Small Sample Size Surveys for GBV Programming: Main Results Report

1st October 2018

By: Maureen Murphy, Junior Ovince*, Pierre Philippe Wilson Registe**, Ulrick Jean-Claude** and Manuel Contreras**

Author Affiliations:

* The Global Women's Institute at George Washington University
** Institut de Formation du Sud

The Global Women's Institute
THE GEORGE WASHINGTON UNIVERSITY

IfoS
Institut de Formation du Sud

Table of Contents

Background

Population-based surveys that help us understand the magnitude and scope of gender-based violence (GBV) in communities are becoming commonplace. Efforts such as the World Health Organization (WHO) Multi-Country Study on Domestic Violence, the International Violence Against Women Survey (IVAWS) and the Demographic Health Survey (DHS) have shown that population-based data collection on GBV is possible and can be done in an ethical manner. Population-based surveys have also been used to measure the outcome and impact level changes in GBV programmes (including changes in knowledge, attitudes and behaviours). Large-scale studies in countries such as Uganda and Senegal have shown the utility of measuring population-level change to demonstrate the effectiveness of GBV programmes (Abramsky, et al., 2014; Diop, et al., 2004). However, these rigorous research practices are rarely applied in humanitarian settings (Hossain & McAlpine, 2017).

High quality surveys that measure the impact of GBV programmes in non-humanitarian settings often involve intensive time, energy and technical engagement of outside researchers. Typically, surveys in development and humanitarian settings utilize multi-stage cluster sampling designs, which often require large sample sizes to accurately estimate rates of GBV and measure changes in knowledge, attitudes or behaviours. In humanitarian settings, organizations often lack the time, resources or expertise to implement these rigorous surveys – effectively eliminating the possibility of measuring programme impact. Despite this, the international GBV prevention and response community has clearly expressed the need for high quality evidence to demonstrate the effectiveness of interventions.[1] In order to meet this need, further exploration of alternative sampling procedures is needed to reduce the barriers to collect high quality population-based data in these settings.

Lot Quality Assurance Sampling

Lot quality assurance sampling (LQAS) has been adapted from the manufacturing industry where it was developed as a way to test a small proportion of products to determine if they are of acceptable quality (Dodge & Roming, 1959). In the 1980s and 90s, this methodology was adapted for use in the public health arena (Robertson, et al., 1997; Smith, 1989). Rather than 'lots' or batches of products, public health researchers are interested in groups of people – whether in a community, a health facility catchment area or other organisational unit. For example, LQAS is often employed in immunization coverage surveys, where after an immunization campaign is

[1] See Outcome 5-4 of the *Call to Action on Protection from Gender-based Violence in Emergencies* which states "Continue to build the evidence base to define effective GBV prevention and response interventions in humanitarian settings"

completed; the population is surveyed to determine if an acceptable proportion of the overall population actually received the vaccination.

Since its adaptation for public health use, LQAS techniques have been gaining in popularity, particularly in public health programmes in development settings. They have been used for programme assessments as well as monitoring and evaluation of programme performance in field such as HIV/AIDS, sexual and reproductive health, growth and nutrition, and water, sanitation and hygiene (WASH) as well as for quality management and after disasters (natural and public health) and in post-conflict settings (Robertson & Valadez, 2006). The LQAS methodology has also been employed in humanitarian settings – but in a much more limited fashion and never, to our knowledge, in the protection sector (for examples from other sectors see: Government of the Republic of South Sudan, Humanitarian Innovation Fund and Liverpool School of Tropical Medicine, 2014; Harding et al., 2017; Pham, Chambers Sharpe, Weiss, & Vu, 2016).

At its core, LQAS designs in public health settings are structured to be programmatic M&E tools. Managers sub-divide their project areas into "lots" or "supervision areas"– based on programmatic relevant criteria (health facility catchment areas, areas covered by one community activist supervisor, etc.) and then select key outcome indicators and set specific targets to be achieved. Periodic small sample size surveys are then carried out to track progress around these outcome indicators. Within each of these supervision areas, progress is assessed by a simple binary choice – the target is met or not. For example, in an immunization campaign perhaps 90% of children under the age of 5 were expected to have received a specific vaccine in order for the campaign to be considered successful. Using LQAS, surveyors would move through a community checking a small sample of immunization cards of children under 5 to determine if that area met the acceptable rate of quality (at least 90% of children under 5 vaccinated). In areas where the proportion of the population was not found to have received the vaccine, targeted follow up vaccination campaigns would be held in these specific areas.

While originally designed to make specific programme decisions in small programme areas, researchers also use LQAS techniques to make overall estimates of population-based indicators by pooling and weighting the data collected via individual supervision areas (Valadez et al., 2003). These methods have even been combined with multi-stage cluster sampling approaches to be applicable for collecting data in large country programmes (Hedt, Olives, Pagano, & Valadez, 2008).

