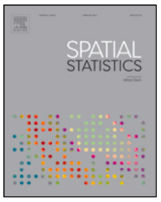
Spatial Statistics 57 (2023) 100772

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta



A zero-dose vulnerability index for equity assessment and spatial prioritization in low- and middle-income countries

C.E. Utazi [a,b,*], H.M.T. Chan [b], I. Olowe [a], A. Wigley [a], N. Tejedor-Garavito [a], A. Cunningham [a], M. Bondarenko [a], J. Lorin [c], D. Boyda [c], D. Hogan [c], A.J. Tatem [a]

[a] WorldPop, School of Geography and Environmental Science, University of Southampton, University Rd, Southampton, SO17 1BJ, UK
[b] School of Mathematical Sciences, University of Southampton, University Rd, Southampton, SO17 1BJ, UK
[c] Gavi, The Vaccine Alliance, Geneva, Switzerland

ARTICLE INFO

Article history:
Received 7 April 2023
Received in revised form 15 August 2023
Accepted 15 August 2023
Available online 21 August 2023

Keywords:
Vaccination coverage
Bayesian inference
INLA-SPDE approach
Demographic and Health Surveys
Multiple Indicator Cluster Surveys

ABSTRACT

Many low- and middle-income countries (LMICs) continue to experience substantial inequities in vaccination coverage despite recent efforts to reach missed communities and reduce zero-dose prevalence. Geographic inequities in vaccination coverage are often characterized by a multiplicity of risk factors which should be operationalized through data integration to inform more effective and equitable vaccination policies and programmes. Here, we explore approaches for integrating information from multiple risk factors to create a zero-dose vulnerability index to improve the identification and prioritization of vulnerable communities and understanding of inequities in vaccination coverage. We assembled geolocated data on vaccination coverage and associated risk factors in six LMICs, focusing on the coverage of DTP1, DTP3 and MCV1 vaccines as indicators of zero dose and under-vaccination. Using geospatial modelling techniques built on a suite of geospatial covariate information, we produced 1×1 km and district level maps of the previously unmapped risk factors and vaccination coverage. We then integrated data from the maps of the risk factors using different approaches to construct a zero-dose vulnerability index to classify districts

* Corresponding author at: WorldPop, School of Geography and Environmental Science, University of Southampton, University Rd, Southampton, SO17 1BJ, UK.
E-mail address: C.E.Utazi@soton.ac.uk (C.E. Utazi).

https://doi.org/10.1016/j.spasta.2023.100772

within the countries into different vulnerability groups, ranging from the least vulnerable (1) to the most vulnerable (5) areas. Through integration with population data, we estimated numbers of children aged under 1 living within the different vulnerability classes. Our results show substantial variation in the spatial distribution of the index, revealing the most vulnerable areas despite little variation in coverage in some cases. We found that the most distinguishing characteristics of the most vulnerable areas cut across the different subdomains (health, socioeconomic, demographic and geographic) of the risk factors included in our study. We also demonstrated that the index can be robustly estimated with fewer risk factors and without linkage to information on vaccination coverage. The index constitutes a practical and effective tool to guide targeted vaccination strategies in LMICs.

Contents

1. Introduction

The expansion of routine childhood immunization has been described as one of the most successful public health interventions, as evidenced by marked reductions in childhood mortality and morbidity globally (Li et al., 2021; Andre et al., 2008; World Health Organization, 2022) following the introduction of the Expanded Programme on Immunization by the World Health Organization (WHO) in 1974 (Keja et al., 1988). In particular, the past decade has witnessed a remarkable progress in the introduction of new vaccines in low- and middle-income countries (LMICs) such as rotavirus vaccines and those protecting against pneumococcal pneumonia, with impressive increases in their coverage. Despite this success, the coverage of basic routine vaccines still fall short of WHO's target in many LMICs (WHO, 2021; WHO and UNICEF, 2022) and increasing the coverage of routine immunization (RI) to reach the last 20% in these countries has been a major global health challenge. Partly occasioned by the COVID-19 pandemic, in 2021 about 25 million children were not fully vaccinated with all three recommended DTP doses, out of which 18 million were zero-dose children — operationally defined as children who did not receive any dose of the

diphtheria, tetanus, pertussis (DTP) containing vaccine. The coverage of DTP1 is often considered an indicator of access to RI services and a proxy for non-receipt of basic vaccines administered through the RI programme. Almost all zero-dose children were estimated to be living in LMICs in the African and South Asian regions, with India, Nigeria, Ethiopia, DRC and Pakistan being among the top 10 countries with the most unprotected children in 2021 (WHO and UNICEF, 2022).

Zero-dose and under-immunized children are often the most marginalized and deprived and are also reported to have the highest rates of vaccine-preventable morbidity and mortality (WHO, 2020). The persistence of these vulnerable populations likely results in continued disease circulation and increased frequency of outbreaks within countries, undermining disease control and elimination efforts. Recognizing this, the World Health Organization's Immunization Agenda 2030 has set a target of 50% reduction in the number of zero-dose children by 2030, relative to pre-pandemic levels (WHO, 2020). Reaching zero-dose and missed communities is also critical to achieving health policy targets set out within the Sustainable Development Goals (SDGs) (United Nations, 2015) and Gavi, the Vaccine Alliance's 2021–2025 Strategy (Gavi, 2020).

To design and implement effective interventions to reach communities at risk of zero dose and under-vaccination, current and reliable evidence is needed to identify where they reside and their most defining characteristics. Previous efforts have either focused on producing spatially detailed estimates of vaccination coverage to uncover inequities, identify areas of low coverage and locate vulnerabilities within the health system (Mosser et al., 2019; Utazi et al., 2018, 2020; Takahashi et al., 2017; Local Burden of Disease Vaccine Coverage Collaborators, 2021; Utazi et al., 2019), or estimating numbers of zero-dose children among key geographically marginalized populations such as those living in remote-rural, urban slum and conflict-affected areas (Wigley et al., 2022). However, beyond these key geographic characteristics, inequities in vaccination coverage are potentiated by many health-related, socioeconomic and demographic factors (Utazi et al., 2022b; Rainey et al., 2011), meaning that substantial proportions of zero-dose children could potentially be found in other settings (Wigley et al., 2022). Also, standalone assessments of coverage maps are limiting as these do not explore important geographical associations with major contributing factors for non- and under-vaccination which can yield actionable insights for vaccination programming. Mapping vaccination coverage to identify vulnerable communities is potentially an effective approach when there are substantial heterogeneities in the spatial distribution of vaccination coverage and when accurate, timely and reliable input data are available. However, in settings where coverage is sub-optimal and has less spatial variation or coverage estimates are of poor quality, accurate identification and prioritization of these vulnerable areas can be hampered greatly, necessitating approaches that utilize more information and which leverage the bidirectional dependence between vaccination coverage and associated risk factors for improved estimation of inequities (areas of low coverage are often the result of the combined effect of these risk factors; on the other hand, the persistence of low coverage areas can exacerbate inequities through increased morbidity and mortality).

Moreover, considering that household surveys which constitute the major source of data used to produce maps of vaccination coverage are conducted typically every 5 years in most LMICs, identifying a subset of risk factors, which could be potentially obtained from other sources such as health information systems, censuses, registries and other surveys, and mapped at subnational scales, could enable more regular mapping of vulnerabilities.

Here, we develop a zero-dose vulnerability index to support efforts towards identifying zero-dose and missed communities in LMICs and designing effective strategies to reach them. The index aims to operationalize the multiplicity of factors associated with zero dose and under-immunization through data integration, and to contribute to a better understanding of the role of geography in these associations through uncovering the most defining characteristics of the most vulnerable areas. We demonstrate that the index can be effectively used to capture geographic vulnerabilities when limited data on the risk factors are available and with or without the availability of information on vaccination coverage. Our methodology provides a mechanism to harmonize data from multiple spatial scales to produce the index at the district level, usually referred to as the second administrative level in most LMICs, which is the target geographic unit for our analysis.

2. Methods

2.1. Vaccination coverage data and associated risk factors

Data on risk factors for zero dose and under-immunization were obtained from the most recent household survey conducted in six countries. These were Demographic and Health Surveys (DHS) conducted in the Democratic Republic of Congo (DRC, 2013–14), Ethiopia (2016), India (2015–16), Pakistan (2017–18), Uganda (2016) and the Multiple Indicator Cluster Survey-National Immunization Coverage Survey (MICS-NICS) conducted in Nigeria in 2021. These countries were selected due to being the focus of previous work (Utazi et al., 2022b, 2019), having large numbers of zero-dose children (WHO and UNICEF, 2022) and being high priority countries for Gavi. From these surveys, we extracted data on 19 risk factors (see Supplementary Table 1) determined in a previous study (Utazi et al., 2022b) as top-ranking predictors of zero dose and under-vaccination. Other variables such as women's participation in decision-making, malnutrition (underweight), use of modern contraception and travel time to the nearest health facility were not included in Utazi et al. (2022b) but were considered to be of interest in this work. Also, conflict was not a top-ranking predictor in Utazi et al. (2022b), but it was included in this work due to current interest in identifying and reaching populations living in conflict areas within the global immunization community. We also extracted information on the coverage of DTP1, DTP3 and MCV1 vaccines, which were used in previous studies to evaluate the odds of zero dose and under-vaccination, to help inform the weights assigned to the risk factors when constructing the index using some approaches (see Section 2.4) and to estimate coverage levels in areas with differentiated levels of vulnerabilities. The risk factors and indicators of coverage are shown in Fig. 1, with detailed definitions provided in Supplementary Table 1. The definitions of the risk factors were calibrated to ensure uniformity in the direction of their relationships with vaccination coverage and to highlight as much as possible areas at greatest risk of zero dose and under-vaccination. Where applicable, our analyses relate to women of reproductive age (i.e., those aged between 15 and 49 years) and children aged between 12 and 23 months.

