Maximizing use of existing data to strengthen program design, evaluation, and impact

HealthBridge Foundation of Canada
Peter R. Berti, Rebecca Brodmann

Université de Montréal
Milena Nardocci, Malek Batal

Center for Creative Initiatives in Health and Population
Tran Hung Minh

University College London
Naomi Saville

Statistics for Sustainable Development
Nicolas Greliche

October 2020

Summary

INTRODUCTION. Non-government and civil society organizations spend substantial time and resources collecting data in order to design and monitor their health interventions. The costs vary depending on the objectives, design, methods, sample size, local context and work required to plan and perform data collection. Depending on the number of questions or objectives, interviews can be long, making it difficult to engage respondents throughout the whole process. In addition, the accuracy of the indicator estimates in NGO-led surveys is often insufficient for project design and monitoring purposes, due to relatively small sample sizes and the inherent high variability of the indicators of interest, which can lead to ineffective program design and poor measurement of outcomes.

However, estimates of numerous health and social indicators in many countries already exist in publically available datasets, such as the Demographic and Health Surveys (DHS, supported by USAID) and the Multiple Indicator Cluster Surveys (MICS, supported by UNICEF). DHS and MICS provide standardized data collected using rigorous methods, well-trained enumerators, and large sample sizes. Among the challenges of using publically available data like DHS and MICS to assess baseline conditions is that they are often designed to be representative at the national, regional and provincial level, but rarely are the data representative of lower levels, such as district and village where non-governmental organizations (NGOs) are often working. In addition, DHS and MICS are collected every three to ten years so there may be a ten-year gap between when the DHS/MICS data were collected, and the baseline conditions that the NGO wants characterized.

Despite the challenges, using publically available data to complement or replace NGOs' primary data collection for project baseline measures and project monitoring would save valuable resources, reducing the burden on the NGO staff and on respondents. Note that we would not consider publically available data suitable for endline evaluations because the NGO work will have tried to facilitate change from those baseline conditions.

HYPOTHESES: We hypothesize that publically available data can provide estimates of baseline conditions similar to those reported in NGO baseline reports when matched as closely as possible for location, year, and season of data collection. Further we hypothesize that the impact of differences in year, geographical level, and season varies across health indicators.

This study respects current research ethics standards and it was approved by the Health Research Ethics Board of the *Université de Montréal* (CERSES-19-030-D).

Four parts to the report: *Part 1* presents results of the analyses comparing data from two sources: NGO baseline indicators compared to estimates calculated using DHS or MICS. *Part 2* presents results of analyses comparing DHS data from different years and geographical levels from a single source. These analyses allowed us to work with a much larger sample size and to compare data from DHS in one year or geographical level vs DHS in a different year or geographical level. *Part 3* involves a simulation which quantifies the impact of sampling error when isolated from other sources of error. *Part 4* of this report presents two case studies performed in Vietnam and Nepal comparing data from one NGO baseline report for each country with estimates calculated using MICS (Vietnam) and DHS (Nepal).

PART 1: Analysis of NGO vs DHS/MICS

The objective of these analyses was to compare the indicator estimates from NGO baseline reports to indicator estimates calculated using publically available data (DHS or MICS). A sample of 46 NGO baseline reports from MNCH projects carried out in 23 low- or middle-income countries were compiled. From the reports, we extracted: country name, NGO name, date of beginning and end of data collection, population of study and inclusion and exclusion criteria, indicator name, sample size (total and used for each indicator), location of data collection, and the indicator estimate. Location data (e.g., region, province, district, or/and village name) was used to code the location by geographical level, where first level represents the smallest geographical division in a country, such as village, town, locality, traditional authority; the second level represents district or district council (a division larger than a village but smaller than the third level); the third level represents province, state, department, county or district (if it refers to a division equivalent to province or state); the fourth level represents region; and the fifth level represents country level data.