The Study

Based on this method's proven utility in other setting and sectors, the research team from the Global Women's Institute (GWI) at the George Washington University and the Institut de Formation du Sud (IFOS) thought that LQAS techniques had the potential to be used to gather reliable, population-based data for GBV programmes in conflict and humanitarian settings.

In order to test whether LQAS would be an appropriate method for collecting population-based data on GBV indicators, researchers conducted a pilot study using the methodology in Marigot Commune in South-east Haiti. This location was selected because in 2017 GWI and IFOS had conducted a baseline study for an impact evaluation of a GBV prevention programme and had utilized multi-stage cluster sampling techniques. It was therefore possible for GWI and IFOS to directly compare the results of both approaches.

For the impact evaluation, Marigot Commune was being utilized as the control community (i.e. the GBV intervention that the impact evaluation is studying is not being implemented in this location). Therefore, it is not expected that any measureable change on GBV attitudes or prevalence would have taken place between the cluster survey conducted in 2017 and the LQAS survey conducted in 2018. After collecting data in the commune utilizing both sampling methods, the research team compared the results of the two approaches for select GBV indicators as well as cost, logistics and human resources considerations.

Methodology

Multi-stage Cluster Sampling

For the multi-stage cluster survey, a total of 1,158 women and girls (aged 15-64) in Marigot Commune were originally surveyed. This sample size was calculated based on available national statistics and local knowledge of the area, which led the research team to estimate the prevalence of past 12 months physical intimate partner violence (IPV) in the programme area as approximately 20% at baseline. The sample was calculated in order to detect a reduction of overall prevalence of past year physical IPV to 14% over the course of a three year project.

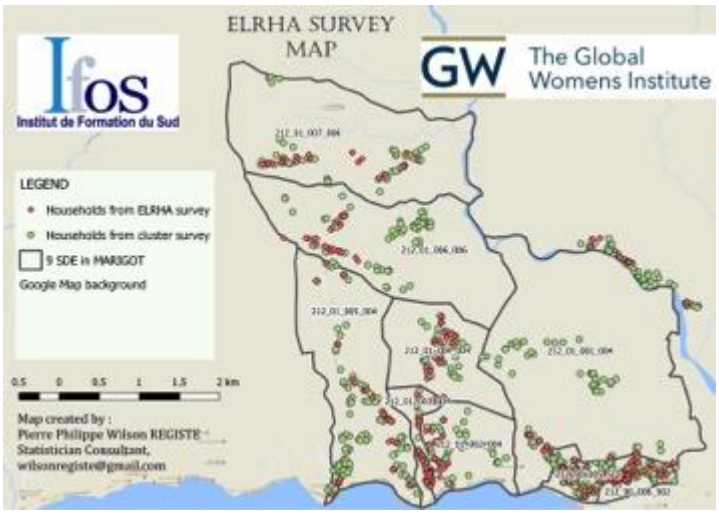
Within Marigot Commune, local census enumeration areas (SDEs) were utilized as clusters. Twenty out of a total of 38 potential clusters were randomly selected through probability proportionate to size (PPS) methodology. Within each cluster, a household listing exercise mapping all accessible households with GPS coordinates was undertaken prior to the survey. All the individual households in the selected SDE were then approached for the interview. Once a household was selected, all eligible women were listed and then a random number generator selected the final respondent. If the selected respondent was not available at the time of the initial interview, up to 3 follow up visits were made. No replacements were selected for women and girls who could not be interviewed. Sixteen female Haitian data collectors conducted the women's survey in this area. Overall, the data collected in Marigot was completed in 22 days.

Lot Quality Assurance Sampling

For the LQAS survey, in order to ensure comparability of the results, the study also targeted locations in Marigot Commune where the cluster survey had been implemented during the prior

year. Nine supervision areas (SAs), based on census enumeration areas, were surveyed during the course of the study. In each SA, the research team aimed to complete at least 19 interviews. This sample size was selected based an LQAS decision rule table, which shows that sample sizes larger than 19 in a single SA generally do not result in more accurate assessments unless the sample is much larger and more costly (Valadez J.J, Weiss, Leburg, & Davis, 2002). To adjust for non-response, a total of 25 households to approach were randomly selected within each SA.

Five Haitian women served as data collectors for the study. Each of these women had previously participated in the multi-stage cluster survey and had good knowledge of the survey tool and ethical considerations due to this previous experience. A three-day refresher training was held to review the data collection tools and survey administration, including the new sampling procedures.



All data collection tools were adapted from the same instrument utilized in the multi-cluster survey. In order to reduce the time burden on the respondents, the survey tool was slightly shortened compared to the original tool utilized in the cluster survey. Most of the reductions occurred after the series of questions regarding experiences of violence to ensure that the respondents experienced the same introductory sections, which were designed to build trust between the data collector and respondent and cover less sensitive subjects (background, education, family, etc.) than the eventual questions on experiences of violence. A screening question was added to the tool to determine if the respondent had ever been married or cohabitated with a man– as the sample was designed to test the difference between the prevalence of IPV as a key indicator and therefore all respondents needed to have had a partner at least once time in their lifetime. Girls between the ages of 15-17 who had dated but never married or lived with a man were excluded from the survey and analysis for ethical reasons related to the need for parental consent from this group.