For the DHS-/MICS-NICS-derived risk factors and vaccination coverage, we first extracted and summarized information on these variables at the survey cluster level. Also, using the (displaced) geo-coordinates of these clusters, we extracted the cluster-level data values of the risk factors whose gridded estimates were available prior to our analyses using approaches described in Perez-Heydrich et al. (2013).

2.2. Geospatial covariate data, covariate selection and population data

To produce high resolution maps of the risk factors (where these were not already available) and vaccination coverage, we assembled a suite of geospatial socioeconomic, environmental and physical covariates. These covariates include nightlight intensity, vegetation index, livestock density, travel time to urban areas and land surface temperature — see Supplementary Table 3. As detailed in previous work (Utazi et al., 2018, 2019, 2020, 2021), these covariates were processed to produce 1×1 km input raster data sets and corresponding cluster level data using standard approaches (Perez-Heydrich et al., 2013). Following previous work (Utazi et al., 2018, 2019, 2020), covariate selection was carried out to determine the best set of covariates for modelling each outcome/risk factor for each country. The covariate selection process, implemented in a non-spatial modelling framework using binomial regression models, included checking the relationships between the covariates and the risk factors and applying the log transformation where necessary to improve linearity, checking for (multi)collinearity, etc. To avoid circularity in the analyses, we ensured that the covariates included none of the risk factors for constructing the index.

To facilitate the estimation of the risk factors and vaccination coverage at the district level, we obtained population estimates for children aged under 1 year old from WorldPop (Tatem, 2017) for each country corresponding to the year of the survey for the country. These data were also used to produce estimates of numbers of children aged under 1 year old living within different vulnerability classes — see Section 3.3.

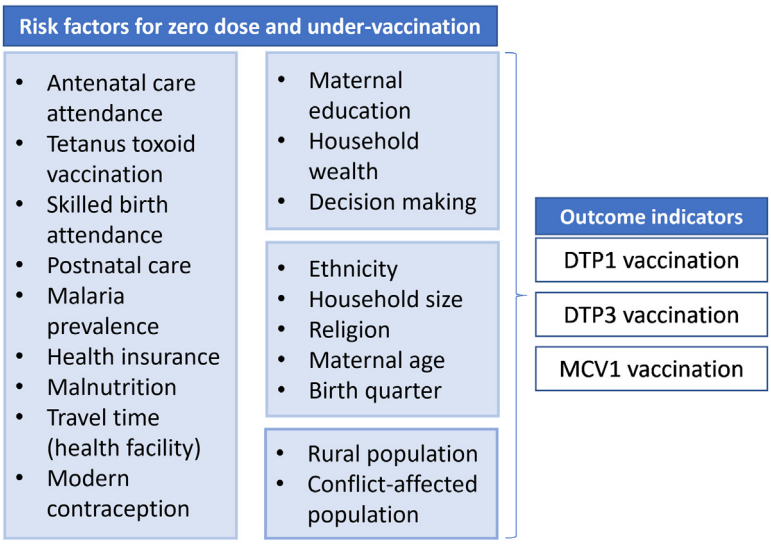


Fig. 1. Risk factors for zero dose and under-vaccination and vaccination coverage indicators considered in the study.

2.3. Geospatial modelling, prediction and validation

We fitted geostatistical models with a binomial likelihood to predict the risk factors and vaccination coverage at 1×1 km resolution. For $i = 1, \ldots, n$, where n is the number of survey locations in a given country, let $Y(\mathbf{s}_i)$ denote the number of individuals possessing a given attribute relating to a risk factor or vaccination coverage at survey location $\mathbf{s}_i$, and $m(\mathbf{s}_i)$ the number of individuals sampled from the location. The geostatistical model is given by

$$\begin{aligned} Y(\mathbf{s}_i)|p(\mathbf{s}_i) &\sim \text{Binomial}\big(m(\mathbf{s}_i), p(\mathbf{s}_i)\big), \\ \text{logit}\big(p(\mathbf{s}_i)\big) &= \mathbf{x}(\mathbf{s}_i)'\boldsymbol{\beta} + \omega(\mathbf{s}_i) + \epsilon(\mathbf{s}_i), \end{aligned} \tag{1}$$

where $p(\mathbf{s}_i), (0 \leq p(\mathbf{s}_i) \leq 1)$ is the true prevalence at location $\mathbf{s}_i$, $\mathbf{x}(\mathbf{s}_i)$ is a vector of covariate information and $\boldsymbol{\beta}$ – the corresponding regression coefficient, $\epsilon(\mathbf{s}_i)$ is an independent and identically distributed (iid) normal error term with variance, σ_ϵ^2, used to model non-spatial residual variation, and $\omega(\mathbf{s}_i)$ is a Gaussian spatial random effect used to capture residual spatial correlation in the model. Further, we assumed that $\boldsymbol{\omega} = \big(\omega(\mathbf{s}_1), \ldots, \omega(\mathbf{s}_n)\big)' \sim N(0, \Sigma_\omega)$, where Σ_ω follows the Matérn covariance function (Matérn, 1986) given by $\Sigma_\omega(\mathbf{s}_i, \mathbf{s}_j) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa\|\mathbf{s}_i - \mathbf{s}_j\|)^\nu K_\nu(\kappa\|\mathbf{s}_i - \mathbf{s}_j\|)$. The notation $\|.\|$ denotes the Euclidean distance between the cluster locations $\mathbf{s}_i$ and $\mathbf{s}_j$, $\sigma^2 > 0$ is the marginal variance of the spatial process, κ is the scaling parameter related to the range r ($r = \frac{\sqrt{8\nu}}{\kappa}$) — the distance at which spatial correlation is close to 0.1, and K_ν is the modified Bessel function of the second kind and order $\nu > 0$. For identifiability reasons, the smoothing parameter, ν, was set equal to unity — see Lindgren et al. (2011).

We assigned a $N(0, 10^3\mathbf{I})$ prior to the regression parameter, $\boldsymbol{\beta}$. We placed a penalized complexity (PC) prior (Simpson et al., 2017) on σ_ϵ such that $p(\sigma_\epsilon > 3) = 0.01$. Similarly, following Fuglstad et al. (2019), a joint PC prior was placed on the covariance parameters of the spatial random effect, $\boldsymbol{\omega}$. These were: $p(r < r_0) = 0.01$ and $p(\sigma > 3) = 0.01$, with r_0 chosen to be 5% of the extent of each country in either the north-south or east–west direction.

For each modelled risk factor and indicator of vaccination coverage, the model described in Eq. (1) was fitted in a Bayesian framework using the integrated nested Laplace approximation — stochastic partial differential equation (INLA-SPDE) approach (Rue et al., 2009; Lindgren et al.,

2011). Using the fitted models, we obtained predictions of the modelled indicators at 1 × 1 km resolution and the district level. The district level estimates were obtained as population-weighted averages taken over the 1 × 1 km grid cells falling within each district. That is, for $i = 1, \ldots, n_A$,

$$p^r(A_i) = \int_{A_i} p^r(\mathbf{s}) \times q(\mathbf{s}) d\mathbf{s} \approx \sum_{j=1}^{m_i} p^r(\mathbf{s}_j) \times q(\mathbf{s}_j), \tag{2}$$

where m_i is the number of grid cells with centroids in area A_i, $q(\mathbf{s})$ is the proportion of the population of district A_i at grid location $\mathbf{s}$ and n_A is the number of districts in each country. We note that if the jth grid cell lies at the boundary of the district, $q(\mathbf{s}_j)$ could also be possibly obtained as the proportion of the district population living within the area of intersection. The accuracy of the modelled estimates was assessed using cross-validation techniques described in previous work (Utazi et al., 2018, 2020, 2022a). We note that both change of support problems (COSPs) (Gotway and Young, 2002) involved in our methodology, i.e., the creation of 1 × 1 km modelled surfaces of some risk factors using model (1) (point-to-point COSP) and aggregation of grid level estimates to the district level (point-to-area COSP, see Eq. (2)) were dealt with using common approaches (Local Burden of Disease Vaccine Coverage Collaborators, 2021; Mosser et al., 2019; Utazi et al., 2021).

