BMJ Open

High-resolution mapping of essential maternal and child health service coverage in Nigeria: a machine learning approach

Yoshito Kawakatsu, Jonathan F Mosser, Christopher Adolph, Peter Baffoe, Fatima Cheshi, Hirotsugu Aiga, D A Watkins, Kenneth H Sherr

ABSTRACT

Background National-level coverage estimates of maternal and child health (MCH) services mask district-level and community-level geographical inequities. The purpose of this study is to estimate grid-level coverage of essential MCH services in Nigeria using machine learning techniques.

Methods Essential MCH services in this study included antenatal care, facility-based delivery, childhood vaccinations and treatments of childhood illnesses. We estimated generalised additive models (GAMs) and gradient boosting regressions (GB) for each essential MCH service using data from five national representative cross-sectional surveys in Nigeria from 2003 to 2018 and geospatial socioeconomic, environmental and physical characteristics. Using the best-performed model for each service, we map predicted coverage at 1 km² and 5 km² spatial resolutions in urban and rural areas, respectively.

Results GAMs consistently outperformed GB models across a range of essential MCH services, demonstrating low systematic prediction errors. High-resolution maps revealed stark geographic disparities in MCH service coverage, especially between rural and urban areas and among different states and service types. Temporal trends indicated an overall increase in MCH service coverage from 2003 to 2018, although with variations by service type and location. Priority areas with lower coverage of both maternal and vaccination services were identified, mostly located in the northern parts of Nigeria.

Conclusion High-resolution spatial estimates can guide geographic prioritisation and help develop better strategies for implementation plans, allowing limited resources to be targeted to areas with lower coverage of essential MCH services.

BACKGROUND

National and subnational coverage of essential maternal and child health (MCH) services in low/middle-income countries (LMICs) are typically estimated using nationally representative cross-sectional surveys, including Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS).

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study used representative data from five national household surveys conducted in Nigeria from 2003 to 2018, including over 150 000 households.
⇒ Health service coverage of 10 essential maternal and child health services were estimated under the best-performing model with a wide range of geospatial data.
⇒ We compared gradient boosting models and generalised additive models for predicting health service coverage by assessing four performance criteria: weighted root-mean-squared error, mean bias error, mean squared error and nominal coverage of 95% prediction intervals.
⇒ Our modelling approach can be applicable to other countries with similar household survey datasets and health services which are associated with geospatial data.
⇒ However, our approach is less applicable for health services like treatments of childhood illnesses, which are more associated with individual and contextual factors.

Because these surveys are generally designed to produce estimates at the national and subnational levels (eg, administrative level 1) exclusively, these estimates can mask district-level and community-level inequities in service coverage. As a result, it can be difficult for policymakers to identify underserved geographic locations and prioritise interventions that match each community’s context and needs.

Although estimates of essential MCH service coverage are rarely available at the lowest administrative level in LMICs, several earlier studies reported service coverage of diphtheria-tetanus-pertussis (DTP) vaccinations, measles vaccinations and contraceptive use at subdistrict or community levels in LMICs.¹-⁴ Prevalence rates of HIV⁵ and


Prepublication history
► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2023-080135).

Received 21 September 2023
Accepted 12 May 2024

© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to Yoshito Kawakatsu; y.kawakatsu.0829@gmail.com
Plasmodium falciparum were also estimated at the subnational level in some LMICs. Finally, the DHS Programme Map Surfaces provides grid-level estimates of coverage for multiple health service estimated using a Bayesian geostatistical approach with a standardised set of covariates across countries. However, these studies focus only on a single survey or a single indicator, whereas studies combining data from multiple surveys to investigate multiple indicators over time are lacking. The DHS Programme Map Surfaces reports recommend the use of additional geospatial covariates and integrated data from multiple surveys to further improve their methodologies and monitor progress towards universal health coverage (UHC) at subnational level.