The data collection team was given a list of the GPS coordinates for 25 selected households and each enumerator had a data collection tablet equipped with GPS to utilize in tracing each selected household. In addition, each enumerator had paper maps previously developed by IFOS with the approximate household location and key local landmarks (roads, rivers, churches, etc.) to aid in the tracing of households. Due to time and logistics, if at least 19 households were completed in an SA, the team did not return to that location again to trace any remaining

households. Overall during data collection, the team visited 199 households and completed interviews with 93.5% (186) of selected households. In only one SA was the expected sample size of 19 completed interviews not achieved. Data collection was completed over the course of 10 days.

Table 1: Completed Surveys by SA – LQAS Survey

		Completion Status			
		1. Not completed	2. HH questionnaire only	3. Women's Questionnaire completed	Total
Supervision Areas	1	0	0	23	23
	2	0	1	21	22
	3	0	2	20	22
	4	0	2	19	21
	5	1	0	22	23
	6	0	0	18	18
	7	0	0	20	20
	8	0	4	23	27
	9	0	3	20	23
	Total	1	12	186	199

To ensure comparability between the two survey methods, adjustments were made during secondary analysis of the data from the cluster survey and a subset of 779 of the completed surveys with ever married or cohabitated women and girls in Marigot was utilized for final comparison. This was done to ensure direct geographical comparison between the areas where each survey was implemented and due to the focus on IPV as the primary outcome of interest (which required a focus on only ever partnered respondents).

Ethical Considerations

For both survey approaches, there were a number of ethical considerations taken into account in study design and implementation. First, safety of the respondents was prioritized through a number of actions. The surveys were never described as a survey on violence against women and girls in the wider community or with the head of households when introducing the purpose of the interview. Instead they were framed as surveys on women's health and life experiences. In addition, only one woman was selected per household – to ensure that no one else in the household knew the true subject matter of the survey. In order to ensure the confidentiality of the collected data, no identifiable information (name, addresses, etc.) was collected through the survey.

Picture 1: Data Collectors and GWI staff prepare for fieldwork

All data collectors were female to ensure that the female respondents felt comfortable disclosing experiences of violence. In addition, the data collectors were trained to stop the interview if anyone else came into the room or to pivot the discussion to topics not related to violence (for example menstrual hygiene). Data collectors were also trained on supportive listening and managing distress, in case any respondent experienced mild distress during the interview. They also learned how to identify more serious signs of distress and were instructed to refer these cases to their supervisors if encountered. All women who participated in the study were offered referral information on available GBV services in the community.

Results

The results section compares the findings from the multi-stage cluster survey conducted in 2017 and the LQAS survey conducted in 2018. This includes a brief examination of some of the key socio-demographic trends between the two surveys as a way to understand the comparability of the samples and then a review of the overall findings on key GBV indicators (prevalence of violence and gender-equitable attitudes).

Socio-demographics

Overall, the selected samples for both the multi-stage cluster and LQAS surveys were very comparable. For example, in Table 2 the age distributions between both sampling methods are similar for most age categories. In addition, for indicators on the respondents' educational status and literacy the sample populations were generally alike in their profiles. For example, about half of respondents in both groups knew how to read and write while just under half of the respondents in both groups had never attended school. These similarities continued for indicators of economic activity, where about 70% of both samples were working in some capacity.

Table 2. Socio-Demographics

	LQAS %	Cluster %
	n =186	n = 779

AGE		
Age Range		
15-24	10.2	12.5
25-34	29.6	28.6
35-44	26.3	23.5
45-54	12.9	17.1
55-64	21.0	18.4
EDUCATION AND LITERACY		
Know how to read and write	52.7	54.4
Education level		
No School	46.2	45.1
Primary School	23.7	28.6
Secondary School/Higher Education	30.1	26.3
ECONOMIC ACTIVITIES		
Working Status		
Working/student	68.3	73.2
Not Working	31.7	26.8
PARTNERSHIP DETAILS		
Partnership Status *		
Currently partnered	93.0	82.7
Formerly partnered	7.0	17.3
Age at First Union (marriage or co-habitation)		
19 or younger	27.4	26.4
20 or older	72.6	73.6

* P <= .05;** P <= .01; ***P <= .001

In addition, partnership characteristics were compared between the two samples. All respondents included in the data analysis were partnered at some point in their lifetime. While respondents in both surveys were typically married at the age of 20 or older (about 70% of each sample), there were differences between the two surveys in the details on these unions. In the

LQAS survey, more respondents reported that they were currently partnered (93%) compared to those in the cluster survey (83%), while more respondents in the cluster survey reported that they had previously been married or had a partner (17% compared to 7%) but were not currently partnered.

