Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Population segmentation using social determinants of health: Implications for tailored maternal and child health programming in Ethiopia

Samson Gebremedhin

Samson.Gebremedhin@aau.edu.et

Addis Ababa University

Ayalew Astatkie

Hawassa University

Abyot Asres

Project HOPE

Fisseha Shiferie

Project HOPE

Asnake Worku

Project HOPE

Gashaw Andargie Biks

Project HOPE

Dawit Abraham

Project HOPE

Frehiwot Nigatu

Project HOPE

Wondwossen Asefa Alemayehu

Project HOPE

Lee Pyne-Mercier

Sonder Collective

Frank DelPizzo

Gates Foundation

Lang Gao

Sonder Collective

Steven Neri

Project HOPE

Michael Kalnoky

Sonder Collective

Jeremy Cooper

Gates Foundation
Chloé Roubert
Sonder Collective
Gwyn Hainsworth
Sonder Collective
Elisabeth Root
Gates Foundation
Tracy Johnson
Gates Foundation

Research Article

Keywords: Population segmentation analysis, precision public health, health equity

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11496606/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Population segmentation using social determinants of health: Implications for tailored maternal and child health programming in Ethiopia

Samson Gebremedhin[1,2*], Ayalew Astatkie[1,3], Abyot Asres[1, 4], Fisseha Shiferie[1], Asnake Worku[1], Gashaw Andargie Biks[1], Dawit Abraham[1], Frehiwot Nigatu[1], Wondwossen Asefa Alemayehu[5], Lee Pyne-Mercier[6], Frank DelPizzo[7], Lang Gao[6], Steven Neri[5], Michael Kalnoky[6], Jeremy Cooper[7], Chloé Roubert[6], Gwyn Hainsworth[6], Elisabeth Root[7], Tracy Johnson[7]

[1]Project HOPE, Addis Ababa, Ethiopia

[2]School of Public Health, Addis Ababa University, Addis Ababa, Ethiopia

[3]School of Public Health, Hawassa University, Hawassa, Ethiopia

[4]School of Public Health, Mizan Tepi University, Mizan Teferi, Ethiopia

[5]Project HOPE, Washington D.C., USA

[6]Sonder Collective, Helsinki, Finland

[7]Gates Foundation, Seattle, USA

*Corresponding author, e mail: samsongmgs@yahoo.com or Samson.Gebremedhin@aau.edu.et

Abstract

Background: Despite substantial improvements in reproductive, maternal, newborn, child health and Nutrition (RMNCH+N) outcomes in Ethiopia over the past two decades, health inequalities persist. Conventional one-size-fits-all approaches often fail to address the complex social determinants that shape health outcomes. This study aimed to segment women of reproductive age (WRA) in Ethiopia based on social determinants of health and develop an evidence base to support precision RMNCH+N programming.

Methods: Context-specific RMNCH+N vulnerabilities were measured in a nationally representative survey of 5,187 WRA. Data included over 400 vulnerability indicators and 100 outcomes. Vulnerability variables were selected and reduced using standardized statistical methods, including exploratory data analysis and principal component analysis. Population segmentation was then performed using latent class analysis, with separate models estimated for urban and rural populations. Segments were ranked using a composite score based on over 40 outcomes. A screening Typing Tool (TT) was developed using a smaller set of variables based on Classification and Regression Tree analysis.

Results: The segmentation analysis identified four urban and five rural segments. Based on a composite RMNCH+N outcomes score, three urban segments (U1.1, U1.2, U1.3) were classified as least vulnerable; two rural segments (R2.1, R2.2) as less vulnerable; one urban segment (U3) and two rural segments (R3.1, R3.2) as more vulnerable; and one rural segment (R4) as most vulnerable. In urban areas, U3 comprised 14.2% of women, while U1.1 was the largest segment (30.3%). In rural areas, R4 comprised 8.5% of WRA, whereas R2.2 was the largest (33.0%). Under-five mortality increased across vulnerability categories, from 8 to 55 deaths per 1,000 live births in urban population, and from 40 to 74 per 1,000 in rural population. The TT achieved over 80% sensitivity and specificity for most segments.

Conclusions: The vulnerability-based population segmentation identified distinct WRA segments with unique social determinant profiles and markedly different RMNCH+N outcomes, providing a foundation for precision programming and targeted interventions to improve health equity and outcomes.

Keywords: Population segmentation analysis, precision public health, health equity

1. Background

Over the past 25 years, Ethiopia has achieved substantial reductions in maternal and child mortality, with the maternal mortality ratio and under-five mortality rate (U5MR) declining by approximately 70% to 80% [1,2]. Despite these gains, Ethiopia remains a high-burden country due to its large population and persistently high levels of maternal and child mortality [3,4]. An estimated 8,000 maternal deaths occur annually, placing the country among the five countries with the highest numbers of maternal deaths worldwide [3]. Similarly, more than 180,000 children die before reaching their fifth birthday each year, accounting for approximately one in every 27 under-five deaths globally [4].

Health inequalities are also an increasing concern in Ethiopia, particularly in reproductive, maternal, newborn, child health and nutrition (RMNCH+N). Evidence shows that RMNCH+N services are disproportionately utilized by wealthier populations, reflecting a persistent pro-rich distribution of care [5–7]. Ethiopia ranks among the five low- and middle-income countries (LMICs) with the greatest wealth-related inequalities in antenatal care (ANC) coverage and quality [8]. Significant geographic disparities between regions and urban-rural populations, as well as social inequalities linked to maternal education, further contribute to inequities in service utilization and child survival [9–11].

Precision public health aims to deliver the right intervention to the right population at the right time using data and advanced analytics [12,13]. Population segmentation supports this approach by dividing heterogeneous populations into relatively homogeneous subgroups, based on multiple sources of inequalities, enabling more effective and equitable tailored interventions [14]. Vulnerability-based segmentation extends this framework by grouping populations according to socio-cultural, economic, geographic, and behavioral vulnerabilities, thereby improving understanding of how social determinants shape health outcomes and facilitating earlier risk identification and prevention than outcome-based approaches [15,16].

One example of a vulnerability-based segmentation approach is the *Pathways* method [15]. Pathways is a vulnerability-based framework grounded in the social determinants of health, recognizing that RMNCH+N outcomes are shaped by interacting individual, social, economic, and environmental factors. It conceptualizes vulnerability across six domains: The woman and her past experience; household relationships; household economics and living conditions; social support; health mental models; and human and natural systems. By integrating these dimensions, the approach identifies population subgroups with distinct vulnerability profiles to inform tailored RMNCH+N interventions [15]. Pathways has recently been applied in India, Nigeria, Kenya, Senegal, and Indonesia [17]. By identifying distinct

vulnerability profiles rooted in the social determinants of health, it enables policymakers and program planners to design targeted, women-centered interventions that address the underlying drivers of RMNCH+N inequities and improve service delivery for vulnerable populations [18].

The current study applies the *Pathways* segmentation approach to segment Ethiopian women of reproductive age (WRA) based on multidimensional vulnerabilities. In addition to generating actionable population segments, the study documents the application of this emerging methodology in Ethiopia and explores its potential to inform precision public health and equitable RMNCH+N service delivery.

2. Methods

2.1. Design

The development of the vulnerability-based population segmentation approach followed six successive steps within a sequential exploratory mixed-methods framework, whereby findings from the qualitative phase informed subsequent quantitative analyses. The steps were: (1) identifying context-specific vulnerabilities through qualitative research; (2) measuring these vulnerabilities using national survey; (3) conducting separate segmentation analyses for urban and rural strata; (4) ranking segments based on RMNCH+N outcomes; (5) developing vulnerability profiles of segments using programmatically relevant characteristics; and (6) creating a Typing Tool (TT) which can classify women into one of the identified segments.

The present manuscript focuses primarily on Steps 2, 3, 4, and 6. Detailed methods and findings related to the identification of context-specific vulnerabilities through qualitative study (Step 1) and the characterization of population segments through vulnerability profiling (Step 5) are provided in the extended technical report available as Supplementary File 1.

