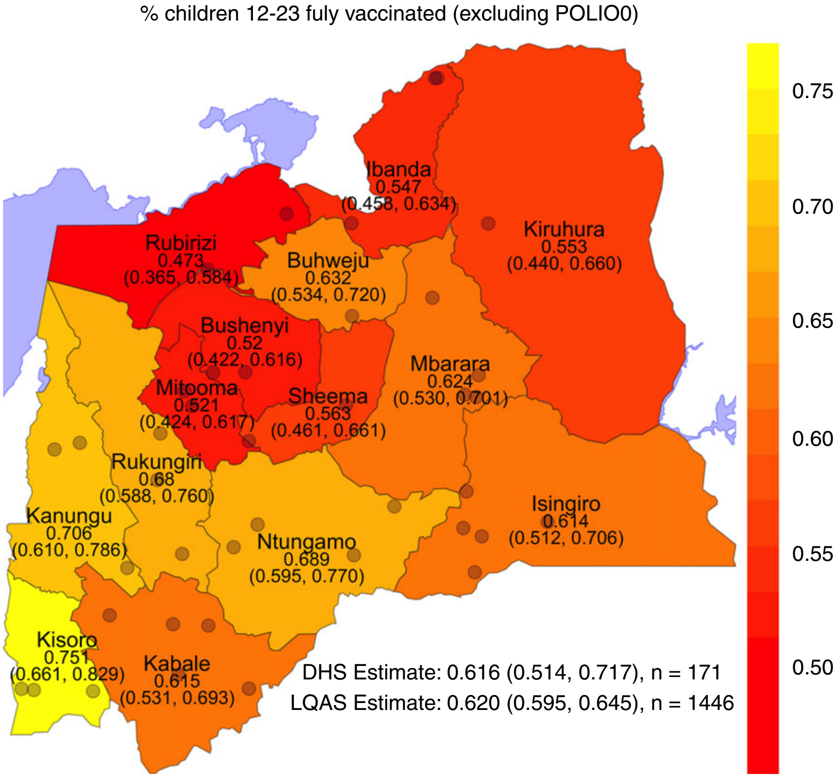

Figure 4 Distribution of the indicator '% of children 12–23 months who are fully vaccinated' under *Definition 2* (1 BCG + 3 DPT + 3 POLIO + MEASLES) across the 14 districts of the south-west region.§,¶ Test for homogeneity of prevalences: $P = 0.002$.
§The 40 clusters that were sampled by the DHS Programme for inclusion in their survey are denoted by translucent circles.
¶See Table A6

teer the name of their antimalarial; if the respondent does not know the name of the antimalarial, they are shown the packages of medications to support their response. In the LQAS survey interview, respondents are also asked to volunteer the name of their antimalarial, but the packets of medications are not shown.

Another discrepancy is '% of households using iodised salt' (Table A7), where the DHS estimate is higher than the LQAS estimate in two circumstances. This difference could reasonably be attributed to the methods used by the interviewer to determine the presence of iodised salt. During the DHS interview, the interviewer asks the respondent for a teaspoonful of cooking salt and performs a chemical test for presence of iodine (Household Questionnaire Item #140). During the LQAS survey interview (Mothers of children 12–23 months Questionnaire Item #514), the interviewer requests the household's salt packet and checks the packaging for indication of iodization. In short, there is no chemical testing and the package may underreport the presence of iodine. We must also take into account that the DHS uses a representative sample of all households whereas the LQAS uses a representative sample of households with mothers of children 12–23 months of age. The former comprises a population with more variation and could include a confounder associated with purchasing of iodised salt. Nevertheless, the populations are not equivalent. When we extract the households with children 12–23 months from the DHS for comparison with the LQAS, the results are closer (95.9% *vs*. 92.2%) but we compare an LQAS sample of $n = 1371$ with a DHS cluster sample of $n = 171$. The power in the LQAS sample to detect small differences may be the reason for this statistically significant but clinically insignificant result.

An additional discrepancy is '% of households with safe water supply' (Table A8), but an explanation for the difference is not as readily available as for the previous three examples. In comparing the available option responses in the two surveys for 'source of drinking water', we see that they are largely the same with two exceptions: 'public tap/standpipe' and 'protected spring'. Both of these safe water sources that are included as

DHS response items but are not LQAS response items. Exclusion of these response items from the numerator of the DHS indicator only further exacerbates the difference between the two prevalence measures.