Nigeria has the highest population and GDP in Africa. However, a high burden of maternal and child mortality in Nigeria significantly contributes to global maternal and child deaths. Previous surveys and studies identified noticeable disparities in health outcomes and disease burden across regions in Nigeria. Also, its diverse sociocultural and economic characteristics influences health-seeking behaviours and service utilisation differently across states and lower administrative areas, necessitating a granular analysis.

There is a critical need for district-level and community-level data on coverage of essential MCH services, so that countries may effectively target interventions, including new health facilities and outreach services, to communities with low coverage of multiple MCH services. Therefore, this study aimed at estimating grid-level coverage of essential MCH services (ie, antenatal care (ANC), facility-based deliveries, childhood vaccinations and treatments of childhood illnesses) in Nigeria using machine learning methods. We further compare the prediction precision of two machine learning techniques, generalised additive models (GAMs) and gradient boosting (GB). We then use these methods to create high-resolution maps to visualise inequities in MCH service coverage across specific locations in Nigeria.

METHODS
Survey data in Nigeria

This study uses data from five nationally representative cross-sectional surveys administered in Nigeria between 2003 and 2018 that include geolocation data of primary sampling units (PSUs): DHS 2003, DHS 2008, DHS 2013, DHS 2018 and MICS 2016–2017. DHS data were extracted from IPUMS DHS. Methodological details on the conduct of these five surveys are published elsewhere.

All five surveys employed stratified two-stage or three-stage cluster sampling. Each PSU for DHS 2003 was composed of one or more enumeration areas (EAs) developed for the 1991 Population Census. PSUs for other DHS and MICS from 2016 to 2017 consisted of one or more EAs developed for the Population and Household Census 2006. MICS 2016–2017 excluded 101 PSUs located in Borno, Yobe and Adamawa States due to insecurity in those regions. For the same reason, DHS 2018 excluded 11 of 27 local government areas (LGAs) in Borno State.

To protect confidentiality of personal geoinformation, the public versions of all five surveys’ data randomly displaced Global Positioning System coordinates of the locations of respondents’ PSUs. Geolocations of urban PSUs were randomly displaced within 2 km buffers, whereas rural PSUs were displaced within either 5 km buffers in 99% of cases or 10 km buffers in 1% of cases. The direction and distance of displacement for each household was randomly determined. A previous study reported that errors in direction and distance due to random household displacement were sufficiently negligible to estimate measles vaccination coverage in a 10 km² grid. Geolocation data for 16 of 3533 PSUs (0.5%) were missing across the four DHSs. Similarly, geolocation data for 1 of 2239 PSUs (0.0004%) was missing in MICS 2016/2017. After initial random displacement, 14 PSUs (1 in DHS 2008 and 13 in MICS 2016/2017) were ‘located’ either in the sea or outside country boundaries. We resampled those PSUs by conducting random displacement again within 5 km until they were within Nigeria’s land boundaries. If there was no appropriate point within 5 km from the original PSUs’ locations, we resampled those PSUs within 10 km. Of these 14 PSUs, 8 were successfully resampled, and 6 cases that could not be appropriately displaced across 10,000 attempts were discarded. Online supplemental file 1 shows the locations of PSUs with at least one target population for ANC and facility-based delivery.

Essential MCH services

Ten essential MCH services were included in this study: (1) ANC; (2) facility-based delivery; (3–8) childhood vaccinations (BCG, first and third DTP/pentavalent (Penta) and oral polio (OPV), and first dose of measles); and (9 and 10) treatments of childhood illnesses for fever/cough and diarrhoea. The definitions and study population of essential services are shown in table 1. The target population for ANC and facility-based delivery are women 15–49 years of age having given a live birth during the last 23 months. Children 12–23 months of age are the study population for childhood vaccinations.