Prevalence of Violence

The core indicators examined for this comparison study was the prevalence of physical and/or sexual intimate partner violence (IPV) amongst ever-partnered (married or cohabitated) respondents. In order to explore the applicability of LQAS techniques for programme M&E of GBV programmes, we first present the overall prevalence of IPV by supervision area (see Table 3). In this example, we use the data collected through the cluster survey as the benchmark for comparison and classification of the results in each SA. We have used the classifications of 'more violence' for areas where the rates of violence exceeded the average of the cluster survey and 'less violence' for areas where the violence was lower than the benchmark. Overall, we found that in almost all SA's more violence was reported utilizing the LQAS method compared to the cluster method.

Table 3. Physical and/or Sexual IPV via collected via LQAS

	Number reporting violence	Number sampled	%	Decision Rule (obtained in cluster survey)	Classification
Supervision Area 1	12	23	52.2	6	More violence
Supervision Area 2	7	21	33.3	6	More violence
Supervision Area 3	8	20	40.0	5	More violence
Supervision Area 4	6	19	31.6	5	More violence
Supervision Area 5	7	22	31.8	6	More violence
Supervision Area 6	5	18	27.8	5	Less violence
Supervision Area 7	9	20	45.0	5	More violence
Supervision Area 8	7	23	30.4	6	More violence
Supervision Area 9	10	20	50.0	5	More violence

This type of classification is similar to how data is typically used when LQAS is employed as a baseline study. In theses cases, researchers classify each individual SA survey as "above" or "below" average of the entire sample. This process helps programme managers prioritize which areas require more immediate support and attention to improve on the selected indicator.

After examining the data for each SA, weighted estimates for the overall study areas and associated confidence intervals were calculated. Prevalence of the lifetime physical and/or sexual IPV was found to be 38.4% (95%CI: 31.6-45.5) using the LQAS method while it is 29.9% (95%CI:

26.8 – 33.2) using the cluster method. Similar trends were found for violence within the past 12 months with 29.5% (95%CI: 23.0-36.0) of respondents in the LQAS survey reporting experiencing physical and/or sexual violence in the past 12 months, while 21.4% (95%CI: 18.7 – 24.4) reported this in the cluster survey.

TABLE 4. PHYSICAL AND/OR SEXUAL IPV FROM MOST RECENT OR CURRENT PARTNER – LIFETIME AND PAST 12 MONTHS

TYPES OF VIOLENCE	LQAS n =186			Cluster n = 779		
	%	95% CI	Standard Error	%	95% CI	Standard Error
LIFETIME						
PHYSICAL VIOLENCE	21.3	15.7 – 27.4	3.01	15.4	13.0 – 18.1	1.29
SEXUAL VIOLENCE	32.9**	26.5 – 40.0	3.45	22.3**	19.5 – 25.4	1.49
PHYSICAL AND/OR SEXUAL VIOLENCE	38.4*	31.6 – 45.5	3.58	29.9*	26.8 – 33.2	1.64
PAST 12 MONTHS						
PHYSICAL VIOLENCE - PAST 12 MONTHS	12.8	8.7 – 18.4	2.46	10.1	8.2 – 12.4	1.08
SEXUAL VIOLENCE - PAST 12 MONTHS	26.9***	21.0 – 33.7	3.26	16.6***	14.1 – 19.3	1.33
SEXUAL AND/OR PHYSICAL VIOLENCE - PAST 12 MONTHS	29.5*	23.0 – 36.0	3.35	21.4*	18.7 – 24.4	1.47

* P <= .05;** P <= .01; ***P <= .001

Overall, while the point estimate for each indicator were higher utilizing the LQAS methodology, the results from both methodologies showed similar trends. For example, sexual IPV was reported more often than physical IPV in both studies and lifetime physical and/or sexual violence was estimated to be about 8-9% points higher than IPV in the past 12 months. While the estimates for rates of sexual IPV were higher in the LQAS survey (and thus the rates of physical and/or sexual violence), the differences between the reported rates of physical IPV were not statistically significant. One other difference between the two methodologies was the overall standard error and confidence interval ranges – which were larger for LQAS methodology across all the indicators due to the smaller sample size employed.

TABLE 5. PSYCHOLOGICAL AND ECONOMIC IPV FROM MOST RECENT OR CURRENT PARTNER – LIFETIME AND PAST 12 MONTHS

TYPES OF VIOLENCE	LQAS n =186	Cluster n = 779

	%	95% CI	Standard Error	%	95% CI	Standard Error
LIFETIME						
ECONOMIC VIOLENCE	21.0	15.7 – 27.4	3	15.0	12.6 – 17.7	1.28
PSYCHOLOGICAL VIOLENCE	43.0	35.7 – 49.9	3.64	35.4	32.1 – 38.8	1.71
PAST 12 MONTHS						
ECONOMIC VIOLENCE	17.9**	12.8 – 23.8	2.82	9.6**	7.7 – 11.8	1.06
PSYCHOLOGICAL VIOLENCE	32.3	26.0 – 39.4	3.44	26.3	23.3 – 29.5	1.58

*P <= .05;** P <= .01; ***P <= .001

For economic and psychological violence, there was no statistical difference on any of the indicators collected by the two methodologies except for economic violence in the past 12 months where data collected via LQAS was higher. As with the previous indicators, the confidence intervals and standard error generated by the LQAS were larger.