One discrepancy worth mentioning concerns an indicator we omit from the final analysis due to lack of definition compatibility, namely '% of mothers of children aged 0–11 months who took iron supplementary tablets for at least 90 days during last pregnancy'. The estimate from DHS data yielded 0.044 [95% CI (0.007, 0.082) with n = 205] while the LQAS data yielded 0.776 [95% CI (0.754, 0.797) with n = 1446]. We believe this discrepancy is caused by the way the questions are asked of the respondents. Within the DHS Women's Questionnaire, respondents are asked 'How many days did you take iron tablets during your last pregnancy?' and provide an integer. Within the LQAS Mothers of children aged 0–11 months questionnaire, respondents are asked 'Did you take iron tablets for at least 90 days during your last pregnancy?' and provide a yes or no. The estimation goals of the two questions are different; the DHS wanted to report an average number of days, and the LQAS wanted a binary classification.

Differences between prevalence estimates, such as those discussed above, do not mean that one estimate is correct and the other is not. Rather, these differences expose differences in questionnaire items and interviewer protocols that can lead to the improvement of both surveys. Prevalence estimates that are similar lend support to the other, leading us to believe that the calculated estimate may be close to reality.

Comparison of costs

It is interesting but difficult to compare the costs of LQAS with those of DHS as the purpose for their respective uses is different. One clear difference in this Ugandan case is that the DHS is designed to measure indicators at a regional level while the LQAS survey utilises the measures at the district level. Hence, many more district-level samples are collected with the LQAS survey. The only financial data in the literature concerning DHS costs come from the 1991, 1994, 1996, 1999 Tanzania surveys [11]. That study took all expected recurrent and non-capital costs, divided by the number of participating households times the national estimate of average household size for 2000–01. Oddly, as this results in a lower cost estimate, all members of the household were considered as participants, rather than just those interviewed. The cost was $19.57 per participant (or $25.25 in 2013 dollars). Using this information to estimate the cost per interview in the 2011 UDHS, which includes a household and a women's survey in the same household, and men in every third sampled house, the cost per interview was $57.94 (or $130.37 per household in 2013 dollars).

The cost data for LQAS come from a detailed cost study in Costa Rica [5] and a comparative assessment from 2002 of three USAID projects in Nepal, Nicaragua and Armenia [12]. LQAS promotes the engagement of District Health Managers as a cost-saving mechanism as their costs are already paid by the Ministry of Health. These in-kind costs are included in this analysis as an LQAS cost. Taking into account that LQAS uses parallel sampling of interviewees (all in different households), the cost per interview is $11.17, using the first index household as the reference (or $29.28 per household in 2013 dollars). In these examples, LQAS was at least 4.5 times less expensive than DHS for each household participating in the survey and 5.2 times less expensive for each interview. We note though the UDHS used a questionnaire more extensive than that of the LQAS survey, and included height and weight measurement, blood specimen collection for on-site anaemia and laboratory vitamin A testing. An extensive questionnaire and biological measurement does increase the costs of a DHS.

Surveys are complementary, not redundant

From the prevalence comparisons, we see that as a secondary by-product the LQAS survey provides very similar information to that of the DHS. Twenty-one of 37 comparisons for the 25 selected indicators failed tests of statistical difference, including important measures of HIV knowledge and sexual behaviour, malarial prophylaxis, child vaccination and nutrition. Seven statistical differences were clinically insignificant resulting in a failure to find meaningful difference in 75.7% of the comparisons. Many of the prevalence estimates that did not agree across the two surveys have reasonable explanations. Other comparisons of LQAS with demographic surveillance systems have proved to have an excellent agreement of results, but in those occasions the indicators were identical [13]. Similarly, reliability studies of LQAS have recently compared data collected by managers who use LQAS results to improve their own programmes with data collected by disinterested data collectors; the concordance of the two data sets was very high [14].

In fact, the information provided by the LQAS survey is a superset of the information provided by the DHS; it provides similar information to that of the DHS, and more. In general, for a fixed sample size, a stratified sampling strategy produces more precise estimates than a cluster sampling strategy. In the case of this particular regional study, where the LQAS survey sample was stratified and the DHS used a cluster sample, for all indicators the LQAS sample size was larger than that of the DHS. This suggests that the LQAS measures are more precise but we do again note that the surveys were designed and conducted for different purposes, so a comparison of sample size is not so straightforward. The large sample sizes also may have led to the statistical differences between the surveys that are not important from a health system management perspective.