The outcome variables for this study are the grid-level utilisation rates for each essential MCH service. First, we created a map with 1 km×1 km cells in urban areas and 5 km×5 km cells in rural areas. In this study, we employed a grid cell-based approach to classify urban and rural areas, adhering to the harmonised definition by Dijkstra and Poelman. Urban areas were defined as grid cells with a population density of at least 300 inhabitants per km² and a total population exceeding 5000, ensuring a consistent classification across Nigeria. Grid cells not meeting these criteria were classified as rural. To accurately capture the heterogeneous nature of urban areas and manage computational efficiency across Nigeria’s vast landscapes, we differentiated spatial resolutions: 1 km×1 km cells...
were used for urban areas to detail the high variability in population distribution and access to services, whereas 5 km×5 km cells were applied in rural areas, balancing the need for detailed analysis and computational feasibility.

For each essential MCH service, survey year and grid cell, we constructed a utilisation rate using the count of children 12–23 months of age or women 15–49 years of age who used that service as the numerator, and the number of individuals categorised into the study population, children 12–23 months of age or women 15–49 years of age depending on the indicator, as the denominator.

Covariate data
Several earlier studies found geospatial socioeconomic, environmental and physical factors to be associated with the spatial distribution of under-five mortality cases and geographical heterogeneity in MCH service coverage.26,27,28 Table 2 shows the geospatial covariates employed in this study, except for longitude and latitude. Household poverty and female and male education attainment by grid cell were extracted from high-resolution data published in earlier studies.26,27 Population density and total population data were extracted from the WorldPop database.28 Road density data were extracted from the Global Roads Inventory Project.29 Meteorological data such as precipitation and evapotranspiration and nighttime light data were extracted from TerraClimate30 and High Resolution Electricity Access.31 We also created three spatial covariates: (1) travel time to the most accessible health facility; (2) travel time to the nearest city; and (3) the number of major roads crossing the household’s grid cell. Travel time to the most accessible health facility was estimated by using the geolocations of health facilities run by state governments, community-based organisations and faith-based organisations in Nigeria32 with the friction surface described in the section of Model performance indicators and selection of final models. Finally, we create high-resolution coverage maps for each essential MCH service using the best prediction model for that service. We also estimated the aggregated service coverage at various administrative levels and in rural and urban areas. The High-resolution map creation section provides methods of creating high-resolution maps and estimating aggregated service coverage.

### Prediction modelling and high-resolution map creation
We compared the prediction performances of two models, GAMs and GB models, using a pooled dataset of five nationally representative cross-sectional surveys. These models are described in the Prediction models section. For each service, we identify the best prediction model based on each model’s performance on four indicators described in the section of Model performance indicators and selection of final models. Finally, we create high-resolution coverage maps for each essential MCH service using the best prediction model for that service. We also estimated the aggregated service coverage at various administrative levels and in rural and urban areas. The High-resolution map creation section provides methods of creating high-resolution maps and estimating aggregated service coverage.

### Prediction models: GAMs and GB models
We consider two modelling strategies from machine learning—GAMs and GB models—for predicting essential MCH service coverage. Both models rely on the same set of covariates but employ different estimation techniques: GAMs emphasise non-parametric estimation of flexible functional forms, whereas GB involves prediction from a weighed ensemble of sequentially constructed models, each of which attempts to better predict cases