For non-partner sexual violence, as with partner violence the overall point estimates were higher in the data collected via the LQAS methodology (21.6% for lifetime violence versus 11.6% in the cluster methodology). However, there was no statistical difference between the rates of non-partner sexual violence in the past 12 months.

TABLE 6. NON-PARTNER SEXUAL ASSUALT – LIFETIME AND PAST 12 MONTHS

TYPES OF VIOLENCE	LQAS n =186			Cluster n = 779		
	%	95% CI	Standard Error	%	95% CI	Standard Error
LIFETIME						
NON PARTNER SEXUAL VIOLENCE	21.6***	16.2 – 28.0	3.03	11.6***	9.5 – 13.9	1.15
PAST 12 MONTHS						
NON PARTNER SEXUAL VIOLENCE	8.2	4.8 – 12.7	2.02	4.6	3.3 – 6.3	.75

* P <= .05;** P <= .01; ***P <= .001

Attitudes

While prevalence of IPV and non-partner sexual assault are two examples of the types of data that can be analysed via LQAS, there are an innumerable number of other indicators that could be collected through this approach. In the GBV sector, one example of this is data related to gender equitable attitudes and acceptance of violence.

Table 7 shows acceptance of gender inequitable roles and violence amongst community members collected via the LQAS and cluster approaches. Overall, the data collected was quite

similar though the respondents generally had slightly more equitable attitudes – both on gender roles and use of violence - in the LQAS survey. For most indicators these differences were statistically significant.

TABLE 7. GENDER ROLES AND ACCEPTANCE OF VIOLENCE

TYPES OF VIOLENCE	LQAS n =186			Cluster n = 779		
	%	95% CI	Standard Error	%	95% CI	Standard Error
GENDER ROLES						
CHANGING DIAPERS, GIVING A BATH, AND FEEDING KIDS IS MAINLY THE MOTHER'S RESPONSIBILITY	78.3***	71.0 – 83.8	3.03	95.5***	93.9- 96.8	.74
A WOMAN'S ROLE IS TAKING CARE OF HER HOME AND FAMILY	85.1***	79.2 – 89.5	2.62	94.4***	92.6-95.8	.83
WOMEN AND MEN SHOULD SHARE AUTHORITY IN THE FAMILY	74.9***	68.0 – 80.5	3.19	85.6***	83.0 – 88.0	1.26
ACCEPTANCE OF VIOLENCE						
IT IS THE ENTIRE COMMUNITY'S RESPONSIBILITY TO PREVENT MEN FROM BEATING THEIR WIVES	81.3	75.0 - 86.2	2.87	76.0	72.9-78.9	1.53
A WOMAN SHOULD ACCEPT VIOLENCE TO KEEP HER FAMILY TOGETHER	16.4*	11.4 - 22.0	2.72	24.9*	22.0-28.0	1.55
IF A WOMAN IS RAPED SHE HAS DONE SOMETHING CARELESS TO PUT HERSELF IN THAT SITUATION	14.4*	10.1 – 20.2	2.58	22.3*	19.5-25.4	1.49

* P <= .05;** P <= .01; ***P <= .001

Discussion

Overall Results and Data Quality

Globally GBV prevalence data are gathered via self-report survey methods, which have considerable limitations, even when all the appropriate ethical, safety and methodological considerations are in place (gender-matched data collectors – respondents, well-trained and non-judgmental data collectors, private locations for interviews, etc.). Even in the most carefully designed studies, self-reported rates of violence under-estimate the true prevalence occurring in the wider population as some people will never feel comfortable speaking about their experiences. Based on this assumption, studies that estimate higher rates of violence are generally seen as closer to the underlying true prevalence rates.

Comparing the prevalence rates of violence between these two survey methods, we can see that the overall trends are the same (more respondents reporting sexual IPV compared to physical, psychological violence more commonly reported than economic violence, etc.) though the overall point estimates collected via the LQAS methodology are higher than the data collected through multi-stage cluster sampling, though this finding was not always statistically significant.

This finding suggests that women participating in the LQAS survey might have felt more comfortable disclosing their experiences of violence during this study compared to the previous cluster survey. While every effort was made by the research team to replicate conditions as closely as possible when delivering both surveys – there were a number of factors that may have contributed to the higher disclosure rates in the LQAS survey. While some of these factors were specific to the delivery of these two surveys, most lessons are applicable for the wider GBV and research communities. These considerations and their implications for future population-based research efforts will be discussed here in detail.

