GRIDDED POPULATION SURVEY CASE STUDY

Author: Dale Rhoda & Caitlin Clary, Biostat Global Consulting
Date: May 10, 2024

POPGRID DATA COLLABORATIVE

OVERVIEW OF VACCINATION COVERAGE SURVEYS

The World Health Organization (WHO) has provided instructions for conducting high-quality household surveys to estimate vaccination coverage since the late 1970s. One application of such surveys is to evaluate the performance of vaccination campaigns, which occur when a Ministry of Health conducts a special effort to vaccinate every eligible child in the country over the course of just a few weeks. These campaigns aim to increase population immunity to highly contagious diseases such as measles. If a campaign immunizes a high proportion of children, then outbreaks and epidemics will be very unlikely among the cohort who received the campaign dose.

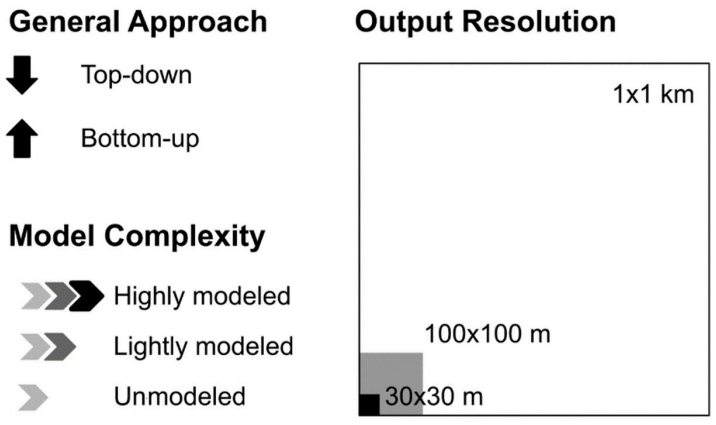
A post-campaign coverage survey (PCCS) is a special household survey where respondents are selected through a rigorously random selection process and asked whether their child received the campaign dose. Some campaigns give children cards to serve as a record that they received the dose. Most campaigns also put a pen mark on the child's fingernail, so the child will not be vaccinated twice and to serve as a reminder to the caregiver, when asked, to say that yes, the child was vaccinated. The survey interviewer asks to see the card and the child's finger and asks the child's caregiver whether they received the campaign dose.

Although campaigns estimate coverage by dividing the number of doses administered by the estimated eligible population, it is a best practice to also conduct an independent PCCS.

OVERVIEW OF GRIDDED POPULATION DATA

Over the last two decades, massive increases in computing power and the availability of high-resolution Earth observation and open geospatial data have fueled the development of gridded population estimates. These modeled population maps are typically described in terms of their general approach (top-down vs. bottom-up), modeling technique (highly, lightly, and unmodeled), and output resolution (Leyk, et al., 2019).

Top-down gridded population models are based on population counts in census or other geographic units that cover the entire population. Generally, top-down datasets are dasymetric, meaning that population disaggregation is informed by auxiliary datasets and that the estimated population in grid cells matches the population of input geographic units.



Bottom-up models use counts of the population in a selection of small areas, or assumptions about household size, to predict population in unmeasured grid cells. Modeling approaches include highly complex statistical or geographic algorithms that vary population based on the presence of human activity as measured with multiple auxiliary spatial datasets. Lightly modeled datasets disaggregate census or similar population counts equally among cells; however, disaggregation is constrained to populated places first, as defined by settlement extents or building footprints (top-down only). Unmodeled gridded population datasets uniformly disaggregate census population estimates to all grid cells (top-down) or directly aggregate point locations to grid cells (bottom-up).

Available multicountry gridded population datasets at the time of this writing include GPWv4.11, LandScan Global and HD, GHS-POP R2023A, WorldPop Global Constrained and Unconstrained, and HRSL.

DATASET	PRODUCER	COVER	YEAR(S)	APPROACH	MODEL	RESOLUTION
GPWv4.11	CIESIN, Columbia	Global	2000 - 2020, in 5 yr intervals	Top-down	Un-modeled	~1x1km
LandScan Global	ORNL	Global	2022	Top-down	Highly	~1x1km
LandScan HD	ORNL	33 Countries	~2020	Bottom-up	Highly	~100x100m
GHS-POP R2023A	EC-JRC	Global	1975 - 2030, in 5 yr intervals	Top-down	Lightly	~100x100m
WorldPop Un-constrained	WorldPop, Southampton	Global	2000 - 2020, in 1 yr intervals	Top-down	Highly	~100x100m
WorldPop Constrained	WorldPop, Southampton	Global	2020	Top-down	Highly	~100x100m
HRSL	Meta (Facebook)	140 Countries	~2020	Top-down	Lightly	~30x30 m