It is indeed true that the purposes and intended use of data generated by the two surveys are different. For example, the DHS is designed to collect information on the population of living mothers with children under five years of age, so there is five years of history in every resulting measure. The LQAS survey is designed to collect information on the population of mothers with younger children such as under one year of age, or 12 to 23 months of age so this survey gives information on health system performance from the recent past. This short time frame lends flexibility to the survey, so questionnaire items can be modified and updated based on the most effective direction of healthcare delivery.

However, the stratification of the LQAS sample allows us to investigate the geographic variability of the regional point estimate, exemplified in Figures 3 and 4. Use of such information, in conjunction with demographic information like the population distribution of Figure 2, provides the structure needed for the evidence-based allocation of resources. The LQAS results provide a further and more granular depiction of variability when considering the classification of subdistrict-level supervision areas according to a coverage target. The subdistrict areas (counties, subcounties and parishes in the case of Uganda) are not presented in Figures 3 and 4, but are the main reason for using LQAS, to empower subdistrict managers to manage by quickly available classification results reflecting the current condition of the area for which they are responsible. This is in contrast to DHS data, which are able to give a single estimate for the region that cannot be disaggregated [1]. Although an analyst could consider, alternatively to LQAS, a design akin to a stratified DHS, the analyst would lose many of the advantages particular to LQAS, including the ease of data collection, the timeliness of results, and relatively low financial and human cost.

To our knowledge, this formal comparison of indicators as calculated using LQAS data and DHS data collected within similar time periods is the first of its kind. However, a comparison on the basis of an emulation was reported in [15]. Our findings are quite similar to other comparisons to the LQAS sampling procedure seen in the M&E literature. For example, Singh *et al.* [16.] report consonance of immunization coverage estimates in a region of India as calculated from data using the LQAS sampling method and from data using the 30-cluster survey method of the World Health Organisation's Expanded Programme on immunisation [17]. Bhuiya *et al.* [13] also report agreement of estimates from LQAS data and 'health and demographic system' data collected in Matlab, Bangladesh.

Several individuals involved in global health policy have commented on the need of data at different levels for policy-making and management [2]. As evidenced by our study and similar studies discussed above, the LQAS methodology provides these multilevel data, whereas the DHS, by nature of its design, cannot. The DHS has built a reputation of providing high-quality data for international comparisons; we have shown that LQAS gives the same accuracy, but is programmatically more relevant [1, 3]. Further, LQAS builds local capacity, because regular data collection will lead to its institutionalisation. Chan *et al.* [2] describe this institutionalisation as 'essential', because it strengthens a country's ability to collect, process, analyse and use health data. Also, by virtue of using local health workers to collect LQAS data, it is cheaper than the DHS.

Conclusion

The LQAS sampling method is a viable, timely, and informative complement to the DHS that can be used in interstitial years. It is more-management oriented because of the quick turnaround of data collection and analysis, allowing for targeted, data-driven decisions to be made quickly. This results in timely and local evidence of the value of the data collected and it might also convince local data gatherers of the value of the data gathering effort and result in higher quality data.

Acknowledgements

We thank Stephen Lwanga, Management Sciences for Health (MSH) Country Director for providing logistical support for this research; William Vargas, Charles Nkolo and Joseph Ouma for their support during the implementation of this research; the District Health Officers and health workers of the south-west region of Uganda for their commitment and dedication to the health needs of the population they serve. We would like to thank all other staff members of the STAR-E LQAS project and MSH staff members without whom this work would not have been possible. This work was supported by the American people through USAID under terms of the Cooperative agreement with Management Sciences for Health. SCA was also supported by NIH Grants 5T32AI007358-24 and 5T32AI007358-25.