2.2. Measuring sources of vulnerabilities

The sources of vulnerability identified through the qualitative study, together with key RMNCH+N indicators, were measured in a national *Pathways* survey conducted in late 2025. The survey included 5,187 women aged 15-49 years who had experienced pregnancy within the preceding 10 years. The sample size was calculated assuming a 95% confidence level, 2% margin of error, design effect of 2, 10% anticipated non-response rate, and RMNCH+N outcome prevalences ranging from 30% to 60%, based on estimates from the 2019 Ethiopia Mini DHS [19].

Participants were selected using a two-stage cluster sampling design from 251 enumeration areas (EAs) distributed across all first-level administrative divisions of Ethiopia. EAs were drawn from the Ethiopian Statistical Service (ESS) sampling frame developed for the recent national census. Within each selected EA, a comprehensive listing of eligible women was done from whom study participants were randomly drawn.

Data was collected electronically using Kobo Toolbox app by trained and experienced interviewers and field coordinators. Anthropometric measurements (height, weight and middle upper arm circumference) were obtained following standardized protocols. Survey instruments were translated into major local languages, underwent formal cognitive testing and piloting prior to field implementation.

Vulnerability was assessed across six *Pathways* domains. The domain *The woman and her past experience* included childhood adversities, socio-demographic characteristics, health information and media exposure, and digital literacy. *Household relationships* captured intimate partner violence, women's empowerment and agency, partner support, and trust. *Household economics and living conditions* included measures of wealth, income, food insecurity, livelihood coping capacity, and time poverty. *Social support* assessed social support networks, social capital, and group participation. *Health mental models* included gender norms, health beliefs, and healthcare-seeking behaviors. Finally, *Human and natural systems* captured geographic, environmental, infrastructural, livelihood, health system, migration, and conflict-related factors.

The survey also measured multiple RMNCH+N outcomes. These included family planning (FP) indicators (contraceptive use and intentions, unmet need, and postpartum FP), maternal and newborn health indicators (ANC, skilled birth attendance, postnatal care (PNC), birth outcomes, kangaroo mother care (KMC), and chlorhexidine cord care), child health indicators (vaccination coverage and infant and young child feeding practices), nutritional indicators (women's and children's anthropometric status and dietary diversity), neonatal and under-five mortality (U5MR), and mental wellbeing of women as measured using the WHO-5 Well-Being Index [20]. Detailed descriptions are provided in Supplementary File 1.

2.3. Population segmentation analysis

The analysis followed a sequential statistical approach in which univariate analysis, exploratory data analysis (EDA), and principal component analysis (PCA) were used for variable screening, selection, and reduction. This was followed by latent class analysis (LCA) to identify distinct population segments. The analysis was guided by a panel of five public health experts (PoE) with expertise in the Ethiopian context and the Pathways segmentation methodology. Further details on each analytical step are provided below.

2.3.1. Univariate analysis

The univariate analysis examined the frequency distributions of both vulnerability and outcome variables with the aim of: (i) ensuring that variable types were appropriate; (ii) identifying variables measured only in sub-samples and recode them where feasible so that vulnerability variables were valid for all observations; (iii) detecting vulnerability variables with sparse categories that could lead to small segment sizes and recoding or collapsing such categories; and (iv) assessing the distribution of variables based on alternative coding schemes to ensure that variables have appropriate frequency distribution for further analysis. Based on the results of the univariate analysis, several data preparation steps were undertaken,

including the creation of composite variables, collapsing categories with low frequencies, transforming numeric variables into categorical variables, and converting system-missing values into user-defined missing categories where necessary.

2.3.2. Exploratory data analysis

The objective of the EDA was to identify the most relevant vulnerability variables across the six Pathways domains that were most strongly associated with several RMNCH+N outcomes, so that these variables could be carried forward to the PCA. At this step, the association between each vulnerability variable and more than 60 outcome variables was assessed individually using predicted probability plots, and odds ratios (ORs) with 95% confidence intervals (CIs). At the conclusion of the EDA, vulnerability variables that demonstrated meaningful associations with multiple outcome variables were retained for further analysis. The EDA was stratified by place of residence (urban vs. rural); consequently, separate sets of variables were shortlisted for each stratum. Beyond statistical considerations, the PoE also evaluated the practical relevance, plausibility, and overall importance of the vulnerability factors across multiple RMNCH+N outcomes, rather than focusing on a single outcome.

2.3.3. Principal component analysis

PCA was subsequently applied to reduce redundancy among variables while preserving data richness. PCA was conducted separately for each of the six domains and stratified by place of residence (urban vs. rural), resulting in a total of 12 separate models. The aim was to retain approximately 4–8 variables per domain within each stratum. The number of principal components (PCs) to retain was guided by the cumulative proportion of explained variance and inspection of scree and variable representation plots. Biplots constructed from the selected components were examined to detect redundancy among variables by considering both the direction (correlation structure) and magnitude (loading strength) of variable vectors. In addition, the quality of representation of variables (cos^2) on the retained PCs was used to evaluate how well each variable was captured by the factor solution. Variables were excluded using an iterative backward elimination process within each domain, whereby variables were removed stepwise until a final set representing distinct dimensions with minimal redundancy was retained. In addition to statistical considerations, the selection process also incorporated the practical relevance of each variable.

2.3.4. Latent class analysis

Finally, population segmentation was conducted using LCA. LCA is particularly well-suited for categorical data used in this analysis, as it is specifically designed to model relationships among discrete variables without requiring assumptions of continuity or normality [21]. It employs a probabilistic framework to assign individuals to latent groups based on observed patterns of vulnerability factors [21]. Separate LCA models were estimated for urban and rural strata, resulting in distinct segmentation solutions for both.

A bottom-up approach was used to develop the final segmentation model. The process began with a small set of core vulnerability factors, typically 2–3 per domain. These factors were identified based on both statistical strength and practical relevance. Additional variables were then introduced iteratively, and their contribution was evaluated in terms of both statistical performance and substantive importance.

Selection of the optimal segmentation solution was guided by a combination of statistical and practical criteria. Multiple LCA models were estimated with the number of classes (K) ranging from 2 to 10, and each solution was reviewed. Statistical considerations included model convergence, goodness-of-fit indicators (e.g., Bayesian Information Criterion [BIC]), stability of the likelihood function (e.g., consistency of maximum likelihood values and evidence of a clear global maximum), and appropriate segment sizes (typically ranging from 10% to 30% of the population). Model stability was further assessed using Sankey plots to examine the consistency of class membership across competing solutions. Practical considerations included the interpretability and programmatic relevance of the segmentation solutions, as well as the presence of meaningful differentiation across segments. To ensure reproducibility, a fixed random seed was applied, and each model was estimated using 100 simulations.

2.4. Ranking of segments

Once the optimal segmentation solution was identified for urban and rural populations, the segments were ranked according to their RMNCH+N outcome profiles. All binary outcome variables were first recoded with 0 indicating a favorable outcome and 1 an unfavorable outcome. Correlations among outcomes were examined, and one variable from each highly correlated pair ($r > 0.90$) was excluded. Segment ranking was performed using a pooled framework that combined all urban and rural segments. For each outcome, weighted prevalence estimates were calculated by segment and normalized using min-max scaling, such that the segment with the lowest prevalence was assigned a value of 0 and the highest a value of 1. The median of the normalized outcome scores was then calculated for each segment.

Segments were classified into four vulnerability levels based on their median normalized scores: least vulnerable (0.00 ≤ score < 0.25), less vulnerable (0.25 ≤ score < 0.50), more vulnerable (0.50 ≤ score < 0.75), and most vulnerable (0.75 ≤ score ≤ 1.00). Segments were labeled using the format U1-U4 and R1-R4, where the prefixes denote urban or rural residence and the number denotes the vulnerability level. When multiple segments fell within the same level, suffixes (e.g., R3.1 and R3.2) were assigned according to their vulnerability scores, with higher suffixes indicating greater vulnerability within the same category.

2.5. Quantitative profiling of the segments

Quantitative profiling involved developing detailed descriptions of the urban and rural segments using key vulnerability indicators, including those considered during the segmentation process. The objective was to generate a deeper understanding of the segments and identify characteristics relevant for designing tailored interventions. The results of the quantitative profiling are provided in Supplementary File 1. The profiling also incorporated a geographic dimension. First, the distribution of urban and rural segments across Ethiopia's administrative regions was examined. Second, Bayesian spatial interpolation was applied to rural segments to visualize their geographic distribution beyond administrative boundaries and identify areas of higher segment concentration. This analysis was not conducted for urban segments because their distribution is not spatially continuous.