DHS (from twenty countries) and MICS (for surveys from three countries: Vietnam, Laos, and South Sudan) indicator estimates were matched with the estimates from the NGO surveys, matching geographic level as closely as possible. In total there were 139 indicators, grouped into 41 subgroups, matched between NGO reports and DHS/MICS. Six of the indicators were derived from numerical variables (e.g., Number of children under 5 years in the house, mean) and 133 were derived from binary variables (e.g., Stunting, %). A total of 2,174 pairs of NGO-DHS/MICS indicators were retained for our analyses. The difference (and absolute difference) between the pairs were calculated, as was the difference between geographical level, the difference between year of data collection, and whether the data were collected in the same season or not. The difference between some key estimates is summarized in the following table, which shows that most indicators had pairs of estimates that were different by 20% or more, and most had less than one-third of the pairs of estimates within 5%.

	Proportion of indicator pairs with absolute difference			
Subgroup	**within 5%**	**within 10%**	**within 15%**	**within 20%**
Child anthropometrics				
Stunting (%)	38	66	83	93
Underweight (%)	31	55	74	81
Child diet				
Ate 4+ food groups (%)	24	47	73	90
Exclusive breastfeeding: 0-6m (%)	15	32	43	52
Initiation of breastfeeding within 1 hour of birth (%)	36	50	64	75
Child health				
Diarrhea in the last two weeks (%)	32	55	68	71
Household characteristics and wealth				
HH has electricity (%)	60	70	85	95
Head of HH is male (%)	56	79	88	92
HH has a car (%)	91	98	100	100
Maternal characteristics/health				
Woman able to read (%)	12	62	75	87
Birth at a health facility/assisted by skilled birth attendant (%)	18	37	53	70
Woman received ANC (%)	24	44	58	75
WASH				
HH has improved drinking water (%)	35	52	67	82
HH has improved sanitation (%)	12	34	45	54

In order to better understand how geographical level and year differences might explain the DHS/MICS-NGO differences, we used ANOVA to partition the variance due to indicator, geographic level difference and year difference. Overall, approximately 15 to 20% of the variance was due to indicator, less than 5% was due to geographic level and year difference, and most of the variance was simply random and not explained by the model.

PART 2: Analysis of DHS vs DHS

We compared DHS data across multiple years or multiple regions, to examine the sources of variation in differences in the estimates due to sampling error, year of data collection or geographical level, but not in methods (since the DHS methods are largely consistent across surveys). DHS data from the seven countries were used: Bangladesh, Ethiopia, Kenya, Malawi, Pakistan, Tanzania, and Zambia, and the same indicators as were used in DHS/MICS vs NGO were retained. DHS indicators from different cycles and geographical levels were matched using different combinations mimicking the actual DHS/MICS-NGO scenarios: indicators from the same level but different years (Scenario 1), indicators from the same year but different levels (Scenario 2), and indicators from different years and levels (Scenario 3). The data from the most recent cycle included was also the data from the lower geographical levels (to mimic the NGO data), and the data from the older cycle was also the data from the higher geographical level (to mimic the DHS data). The absolute difference between the two DHS estimates is summarized in the following table.

	Absolute difference between estimates			
Variable	**Mean**	**SD**	**Median**	**Max**
Year difference				
0-2.5 year	9.4	10.2	6.1	93.3
3.0-5.0 years	8.7	9.7	5.5	84.1
5.5-6 years	13.0	14.1	8.6	97.6
Geographical level difference				
0	9.6	11.4	5.7	97.4
1	9.8	10.8	6.3	95.6
2+	11.2	12.3	7.2	97.6