References

1. Boerma JT, Stansfield SK. Health statistics now: are we making the right investments? *Lancet* 2007: **369**: 779–786.
2. Chan M, Kazatchkine M, Lob-Levyt J *et al.* Meeting the demand for results and accountability: a call for action on health data from eight global health agencies. *PLoS Med* 2010: **7**: e1000223.
3. Murray CJL, Frenk J. Health metrics and evaluation: strengthening the science. *Lancet* 2008: **371**: 1191–1199.
4. Barrere B, Fishel J, McInturff S *et al.* The Demographic and Health Surveys (DHS) Program. 2014.
5. Valadez JJ. *Assessing Child Survival Programs in Developing Countries: Testing Lot Quality Assurance Sampling*. Harvard University Press: Cambridge, 1991.
6. Uganda Bureau of Statistics III. *Uganda Demographic and Health Survey 2011*. Uganda Bureau of Statistics and ICF International Inc.: Kampala, Uganda and Calverton, Maryland, 2012.
7. Pagano M, Valadez JJ. Commentary: Understanding practical lot quality assurance sampling. *J Epidemiol* 2010: **39**: 69–71.
8. StataCorp. *Stata: Release 13. Statistical Software*. College Station, TX: StataCorp LP, 2013.
9. StataCorp. Stata 13 Base Reference Manual. College Station, TX: Stata Press, 2013.
10. Dean N, Pagano M. Evaluating confidence interval methods for binomial proportions in clustered surveys. *Journal of Survey Statistics and Methodology*, in press.
11. Rommelmann V, Setel PW, Hemed Y *et al.* Cost and results of information systems for health and poverty indicators in the United Republic of Tanzania. *Bull World Health Organ* 2005: **83**: 569–577.
12. Grundmann C. The Costs of Using LQAS for Project Management, Monitoring and Evaluation. Washington, DV: NGO Networks for Health Project, 2002 2002. Report No.
13. Bhuiya A, Hanifi SMA, Roy N, Streatfield PK. Performance of the lot quality assurance sampling method compared to surveillance for identifying inadequately-performing areas in Matlab, Bangladesh. *J Health Popul Nutr* 2007: **25**: 37–46.
14. Beckworth CA, Davis RH, Faragher B & Valadez JJ. Can health workers reliably assess their own work? A test-retest study of bias among data collectors conducting a Lot Quality Assurance Sampling survey in Uganda. Health policy and planning. 2014.
15. Biedron C, Pagano M, Hedt BL *et al.* An assessment of lot quality assurance sampling to evaluate malaria outcome indicators: extending malaria indicator surveys. *Int J Epidemiol* 2010: **39**: 72–79.
16. Singh J, Jain DC, Sharma RS, Verghese T. Evaluation of immunization coverage by lot quality assurance sampling compared with 30-cluster sampling in a primary health centre in India. *Bull World Health Organ* 1996: **74**: 269–274.
17. Henderson RH, Sundaresan T. Cluster sampling to assess immunization coverage: a review of experience with a simplified sampling method. *Bull World Health Organ* 1982: **60**: 253–260.

Appendix
Tables A1 to A10 of Statistical Results

Table A1 Comparison of *HIV Counselling and Testing* indicators

Indicator	LQAS estimate	DHS estimate	Comparison (P-value)
% of individuals who were counselled and received an HIV test in last 12 months and know their results.	Women (n = 1445): 0.440 (0.414, 0.465)	Women (n = 1097): 0.388* (0.351, 0.425)*	0.024
	Female youth (n = 781): 0.350 (0.317, 0.384)	Female youth (n = 451): 0.353 (0.306, 0.400)	0.921
	Men (n = 1446): 0.294 (0.271, 0.318)	Men (n = 291): 0.217 (0.166, 0.268)	0.006
	Male youth (n = 633): 0.204 (0.174, 0.237)	Male youth (n = 116): 0.161 (0.081, 0.240)	0.335
% of mothers of children 0–11 months who were counselled and received an HIV test during the last pregnancy and know the results.	Mothers (n = 1446): 0.870 (0.852, 0.886)	Mothers (n = 205): 0.820 (0.746, 0.893)	0.197

*The value is as reported in the 2011 UDHS Final Report. If a quantity is unmarked, it was calculated by the authors for this study. 'Women' are 15–49 years of age, 'men' are 15–54 and 'youth' are 15–24. Prevalences are compared using a two-sample two-sided Z-test of proportions.