Picture 2: IFOS trains data collectors on use of the tablets

First, key to the implementation and success of the LQAS survey was the skill and experience of the data collection team. To reduce costs and simplify logistics, GWI and IFOS recruited data collectors (5 Haitian women) from the wider pool of 16 Haitian women who had served as data collectors during the cluster survey in 2017 rather than recruiting a completely new data collection team. This could

have influenced the quality of the second survey in a number of ways. First, IFOS – who managed the data collector selection process – wanted to ensure the highest quality data was collected and therefore picked some of the highest performing data collectors from the previous study. In addition, these women had been previously participated in in-depth training and collected data utilizing a very similar data collection tool during the previous survey. This could have resulted in a more highly skilled data collection team for the LQAS survey, which may have led to respondents feeling more comfortable disclosing their experiences of violence during the second survey.

In addition, the smaller research team was considerably easier to manage and supervise compared to the larger pool of researchers from the cluster survey. This allowed for closer monitoring of quality and more individualized support from GWI and IFOS to the data collectors. The need to have fewer data collectors meant that lower performing data collectors were not included in the final survey administration team. While in this case, the highest performing data collectors were able to be identified from the previous survey, in other scenarios this could be achieved by having a larger pool of data collectors participate in the training and pilot, with only the highest performing amongst them continuing on to participate in the actual survey. While this is already a practice that occurs in some surveys around the world, the need for a much smaller pool of final selected data collectors may simplify this process and allow for only the highest quality trainees to collect data during the actual survey implementation.

Another potential factor affecting the quality of survey implementation was the shorter timeframe overall for data collection. While data for the cluster survey was collected over 22 days, all data was collected for the LQAS over the course of 10 days. While this has implications of the budget and logistics considerations – which will be discussed below – it also may have an effect on data quality. There is potentially a fatigue factor that plays into lengthy data collection exercises no matter the quality of the data collectors and design of the study. The implication is that shorter data collection exercises might lead to higher quality data. In addition, the shorter questionnaire utilized by the LQAS survey may have also reduced overall fatigue of the data collectors.

While prevalence data was higher in the LQAS survey compared to the cluster survey, respondents had slightly more gender equitable attitudes in the LQAS survey. This could be explained by the gap of 1 year between data collected in the cluster survey and the LQAS survey, which may have allowed for some attitudinal change to occur within the group. In addition, there might have been some bias introduced into the study population who had previously answered questions related to gender and violence during the cluster survey. This previous data collection activity could have triggered the population to think more about these issues and led to some attitudinal changes by virtue of participating in the survey.

While overall the data collected via the LQAS methodology appears to be of good quality, there remain some limitations of this approach. For one, data collected using LQAS sampling strategies

has larger standard errors and therefore wider confidence intervals surrounding the point estimates compared to cluster surveys. In addition, the small sample size is designed to answer questions with bi-nominal (e.g. yes/no, agree/disagree, correct/not correct) answer categories and not more complex analytical questions. As such, it is an appropriate method to understand key outcome indicators (e.g. prevalence, attitudes, etc.) but further more complex analysis, for example on how potential drivers of violence are associated with experiences of violence, are not possible with this method.

Cost and Logistic Considerations

As noted above, cost savings were achieved by utilizing previously trained data collectors who only needed a refresher training on the data collection tools and procedures rather than a full 2-3 week training. As GBV research involves a considerable number of ethical considerations in order to be accomplished safely, the research team would recommend a full training prior to any study that inquires about personal experiences of violence no matter what sampling methodology is utilized. Research that focuses on gender-attitudes or overall acceptance of violence may require less lengthy training periods.

In order to have the most direct comparisons, we have focused explicitly on data collection costs for budgetary analysis– assuming that in a typical setting the training and researcher support costs (e.g. research protocol development, tool development, analysis and report writing) would generally be the same no matter the sampling methodology. Overall, the LQAS approach showed considerable cost savings – with overall costs only 25% of the costs of collecting data in Marigot Commune using the multi-cluster model - (see Table 8) due to the smaller number of data collectors and reduced days of data collection.

Table 8: Direct Field Work Costs - Budget Comparisons (all figures are in USD)

Categories	LQAS	Cluster[2]
Human Resources (including data collectors, supervisors and the services of a local statistician)	11,760	50,105
Travel (hiring of vehicles, driver)	3,960	11,952
Equipment (Tablets, software, etc.)	830	3,481
Total	16,550	65,538

[2] These figures are specific estimates based on the sub-sample utilized for analysis in this study.