2.6. Typing Tool development

The TT is a simplified screening instrument designed to enable frontline health workers (FHWs) to classify women into population segments so that they can deliver tailored interventions. The tool can also be used to recruit participants from specific segments for research purpose or classify respondents in existing datasets for secondary analyses. The tool was developed using Classification and Regression Tree (CART) analysis based on a reduced set of vulnerability variables from the segmentation model. Decision trees were optimized to achieve high classification performance, targeting sensitivity and specificity of at least 80% across all segments.

TT development followed a top-down variable reduction approach by which the full set of segmentation variables, and predictors with low importance scores were iteratively removed while monitoring classification performance. The process continued until an acceptable balance between accuracy and practicality was achieved. The TT's validity was assessed by comparing observed and predicted segment membership using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV),

and balanced accuracy. The final TTs were based on pruned decision trees selected using the complexity parameter (CP) criterion to balance predictive accuracy and simplicity.

2.7. Measurement of vulnerability variables

The survey measured nearly 450 vulnerability variables across the six pathways vulnerability domains, here we only describe variables used for the actual segmentation analysis and TT development.

2.7.1. The Woman and Her Past Experiences

Women's background characteristics included the place where they primarily grew up (urban or rural), educational attainment (no formal education, primary education, or secondary/tertiary education), age at first marriage (<18 years vs. ≥18 years), number of children under five years of age (0-1 vs. ≥2), and access to a mobile phone, defined based on ownership or presence of another mobile phone at home the women can use (yes/no). Exposure to mass media was assessed using a composite score derived from frequency of exposure to four media sources: internet, radio, television, and newspapers or magazines. Item scores were summed to generate a total score ranging from 0 to 12 and categorized as no exposure (0), low exposure (1-4), and high exposure (5-12).

2.7.2. Household Relationships

Exposure to intimate partner violence (IPV) was assessed using DHS-based questions capturing lifetime experience of emotional, physical, and sexual violence [22]. Perceived partner support was measured using a composite score based on support received in healthcare, economic activities, domestic responsibilities, and emotional well-being. Women's agency within the household was captured through two dimensions – Empowerment in household leadership and Economic autonomy and control over earnings. Empowerment in household leadership was assessed using a composite measure of women's participation in major household decisions and influence over household affairs and was categorized as low (<2), medium (2-2.9), or high (3-4). Economic autonomy and control over earnings were measured using indicators of women's ability to manage and influence financial decisions regarding personal and household income and similarly categorized as low, medium, or high.

2.7.3. Household Economics and Living Conditions

Household wealth was measured using a composite index based on 39 household asset and living-condition indicators as commonly done in DHS. Wealth scores were generated using PCA and classified into five wealth quintiles. Annual household income per capita was calculated based on income over the last 12 months and categorized into five income quintiles. Missing income values were imputed using the

median imputation approach. Finally, households were classified as living below or above the national poverty line based on the annual per capita household income [23]. Respondents also reported their perceived current financial situation, which was classified as able to meet all or most household needs, barely making ends meet, or experiencing severe poverty. Household food deprivation was measured using the Household Hunger Scale [24].

2.7.4. Social Support

Social support was assessed using multiple complementary measures. Perceived social support was measured using the Oslo Social Support Scale (OSSS-3) [25]. The Support Breadth Index measured women's perceived ability to obtain support from relatives, friends, neighbors, and religious groups. The index was adapted from items included in the European Social Survey [26] and categorized as low (0-2), medium (3-5), or high (6-8). Social capital was measured using the World Bank's Social Capital Assessment Tool and, more specifically, the Integrated Questionnaire for the Measurement of Social Capital (SC-IQ) [27]. A Composite Social Capital Score was also calculated based on respondents' perceived access to social capital across six domains during emotional needs, economic needs, health needs, influence within the community, needs during social ceremonies, and support during emergencies. Diversity of group membership was measured based on membership in community-based organizations and categorized as None, Small (1 group), and Medium or large (2 or more groups).

2.7.5. Health Mental Models

Health mental models included individual beliefs, attitudes, and normative influences related to health behaviors. Acceptance of wife beating was measured using standard attitudinal items commonly used in DHS and classified as no justification versus some justification. Women's knowledge of modern family planning methods was assessed as knowledge of at least one modern method (yes/no). Health beliefs were measured using a composite Health Belief Model (HBM) score derived from the six constructs of perceived susceptibility, perceived severity, perceived benefits, perceived barriers, cues to action, and self-efficacy. Each construct was rated out of 4, a composite score ranging from 0 to 24 was generated and categorized as low (0-8), medium (9-16), or high (17-24). Descriptive and injunctive gender norms were measured using a modified G-NORM (Gender Norms Scale) [28] and categorized into tertiles representing egalitarian, intermediate, and restrictive norms. Preferred place of delivery was also included as an indicator and classified as home versus health facility delivery.

2.7.6. Human and Natural Systems

Contextual and environmental conditions were captured through measures of the local setting and access to services. Enumeration areas were classified as major cities, smaller towns or peri-urban areas, and rural areas. Predominant livelihood patterns were categorized as urban livelihoods, agrarian/agropastoral livelihoods, or pastoral livelihoods (nomadic or sedentary). Physical access to infrastructure was assessed using travel time to the nearest all-weather road (<30 minutes vs. >30 minutes). Access to health services was measured by walking time to the nearest health facility and categorized as <15 minutes, 16-30 minutes, or ≥30 minutes. Household access to improved drinking water was classified according to standard WHO/UNICEF Joint Monitoring Program criteria as improved or unimproved [29].

2.8. Data management and analysis

Survey data were cleaned using SPSS version 32 and analyzed using R statistical software. All analyses employed a weighted approach to account for the survey's multistage sampling design. Sampling weights were derived by calculating design weights and selection probabilities at both the primary and secondary sampling unit levels. Post-stratification adjustments were then applied to align the weighted sample with the national population distribution according to urban-rural residence and level 1 administrative regions. Finally, the weights were normalized so that the sum of the weighted observations equaled the unweighted sample size. Analysis related to population segmentation, including EDA, PCA, LCA, and CART modeling, was conducted using the R statistical software. These analyses were implemented by customizing standardized scripts developed by the Institute for Disease Modeling (IDM) of the Gates Foundation.

3. Results

3.1. Socio-demographic characteristics

Table 1 summarizes the basic characteristics of the WRA included in the segmentation analysis. A total of 5,187 WRA who had experienced at least one pregnancy during the 10 years preceding the survey were included. After applying sampling weights, 76.0% of respondents (n = 3,941) were from rural areas, while 24.0% were from urban areas.

The mean (± SD) age of respondents was 30.1 (± 6.8) years, and nearly half (48.6%) were aged 25-34 years. The median number of live births was 3, and almost two-thirds of women (63.3%) had three or more live births. Approximately two in five women (41.8%) had no formal education. The vast majority (90.8%) were married or living with a partner. Among these women, 34.3% of their partners had no formal education. Nearly half of all households (50.9%) had five or more members. According to WHO and UNICEF definitions [29], 21.6% of households used improved sanitation facilities, while 68.3% had access to improved drinking water sources. Based on reported household income, 44.2% of households lived below the national poverty line [21] (Table 1).