<table>
<thead>
<tr>
<th>Health service</th>
<th>Definition</th>
<th>Target population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four or more antenatal care visits</td>
<td>Antenatal care four or more times by trained health personnel (ie, doctor, nurse, midwife, auxiliary midwife) during pregnancy at the point of survey</td>
<td>Women aged 15–49 years with a last birth in the last 23 months</td>
</tr>
<tr>
<td>Facility-based delivery</td>
<td>Delivery at public or private health facility in the last 23 months at the point of survey</td>
<td>Women aged 15–49 years with live births in the last 23 months</td>
</tr>
<tr>
<td>Childhood immunisation</td>
<td>Children who had received one dose of BCG vaccine</td>
<td>Children aged 12–23 months</td>
</tr>
<tr>
<td></td>
<td>Children who had received first and third dose of diphtheria-tetanus-pertussis vaccine or pentavalent vaccine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children who had received first and third dose of oral polio vaccine</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Children who had received first dose of measles vaccine</td>
<td></td>
</tr>
<tr>
<td>Treatment for common childhood illness</td>
<td>Children under five with fever/cough and diarrhoea in the last 2 weeks for whom care was sought at a health facility</td>
<td>Children aged 0–59 months</td>
</tr>
<tr>
<td>Category</td>
<td>Year*</td>
<td>Data type</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>Poverty</td>
<td>2010</td>
<td>Continuous</td>
</tr>
<tr>
<td>Male education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td>2003–2018</td>
<td>Continuous</td>
</tr>
<tr>
<td>Total population</td>
<td>2003–2018</td>
<td>Continuous</td>
</tr>
<tr>
<td>Time to health facility</td>
<td>2015</td>
<td>Continuous</td>
</tr>
<tr>
<td>Time to city</td>
<td>2015</td>
<td>Continuous</td>
</tr>
<tr>
<td>Number of main roads crossing a grid</td>
<td>2021</td>
<td>Continuous</td>
</tr>
<tr>
<td>Road density</td>
<td>2018</td>
<td>Continuous</td>
</tr>
<tr>
<td>Precipitation</td>
<td>2003–2018</td>
<td>Continuous</td>
</tr>
<tr>
<td>Evapotranspiration</td>
<td>2003–2018</td>
<td>Continuous</td>
</tr>
<tr>
<td>Nighttime lights</td>
<td>2012–2018</td>
<td>Continuous</td>
</tr>
<tr>
<td>Malaria prevalence</td>
<td>2003–2018</td>
<td>Continuous</td>
</tr>
<tr>
<td>Childhood diarrhoea</td>
<td>2003–2017</td>
<td>Continuous</td>
</tr>
</tbody>
</table>

*Data availability between 2003 and 2018.
that were poorly predicted by the previous model in the sequence.

**Generalised additive models**

We estimated grid-level quasi-binomial GAMs for each essential MCH service using the mgcv package in R. The functional form of the models allows for non-linear relationships between the outcome and covariates, and is given by:

\[
\log \left( \frac{y_{ijt}}{\mu_{ijt}} \right) = \beta_0 + \beta_1 \text{Urban}_{ijt} + \text{State}_{ij}\cdot \text{year}_{ijt} + \text{(Interaction)} + \text{ Poverty}_{ijt} + \\
\text{(Female sổdu) + (Male sokdu) + (Pop sokdu) + (Pop_total) + (No_road) + (Road_den) + (Time_city) + (Time_HF) + (Rain) + (Dry) + (Light) + (Male sex) + (Male sex + (lat, long) is an isotropic smooth of latitude and longitude on the sphere with second derivative penalty. In all other cases, is the count of cluster locations with non-zero service coverage outcomes followed the Tweedie distribution. We tested a sequence of variance power between 1 and 2 by running iterative models and selected the best power based on the negative log-likelihood. Moreover, we conducted a grid search to select the best set of hyper-parameters based on negative log-likelihood for Tweedie regression. In the grid search, we tested the following parameters and their ranges: (1) the maximum depth of a tree, 2–10 by increments of 2; (2) the minimum sum of instance weight needed in a child, 1–3 by increments of 0.5; (3) the subsample ratio of columns when constructing each tree, 0.5–1.0 by increments of 0.1; (4) the subsample ratio of the training instances, 0.5–1.0 by increments of 0.1; and (5) step size shrinkage used in update to prevent overfitting, 0.01–0.3 (specifically, we tested 0.01–0.1 by increments of 0.01, as well as 0.2 and 0.3).