Picture 3: A data collector consults a paper map in the field

From a logistics perspective, there are some advantages and some disadvantages to the LQAS approach. For this survey, the research team had the added advantage having previously completed a household listing exercise in the Commune in preparation for the original multi-cluster survey. This allowed the research team use the database of household GPS coordinates to randomly select households to participate in the survey. Data collectors would manually enter the full GPS coordinates of each house in the area at the beginning of the day. They would then use Google maps (offline with previously downloaded maps) and follow their "blue dot" representing their current location to the selected households. This approach had mixed success. In general most data collectors were able to find the general location of the household using Google maps but the specificity of the app was often not good enough to ensure that exact house noted by the coordinates was found. In these cases, the data collectors selected the house closest to the Google maps marker. In addition, the lack of detailed maps on Google maps in more rural areas sometimes meant that there no indication of roads, waterways or other landmarks – just a grey screen with the target location and blue dot representing the data collector. In these cases, data collectors had to cross check with paper maps developed by IFOS that placed the GPS coordinates in relation to key local landmarks.

Other logistical challenges occurred in the distance between households. In multi-stage cluster sampling, the final households would typically be selected using a systematic sample within each cluster. In LQAS sampling, the distance between households can be much greater. This caused challenges for the data collection where fieldworkers had to walk long distances between households. In addition, safety considerations meant that often data collectors would need to work in pairs to find houses in more remote areas. Despite this, most data collectors were able to complete the expected quota of 4 completed surveys per day per person.

Implications for GBV Research in Humanitarian Settings

Overall, the research team found LQAS to be a relatively quick and effective methodology for collecting population-level data on GBV indicators – including both prevalence and attitudes. Marigot Commune in southeast Haiti was selected as the location for the pilot due to its rough terrain that mimics the harsh conditions one would find in a humanitarian emergency. Data collectors often had to walk for multiple hours up and down local mountains to access the households selected for the survey. Despite these harsh conditions, the quality of the data collected via LQAS appeared to be high. For example, the Haiti Enquête Mortalité et Utilisation des Services (EMMUS-VI) survey from 2016 and 2017 found that 23.5% of ever partnered women

aged 15-49 reported ever experiencing physical and/or sexual violence. Using the LQAS approach, an overall estimate of 38.49% (95% CI: 31.6-45.5%) was produced. This suggests LQAS could be a reliable approach to get quick and accurate data in humanitarian settings.

Picture 4: A data collector walks to find an assigned household.

Logistics constraints related to the lack of existing maps and the need to randomly select households, and then respondents from the overall population at large are some of the biggest constraints to using this approach in humanitarian settings. Time is required to create maps in areas where this approach will be utilized. However, the rapid increases in GPS technology and large-scale efforts from humanitarian actors such as OCHA, the Red Cross, and the Humanitarian Open Street Map Project in mapping efforts in humanitarian settings are reducing these barriers. In addition, GPS enabled mobile phones mean that a small team on motorbikes can do a household mapping exercise relatively quickly and cheaply. In addition, in settings where the overall population is required to be registered (government refugee registration cards, ration lists, etc.) these sources can be used as alternatives for respondent selection assuming that some indicator of geographic location (address, neighbourhoods, villages) can be used to sort potential respondents prior to selection.

There also have been multiple adaptions of these approaches to areas were detailed maps were not available. For example, in some uses of LQAS, when general population figures are available but specific maps not, researchers have used probability proportional to size (PPS) sampling to pick the total number of interviews needed per community. Once the data collection team arrives in a community, they sub-divide the community randomly selecting smaller and smaller sub-divisions (for example, the left side of the river or the right) until an area small enough to manually map to select a house is selected (Pham, Chambers Sharpe, Weiss, & Vu, 2016). Once this small area is mapped, the final household for interview is randomly selected. Similarly, LQAS methodologies have been combined with cluster sampling techniques to gather actionable information for sub-regions when the survey needs to cover a wide geographic area (Hedt, Olives, Pagano, & Valadez, 2008).

The safety and security of both respondents and data collectors is another key consideration when utilizing LQAS methodologies in humanitarian settings. From the perspective of the respondents, there may be considerable ethical advantages of using LQAS rather than large-scale

cluster sampling strategies. A key ethical principle of GBV is to reduce knowledge of the study topic amongst participants who are not directly participating as respondents might be subject to negative consequences such stigma or even violence from an abusive partner if it becomes known that they participated in a survey on violence. By utilizing small sample size approaches such as LQAS the overall footprint of the study is much smaller compared to large multi-stage cluster surveys. Fewer data collectors are utilized and in general each SA can be completed in about a day. This ability to quickly finish a geographical area may minimize local interest in what the study is about and reduce the likelihood that the overall topic of the survey – experiences of GBV – becomes widely known in the community. In addition, the wider distances between households may reduce the likelihood that neighbours talk to each other about their experiences participating in the study – again increases confidentially.

Conversely, the wide geographical spread of the respondents may have potential negative effects on the data collection team. Rather than being together in a group in a small area, individual data collectors may need to travel long distances in order to find households far from others. It is important to consider this in field work planning and ensure that appropriate safety protocols are in place, even if this means hiring more data collectors to have pairs that work together to travel to far flung locations.