323 Table 1: Socio-demographic characteristics of women of reproductive age, Ethiopia, 2025.

Socio-demographic characteristics	Frequency (n = 5,187)	Percent
Place of residence		
Urban	1,246	24.0
Rural	3,941	76.0
Maternal age (years)		
15 to 24	988	19.0
25 to 34	2,521	48.6
35 or above	1,679	32.4
Number of livebirths		
< = 2	1,901	36.6
3 to 4	1,486	28.6
5 or above	1,801	34.7
Women's educational status		
No formal education	2,167	41.8
Primary	2,004	38.6
Secondary or above	1,017	19.6
Religion		
Orthodox	2,129	41.0
Muslim	1,826	35.2
Protestant	1,145	22.1
Others	87	1.7
Marital status		
Partnered	5,166	90.8
Divorced or separated	307	5.9
Widowed	149	2.9
Never in union	21	0.4
Partner's educational status (n= 4,710)		
No formal education	1,617	34.3
Primary	1,754	37.2
Secondary or above	1,294	27.5
Not sure	45	1.0
Household size		
Small (1 to 4 members)	2,029	39.1
Medium or large (5 or above members)	3,158	60.9
Households with improved:		
Drinking water source	3,541	68.3
Excreta disposal	1,120	21.6
Roof	3,918	75.5
Floor	859	16.6
Wall	368	7.1
Sex of the household head		
Male	4691	90.4
Female	496	9.6
Household poverty based on income/person/year		
Above national poverty line	2,897	55.8
Below national poverty line	2,290	44.2

3.2. Status of key RMNCH+N indicators

Modern contraceptive prevalence among in-union women was 47.1%, while unmet need for FP was 26.6%. Among women with recent history of birth, only 55.8% completed four or more ANC visits, and 4.8% had eight or more visits. Skilled birth attendance was 60.7%, while only 22.7% of women received PNC within 24 hours of delivery and 3.0% completed four or more PNC visits. Coverage of essential newborn care was similarly low, with 21.6% of newborns receiving PNC within 24 hours, 24.7% receiving chlorhexidine cord care, and 35.2% of low birth weight or premature infants benefiting from KMC.

Early initiation of breastfeeding (82.4%) and exclusive breastfeeding among infants under six months (77.5%) were relatively high; however, only 17.4% of children aged 6-24 months achieved minimum dietary diversity and 14.6% met the minimum acceptable diet. Immunization coverage was 79.4% for Penta 1 to 63.4% for Penta 3. Neonatal and U5MR rates were 23 and 47 deaths per 1,000 live births, respectively. Among children under five, 32.1% were stunted, 7.7% wasted, and 16.9% underweight. Additionally, 23.7% of non-pregnant women were undernourished, while 45.2% scored at or below the WHO-5 screening threshold for poor mental wellbeing (Figure 1).

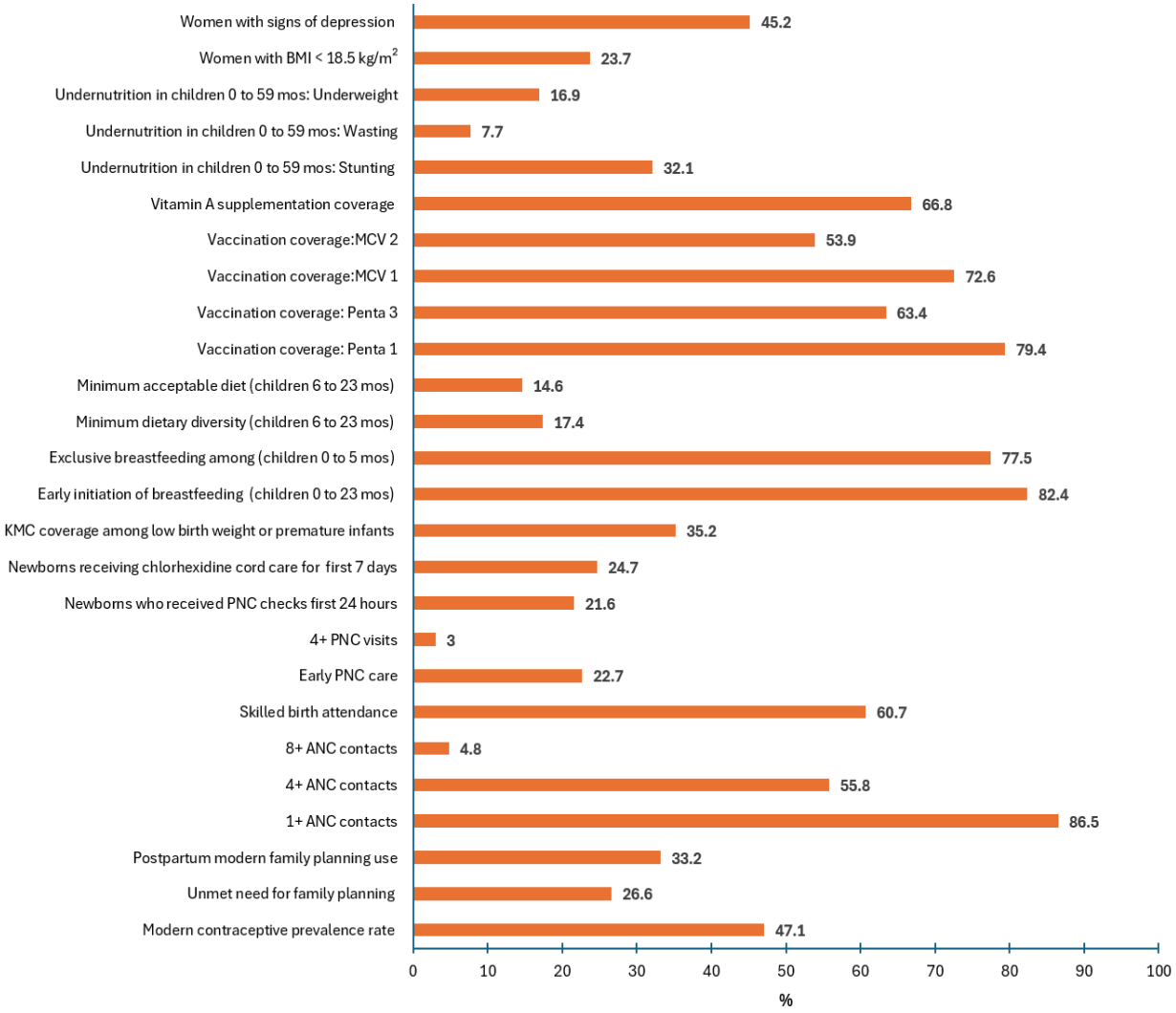


Figure 1: Summary of RMNCH+N indicators, Ethiopia, 2025.

3.3. Reduction of vulnerability variables

Univariate analyses were initially conducted on 449 potential vulnerability variables, ranging from 25 variables in the Social Support domain to 130 in the Household Economics and Living Conditions domain. Based on these analyses, 289 variables were retained for EDA. The EDA assessed associations between each vulnerability variable and 60 RMNCH+N outcomes separately for urban and rural strata, using ORs and predicted probabilities with 95% CIs. At the end of this process, 140 vulnerability variables were retained. PCA was subsequently applied within each domain and stratum to reduce redundancy while preserving the richness of the dataset. This resulted in the selection of 34 variables for the urban stratum and 38 variables for the rural stratum, with four to seven variables retained per domain (Table 2).

Table 2: Selection and reduction of vulnerability variables through the process of univariate, exploratory and principal component analysis.

Domains of vulnerability	Vulnerability variables					
	Initial list	After univariate analysis	After EDA		After PCA	
			Urban	Rural	Urban	Rural
The woman and her past experiences	74	59	19	20	7	7
Household relationships	61	39	17	18	5	7
Household economics and living conditions	130	78	37	32	6	6
Social support	25	20	7	11	4	5
Health and mental models	97	90	24	26	6	6
Natural and human systems	62	38	14	22	6	6
Total	**449**	**289**	**118**	**129**	**34**	**37**

3.4. Prioritizing outcome variables

Among the 128 outcome variables, 68 were excluded for the following reasons: variable type different from binary nature needed for EDA (n = 26), limited practical relevance as judged by the PoE (n = 22), redundancy with other outcome variables (n = 11) or already included as a part of another composite outcome (n = 8). Ultimately the following 60 variables were selected as priority outcome variables (Table 3).