2.4. Constructing the vulnerability index

Here, we seek to create a zero-dose vulnerability index to summarize patterns in vulnerabilities to zero dose and under-vaccination across multiple risk factors and to enable spatial prioritization and a more compact and comparable equity assessment using these factors. The vulnerability index was constructed using five methods which are an adaptation of approaches used in similar contexts (OECD, 2008; Macharia et al., 2020). The distinction among these methods lies in how the risk factors were weighted to construct the index. The methods were selected to enable us to evaluate the robustness and sensitivity of the index (i) when using only data on the risk factors to create the index, (ii) when drawing insights from information on vaccination coverage and (iii) when using a limited number of the risk factors.

To construct the index, we first rescaled or normalized the risk factors to a common scale ranging from 0 (lowest risk) to 100 (highest risk) using the formula: $\tilde{v}(\mathbf{s}) = \left(v(\mathbf{s}) - v^0_{min}(\mathbf{s}) / (v^0_{max}(\mathbf{s}) - v^0_{min}(\mathbf{s}))\right) \times 100$, where $\tilde{v}(\mathbf{s})$ is the rescaled value, $v(\mathbf{s})$ – the value to be rescaled, $v^o_{min}(\mathbf{s})$ – the minimum value on the old scale and $v^o_{max}(\mathbf{s})$ is the maximum value on the old scale. Let $\tilde{v}_1(\mathbf{s}), \tilde{v}_2(\mathbf{s}), \ldots, \tilde{v}_k(\mathbf{s})$ denote the k rescaled risk factors. In our analysis, k ranges from 17 risk factors in Nigeria and Pakistan to 19 risk factors in other countries (Supplementary Table 2). The five approaches for constructing the index are described as follows.

Direct weighting: This approach combines information on all the risk factors to construct the index, independent of vaccination coverage. Here, the risk factors can be combined using expert-derived weights (from health practitioners, policy-makers or researchers) or other predetermined weights, hence offering some flexibility as to how information is integrated to create the index. The composite index derived through this method is given by:

$$VI_{direct}(\mathbf{s}) = \frac{\sum_{i=1}^{k} w_i \tilde{v}_i(\mathbf{s})}{\sum_{i=1}^{k} w_i}. \tag{3}$$

In the simplest case in which no prior or locally-derived expert information is available on the relative importance of the risk factors, equal weights can be used, e.g., $w_i = 1 \ \forall i$ such that $\sum_{i=1}^{k} w_i = k$. This is the option adopted when using this approach in our work.

Group-based weighting: This approach ensures a more balanced structure in the index, but it is also constructed independent of information on vaccination coverage. The risk factors were first grouped into three subdomains, namely health-related (8 risk factors), socioeconomic (3 risk factors) and demographic/geographic (7) risk factors — see Supplementary Table 1. The indicators within each group were then weighted and averaged using predetermined weights. Finally, the resulting scores from each group were averaged to obtain the index as given below:

$$VI_{group} = \frac{VI_{SE} + VI_H + VI_{DG}}{3}, \tag{4}$$

where

$$VI_H = \frac{\sum_{i=1}^{r} w_i^1 \tilde{v}_i(\mathbf{s})}{\sum_{i=1}^{r} w_i^1};\ VI_{SE} = \frac{\sum_{i=1}^{s} w_i^2 \tilde{v}_i(\mathbf{s})}{\sum_{i=1}^{s} w_i^2};\ \text{and } VI_{DG} = \frac{\sum_{i=1}^{t} w_i^3 \tilde{v}_i(\mathbf{s})}{\sum_{i=1}^{t} w_i^3}.$$

In Eq. (4), r, s, and t denote the numbers of health-related, socioeconomic and demographic/geographic risk factors, respectively, while $w_1^1, \ldots, w_r^1$, $w_1^2, \ldots, w_s^2$, $w_1^3, \ldots, w_t^3$ denote the respective weights applied to the risk factors within their subdomains. As in the direct weighting approach, we adopted equal weights within the groups, setting these equal to one in each case.

Regression rank-based weighting I and II: Here, the relationships between the risk factors and vaccination coverage indicators were exploited to generate the weights used to construct the index; hence, ensuring that more important risk factors contribute more towards the index. We first ranked the risk factors using their predictive R^2 statistics by fitting simple binomial regression models to cluster-level data using coverage indicators characterizing the likelihood of zero dose and under-vaccination (i.e., DTP1, DTP3 and MCV1 coverage) as outcome variables. These simple binomial regression models are reduced versions of model (1) including only a single covariate and excluding the spatial and iid random effects. We calculated the predictive R^2 statistics in each case following a Monte Carlo cross-validation exercise in which 80% of the data were used for model-fitting and the remaining 20% used for validation. We note that other cross-validation schemes are possible.

With regression rank-based weighting I, the ranks obtained for each risk factor were averaged over all three outcome variables to obtain the weight applied to the risk factor when constructing the index. The index under this approach is given by

$$VI_{RBI}(\mathbf{s}) = \frac{\sum_{i=1}^{k} w_i^{r1} \tilde{v}_i(\mathbf{s})}{\sum_{i=1}^{k} w_i^{r1}}, \tag{5}$$

where w_i^{r1} represents the regression rank-based weights.

With regression rank-based weighting II, the risk factors were grouped into four classes based on the quartiles of their predictive R^2 statistics for each outcome variable. Ranks ranging from 1 (the bottom 25% of the risk factors) to 4 (the top 25% of the risk factors) were then assigned to the classes accordingly. These ranks were averaged over all three outcome variables for each risk factor to derive the corresponding weight used to construct the index. The index obtained by using this approach is given by

$$VI_{RBII}(\mathbf{s}) = \frac{\sum_{i=1}^{k} w_i^{r2} \tilde{v}_i(\mathbf{s})}{\sum_{i=1}^{k} w_i^{r2}}, \tag{6}$$

where w_i^{r2} represents the regression rank-based weights produced using the quartiles of predictive R^2 statistics of the risk factors. The weights derived by analysing cluster-level data using both approaches were extrapolated to the gridded values of the risk factors to calculate the index at 1×1 km resolution.

Factor analysis approach: This approach is particularly effective for dimension reduction (i.e., to explore the possibility of using a smaller number of risk factors to create the index to encourage parsimony) and to minimize the effect of (multi)collinearity among the risk factors (OECD, 2008; Johnson and Wichern, 2002). The risk factors were first modelled as linear combinations of latent variables, known as factors, plus error terms. Letting p denote the number of common latent factors, the factor model can be expressed in matrix notation as:

$$\mathbf{V} = \boldsymbol{\alpha}\mathbf{F} + \boldsymbol{\epsilon}$$

where $\mathbf{V}$ is a $k \times n$ observation matrix containing the values of the risk factors at the cluster level scaled to have zero mean and unit variance, $\boldsymbol{\alpha}$ is a $k \times p$ matrix of factor loadings, $\mathbf{F}$ is a $p \times n$ matrix of uncorrelated common factors each with zero mean and unit variance, and $\boldsymbol{\epsilon}$ is a $k \times n$ matrix of zero-mean iid error terms. We employed the principal components analysis approach to extract the factors, the scree plot approach (OECD, 2008; Johnson and Wichern, 2002) to choose the number of factors necessary to represent the data, and varimax rotation to facilitate a clear

pattern of factor loadings, all of which are common choices for factor analysis (OECD, 2008). Post model-fitting, the factor loadings were squared and scaled to sum to unity. Among the G factors retained, only variables whose squared and scaled factor loadings were greater than 10% were used in the analysis. These factors were treated as individual subdomains or groups of the risk factors and the (squared and scaled) loadings used to weight the retained variables, as in the group-based weighting approach, to obtain an index $VI_g^f (g = 1, \ldots, G)$ for each group. The composite index was then constructed as a weighted average of the group-based indices, with the weights $w_1^f, \ldots, w_G^f$ being the proportion of the total variance explained by the factors. This is given by

$$VI_{factor} = \frac{\sum_{g=1}^{G} w_g^f VI_g^f}{\sum_{g=1}^{G} w_g^f}. \tag{7}$$

As in the regression-based approaches, the within-group weights for the selected variables and the weights for the factors/groups were also extrapolated to the rescaled gridded values of the risk factors to calculate the index at 1×1 km resolution.

In each case, we rescaled the indices produced through using these methods to range between 0 and 100, and obtained their district level estimates as population-weighted averages taken over the 1×1 km grid cells falling within each district, as before. We then defined a five-level gradient of vulnerability ranging from 1 (least vulnerable) to 5 (most vulnerable) using the quantiles of the district level estimates of each index. We quantified the similarities/differences between the spatial distributions of the indices (on the continuous scale) produced using these approaches at the district level for each country using the Pearson correlation coefficient and the Manhattan distance metric given by $D(VI^a, VI^b) = \frac{\sum_{i=1}^{n_A} |VI_i^a - VI_i^b|}{n_A}$, where VI_i^a and VI_i^b are the vulnerability index estimates for district i when using methods a and b, respectively, and n_A is the number of districts in the country. When determined for all plausible pairs of methods, we chose the method that had the lowest average Manhattan distance (VI_{RBI} and VI_{RBII} were individually compared with other methods) as the most representative of all the methods investigated. We also computed the correlations between the district level estimates of vaccination coverage (DTP1, DTP3 and MCV1) and the vulnerability indices to understand the similarities/differences between the vaccination coverage maps and the vulnerability indices explored here with respect to identifying areas at risk of non- and under-vaccination.