**Model performance indicators and selection of final models**

We compared the performance of the GAM and GB models for each MCH service on four performance criteria, each of which was computing using 5-fold cross-validation:

1. Weighted root-mean-squared error (WMSRE) of proportions of essential MCH service utilizations,

\[
\text{WMSRE} = \sum \left( d_i \times \sqrt{\left( \frac{P_i - \hat{P}_i}{\hat{P}_i} \right)^2} / \sum d_i \right);
\]

2. Mean bias error,

\[
\text{MBE} = \frac{1}{n} \sum (P_i - \hat{P}_i);
\]

3. Mean squared error,

\[
\text{MSE} = \frac{1}{n} \sum (P_i - \hat{P}_i)^2;
\]

4. Nominal coverage of 95% prediction intervals,

\[
\text{nCoverage} = 100 \times \frac{n}{\sum_{i=1}^{n} I(l_i \leq \hat{y}_i \leq u_i) / n}.
\]

\(P_i\) and \(\hat{P}_i\) denote observed and predicted proportions at cluster \(i\), \(d_i\) is the size of the study population at cluster \(i\), and \(n\) is the count of cluster locations with non-zero study populations. \(l_i\) and \(u_i\) are the lower and upper limits of the prediction interval and \(I(\cdot)\) is an indicator function. Nominal coverage indicates the proportion of the estimated numbers of study population having used essential maternal and child services (\(y_i\)) within 95% CIs of predicted services coverage. Due to erratic behaviour around the endpoints (ie, 0 and 1) using binomial probabilities, whenever observed values were exactly 0 or 1, we set the estimated lower and upper CIs in GAMs to either 0 or 1, respectively. We employed methods employed in an earlier study on cluster-level estimates of measles vaccinations to calculate the nominal coverage of 95% prediction intervals.

To determine a final model for each MCH service, we first checked whether either the GAM or GB model was dominant on all the four measures of goodness of fit. If either model was uniformly superior on these four metrics, we selected it as the prediction model for the
service. When the results of the four metrics disagreed on the best model, we further examined whether one of the models had WRMSE at least 0.01 units lower, MSE at least 0.005 units lower, MBE at least 0.01 units lower, or nCoverage at least 1% closer to 95%. Whichever model met a greater number of these criteria than the others was selected as the prediction model for that service. There was no case that both the GAM and GB model met the same number of these criteria.

High-resolution map creation
Using the best predicting model, either GAM or GB, for each service, we created high-resolution coverage maps for essential MCH services from 2003 to 2018. We also computed 95% CIs of predicted grid-level service coverage using the estimated SE for each predicted value for GAMs or bootstrapped SEs for GB models. In addition, to ensure the logical sequence of vaccine coverage across vaccination doses, we checked the number of grids that have a higher coverage for a subsequent dose (Penta3/OPV3) than that for an initial dose (Penta1/OPV1). We have not found such grids over years.

We estimated aggregate predicted proportions aggregated at various administrative levels by calculating weighted mean values and using as weights in each grid location the proxy target population from WorldPop. For ANC and facility-based delivery, we used the sum of the number of children 0–12 months of age in the year of estimation and in the prior year as a proxy for the number of women 15–49 years of age with a live birth during the last 23 months. For childhood vaccinations, we used the total number of children 0–12 months of age in the prior year as a proxy for the number of children aged 12–23 months. To estimate gaps in MCH service coverage between rural and urban areas over time, the difference-in-difference (DiD) estimator for each essential MCH service was calculated by taking the difference in an MCH service coverage of urban areas between 2018 and 2003 and subtracting the difference for rural areas between 2018 and 2003. 95% CIs for the DiD estimates were computed using bootstrapped SEs. A bivariate choropleth map is used to display mean MCH coverage.

Following previous work,35 we treated as missing values of MCH service coverage for the grid locations which include either lakes or very low population (less than 10 people per km²). We did not estimate predicted values for grid locations lacking covariate data.

Patient and public involvement
Patients or the public were not involved in the design, conduct, reporting or dissemination plans of our research.