Table A2 Comparison of *HIV PMTCT* indicators

Indicator	LQAS estimate	DHS estimate	Comparison (*P*-value)
% of mothers of children 0–11 months who were counselled for 'prevention of mother-to-child transmission' services during last pregnancy.	Mothers (*n* = 1446): 0.913 (0.898, 0.927)	Mothers (*n* = 205): 0.785 (0.701, 0.868)	0.003

Table A3 Comparison of *HIV Knowledge and Sexual Behavior* indicators

Indicator	LQAS estimate	DHS estimate	Comparison (*P*-value)
% of individuals who had sex with more than one sexual partner in the last 12 months.	Women (*n* = 1445): 0.029 (0.021, 0.039)	Women (*n* = 1097): 0.005* (0.001, 0.009)*	0.496
	Female youth (*n* = 781): 0.028 (0.019, 0.042)	Female youth (*n* = 451): 0.002 (0, 0.006)	<0.001
	Men (*n* = 1446): 0.111 (0.096, 0.128)	Men (*n* = 291): 0.155 (0.106, 0.204)	0.095
	Male youth (*n* = 633): 0.072 (0.056, 0.092)	Male youth (*n* = 116): 0.033 (0, 0.067)	0.058
% of individuals who have had sexual intercourse with a non-marital or non-cohabitating sexual partner.	Women (*n* = 1445): 0.086 (0.072, 0.101)	Women (*n* = 1097): 0.010 (0.004, 0.016)	<0.001
	Female youth (*n* = 781): 0.066 (0.051, 0.086)	Female youth (*n* = 451): 0.005 (0, 0.011)	<0.001
	Men (*n* = 1446): 0.198 (0.178, 0.219)	Men (*n* = 291): 0.069 (0.035, 0.102)	<0.001
	Male youth (*n* = 633): 0.111 (0.096, 0.146)	Male youth (*n* = 116): 0.016 (0, 0.040)	<0.001
% of individuals who have had sexual intercourse with a non-marital or non-cohabitating sexual partner in the last 12 months and used a condom at last higher-risk sex.	Women (*n* = 128): 0.317 (0.243, 0.402)	Women (*n* = 11): 0.308 (0.012, 0.605)	0.957
	Female youth (*n* = 54): 0.505 (0.376, 0.634)	Female youth (*n* = 2): 0.551 (0, 1)	0.894
	Men (*n* = 186): 0.426 (0.357, 0.498)	Men (*n* = 20): 0.308 (0.088, 0.527)	0.372
	Male youth (*n* = 82): 0.462 (0.358, 0.569)	Male youth (*n* = 2): 0	n/a
% of youth 15–24 years who have had sexual intercourse before the age of 15.	Female youth (*n* = 781): 0.045 (0.032, 0.061)	Female youth (*n* = 451): 0.054 (0.028, 0.080)	0.532
	Male youth (*n* = 633): 0.076 (0.058, 0.099)	Male youth (*n* = 116): 0.062 (0.001, 0.115)	0.641
% of men who are circumcised.	Men (*n* = 1446): 0.102 (0.087, 0.119)	Men (*n* = 291): 0.088 (0.044, 0.132)	0.561
	Male youth (*n* = 645): 0.072 (0.055, 0.095)	Male youth (*n* = 116): 0.099 (0.022, 0.176)	0.501

*The value is as reported in the 2011 UDHS Final Report. If a quantity is unmarked, it was calculated by the authors for this study. 'Women' are 15–49 years of age, 'men' are 15–54 and 'youth' are 15–24. Prevalences are compared using a two-sample two-sided *Z*-test of proportions.

Table A4 Comparison of *Malaria* indicators

Indicator	LQAS estimate*	DHS estimate	Comparison (*P*-value)
% of children 0–59 months who had fever in the two weeks preceding the survey and received treatment with ACT within 24 h of onset of fever.	0–11 months (*n* = 1353): 0.044 (0.031, 0.064) 12–23 months (*n* = 752): 0.090 (0.071, 0.113) Weighted average for comparison: 0.070 (0.056, 0.084)	0–23 months (*n* = 49): 0.068 (0, 0.138)	0.961
% of mothers of children 0–11 months who received two of more doses of SP/Fansidar during their last pregnancy.	Mothers (*n* = 1446): 0.649 (0.635, 0.684)	Mothers (*n* = 205): 0.267 (0.191, 0.343)	<0.001
% of children 0–59 months who slept under an ITN the night preceding the survey.	0–11 months (*n* = 1446): 0.658 (0.633, 0.682) 12–23 months (*n* = 1446): 0.657 (0.632, 0.681) Weighted average for comparison: 0.657 (0.639, 0.675)	0–23 months (*n* = 412): 0.413 (0.346, 0.481)	<0.001

*Weighted averages were calculated based on the proportion of the aggregated sample that belonged to a particular group.