Table 3: Prioritized list of RMNCH+N outcome variables

Domains	Outcome variables
Contraception	Current use of long-acting contraceptives
	Current use of modern contraceptives
	Postpartum contraceptive use (among women who gave birth in last 12 months)
	Intention to use a contraceptive within the next 12 months or at any time in future
Fertility and unmet need	Unmet need for family planning
	Grand multiparity
Maternity service utilization	ANC 1+, 4+, and 8+ coverage (among women with pregnancy history in last 5 years)
	Early ANC initiation (among women who had one ANC)
	Skilled birth attendance, birth at health facility (among women with birth in last 5 years)
	Early PNC and PNC 4+coverage
Pregnancy loss or abortion	Pregnancy loss in last 5 years
	Abortion history (induced or spontaneous)
Newborn health	Low birth weight (based on reported birth size or birth weight)
	Prematurity
	Early PNC for the newborn
	Chlorohexidine applied on cord for at least 7 days
	KMC for premature and low birth weight babies
Infant and young child nutrition	Timely initiation of breastfeeding
	Exclusive breastfeeding (infants 0 to 5 months)
	Currently breastfed (children 0 to 23 months)
	Continued breast feeding (children 12 to 23 months)
	Timely introduction of complementary foods (children 6 to 8 months)
	Adequate dietary diversity, minimum acceptable diet, minimum meal frequency (children 6 to 23 months)
Vaccination	Penta-1 to -3 dropout (children 12 to 35 months)
	BCG, Penta 1, Penta 3, MCV 1, MCV 2 coverages (children 12 to 35 months)
	Vitamin A supplementation coverage (children 6 to 59 months)
Morbidity	Maternal illness in the last 1 month (any)
	Child illness in the last 1 month (any)
	Child morbidity secondary to diarrhea, fever, and fever with cough
Care seeking	Formal care seeking for child's and women's recent morbidity
Child undernutrition	Child thinness based on middle upper arm circumference
	Stunting, underweight, wasting (under five children)
	Global acute malnutrition (under five children)
Maternal undernutrition	Maternal thinness based on MUAC
	Maternal chronic energy deficiency (BMI < 18.5)
	Maternal height < 150 cm
Obesity	Overweight or obesity in children and in women
Child mortality	Neonatal, infant, and under-five mortality rate (based on births and deaths in last 5 year)
General well being	Women's general mental wellbeing

3.5.

3.6. Population segmentation analysis

3.6.1. Variable selection for segmentation

As described in the Methods section, a bottom-up approach was used to develop the final segmentation model. Starting with two to three key variables from each domain, additional variables were added sequentially and retained only if they improved model fit and contributed to meaningful and interpretable segments. All variables retained after PCA were evaluated, while selected variables excluded during PCA were reassessed based on their programmatic relevance. Ultimately, 36 variables were included in the final segmentation solution, comprising 28 variables each for the urban and rural strata (Table 4).

Table 4: Vulnerability variables used for generating the final WRA segmentation solution

Domains and vulnerability variables	Coding scheme	Strata	
		Urban	Rural
The woman and her past experience			
Place where the woman primarily grows up	1 = Urban; 2 = Rural	✓	X
Women's mass media exposure	0 = No exposure (Score: 0); 1 = Low exposure (Score: 1 to 4); 2 = High exposure (Score: 5 to 12)	✓	X
Women's mass media exposure	0 = No exposure (Score: 0); 1 = Some exposure (Score: 1 or above)	X	✓
Number of under-five children	0 = 0 or 1; 1 = 2 or above	✓	✓
Educational status	1 = No formal education; 2 = Primary education; 3 = Secondary or tertiary education	✓	✓
Age at first marriage	1 = < 18 years; 2 = 18 or above years or NA	✓	✓
Woman has access to mobile phone to use	0 = No; 1 = Yes	✓	✓
Household relationship			
Reported any intimate partner violence (IPV)	0 = No; 1 = Yes	✓	X
Reported any physical IPV	0 = No; 1 = Yes	X	✓
Partner support score	1 =Low (0 to 5); 2 = Medium (6 to 10); 3 = High (11 to 16)	✓	✓
Empowerment in household leadership	1 = Low (< 2); 2 = Medium (2 to 2.9); 3 = High (3 to 4)	✓	✓
Economic autonomy & control over earnings	1 = Low (< 2); 2 = Medium (2 to 2.9); 3 = High (3 to 4)	✓	✓
Household economics and living conditions			
House with improved wall, floor, roof	0 = No; 1 = Yes	✓	X
Household saved money in the past year	0 = No; 1 = Yes	X	✓
Household size	1 = 1 – 4 members; 2 = 5 – 7 members; 3 = 8+ members	X	✓
Household's annual income/person quintile	1 = Poorest or Poorer; 2 = Middle; 3 = Richer or Richest	✓	✓
Household hunger score	1 = Little to no hunger; 2 = Moderate or severe hunger	✓	✓
Current financial situation of the household	1 = Able to meet all/most needs; 3 = Barely make ends meet; 4 = Experiencing severe poverty	✓	✓
Household members have saving account	0 = No; 1 = Yes	✓	✓
Household connected to electricity	0 = No; 1 = Yes	✓	✓
Wealth index quintiles	1 = Poorest; 2 = Poorer; 3 = Middle; 4 = Richer; 5 = Richest	✓	✓
Social support			
Support Breadth Index score	1 = Low (0 to 2); 2 = Medium (3 to 5); 3 = High (6 to 8)	✓	X
Has social capital in health domain	0 = No; 1 = Yes	✓	X

Oslo Social Support Scale	1 = Poor; 2 = Moderate; 3 = Strong	X	✓
Social Capital Score	1 = Low (0 to 3); 2 = High (4 to 6)	X	✓
Women's diversity of group membership	1 = None (0); 2 = Small (2); 3 = Medium or high (3 or above)	✓	✓
Health mental models			
Wife beating attitude	0 = No justification to beat wife; 1 = Some justification to beat wife	✓	X
Preferred place of delivery	1 = At home; 2 = Health facility	X	✓
Woman knows at least modern FPs methods	0 = No; 1 = Yes	✓	✓
Woman's Health belief model composite score	1 = Low (0 to 8); 2 = Medium (9 to 16); 3 = High (17 to 24)	✓	✓
Woman's gender norm score	1 = Egalitarian (tertile 1); 2 = Middle (tertile 2); 3 = Restrictive (tertile 3)	✓	✓
Human and natural systems			
Characteristics of the village	1 = Major cities (regional/zonal capital); 2 = Smaller towns/peri-urban areas; 3 = Rural area	✓	X
Livelihood pattern of the locality	1 = Urban livelihood; 2 = Agrarian/agropastoral; 3 = Nomadic or sedentary pastoralism	X	✓
Distance to the nearest all-weather road	1 = < 30 minutes or not sure; 2 = 30+ minutes	✓	✓
Walking distance to nearest health facility	1 = < 15 minutes; 2 = 16-30 minutes; 3 = 30+ minutes	✓	✓
Household has access to improved water source	0 = No; 1 = Yes	✓	✓

3.6.2. Selection of the optimal segmentation solution for urban strata

Figure 2 presents the fit statistics for the urban segmentation analysis. Although the BIC indicated that the 4-, 5-, and 6-class solutions provided the best fit, the 4-class solution showed the clearest global maximum likelihood. Moreover, mean posterior probabilities remained largely unchanged beyond four classes, suggesting little improvement in classification quality with additional classes (Figure 2).