RESULTS
Model selection and validation
Online supplemental file 2 reports the performance of the GAMs and GB models. Across a range of essential MCH services, GAMs matched or outperformed GB models based on all four performance indicators, as shown in online supplemental file 2. In particular, for ANC, facility-based delivery, BCG, first and third Penta, GAMs performed better than GB models on all four measures: MBEs, MSEs, WRMSEs and nominal 95% coverage. For third OPV and measles vaccinations, GAMs outperformed GB models in terms of MBE and the nominal coverage, while obtaining similar MSEs and WRMSEs. For first OPV, GAMs outperformed GB models on MBE, while MSE, WRMSE and nominal coverage were similar across both approaches. Finally, due to low predictive power and wide prediction intervals in both GAM and GB models, we did not select final models or create coverage maps for treatments of childhood illnesses.

The final predictive models had uniformly low systematic prediction error, with a range of MBE from 0.001 to 0.016 across 10 essential MCH services. The range of MSEs was similarly narrow (0.057–0.068).

To check model fitting, we further compared the observed and estimated share of study populations using essential MCH services (see online supplemental file 3a). Also, online supplemental file 3b includes in-sample and out-of-sample plots for observed vs predicted probabilities for each essential MCH service. Online supplemental file 4 provides plots of smoothed functional forms from the GAMs. GB model estimates of the importance of each feature for predicting each MCH service based are provided in online supplemental file 5. As shown in online supplemental file 4, smoother fitted by the GAMs show different slopes by MCH service types. Online supplemental file 6 compares high-resolution maps of coverage of six essential MCH services based on GAMs and GB models.

Geographic inequality in essential MCH service coverage
High-resolution maps of coverage from 2003 to 2018 of 10 essential MCH services, excepting treatments of childhood illnesses, have been stored in Figshare. High-resolution maps of 6 of 10 essential MCH services (ANC, facility-based delivery, BCG, first Penta, third Penta and measles vaccination) as of 2018 are shown in figure 1 to highlight geographic inequality in service coverage across states and LGAs. Overall, MCH service coverage in rural areas was lower than urban areas across service types. Likewise, MCH service coverage was also lower in northern parts of Nigeria, though higher coverage can be found in some major cities and towns in those states.

Figure 2 shows LGA-level coverage estimates for 8 of 10 essential MCH services by state as of 2018. Online supplemental file 7 also provides mean, minimum and maximum LGA-level coverage of eight essential MCH services by state. Overall, the LGA-level inequalities of third OPV, third Penta and measles were narrower, compared with BCG, first OPV and first Penta. States with either high or low service coverage exhibited reduced IQRs at the LGA level, suggesting diminished intra-state disparities. Conversely, states with
Figure 1  Predicted service coverage of (A) antenatal care, (B) facility-based delivery, (C) BCG vaccination, (D), first pentavalent vaccination, (E) third pentavalent vaccination and (F) measles vaccination in 2018.

Moderate coverage revealed more substantial IQRs, indicative of increased variability in service utilisation across LGAs. The three states with the greatest LGA-level inequality in terms of ANC were Kwara (IQR: 54.2%–82.6%), Borno (IQR: 16.6%–40.8%) and Bayelsa (IQR: 23.4%–46.3%), whereas the top three states for LGA-level inequality in BCG vaccination were Kaduna (IQR: 53.4%–88.1%), Gombe (IQR: 45.3%–78.4%) and Borno (IQR: 35.7%–60.4%). Overall, we find that Borno, Gombe, Kaduna, Kwara and Niger states showed consistently high inequality across LGAs, with each of these states ranking among the top three states in terms of IQR for at least four of eight MCH services. In terms of inequality across the 10 MCH services shown in figure 2, we find the highest LGA-level IQR for facility-based delivery coverage, followed by BCG, first OPV and ANC. The LGA-level IQR for measles vaccination was the lowest of the eight services, although its mean coverage was the lowest, as well.