Table A5 Comparison of *Family Planning & Reproductive Health* indicators

Indicator	LQAS estimate	DHS estimate	Comparison (*P*-value)
% of currently married women who are using any family planning method.	Women (*n* = 1158): 0.487 (0.458, 0.516)	Women (*n* = 1158): 0.296* (0.242, 0.350)*	<0.001
% of mothers of children 0–11 months who attended ANC at least four times during their last pregnancy.	Mothers (*n* = 1446): 0.466 (0.441, 0.492)	Mothers (*n* = 205): 0.487 (0.392, 0.581)	0.677
% of mothers of children 0–11 months who delivered their last baby in a health facility.	Mothers (*n* = 1446): 0.668 (0.644, 0.692)	Mothers (*n* = 205): 0.544 (0.433, 0.655)	0.034
% of mothers of children 0–11 months who were assisted by a skilled health worker during their last delivery.	Mothers (*n* = 1446): 0.645 (0.618, 0.672)	Mothers (*n* = 205): 0.562 (0.448, 0.677)	0.161

*The value is as reported in the 2011 UDHS Final Report. If a quantity is unmarked, it was calculated by the authors for this study. 'Women' are 15–49 years of age, 'men' are 15–54 and 'youth' are 15–24. Prevalences are compared using a two-sample two-sided *Z*-test of proportions.

Table A6 Comparison of *Child Health* indicators

Indicator	LQAS estimate	DHS estimate	Comparison (*P*-value)
% of children 12–23 months who are fully vaccinated. Definition 1 (1 BCG + 3 DPT + 4 POLIO + MEASLES)	12–23 months (*n* = 1446): 0.286 (0.264, 0.310)	12–23 months (*n* = 171): 0.271 (0.190, 0.353)	0.729
% of children 12–23 months who are fully vaccinated. Definition 2 (1 BCG + 3 DPT + 3 POLIO + MEASLES)	12–23 months (*n* = 1446): 0.620 (0.595, 0.645)	12–23 months (*n* = 171): 0.616* (0.514, 0.717)	0.940
% of children 0–11 months with diarrhoea in the last two weeks receiving oral rehydration therapy (ORT).	0–11 months (*n* = 393): 0.176 (0.141, 0.216)	12–23 months (*n* = 46): 0.231 (0.091, 0.371)	0.454

Table A7 Comparison of *Nutrition* indicators

Indicator	LQAS estimate	DHS estimate	Comparison (*P*-value)
% of children under six months of age who are exclusively breastfed.	0–5 months (n = 783): 0.540 (0.503, 0.576)	0–5 months (n = 110): 0.531 (0.412, 0.650)	0.887
% of children 12–23 months receiving vitamin A supplementation in the last six months.	12–23 months (n = 1446): 0.656 (0.631, 0.680)	12–23 months (n = 171): 0.545 (0.456, 0.635)	0.020
% of households using iodised salt.	Households with mothers of children 12–23 months (n = 1372): 0.9218 (0.907, 0.935)	Households[(1)] (n = 1049): 0.984 (0.975, 0.993) Households[(2)] (n = 1128): 0.915 (0.894, 0.937) Children[(3)] (n = 171): 0.959 (0.929, 0.989) [(1)]out of houses that had salt that was tested (denominator includes only houses that had salt that was tested) – DHS uses this. [(2)]out of all non-missing values (denominator includes houses with no salt, and with untested salt). [(3)]out of all children 12–23 months.	[(1)] <0.001 [(2)] 0.609 [(3)] <0.027
% of mothers of children 0–11 months who received vitamin A supplementation within 2 months after delivery.	Mothers (n = 1446): 0.507 (0.482, 0.533)	Mothers (n = 205): 0.294 (0.203, 0.385)	<0.001

Table A8 Comparison of *Water and Sanitation* indicators

Indicator	LQAS estimate	DHS estimate	Comparison (*P*-value)
% of households with safe water supply.	Households (n = 1445): 0.634 (0.609, 0.658)	Households[(1)] (n = 1128): 0.311 (0.252, 0.370) Households[(2)] (n = 1128): 0.430 (0.350, 0.510) [(1)]LQAS safe water is piped, protected well, borehole, rainwater, tanker/truck, bottled water. [(2)]DHS also includes public tap/standpipe and protected spring, which is not in the LQAS questionnaire (may be classified differently within LQAS).	[(1)] <0.001 [(2)] 0.005
% of households with latrine or toilet.	Households (n = 1445): 0.970 (0.959, 0.977)	Households (n = 1128): 0.978 (0.964, 0.992)	0.381