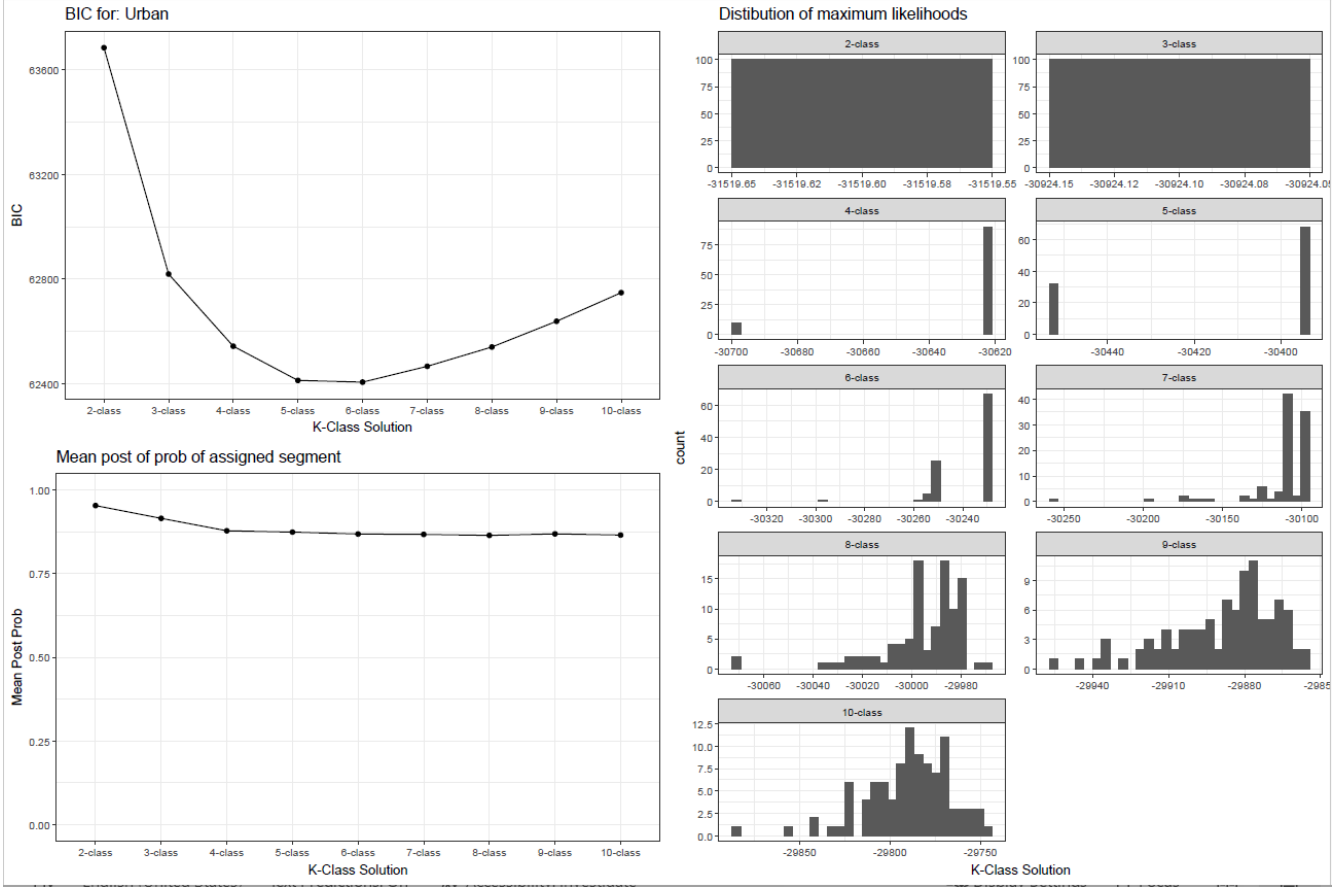


Figure 2: Fit statistics and distribution of maximum likelihoods for the urban WRA segmentation solution

Figure 3 shows the segment size distributions across solutions with 2 to 10 classes. Among the candidate models (4-, 5-, and 6-class), the 4-class solution produced the most balanced segmentation, with segment sizes ranging from 14% to 30%. In contrast, the 5- and 6-class solutions each included at least one segment comprising less than 10% of the sample. Taken together, the fit statistics and segment size distributions supported the 4-class solution as the optimal urban segmentation model (Figure 3).

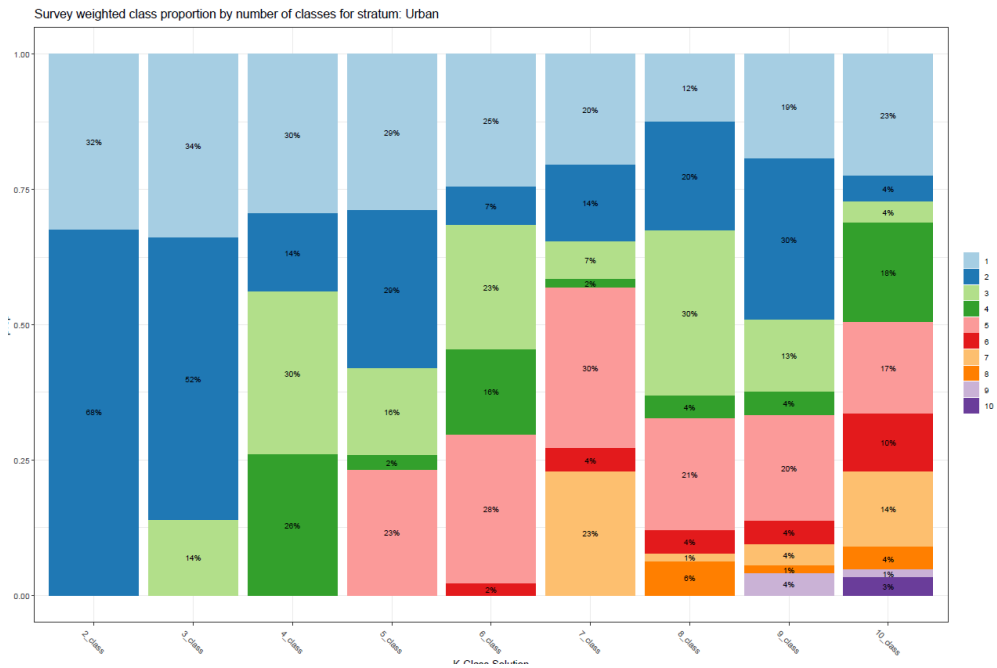


Figure 3: Segment size distribution across urban segmentation solutions

Segment membership across solutions containing 2 to 7 classes was examined using a Sankey plot, which showed largely stable membership and few transitions between segments, indicating a high degree of consistency in segment allocation (Figure 4).

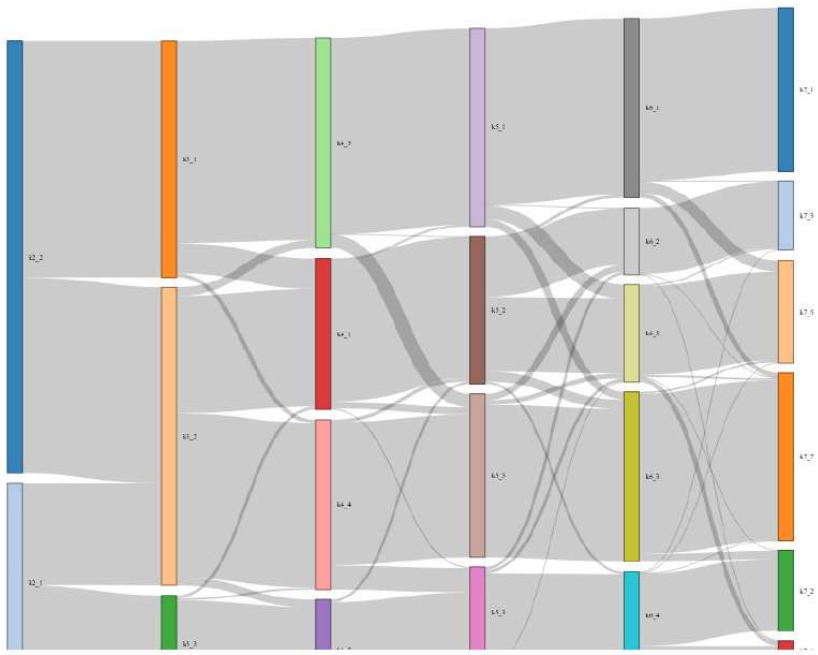


Figure 4: Stability of segment membership across urban segmentation solutions

3.6.3. Selection of the optimal segmentation solution for rural strata

Figure 5 presents the fit statistics for the rural LCA models. Although BIC values tend to decline up to 9-class solution, the 5-class solution achieved the clearest global maximum likelihood. Mean posterior probabilities also changed little beyond six classes, indicating limited improvement with additional classes (Figure 5).