3. Results

3.1. 1×1 km modelled estimates of zero-dose and vaccination coverage indicators

Modelled estimates of some of the risk factors and vaccination coverage indicators are shown in Fig. 2 for DRC at 1×1 km resolution. The associated uncertainties are shown in Supplementary Figure 1. For other countries (and other risk factors for DRC), these maps are shown in Supplementary Figures 2-25 (only the point estimates are shown).

For DRC, these maps reveal strong heterogeneities in the spatial distributions of most of the risk factors and vaccination coverage. There were also strong similarities in the spatial distributions of all three coverage indicators, with the lowest coverage levels and greatest heterogeneities estimated for DTP3 coverage, indicating a lack of health system continuity (compared to DTP1) and potentially effective vaccination campaigns to boost MCV1 coverage. Furthermore, there were similarities between the spatial distributions of some of the risk factors – particularly the health-related factors such as skilled birth attendance, antenatal care attendance and travel time to the nearest health facility – and the coverage estimates. However, there were also remarkable differences between these, demonstrating the importance of data integration to identify areas at risk of zero dose and under-vaccination. For example, the Kongo-Central province is shown to have higher coverage levels, but it also had among the highest levels of malnutrition and lack of postnatal care attendance, both of which were prevalent in many parts of the country. Some risk factors such as lack of health insurance, lack of post-natal care attendance and non-use of modern contraception generally had high prevalence and little variation across the country. There was also a concentration of

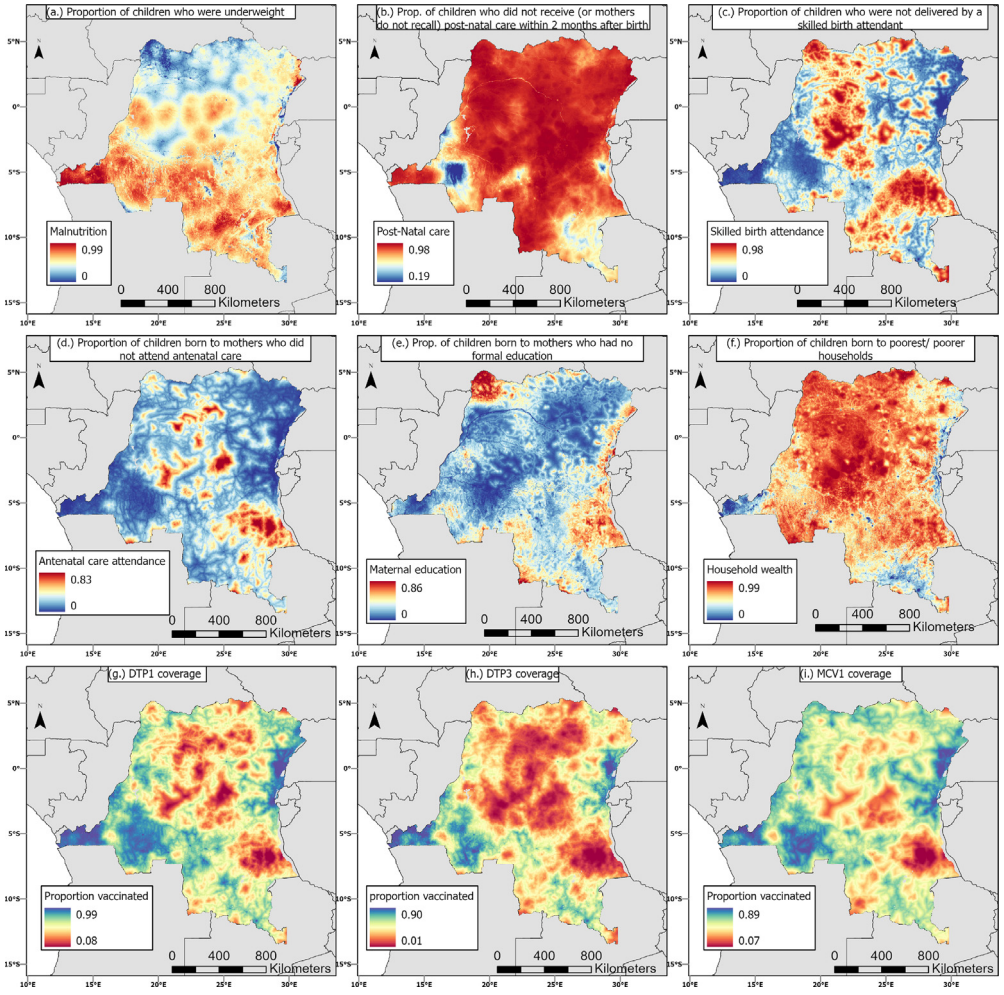


Fig. 2. 1 × 1 km modelled estimates of some risk factors for zero dose and under-vaccination and the coverage of DTP1, DTP3 and MCV1 vaccines in DRC. Corresponding uncertainty estimates are shown in supplementary Figure 1.

conflict-affected areas (according to the definition used here) in the eastern half of the country. Moderate to substantial geographical inequities were observed in all other risk factors as shown in supplementary Figures 2–5.

When considering other countries, we observed substantial heterogeneities in the spatial distributions of vaccination coverage in Ethiopia, Nigeria and Pakistan (supplementary Figures 26–30). However, the coverage maps for Uganda and (to a lesser extent) India showed relatively less spatial heterogeneities (see supplementary Figures 27 and 30). Despite this, we estimated substantial heterogeneities in most of the associated risk factors in all five countries including Uganda, as in DRC — supplementary Figures 2–25.

3.2. 1 × 1 km and district level estimates of the zero-dose vulnerability index and decomposition of associated risk factors

In the regression-based weighting methods, the top-ranking risk factors which were assigned higher weights in the construction of VI_{RBI} and VI_{RBII} were: lack of antenatal care attendance, lack of skilled birth attendance, non-receipt of tetanus toxoid vaccination before birth, poor maternal literacy level and lack of use of modern contraception. These risk factors were selected in the top 5 variables in at least 4 of the 6 countries (supplementary Tables 4–9). Other less frequent top-ranking risk factors were ethnicity, religion, large household sizes, poor households and malaria prevalence.

With the factor analysis approach, between 2 and 4 factors were used to represent the data in the study countries, with the number of risk factors with sizeable loadings ($\geq$ 10%) in the estimated factors having a mode of 3. Factor loadings were all positive, indicating positive associations among the risk factors. In all, the numbers of risk factors used to construct the index using this approach ranged between 6 (DRC) and 11 (Uganda) (supplementary Tables 10–16). Lack of antenatal care attendance, lack of skilled birth attendance and poor households were selected to construct the index in all six countries. Non-receipt of tetanus toxoid vaccination and poor maternal literacy level were each selected in five countries. Ethnicity, large household sizes and religion were each selected in four countries while malaria prevalence and use of modern contraception were each selected in three countries. Other selected variables using this method which occurred in one or two countries included non-participation in decision making, lack of post-natal care, higher travel times to the nearest health facility, lack of insurance and being born to younger/older mothers.

The average Manhattan distances for the indices (supplementary Tables 18–23) show that VI_{direct} was the most representative index for India and Uganda, while VI_{RBII} was the most representative index for DRC, Ethiopia, Pakistan and Nigeria, likely reflecting the stronger relationships between vaccination and some of the indicators in the latter case. We also observed that the vulnerability indices created using the various approaches were most similar for Ethiopia (average Manhattan distance ranged between 2.44 and 3.06) and Pakistan (average Manhattan distance ranged between 5.04 and 6.52). In all, the average Manhattan distances obtained for each country showed that the approaches produced reasonably similar vulnerability index estimates, despite being based on different weighting schemes and utilizing different numbers of risk factors (in the case of factor analysis). This is also corroborated by the very high correlations between the index surfaces shown in the plots in supplementary Figure 27. These plots reveal that while the index surfaces generally tend to be highly correlated, there were also high but less strong correlations between these and maps of vaccination coverage indicators, suggesting some differences in their spatial distributions. These differences between the index and the coverage maps were more pronounced in Uganda and India (to a lesser extent) and Nigeria (DTP3 only), where we had estimated relatively less heterogeneities in vaccination coverage in some cases (see Section 3.3 and supplementary Figure 48). However, the strong correlations estimated between the indices and vaccination coverage in DRC, Ethiopia, Pakistan and Nigeria demonstrate that these risk factors could be sufficiently used to capture inequities in the risk of non- and under-vaccination independent of and in combination with information on vaccination coverage.