Use of LQAS for population-based outcome monitoring

In addition to allowing for overall prevalence estimates, LQAS is a useful and relatively simple monitoring tool that can assess population-level change over time. As a monitoring tool, LQAS supervision areas are routinely sampled and assessed against a pre-determined benchmark to track population-level change. This process could improve the ability of NGOs, governments and UN agencies to routinely measure longer-term behaviour change indicators tied to GBV prevention programmes or assess the overall coverage of GBV response programmes. For example, data on a select number of GBV attitude indicators could be collected at a project baseline. GBV practitioners and the M&E team could then set expected targets for change for each monitoring period (for example the expected change in a 6 month or one year period). Data could then be collected utilizing LQAS approaches from each SA and assessed as either having met or not met the target.. Using this approach, GBV programme managers would be able to assess which locations are progressing more quickly on attitude/social norms change indicators and can target lagging areas for additional support. Members of the GBV sub-cluster could even work together to jointly conduct LQAS monitoring in their programme areas to understand population-level changes across their implementation areas. As the GBV community in the humanitarian communities continues to advocate to incorporate more long-term prevention and social norms change programming, LQAS could be a strong option for monitoring and evaluating population-level changes in a rigorous, cost-effective and timely fashion.

Conclusions

Overall, the research team found a number of advantages of LQAS methodologies compared to more traditional, multi-stage cluster sampling strategies. First, the ability to generate specific, actionable data for small sub-areas that allows programme managers to make timely decisions based on evidence is a key advantage of LQAS. In a traditional multi-stage cluster survey, it is generally not possible to breakdown the data into smaller sub-areas and only one overall point estimate is calculated for each indicator. In addition, in general, LQAS typically rely on smaller sample sizes compared to multi-stage cluster surveys, which can lead to cost savings for researchers and practitioners. Finally, there is a plethora of relatively user-friendly training and open-source support materials available to help programme managers use and adapt these methods in their own settings.

Despite these advantages, there are also some limitations to the LQAS approach. The small sample size results in wider confidence intervals compared to large-scale surveys and does not allow for complex data analysis due to the small sample size. As such, the LQAS approach is not a substitute for larger GBV research studies answering more complex research questions – such as understanding the drivers of GBV. In addition, logistics challenges, particularly regarding lack of maps and distance between households could increase time and logistical complexity of an LQAS survey. However, despite these considerations, LQAS has considerable promise in increasing the ability of GBV programme managers to routinely monitor and assess population-based GBV indicators. Systematically implementing routine LQAS surveys could be the missing link in routine population M&E for GBV programmes.

Bibliography

Abramsky, T., Devries, K., Ligia, K., Nakuti, J., Kyegombe, N., Starmann, E., et al. (2014). Findings from the SASA! Study: a cluster randomised controlled trial to assess the impact of a community mobilisation intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda. BMC Medicine .

Diop, N., Faye, M., Moreau, A., Cabral, J., Benga, H., Cisse, F., et al. (2004). The TOSTAN program: evaluation of a community based education program in Senegal. Washington DC: Population Council.

Dodge, H., & Roming, H. (1959). Sampling Inspection Tables. In J. Wiley, Single and Doube Sampling. New York.

Government of the Republic of South Sudan, Humanitarian Innovation Fund and Liverpool School of Tropical Medicine. (2014). LQAS Household Survey Awerial IDP Settlement 2014.

Harding, E., Beckworth, C., Fesselet, J.-F., Lenglet, A., Lako, R., & Valadez, J. (2017). Using lot quality assurance sampling to assess water, sanitation and hygiene services in a refugee camp setting in South Sudan: a feasibility study. BMC Public Health .

Hedt, B., Olives, C., Pagano, M., & Valadez, J. J. (2008). Large Country-Lot Quality Assurance Sampling: A New Method for Rapid Monitoring and Evaluation of Health, Nutrition and Population Programs at Sub-National Levels. Washington DC: The World Bank.

Hossain, M., & McAlpine, A. (2017). Gender Based Violence Research Methodologies in Humanitarian Settings: An Evidence Review and Recommendations. Cardiff: Elhra.

Pham, K., Chambers Sharpe, E., Weiss, W., & Vu, A. (2016). The use of a lot quality assurance sampling methodology to assess and manage primary health interventions in conflict-affected West Darfur, Sudan. Population Health Metrics .

Robertson, S., & Valadez, J. (2006). Global review of health care surveys using lot quality assurance sampling (LQAS), 1984-2004. Social Sciene & Medicine , 1648-1660.

Robertson, S., Anker, M., & al, e. (1997). The lot quality technique: a global review of applications in the assessment of health services and disease. World Health Statistics Quarterly , 199-209.

Smith, G. (1989). Development of rapid epidemiological assessment methods to evaluate health status and delivery of health services. International Journal of Epidemiology , S2-S15.

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

Valadez, J., Weiss, W., Leburg, C., & Davis, R. (2002). Assessing Community Health Programs: A Participant's Manual and Workbook: Using LQAS for Baseline Surveys and Regular Monitoring. Retrieved from http://www.coregroup.org/working_groups/LQAS_Participant_Manual_L.pdf