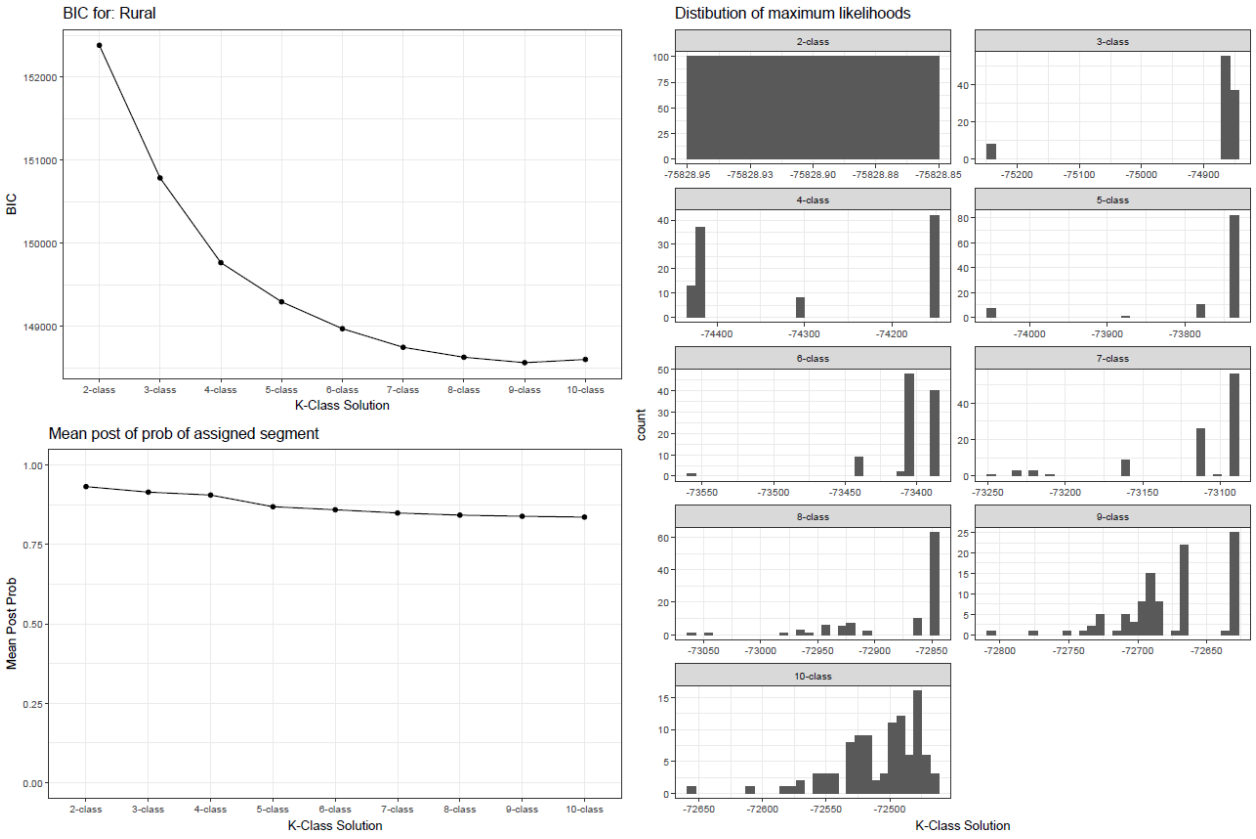


Figure 5: Fit statistics and distribution of maximum likelihoods for rural WRA segmentation solution

Figure 6 presents segment size distributions for solutions with 2 to 10 classes. Among the candidate models, none achieved the ideal segment-size range of 10-30% for all segments. However, the 5-class solution showed the most balanced distribution, with segment sizes ranging from 8% to 33%, whereas the 4-class solution included two segments exceeding 30% of the sample and the 6-class solution produced segment sizes ranging from 8% to 34%. Together with the fit statistics, these findings supported the 5-class solution as the optimal model for the rural stratum (Figure 6).

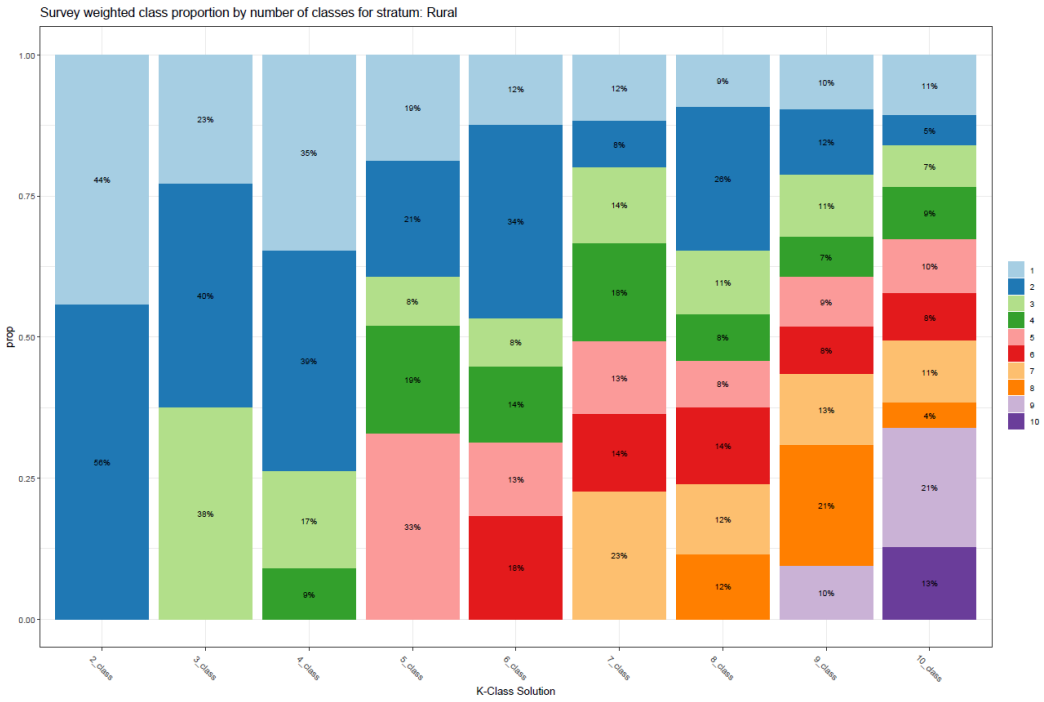


Figure 6: Segment size distribution across rural WRA segmentation solutions

The Sankey plot showed that segment membership remained largely stable across the 2- to 7-class solutions, particularly up to the 5-class model. This consistency supports the stability of the underlying segmentation structure and the selection of the 5-class solution as the optimal model (Figure 7).

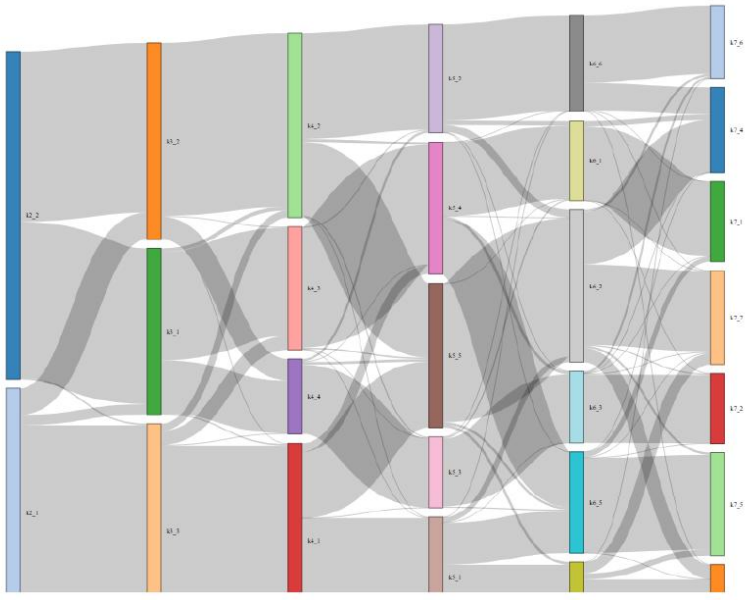
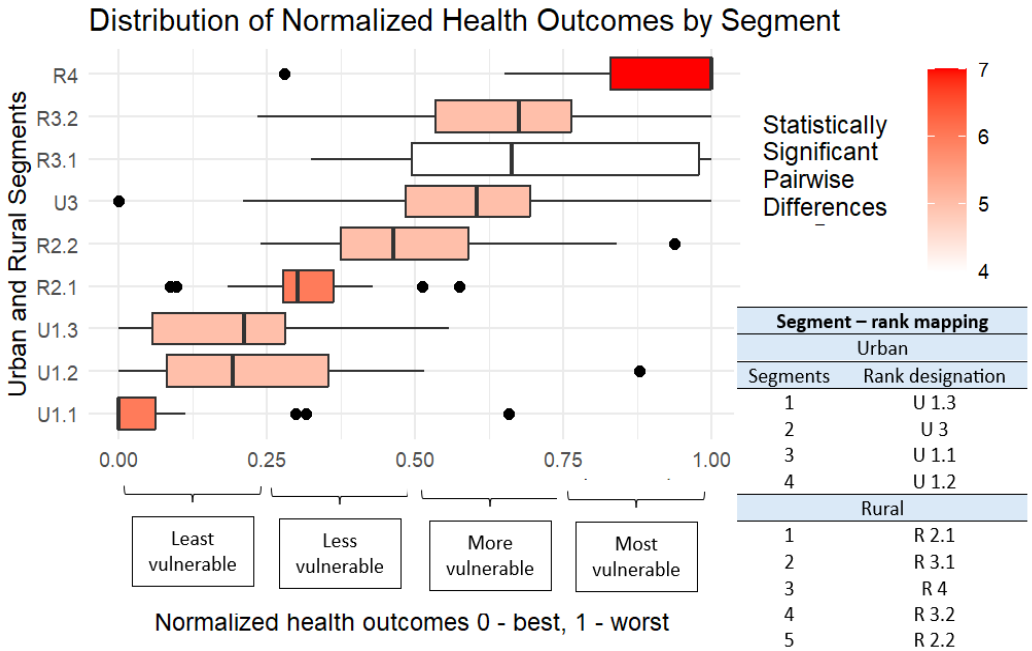