The vulnerability index surfaces produced using all five methods are displayed in Fig. 3 for DRC and supplementary Figures 32–36 for other countries. In each of these countries, these figures show that the index surfaces produced using the methods are very similar, although some minor differences exist. These strong similarities among the index surfaces at 1 × 1 km resolution further corroborate that the index can be created without recourse to data on vaccination coverage, with fewer risk factors and without more complicated index measures. For example, for Ethiopia, all the index surfaces captured the east–west divide in vulnerabilities to zero dose and under-immunization, but the factor analysis approach appeared to have produced the strongest evidence of areas of high vulnerability in the eastern part of the country.

In Fig. 4, we explore the patterns in the index surfaces at the district level across all six countries. Shown are the index classes corresponding to the method that is most representative of other methods for each country as noted previously. These maps reveal strong spatial patterns in the distribution of vulnerabilities within each country. In DRC, the most vulnerable districts

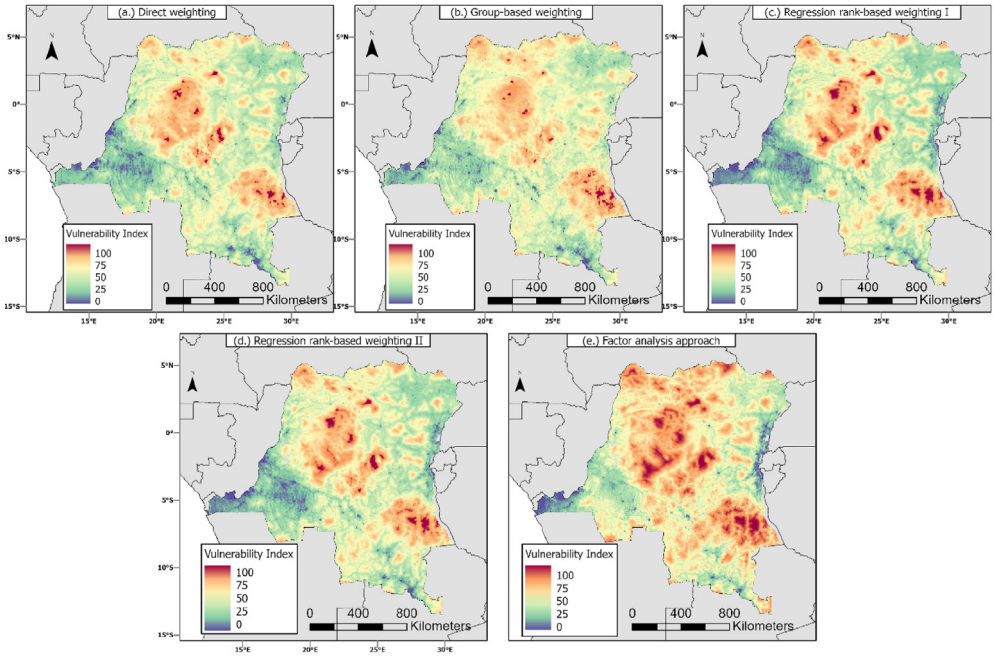


Fig. 3. 1 × 1 km zero-dose vulnerability index maps for DRC produced using the different approaches investigated.

are concentrated within Tanganyika, Sankuru, Tshuapa, Sud-Ubangu and Nord-Ubangi provinces. In Ethiopia, there is a clear east–west divide in vulnerabilities with the most vulnerable areas located within the eastern regions of Somali and Afar, although there are 'more vulnerable' (class 4) areas in the eastern regions. For India, the most vulnerable areas are located mostly in the northeastern region and northern parts of the central and eastern regions. There are also some districts in the northern region which had the highest vulnerability level.

In Nigeria, there is an apparent north-south divide in vulnerability, with the most vulnerable districts concentrated in the northeastern and northwestern parts of the country, but there are also more vulnerable and vulnerable areas (class 3) in other regions. For Pakistan, the most vulnerable districts are concentrated in the Balochistan region, although a few of these are in other regions. For Uganda, clusters of the most vulnerable districts can be found mostly in the northern, eastern and western regions.

Characterizations of the zero-dose vulnerability classes in terms of their most defining risk factors are explored in Fig. 5 and supplementary Figures 37–41. The most defining risk factors for the vulnerability classes were identified as the risk factors which had clear downward trends when moving from the most vulnerable to the least vulnerable areas and/or which had considerably higher values for the most vulnerable areas relative to other areas, as provided in Table 1. We also determined the most distinguishing factors for the most vulnerable areas as those which had considerably higher values for these areas relative to other areas. These plots and Table 1 further affirm the multiplicity of factors that contribute to differentiating between the levels of vulnerability, cutting across the health, socioeconomic and demographic/geographic subdomains investigated here. For Nigeria, Fig. 4 and Table 1 show that there was a total of 11 risk factors that most defined the vulnerability classes. Out of these, higher prevalence of lack of antenatal care attendance, low maternal literacy level, large household sizes, Hausa/Fulani/Kanuri ethnic group and Islamic religion were the most distinguishing factors for the most vulnerable areas. The factors

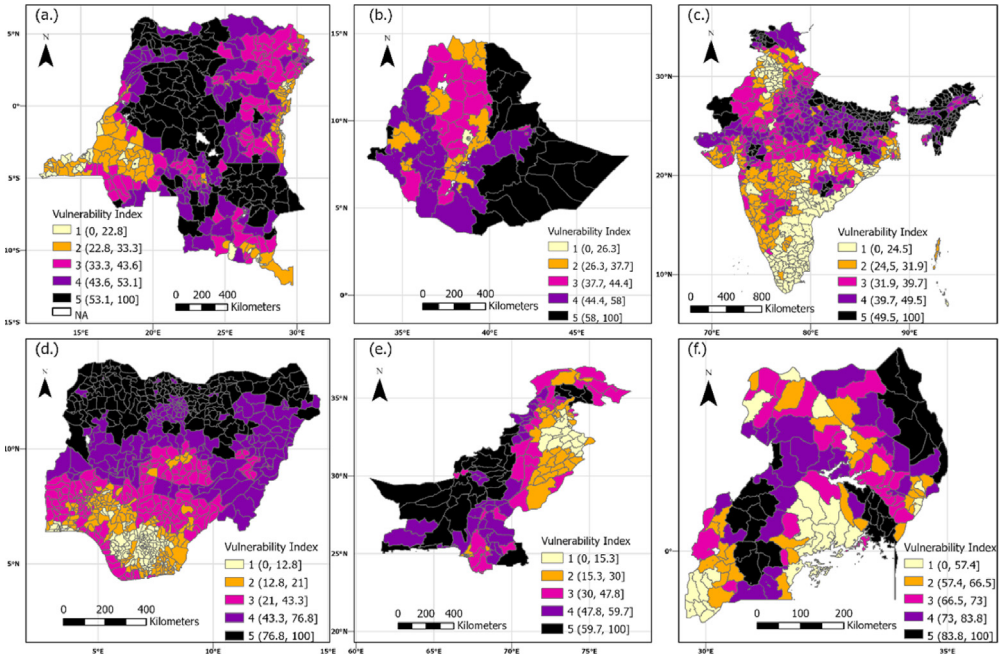


Fig. 4. District level estimates of the vulnerability index grouped into 5 classes ranging from the least vulnerable (class 1) to the most vulnerable (class 5) areas for all six study countries.

that contributed the least to defining the vulnerability classes were lack of health insurance, non-receipt of tetanus toxoid vaccination and lack of postnatal care attendance — the values of these factors were very similar across the different vulnerability classes.

The highest numbers of the most defining risk factors for all vulnerability classes were observed in Ethiopia (15 factors) and India (16 factors) (see Table 1), indicating greater capacity of the risk factors to discriminate between the different levels of vulnerabilities and to identify these most vulnerable areas in both countries. Other countries had between 10 and 11 risk factors. The most frequent risk factors (occurring in at least four countries) among the most defining risk factors were lack of antenatal care attendance, lack of skilled birth attendance, non-receipt of tetanus toxoid vaccination, travel time to the nearest health facility, poor households, low maternal literacy level, birth quarter, lack of use of modern contraception, ethnicity and higher proportion of rural population. Out of these, the most frequent risk factors (occurring in at least four countries) distinguishing the most vulnerable areas were lack of antenatal care attendance, travel time to the nearest health facility and ethnicity.

Interestingly, some of the factors exhibited negative trends, meaning that the highest values of these were seen in the least vulnerable areas. These were lack of use of modern contraception, lack of postnatal care attendance in DRC and being born to younger/older mothers in Nigeria.