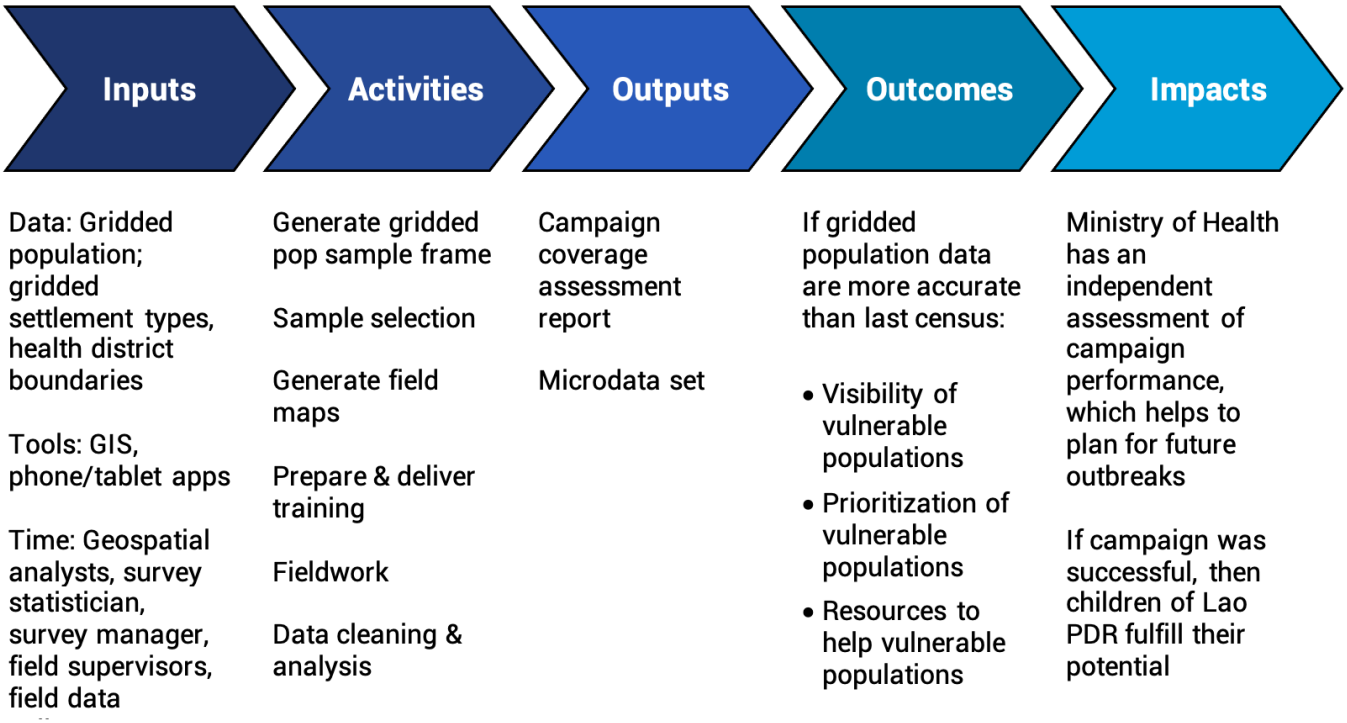
CASE STUDY

Gavi, the Vaccine Alliance, is financing a measles vaccination campaign in Lao PDR in the first half of 2024 and is funding ORB International and Biostat Global Consulting (BGC) to conduct a PCCS. The survey goals are to a) estimate national campaign coverage among the target population and b) oversample the population in health districts where measles outbreaks would be very likely if the campaign is not fully successful. The districts with populations of special concern were identified using administrative vaccination coverage data as well as historical data on measles surveillance, vaccine hesitancy, and cross-border traffic through the district.

To construct a survey sampling frame for the PCCS, BGC used Lao PDR health district boundary coordinates obtained by the WHO, GHS-POP R2023A gridded population data, and GHS-SMOD R2023a gridded settlement type data. The gridEZ R algorithm (https://github.com/caitlinbclary/gridEZ) was used to produce the sampling frame. Originally developed by Claire Dooley and Dana Thomson this algorithm has been recently updated to remove dependencies on deprecated spatial packages in R. The gridEZ algorithm was employed to create primary sampling units (PSUs) covering the study area, with a specified target population and maximum geographic size. The data inputs to gridEZ (boundaries, gridded population, gridded settlement types) enable the algorithm to produce PSUs that respect health district boundaries, have an estimated population close to the target, and have homogenous settlement types where possible.

The collection of 130 PSUs will be drawn from every province of the country in rough proportion to population counts, with adjustments made to accomplish the oversampling. Survey weights will be employed to account for unequal probability of selection.

PATHWAYS TO IMPACT



POTENTIAL POSITIVE IMPACTS

Time is of the essence with a PCCS — we want to visit the households while the memory of the campaign is fresh in the mind of the children's caregivers. Using the gridded population-based sample frame allowed us to generate PSU data directly in a format that is compatible with the survey implementer's field tools and to adapt where necessary to achieve an efficient workflow in the field. Vulnerable populations are very much at the heart of the evaluation goal for this work because they are likely to experience outbreaks. To the degree that the sample frame captures their locations well, they will be incorporated into the sample and contribute their important data and perspective on the campaign and on the status of measles vaccination in the country.

POTENTIAL NEGATIVE IMPACTS

The fieldwork is still in the future, so no negative impacts are documentable yet.

SUMMARY

Gridded population data facilitated development of a customized up-to-date survey sample frame and efficient handoff of PSU coordinates to the fieldwork team. The frame is being used to oversample vulnerable populations nested within a nationally representative survey. Both flexibility and efficiency are beneficial, especially for collecting the data very soon after the campaign to mitigate concerns about respondent recall problems and biases.

REFERENCES

World Health Organization. World Health Organization vaccination coverage cluster surveys: reference manual. No. WHO/IVB/18.09. World Health Organization, 2018. (**https://iris.who.int/handle/10665/272820**)

Leyk, S.; Gaughan, A.E.; Adamo, S.B.; de Sherbinin, A.; Balk, D.; Freire, S.; Rose, A.; Stevens, F.R.; Blankespoor, B.; Frye, C.; et al. Allocating people to pixels: A review of large-scale gridded population data products and their fitness for use. Earth Syst. Sci. Data Discuss. 2019, 11, 1385–1409.

Thomson, D.R. 2022. Designing and Implementing Gridded Population Surveys (D.A. Rhoda, Ed.). Dana Thomson Consulting. (**www.gridpopsurvey.com**)