VaxPop: Vaccination coverage estimation to support immunization programming in low- and middle-income countries

Background

Accurate, spatially detailed datasets on vaccination coverage and related indicators are required by immunization programs to plan and implement interventions and monitor progress towards key development goals such as the SDGs and Immunization Agenda 2030. In many countries where routine data sources can often be incomplete and unreliable, subnational estimates of vaccination coverage are produced using nationally representative surveys such as the Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS). However, most surveys are designed to be representative at the national and provincial levels, due to the high cost of more intensive sampling to produce estimates at lower administrative levels, such as the district level. The VaxPop project addresses this challenge by using state-of-the-art statistical and machine learning methods to produce gridded estimates of indicators of vaccination coverage, which can then be aggregated to operationally relevant administrative levels and integrated with other ancillary data sets for programmatic use.

Method - Data

- Data for mapping vaccination coverage are obtained from household surveys such as DHS, MICS, national immunization coverage surveys (NICS) and post-campaign coverage surveys (PCCS). The most recent survey is typically used.
- These are supplemented with geospatial covariate data obtained from a variety of sources, including those derived from surveys.

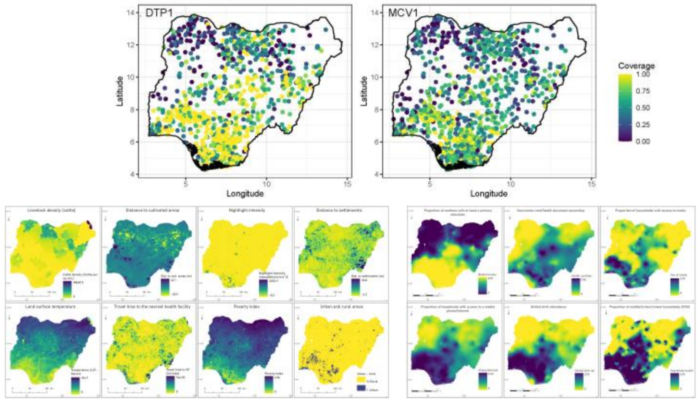

Figure 1: Example vaccination coverage and geospatial covariate data

Method – Data processing and covariate selection

- Individual level cluster level data are extracted and summarized at the cluster level.
- Evidence of vaccination often includes information from vaccination cards and caregiver recall
- Geospatial covariate data layers are processed and harmonized to 1x1 km resolution.
- Wherever applicable, the displacement of the survey cluster locations is taken into account when extracting covariate values at the cluster level.
- Covariate selection is undertaken in a non-spatial framework to resolve multicollinearity and identify important covariates for predicting coverage.

Method – Statistical analyses

- A geostatistical model is typically used in the analyses to model and predict coverage.

$$Y(\boldsymbol{s}) \sim \text{Binomial}\big(m(\boldsymbol{s}), p(\boldsymbol{s})\big)$$

$$\text{logit}\big(p(\boldsymbol{s})\big) = \mathbf{x}(\boldsymbol{s})^T\boldsymbol{\beta} + \omega(\boldsymbol{s}) + \epsilon(\boldsymbol{s})$$

- The model leverages the relationships between vaccination coverage and geospatial covariates to improve the predictions.
- The model also includes random effects to account for spatial dependence and clustering (or non-spatial variation) in the data.
- The model is usually fitted in a Bayesian framework using the INLA-SPDE approach

Results

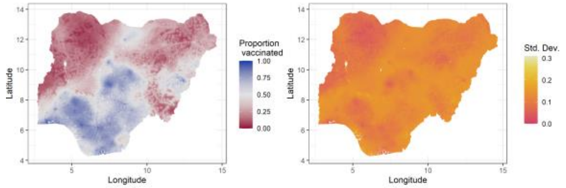

Figure 2: 1x1 km estimates of coverage and associated uncertainties

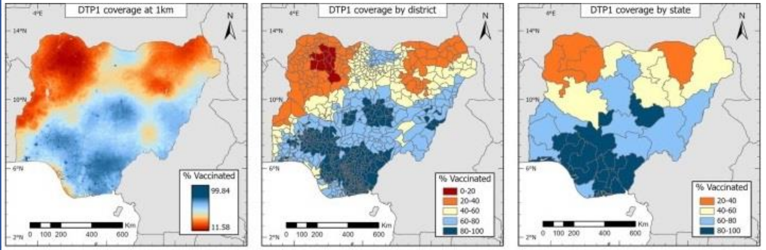

Figure 3: Aggregation of grid level coverage estimates to the district and provincial levels through weighting using population data

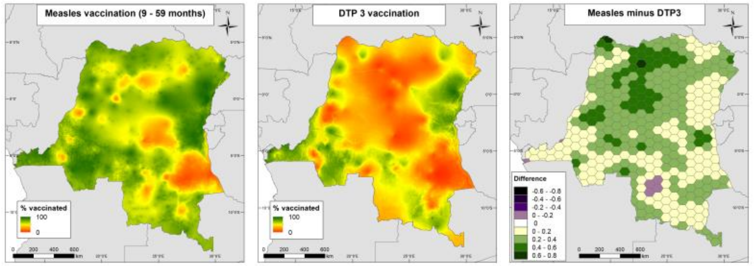

Figure 4: Comparison of MCV1 and DTP3 coverage to assess the effects of vaccine delivery mechanisms (i.e., routine immunization and campaigns)

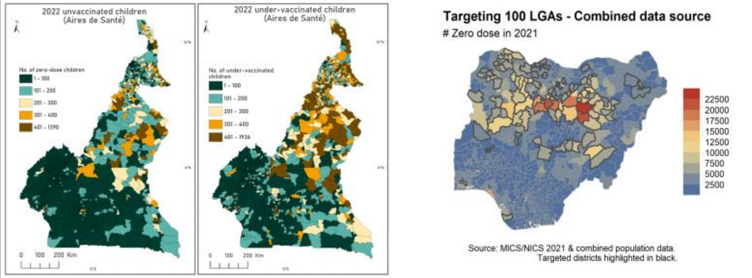

Figure 5: Use cases – estimates of numbers of zero-dose and under-vaccinated children and prioritization of districts for routine immunization interventions through integration with disease surveillance data

Reference

Utazi, C.E. *et al.* Mapping vaccination coverage to explore the effects of delivery mechanisms and inform vaccination strategies. *Nat Commun* **10**, 1633 (2019).