3.3. Exploration of the distribution of under 1s and inequities in vaccination coverage among the vulnerability groups

We estimated wide variations in vaccination coverage among the different vulnerability classes, which were more pronounced in DRC, Ethiopia, Nigeria and Pakistan (Fig. 6 and supplementary Table 24), reflecting substantial heterogeneities in the spatial distribution of coverage in these countries. As expected, the highest coverage levels were estimated for the least vulnerable areas, which appeared to reduce gradually (India and Uganda) or substantially when progressing through

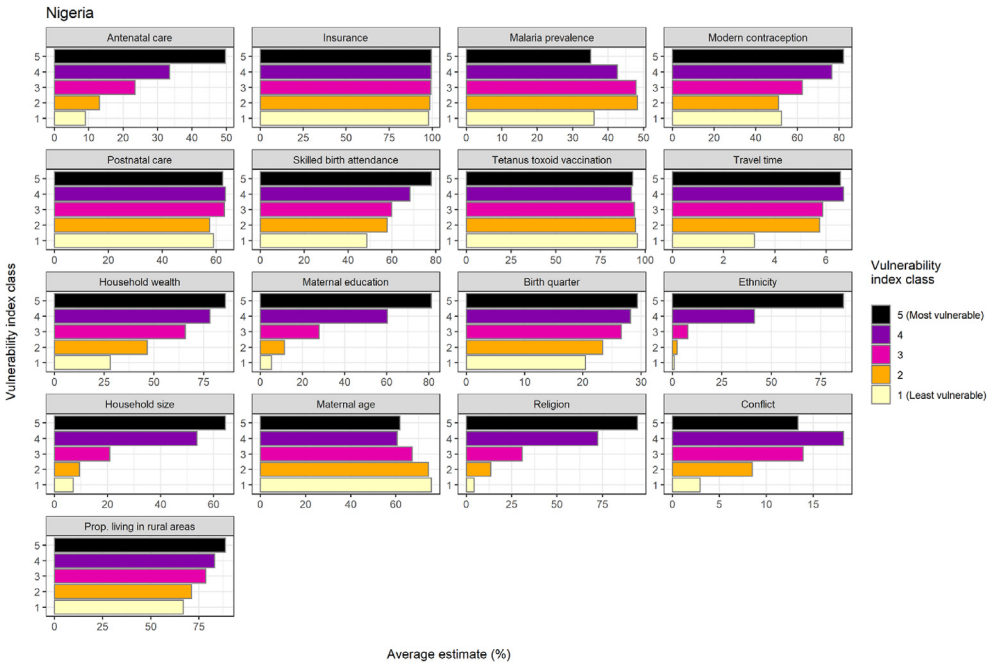


Fig. 5. Distributions of risk factors for zero dose and under-immunization according to different levels of vulnerabilities at the district level estimated through using the zero-dose vulnerability index for Nigeria.

other classes to the most vulnerable areas for which the lowest coverage levels were estimated. In DRC, the differences in coverage between the least and most vulnerable areas (i.e., 20% most privileged and 20% most disadvantaged districts) were 80%, 61% and 64% for DTP1, DTP3 and MCV1, respectively. This implies that the most vulnerable areas will need to experience dramatic improvements in coverage to attain the coverage levels estimated in the least vulnerable areas. For India where the variability in coverage among the vulnerability classes was relatively lower, 45%, 44% and 41% increases in coverage, respectively, were needed in the most vulnerable areas to reach the coverage levels in the least vulnerable areas.

For other countries, the respective increases in DTP1, DTP3 and MCV1 coverage needed to bridge the gaps between the least and most vulnerable areas were: Ethiopia — 45%, 34%, 49%; Nigeria — 78%, 70% and 67%; Pakistan — 45%, 42% and 62%; and Uganda — 40%, 32% and 49%. The downward trends in coverage observed in most cases in Uganda and India, despite both countries having relatively less spatial variation in coverage are also an indication that the index maps accurately identified the vulnerability patterns in both countries. We further explored the differences in vulnerability classification between the index maps and vaccination coverage in supplementary Figures 42–47, which show the most mismatches for Uganda and India.

Estimates of numbers of children aged under 1 (under 1s) living in areas with different levels of vulnerability were fairly evenly distributed among these areas in DRC, India, Nigeria and Uganda (Fig. 6(b) and (c)). Notably, for Nigeria and India, the highest proportions of under 1s were estimated to be living in the more and most vulnerable areas. However, for Ethiopia, both the least and most vulnerable areas had the lowest estimates of the proportions of under 1s, and for Pakistan, the lowest proportion of under 1s was estimated in the most vulnerable areas (with a steady decline in proportion from class 2 to class 5), suggesting sparse population distributions in these cases. Furthermore, Fig. 6(c) also revealed that the numbers of children aged under 1 living in the most vulnerable areas were 665 805, 279 222, 4 495 184, 1 700 172, 232 234 and 272 446 for DRC,

Table 1
A summary of the most defining risk factors for the vulnerability classes in the study countries at the district level. The most defining risk factors for the most vulnerable areas are written in bold letters.

Country	Most defining risk factors	Number of risk factors
DRC	**Antenatal care**, malaria prevalence, **skilled birth attendance**, **tetanus toxoid vaccination**, travel time, household wealth, maternal education, birth quarter, **ethnicity** and proportion living in rural areas	10
Ethiopia	**Antenatal care**, health insurance, modern contraception, postnatal care, skilled birth attendance, **tetanus toxoid vaccination**, **travel time**, **household wealth**, **maternal education**, **birth quarter**, **ethnicity**, **household size**, **religion**, **conflict** and proportion living in rural areas	15
India	**Antenatal care**, health insurance, **malaria prevalence**, modern contraception, postnatal care, **skilled birth attendance**, **tetanus toxoid vaccination**, **travel time**, household wealth, maternal education, **birth quarter**, **ethnicity**, **maternal age**, **religion**, **conflict** and proportion living in rural areas	16
Nigeria	**Antenatal care**, modern contraception, skilled birth attendance, travel time, household wealth, **maternal education**, birth quarter, **ethnicity**, **household size**, **religion** and proportion living in rural areas	11
Pakistan	**Antenatal care**, **malaria prevalence**, modern contraception, postnatal care, **skilled birth attendance**, tetanus toxoid vaccination, **travel time**, household wealth, maternal education, maternal age and **conflict**	11
Uganda	Antenatal care, decision making, malaria prevalence, malnutrition (underweight), modern contraception, skilled birth attendance, **travel time**, **household wealth**, **maternal education**, **ethnicity**, and proportion living in rural areas	11

Ethiopia, India, Nigeria, Pakistan and Uganda, respectively (see supplementary Table 25); while the respective numbers living in the more vulnerable areas were 643 499, 1 050 513, 5 423 568, 1 790 904, 926 932 and 271 600. When combined, the populations living in the more and most vulnerable areas constituted 39%, 40%, 42%, 48%, 21% and 36% of all children aged under 1 in these countries, respectively. Lastly, combining the distributions of under 1s with information on vaccination coverage, on average, about 524 936, 126 828, 2 062 143, 1 327 731, 105 315 and 162 478 children aged under 1 will need to be vaccinated with DTP1 in the most vulnerable areas in the respective countries to achieve the coverage levels seen in the least vulnerable areas. Similar figures can also be obtained for DTP3 and MCV1.

4. Discussion

Reaching zero-dose children and missed communities to improve coverage levels, reduce inequities in coverage and accelerate progress towards the goal of 'leaving no one behind' is currently a critical priority within the global immunization community. The design and implementation of intervention programmes targeting these vulnerable populations should be guided by operationally relevant and spatially detailed data identifying where they reside and their most defining characteristics. Vulnerabilities to zero dose and under-immunization are characterized by multiple risk factors which often exhibit substantial geographical variation within countries. The zero-dose vulnerability index developed in our work operationalized this notion by integrating data on a wide range of top-ranking risk factors for zero dose and under-immunization to characterize inequities in vulnerabilities and highlight priority areas (districts) for interventions in priority LMICs.

Our results indicate that the index is a potentially more useful metric for identifying and prioritizing zero-dose and missed communities, particularly in instances where there is little or no heterogeneity in vaccination coverage. This is evidenced in Uganda where we found relatively less