Table A9 Summary of comparisons, using indicators from Tables 1 through 4*

	Indicator	Subpopulations	Result at 5%	Absolute difference
Table A1: HIV Counselling and Testing	% of individuals who were counselled and received an HIV test in last 12 months and know their results.	Women	Different	0.052
		Female youth	Same	0.003
		Men	Different	0.077
		Male youth	Same	0.043
	% of mothers of children 0–11 months who were counselled and received an HIV test during the last pregnancy and know the results.		Same	0.050
Table A2: HIV PMTCT	% of mothers of children 0–11 months who were counselled for 'prevention of mother-to-child transmission' services during last pregnancy.		Different	0.128
Table A3: HIV Knowledge and Sexual Behavior	% of individuals who had sex with more than one sexual partner in the last 12 months.	Women	Same Different	0.024
		Female youth	Same	0.026
		Men	Same	0.044
		Male youth		0.039
	% of individuals who have had sexual intercourse with a nonmarital or noncohabitating sexual partner.	Women	Different	0.076
		Female youth	Different	0.061
		Men	Different	0.129
		Male youth	Different	0.095
	% of individuals who have had sexual intercourse with a nonmarital or noncohabitating sexual partner in the last 12 months and used a condom at last higher-risk sex.	Women	Same	0.009
		Female youth	Same	0.046
		Men	Same	0.118
		Male youth	n/a	
	% of youth 15–24 years who have had sexual intercourse before the age of 15.	Female youth	Same	0.009
		Male youth	Same	0.014
	% of men who are circumcised.	Men	Same	0.014
		Male youth	Same	0.027
Table A4: Malaria	% of children 0–23 months who had fever in the two weeks preceding the survey and received treatment with ACT within 24 h of onset of fever.		Same	0.002
	% of mothers of children 0–11 months who received two of more doses of SP/Fansidar during their last pregnancy.		Different	0.382
	% of children 0–23 months who slept under an ITN the night preceding the survey.		Different	0.244

*Two indicators are concluded to be the 'same' if the hypothesis test of proportion equality failed to reject at the 5% level. Otherwise, the indicators are concluded to be 'different'. Refer to the indicated table for more detailed information on that indicator, including point estimates, confidence intervals, and *P*-values.

Table A10 Summary of comparisons, using indicators from Tables 5 through 8*

	Indicator	Subpopulations	Result at 5%	Absolute difference
Table A5: Family Planning	% of currently married women who are using any family planning method.		Different	0.191
	% of mothers of children 0–11 months who attended ANC at least four times during their last pregnancy.		Same	0.021
	% of mothers of children 0–11 months who delivered their last baby in a health facility.		Different	0.124
	% of mothers of children 0–11 months who were assisted by a skilled health worker during their last delivery.		Same	0.083
Table A6: Child Health	% of children 12–23 months who are fully vaccinated. *Definition 1 (1 BCG + 3 DPT + 4 POLIO + MEASLES)*		Same	0.015
	% of children 12–23 months who are fully vaccinated. *Definition 2 (1 BCG + 3 DPT + 3 POLIO + MEASLES)*		Same	0.004
	% of children 0–11 months with diarrhoea in the last two weeks receiving oral rehydration therapy (ORT).		Same	0.055
Table A7: Nutrition	% of children under six months of age who are exclusively breastfed.		Same	0.009
	% of children 12–23 months receiving vitamin A supplementation in the last six months.		Different	0.111
	% of households using iodised salt.		Different	0.037
	% of mothers of children 0–11 months who received vitamin A supplementation within 2 months after delivery.		Different	0.213
Table A8: Water and Sanitation	% of households with safe water supply.		Different	0.204
	% of households with latrine or toilet.		Same	0.008

*Two indicators are concluded to be the ‘same’ if the hypothesis test of proportion equality failed to reject at the 5% level. Otherwise, the indicators are concluded to be ‘different’. Refer to the indicated table for more detailed information on that indicator, including point estimates, confidence intervals and *P*-values.

Corresponding Author Joseph J. Valadez, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK. Tel.: +44 151 705 3193; E-mail: joseph.valadez@lstmed.ac.uk