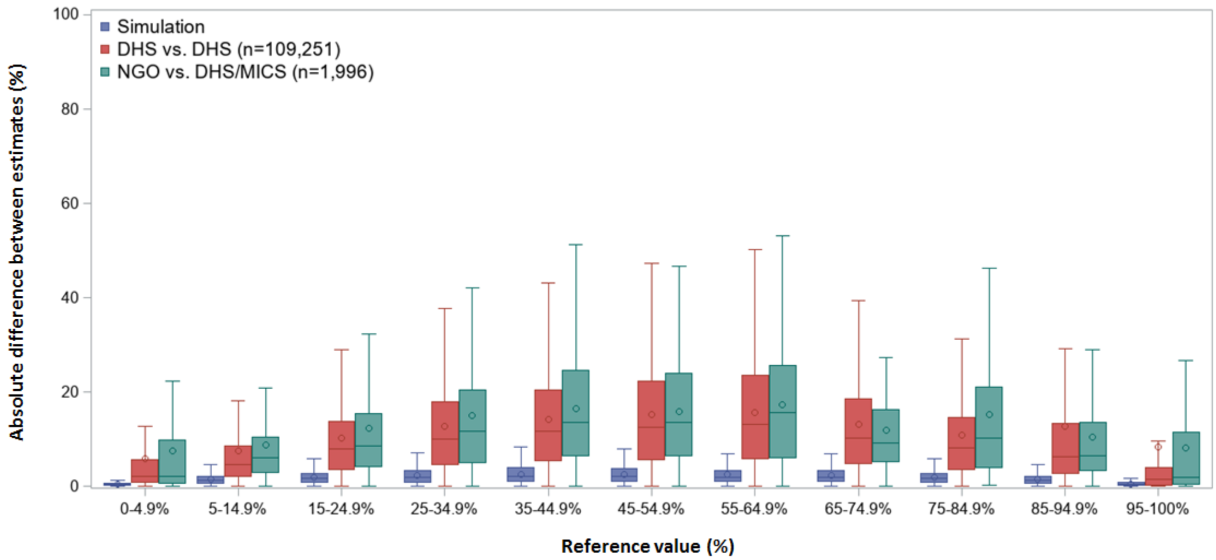
PART 3: Sampling error simulation

We simulated a situation where the only source of imprecision of the indicator's measures would be from sampling error. The simulation samples from a "true" prevalence (p) of 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 99%. We assumed an n of 500, which was the approximate sample size of both DHS and NGO samples in our data set. We then generated a "Baseline Estimate 1" by drawing randomly from a binomial distribution with mean n*p and variance np(1-p). A "Baseline estimate 2" was generated in the same way, and the difference between the first and second estimate was calculated. We ran 1,000 iterations to estimate the distribution of the differences.

The differences between simulated Baseline estimates 1 and 2 was much smaller than the differences between DHS/MICS – NGO, or DHS-DHS.

COMPARISON of results from Parts 1, 2 and 3:

A boxplot of the absolute differences between the two estimates (Part 1: DHS/MICS – NGO; Part 2: DHS – DHS; Part 3: Simulated estimate 1 – Simulated estimate 2) is shown in the following boxplot. The DHS/MICS – NGO difference is largest when the indicator being estimated is between 35 and 65%, and smaller at the extremes. The DHS – DHS differences are very similar to the DHS/MICS – NGO differences, suggesting that it is NOT methodological differences that are responsible for the DHS/MICS – NGO differences (since DHS's methods are largely consistent across years and countries and the results are still similar to DHS/MICS – NGO differences). The DHS – DHS differences are probably only in small part due to random sampling, as the simulation indicates that sampling error contributes only a small amount to the difference in the estimates.



DISCUSSION:

Our interest in carrying out this research was with the hope that if successful, that is if DHS/MICS and NGO estimates were similar, then NGOs could forego baseline data collection and use as a substitute DHS or MICS estimates, or estimates from some other publically available dataset, saving the NGO time and money, and reducing respondent burden. While we cannot give a blanket recommendation that DHS and MICS could always replace NGO baseline surveys, there are at least some situations where DHS/MICS could be used to the NGO's advantage:

- When the estimate is expected to be less than 15% or above 85%;
- When the DHS/MICS data were collected within the past year and the sample size is large (e.g., greater than 500);
- When the indicator of interest is one of the few with consistent similarity between DHS/MICS and NGO estimates;
- When the NGO has tolerance for estimates of low or unknown accuracy and does not need estimates for all 100+ indicators.

We had hypothesized that publically available data can provide estimates of baseline conditions similar to those reported in NGO baseline reports when matched as closely as possible for location, year, and season of data collection and that the impact of differences in year, geographical level, and season varies across health indicators. We found that as year difference increased, the mean difference between estimates slightly increased, and estimates derived data from lower geographical levels (such as village or district from NGO and province for DHS/MICS) contributed to a higher mean absolute difference between estimates. In general, larger sample sizes were obtained at higher geographical levels and the larger the sample size from NGO or DHS/MICS, the smaller the mean absolute difference between estimates. Whether the seasons of data collection were matched or different did not make a measurable difference to the similarity between estimates.