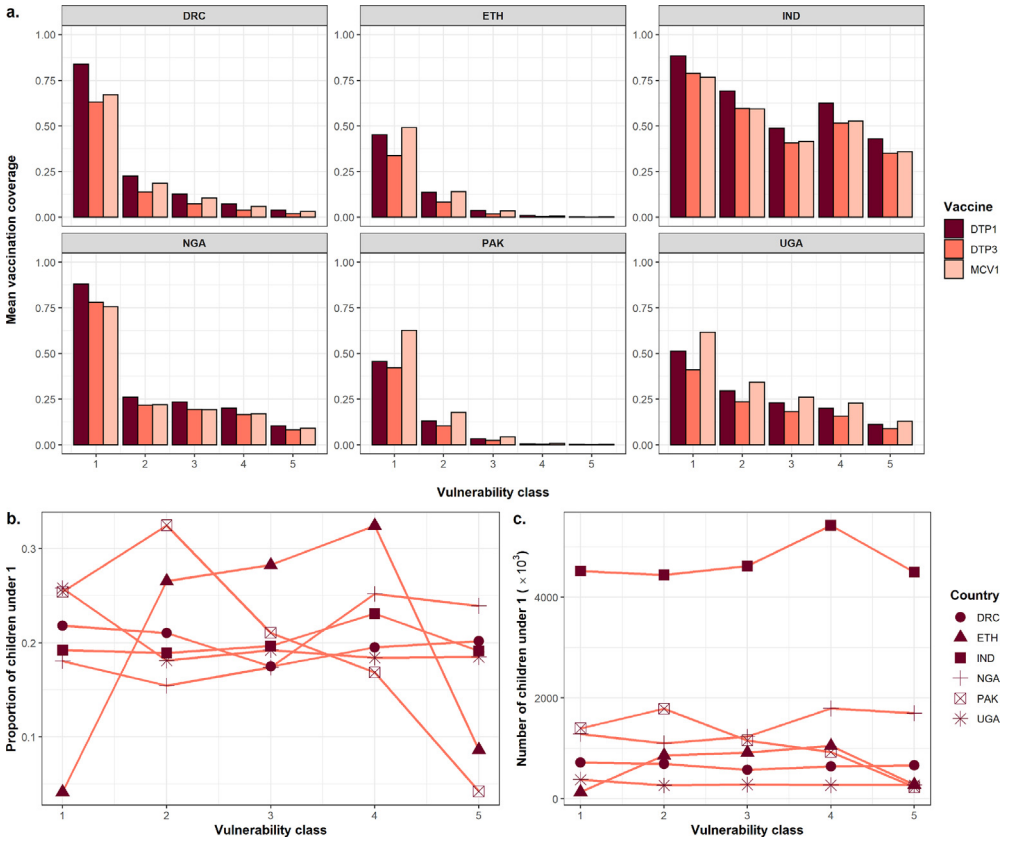


Fig. 6. (a) Differences in the coverage of DTP1, DTP3 and MCV1 vaccines among areas with different levels of vulnerability in the study countries. (b–c) Estimates of the proportions and numbers of children aged under 1 year living in these areas.

geographical variation in DTP1 and DTP3 coverage but stronger geographical disparities in most of the associated risk factors and the ((un)scaled) index — supplementary Figures 14–17 and 49. These stronger disparities in the index surfaces resulted in weaker correlations between the index and vaccination coverage in Uganda. This implies that relying only on estimates of vaccination coverage for geographical prioritization and identifying zero-dose and missed communities in instances where there is little variation in the coverage estimates may not be a highly effective approach.

The five methods for constructing the zero-dose vulnerability index produced nearly identical results in all six study countries, demonstrating that the index can produce robust vulnerability classifications and serve as a reliable prioritization tool when constructed independent of data on vaccination coverage and using limited data on the risk factors. We recall that the regression-based approaches were the only approaches that utilized information on vaccination coverage to derive the weights used to create the index, while the direct weighting, group-based and factor analysis approaches utilized information on the risk factors only. Furthermore, the factor analysis approach utilized between 6 (DRC) and 11 (Uganda) risk factors to construct the index out of at least 17 variables in each case. The most important risk factors identified using the factor analysis approach were: lack of antenatal care attendance, lack of skilled birth attendance, poor households, non-receipt of tetanus toxoid vaccination, poor maternal literacy level, large household sizes, ethnicity and religion, all of which were selected to construct the index in at least four of the six countries. These results have important implications for geographical prioritization in data-poor

settings where up-to-date or reliable data on vaccination coverage may not be readily available. In such circumstances, the index can be constructed if information on these risk factors can be obtained from other sources as highlighted previously, and the classifications produced can be used to strengthen the capacity of immunization programmes to vaccinate priority populations, which is critical to reducing zero-dose prevalence and reaching missed communities. Due to ease of implementation and interpretability and given the similar results we obtained when using various approaches to construct the index, the direct approach may be most practical approach in these settings.

Our analysis revealed dramatic differences in coverage between the least vulnerable areas and other areas, particularly in DRC, Ethiopia, Nigeria and Pakistan, where coverage was found to be more heterogeneous, further validating the discriminatory power of the index. The dramatic decreases in coverage observed when moving from the least vulnerable areas to other areas in these four countries potentially demonstrate that efforts geared towards improving coverage levels should seek to address the much poorer coverage levels in all other vulnerability classes, although zero-dose and missed communities are more likely to be concentrated in areas with higher vulnerability levels. Furthermore, an analysis of the distribution of under 1s among areas with different levels of vulnerability revealed that targeting the most vulnerable areas is likely to yield some impact in DRC, India, Nigeria and Uganda. The sparse population distributions in the most vulnerable areas in Ethiopia and Pakistan mean that better use of resources and greater impact will be achieved in both countries by prioritizing both the more and most vulnerable areas.

By further decomposing the vulnerability classes into their major contributory factors, our analyses produced richer actionable insights than studies determining vulnerabilities using vaccination coverage alone (Utazi et al., 2018; Takahashi et al., 2017; Local Burden of Disease Vaccine Coverage Collaborators, 2021; Utazi et al., 2020) or in combination with key community characteristics only (Wigley et al., 2022). The major contributory factors in differentiating between the vulnerability classes were shown to cut across the different subdomains (health, socioeconomic, demographic/geographic) of the risk factors, indicating their combined discriminatory power and corroborating the multiplicity of factors that define inequities in vaccination coverage (Rainey et al., 2011; Utazi et al., 2022b; Santos et al., 2021). Common (occurring in at least four countries) health-related risk factors characterizing the most vulnerable areas were: lack of antenatal care, lack of skilled birth attendance, higher travel time to the nearest health facility, non-use of modern contraception and non-receipt of tetanus toxoid vaccination. For the socioeconomic factors, these were: low maternal literacy level and poor households; for demographic factors — ethnicity and birth quarter; for geographic factors — higher proportions of rural populations. Among these, the most frequent (occurring in at least four countries) risk factors distinguishing the most vulnerable areas were lack of antenatal care attendance, higher travel time to the nearest health facility and ethnicity. Although previous studies utilizing national or subnational individual level data have established associations between these risk factors and vaccination coverage as highlighted previously (Acharya et al., 2018; Aheto et al., 2022; Okello et al., 2022; Utazi et al., 2022b), our work makes the additional contribution of establishing some of these (Table 1) as being more influential in the most vulnerable areas. These findings highlight the need for strategies that focus on providing integrated primary health care services, as well as programmes which address the socio-economic/demographic/geographic barriers, to reduce the general inequities seen in these most vulnerable areas.

Further analyses are required to test which subsets of the risk factors (either those identified through the factor analysis approach or through the decomposition of the vulnerability classes into their major contributory factors), particularly those with already available gridded estimates such as travel time to the nearest health facility and malaria prevalence, can be used to create plausible estimates of the index in limited data settings. We have used data obtained mostly from household surveys in our work. We will also explore using routine information sources to better understand spatio-temporal patterns in the index, although we note that this will likely constrain the analysis to the district level entirely and the outputs produced may be subject to changes in administrative boundaries. It is also possible to explore the vulnerability classes produced by the various subdomains of the risk factors we investigated, particularly when using the group-based

approach to construct the index. Such analysis of the most vulnerable areas for each subdomain of the risk factors could yield additional insights for vaccination programming (see, e.g., Macharia et al., 2020).

Our work is subject to some limitations. First, we did not consider many supply-side factors (e.g., vaccine stock-outs, non-availability of vaccination staff, long waiting times at vaccination centres Mahachi et al., 2022) and other data (e.g., routine data on the locations and frequencies of disease outbreaks) that could provide additional information on the performance of the health systems within these countries to further refine the vulnerability classification. Secondly, uncertainty estimates were only available for the risk factors whose modelled surfaces were created as part of our analyses; hence, we utilized only the point estimates of the risk factors to construct the index as the associated uncertainties cannot be fully accounted for and were also not of immediate interest. We, however, note that statistically, obtaining uncertainty estimates for the index from the corresponding uncertainty estimates of the modelled indicators is straightforward. For example, with the direct method, the variance of VI_{direct} can be easily calculated as: $VI_{direct}(\mathbf{s}) = \frac{\sum_{i=1}^{k} w_i^2 Var(\tilde{v}_i(\mathbf{s}))}{(\sum_{i=1}^{k} w_i)^2}$. This is a straightforward way to propagate the uncertainties in the (modelled) risk factors to the index. Thirdly, our analyses are tied to the respective survey years in the study countries which may be a concern in contexts where the survey data are a bit dated. A potential approach that could be used to rescale or normalize the risk factors is standardization using z-scores (OECD, 2008). We did not test the sensitivity of our methodology to different rescaling approaches, although we considered the approach we used suitable as it improves the contribution of risk factors lying within a small interval to the index, which is desirable. In our implementation of the INLA-SPDE approach, we did not undertake any sensitivity analyses to investigate the effect of choice of mesh parameters used in our prediction models on our analyses. Finally, our analyses leveraged the associations between the risk factors and vaccination coverage and did not determine any causal relationships between these variables as our data were not suitable for such analyses.

Reducing inequities in vaccination coverage through reaching zero-dose children and missed communities calls for greater geographical precision in the estimation of their spatial distribution and better understanding of their most defining characteristics. The zero-dose vulnerability index developed in our work presents a tool that can be used by governments and global health practitioners to plan and implement appropriate interventions targeting moderately vulnerable to the most vulnerable populations. The heterogeneities in the spatial distributions of the index suggest a need for timely prioritization of these vulnerable populations to facilitate progress towards global immunization goals.