Figure 2  Predicted local government area (LGA)-level coverage of eight essential maternal and child health services in 2018 by state. Orange dots mark the mean coverage within each state. Grey dots show service coverage for each LGA within the state. The distribution of these grey dots is summarised by a violin plot.
Priority LGAs with lower coverage of essential MCH services

Figure 3 is an LGA-level bivariate choropleth map that two-dimensionally shows service coverage of childhood vaccinations (BCG, first and third DTP/pentavalent (Penta) and OPV, and first dose of measles) and maternal health (ANC and facility-based delivery). Coverage level was classified by tertile (ie, higher 33%, middle 33% and lower 33%). The map indicates that a majority of LGAs located in northern part of the country had lower childhood vaccination and maternal health service coverage, excepting some urban LGAs which had higher coverage for either or both types of services. The southwest, including districts in Ogun and Oyo states, had higher maternal health service coverage and lower childhood vaccination coverage. In contrast, the southeastern and central areas of the country had lower maternal service coverage and higher childhood vaccination coverage. Finally, those states in which 80% of LGAs were categorised as high coverage for both types of services. The dark purple areas in figure 3, comprising Abia, Anambra, Edo, Ekiti, Enugu, Imo, Lagos and Osun, can be found in the southern part of Nigeria and had relatively greater population sizes and density than other areas. Online supplemental file 9 provides LGA-level mean coverage of childhood vaccinations and maternal health services as of 2018, by state.

DISCUSSION

The global movement for UHC highlights the importance of identifying subnational inequities in health service coverage in order to better understand the characteristics of underserved populations and thereby correct maldistribution of health resources and services. This movement has led researchers to conduct studies on the heterogeneity of health service coverage within and between subnational levels and to identify the most vulnerable areas for priority setting. Responding to this need, this study estimated grid-level coverage of essential MCH services in Nigeria using publicly available geographic datasets and comparing the performance of two statistical approaches (GAMs and GB).

Cross-validation performance confirms the strong predictive power of the final models of essential service coverage, with the exception of treatments for childhood illnesses. MSE and MBE values of the final model for measles vaccination were slightly better than those reported in an earlier study on grid-level coverage estimates of measles vaccination that employed Bayesian multivariate spatial-temporal modelling. GAM with a shrinkage approach might be applied to a multicounty or global model, to simultaneously account for differing drivers and non-linear relationships. GAM estimation is also less computationally intensive than GB models, particularly to find the best set of hyper-parameters.

The importance of each feature for predicting MCH service coverage, shown in online supplemental file 5, varied by MCH service types. This suggests that further...
study may need to select an appropriate set of geospatial covariates for predicting coverage of specific health services. Additionally, the results for some predictors should be interpreted with caution. For example, education attainment is a model-based estimate generated using geospatial covariates such as night light, access to roads, population and aridity. The consistent importance of male and female education attainment in predicting coverage in GB models might therefore reflect the influence of the other geospatial covariates that were used to generate education status estimates, rather than educational status itself.

One avenue for increasing predictive power is to use health administrative data on essential MCH service utilisations reported from health facilities. Almost all countries have health management information systems that store time-series data on the delivery of key health services from public and private health facilities. For example, Nigeria launched the web-based software District Health Information System version 2 (DHIS2) in 2010. Childhood vaccination data have been collected on a monthly basis since 2014. National-level health administrative data were used for estimating national coverage of childhood vaccination services by the WHO and UNICEF and by a recent publication in the *Lancet*. In addition, DHIS2 contains the aggregated count of clients/patients who receive respective health services at the lowest administrative level every month. Using these lower-level health administrative data may increase predictive power. However, the quality of health administrative data from health facilities is questionable in many countries at present.