However, the partition of variance analyses showed that DHS/MICS and NGO estimates differed, for the most part, in unpredictable ways, and different geographical levels and years difference explained only a small part of the variation, which indicates that variability between estimates was mostly due to random noise or other factors not captured in the models.

Various graphs and statistical models were constructed and examined to try to understand how the differences between estimates varied with differences in year, geographic level, and season. There were no models that provided consistent, coherent explanations of the role of these effects in the difference of estimates.

We hypothesize that large differences between estimates from NGO baseline reports and publically available data might be due to three main reasons: (i) they are not measuring the same underlying true value; (ii) they are not measuring the indicator in the same way; or (iii) the NGO or DHS/MICS teams are measuring the indicator with high technical error of measurement.

(i) It is possible that NGOs' estimates are of different populations (DHS/MICS are nationally representative, whereas NGOs often try to target more marginalized villages); resulting in NGO estimates appearing worse off than the DHS/MICS estimates. However, by comparing indicators related to household wealth, we observed that the mean differences for these indicators were smaller than most indicators. Still, users of DHS/MICS data would need to keep in mind the possibility of different levels of wealth between their target population and DHS/MICS populations.

(ii) Different methods employed while sampling, collecting, processing and analyzing data might also have contributed to the high differences between DHS/MICS and NGO estimates. Many of the NGO baseline reports did not provide details of sampling or analytical methods used, so we were not sure how to match analyses with DHS/MICS survey data.

(iii) Several indicators related to maternal and child health included in this study have not been validated and some have been shown to have low validity, such as maternal report of skilled birth attendance. Low accuracy in reporting can result in bias in unpredictable direction and dimension, resulting in large differences between estimates.

Whatever the cause of the large differences between estimates was, it was not possible to know which of the data sources (DHS/MICS or NGO) provided the most accurate estimation of the true prevalence in the NGOs target population. Furthermore, while we have been comparing DHS/MICS and NGO point estimates, the indicators are measured with error and a typical 95% confidence interval for both DHS/MICS and for NGO estimates would of +/- 5% to 10%.

Much of the analyses in this report document the differences between NGO and DHS/MICS estimates and the lack of identifiable patterns in those differences. However, depending on the objectives, publically available data can be useful in some cases to complement or replace baseline studies and to obtain estimates of important indicators. The case studies from Nepal and Vietnam had many indicators where the DHS/MICS and NGO estimates were similar.

Large population-based health surveys can also be a valuable source of data to understand complex relationships, such as poverty and health, and when planning interventions, since they provide highly standardized data across different countries and regions. Using secondary data can also be interesting in situations of budget or mobility restraint, such as during the COVID-19 pandemic with limited data collection opportunities. However, when using DHS/MICS data, the user must keep in mind the differences between DHS/MICS estimates and NGO estimates highlighted in this report.

CONCLUSION

Our first hypothesis was “that publically available data can provide estimates of baseline conditions similar to those reported in NGO baseline reports when matched as closely as possible for location, year, and season of data collection”. Our answer to this, in brief, is that publically available data can be used, if the NGO is tolerant of imprecise estimates.

Our second hypothesis was “that the impact of differences in year, geographical level, and season varies across health indicators”. Our answer to this is yes (except for season), but the variation due to year, geographic level and health indicator is only a small part of the total variation, and even after accounting for those sources of variation there will still be differences between the estimates from the two data sources.

While an NGO may use the evidence presented here to justify forgoing their own baseline survey, they should keep in mind that DHS and MICS provide estimates for only some of the indicators of interest to the NGO. On average, we estimated 18 of the NGO’s indicators using DHS/MICS, but NGOs were often reporting 100+ estimates, which may justify their carrying out of a full baseline survey. Furthermore, being in the field can provide valuable insights for project design and implementation.