Code availability

The R code used in the analysis is available at: https://github.com/wpgp/vipaper.

Funding

This work was supported by the Bill & Melinda Gates Foundation, United States and Gavi, the Vaccine Alliance [Grant Number INV-002397 awarded to A.J.T, C.E.U and N.T.-G.].

Declaration of competing interest

J. L., D. B. and D. H. work for Gavi, the Vaccine Alliance. The authors declare no other competing interests.

Data availability

All the data used in this work are publicly available via the sources referenced in the article.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spasta.2023.100772.

References

Acharya, P., Kismul, H., Mapatano, M.A., Hatløy, A., 2018. Individual-and community-level determinants of child immunization in the Democratic Republic of Congo: a multilevel analysis. PLoS One 13 (8), e0202742.

Aheto, J.M.K., Pannell, O., Dotse-Gborgbortsi, W., Trimner, M.K., Tatem, A.J., Rhoda, D.A., Cutts, F.T., Utazi, C.E., 2022. Multilevel analysis of predictors of multiple indicators of childhood vaccination in Nigeria. PLoS One 17 (5), e0269066.

Andre, F.E., Booy, R., Bock, H.L., Clemens, J., Datta, S.K., John, T.J., Lee, B.W., Lolekha, S., Peltola, H., Ruff, T., et al., 2008. Vaccination greatly reduces disease, disability, death and inequity worldwide. Bull. World Health Organ. 86, 140–146.

Fuglstad, G.-A., Simpson, D., Lindgren, F., Rue, H., 2019. Constructing priors that penalize the complexity of Gaussian random fields. J. Amer. Statist. Assoc. 114 (525), 445–452.

Gavi, 2020. Gavi strategy 5.0, 2021–2025. URL https://www.gavi.org/our-alliance/strategy/phase-5-2021-2025.

Gotway, C.A., Young, L.J., 2002. Combining incompatible spatial data. J. Amer. Statist. Assoc. 97 (458), 632–648.

Johnson, R.A., Wichern, D.W., 2002. Applied Multivariate Statistical Analysis, fifth ed. Prentice Hall, Upper Saddle River, NJ.

Keja, K., Chan, C., Hayden, G., Henderson, R.H., 1988. Expanded Programme on Immunization. World Health Statist. Q. Rapport trimestriel de statistiques sanitaires mondiales 41 (2), 59–63.

Li, X., Mukandavire, C., Cucunuba, Z.M., Londono, S.E., Abbas, K., Clapham, H.E., Jit, M., Johnson, H.L., Papadopoulos, T., Vynnycky, E., et al., 2021. Estimating the health impact of vaccination against ten pathogens in 98 low-income and middle-income countries from 2000 to 2030: A modelling study. Lancet 397 (10272), 398–408.

Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. 73 (4), 423–498.

Local Burden of Disease Vaccine Coverage Collaborators, 2021. Mapping routine measles vaccination in low-and middle-income countries. Nature 589 (7842), 415–419.

Macharia, P.M., Joseph, N.K., Okiro, E.A., 2020. A vulnerability index for COVID-19: spatial analysis at the subnational level in Kenya. BMJ Global Health 5 (8), e003014.

Mahachi, K., Kessels, J., Boateng, K., Achoribo, A.E.J.B., Mitula, P., Ekeman, E., Lochlainn, L.N., Rosewell, A., Sodha, S.V., Abela-Ridder, B., et al., 2022. Zero- or missed-dose children in Nigeria: Contributing factors and interventions to overcome immunization service delivery challenges. Vaccine 40 (37), 5433–5444.

Matérn, B., 1986. Spatial Variation, second ed. Springer, New York, NY, http://dx.doi.org/10.1007/978-1-4615-7892-5.

Mosser, J.F., Gagne-Maynard, W., Rao, P.C., Osgood-Zimmerman, A., Fullman, N., Graetz, N., Burstein, R., Updike, R.L., Liu, P.Y., Ray, S.E., et al., 2019. Mapping diphtheria-pertussis-tetanus vaccine coverage in africa, 2000–2016: A spatial and temporal modelling study. Lancet 393 (10183), 1843–1855.

OECD, 2008. Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD publishing, http://dx.doi.org/10.1787/9789264043466-en.

Okello, G., Izudi, J., Ampeire, I., Nghania, F., Dochez, C., Hens, N., 2022. Two decades of regional trends in vaccination completion and coverage among children aged 12-23 months: an analysis of the Uganda Demographic Health Survey data from 1995 to 2016. BMC Health Serv. Res. 22, 1–17.

Perez-Heydrich, C., Warren, J.L., Burgert, M.E., 2013. Guidelines on the Use of DHS GPS Data. Spatial Analysis Reports 8, ICF International, Calverton, Maryland, USA.

Rainey, J.J., Watkins, M., Ryman, T.K., Sandhu, P., Bo, A., Banerjee, K., 2011. Reasons related to non-vaccination and under-vaccination of children in low and middle income countries: Findings from a systematic review of the published literature, 1999–2009. Vaccine 29 (46), 8215–8221.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J. R. Stat. Soc. Ser. B Stat. Methodol. 71 (2), 319–392.

Santos, T.M., Cata-Preta, B.O., Victora, C.G., Barros, A.J., 2021. Finding children with high risk of non-vaccination in 92 low-and middle-income countries: A decision tree approach. Vaccines 9 (6), 646.

Simpson, D., Rue, H., Riebler, A., Martins, T.G., Sørbye, S.H., 2017. Penalising model component complexity: A principled, practical approach to constructing priors. Statist. Sci. 32 (1), 1–28.

Takahashi, S., Metcalf, C.J.E., Ferrari, M.J., Tatem, A.J., Lessler, J., 2017. The geography of measles vaccination in the african great lakes region. Nat. Commun. 8 (1), 15585.

Tatem, A.J., 2017. WorldPop, open data for spatial demography. Sci. data 4 (1), 1–4.

United Nations, 2015. Transforming our world: The 2030 agenda for sustainable development. URL http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.

Utazi, C.E., Aheto, J.M.K., Chan, H.M.T., Tatem, A.J., Sahu, S.K., 2022a. Conditional probability and ratio-based approaches for mapping the coverage of multi-dose vaccines. Stat. Med. 41 (29), 5662–5678.

Utazi, C.E., Nilsen, K., Pannell, O., Dotse-Gborgbortsi, W., Tatem, A.J., 2021. District-level estimation of vaccination coverage: Discrete vs continuous spatial models. Stat. Med. 40 (9), 2197–2211.

Utazi, C.E., Pannell, O., Aheto, J.M., Wigley, A., Tejedor-Garavito, N., Wunderlich, J., Hagedorn, B., Hogan, D., Tatem, A.J., 2022b. Assessing the characteristics of un-and under-vaccinated children in low-and middle-income countries: A multi-level cross-sectional study. PLOS Glob. Public Health 2 (4), e0000244.

Utazi, C.E., Thorley, J., Alegana, V.A., Ferrari, M.J., Takahashi, S., Metcalf, C.J.E., Lessler, J., Cutts, F.T., Tatem, A.J., 2019. Mapping vaccination coverage to explore the effects of delivery mechanisms and inform vaccination strategies. Nat. Commun. 10 (1), 1633.

Utazi, C.E., Thorley, J., Alegana, V.A., Ferrari, M.J., Takahashi, S., Metcalf, C.J.E., Lessler, J., Tatem, A.J., 2018. High resolution age-structured mapping of childhood vaccination coverage in low and middle income countries. Vaccine 36 (12), 1583–1591.

Utazi, C.E., Wagai, J., Pannell, O., Cutts, F.T., Rhoda, D.A., Ferrari, M.J., Dieng, B., Oteri, J., Danovaro-Holliday, M.C., Adeniran, A., et al., 2020. Geospatial variation in measles vaccine coverage through routine and campaign strategies in Nigeria: Analysis of recent household surveys. Vaccine 38 (14), 3062–3071.

WHO, 2020. Immunization agenda 2030: A global strategy to leave no one behind. URL https://www.who.int/immunization/immunization_agenda_2030/en/.

WHO, 2021. Implementing the Immunization Agenda 2030: a Framework for Action Through Coordinated Planning, Monitoring & Evaluation, Ownership & Accountability, and Communications & Advocacy. World Health Organization [Internet], Genebra.

WHO and UNICEF, 2022. 2021 WHO and UNICEF estimates of national immunization coverage (WUENIC). URL https://www.who.int/publications/m/item/progress-and-challenges.

Wigley, A., Lorin, J., Hogan, D., Utazi, C.E., Hagedorn, B., Dansereau, E., Tatem, A.J., Tejedor-Garavito, N., 2022. Estimates of the number and distribution of zero-dose and under-immunised children across remote-rural, urban, and conflict-affected settings in low and middle-income countries. PLOS Glob. Public Health 2 (10), e0001126.

World Health Organization, 2022. Immunization coverage: Key facts 2020. https://www.who.int/news-room/fact-sheets/detail/immunization-coverage.