This study clarifies that national and subnational MCH service coverage measures mask significant spatial heterogeneity within Nigeria, which may result in poor decision-making on geographical prioritisation at the LGA and local levels. Likewise, inequities in MCH service coverage between LGAs varied by state and type of service. However, the distribution and use of health resources in Nigeria remain suboptimal in terms of demography, disease burden and pre-existing endowments of health resources. The high-resolution maps generated by this study could be an input into the decision-making process for better geographical prioritisation of health resources. Maps could help identify LGAs that may be significantly lagging behind in multiple indicators, thus requiring broad-scale interventions. In addition, the granular-level estimate may illuminate pockets of underservice within LGAs, thereby offering a more nuanced understanding that would otherwise be masked in broader LGA-level data. To operationalise the use of our high-resolution maps in decision-making, a collaborative framework with government health agencies is essential. We recommend the development of an interactive, user-friendly interface for real-time exploration of data, coupled with regular updates to maintain relevance. Training sessions for local health officials can further facilitate the accurate interpretation and application of these maps. Pilot studies may serve as an initial platform to assess the utility and effectiveness of this approach in guiding resource allocation and healthcare planning.

Whereas prioritising based on a single indicator may make sense for achieving specific goals, such as the Immunization Agenda 2030, broader agendas, such as the Sustainable Development Goals, require prioritisation based on multiple indicators. Accordingly, we recommend that criteria for prioritising geographic locations should not rely on a single indicator of coverage, but rather consider coverage rates for multiple health services in order to ensure optimal provision. In particular, the study highlights within-locality differences in coverage of childhood vaccinations and maternal health services, with some localities performing markedly better in one area or the other.

According to the Lancet Commissions, Nigeria needs to distribute available health resources more equitably through increasing resource management and strategic purchasing capacities. Routinely updated high-resolution maps could also support micro-planning of supplementary activities such as outreach services and childhood vaccination, especially if such mapping is connected with a health information management system such as DHIS2. Specifically, mapping can help estimate the number of individuals in need of services at the community level and thereby inform budgeting and resource allocation.

There are several limitations to this study. Our modelling approach is less applicable to estimating the coverage of health services like treatments of childhood illnesses, which are more associated with individual and contextual characteristics than spatial covariates. Moreover, several other factors likely contribute to the models’ weaker performance in this area. For instance, treatment-seeking behaviours for childhood illnesses can be highly variable, influenced by cultural and religious beliefs, gender dynamics and financial constraints. Additionally, health awareness and severity and frequency of disease can vary substantially, making it challenging to predict coverage effectively. These complexities are not easily captured by spatial covariates alone, resulting in reduced model performance for this service. Inequality between relatively affluent urban areas and their adjacent peri-urban areas suffering higher poverty rates were not captured in the maps we created due to limitations of the available input data. Moreover, our study investigated inequality not by income, educational status, and other factors, but by geospatial conditions only. Due to the data availability issue, we used some temporal mismatched data in our modelling. While the temporal mismatch is not ideal, the alternative of excluding these variables could result in a less robust model, losing potentially significant predictors.

**CONCLUSION**

This study visualised significant geographical inequities of essential MCH service utilisations in Nigeria.
The high-resolution maps herein provide health policymakers/planners with guidance for geographic prioritisation of specific MCH services. These estimates serve as a resource to further develop implementation strategies for maximising limited resources. Strengthening routine service coverage, optimally, is non-commercial; it may permit others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID IDs
Yoshito Kawakatsu http://orcid.org/0000-0003-1853-533X
Jonathan F Mosser http://orcid.org/0000-0001-9093-3186
D A Watkins http://orcid.org/0000-0001-6341-9595

REFERENCES


Open access

Fagbamigbe AF, Nnanatu CC. Modelling the spatial distribution and the factors associated with under-five mortality in Nigeria. *Spat Demogr* 2022;10:255–82.


Barok M. High resolution electricity access indicators dataset. Center for Political Studies UoM, 2021.


Chen T, Guostrin C. XGBoost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016:785–94.