Figure 7: Stability of segment membership across rural WRA segmentation solutions

3.7. Labeling and joint ranking of urban and rural segments

After identifying five rural and four urban segments as the optimal segmentation solutions, the segments were ranked using the approach described earlier. Among the four urban segments, three (U1.1, U1.2, and U1.3) fell within the first, or least vulnerable, quarter. The remaining urban segment (U3) fell within the third, or more vulnerable, quarter. Among rural segments, two (R2.1 and R2.2) fell within the second, or less vulnerable, quarter. Two others (R3.1 and R3.2) fell within the third, or more vulnerable, quarter, while the remaining segment (R4) fell within the fourth, or most vulnerable, quarter (Figure 8).



Segment – rank mapping	
Urban	
Segments	Rank designation
1	U 1.3
2	U 3
3	U 1.1
4	U 1.2
Rural	
1	R 2.1
2	R 3.1
3	R 4
4	R 3.2
5	R 2.2

Figure 8: Joint ranking of urban and rural segments

3.8. Comparison of rural and urban segments based on selected RMNCH+N outcomes

In urban areas, the U5MR varied considerably across segments, ranging from 8 deaths per 1,000 live births in segment U1.1 to 55 deaths per 1,000 live births in segment U3. The U5MR was identical in segments U1.2 and U1.3, at 24 deaths per 1,000 live births. A similarly pronounced pattern was observed across the rural segments. The U5MR ranged from 40 deaths per 1,000 live births in segment R2.1 to 74 deaths per 1,000 live births in segment R4 (Table 5).

Table 5: Mortality among children under five years of age in Ethiopia, 2025

Segments	U5MR
U 3	55
U 1.3	23
U 1.2	23
U 1.1	8
Urban settings	**24**
R 4	74
R 3.2	67
R 3.1	51
R 2.2	44
R 2.1	40
Rural settings	**53**
National	**47**

The urban and rural segments were compared across approximately 20 RMNCH+N outcomes. In both strata, a clear gradient pattern was observed for several key indicators, including maternal health service utilization, vaccination coverage, and vitamin A supplementation. Additional gradients were evident for under-five nutritional status in urban areas; and for contraceptive use, care-seeking for childhood illnesses, and child undernutrition in rural areas. In contrast, the pattern was less consistent for minimum acceptable diet, child morbidity, and, in rural areas, childhood stunting (Figure 9).

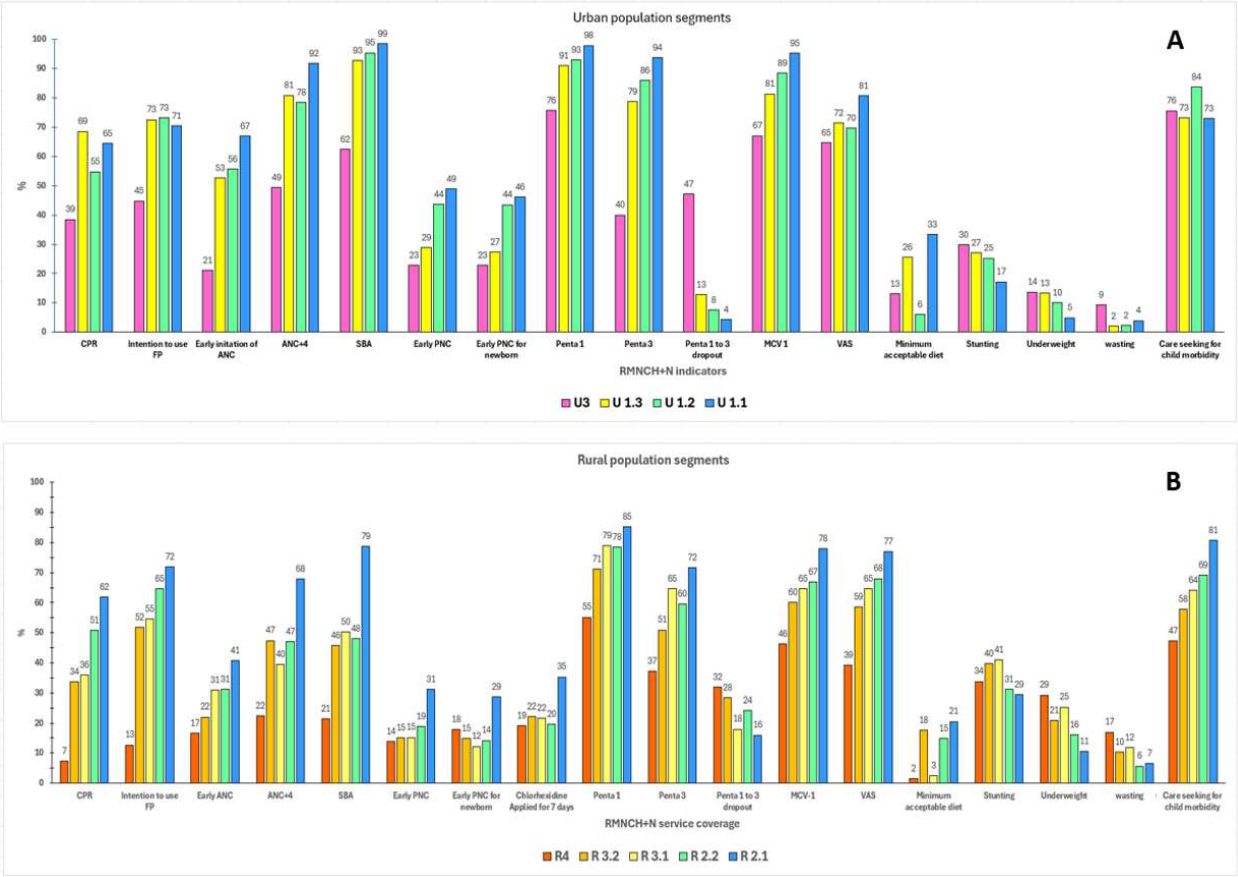


Figure 9: Variations in key RMNCH+N outcomes across urban (Fig. A) and rural (Fig. B) segments.

3.9. Proportion of urban and rural segments at national level

In urban areas, the least vulnerable segment (U1.1) was the largest, comprising 30.3% of women, followed by U1.3 (29.3%) and U1.2 (25.9%), while the most vulnerable segment (U3) accounted for 14.2%. In rural areas, R2.2 was the largest segment (33.0%), followed by R3.1 (20.7%), R3.2 (19.1%), and R2.1 (18.7%), whereas the most vulnerable segment (R4) represented 8.5% of women.

3.10. Geographic profiling of segments

The distribution of segments varied considerably across regions. In urban areas, the more vulnerable segment (U3) was most prevalent in SERS (60.3%), Afar (48.1%), Gambella (40.0%), CERS (37.1%), and Somali (34.6%), whereas the least vulnerable segment (U1.1) was concentrated in Harari (62.5%), Dire Dawa (55.6%), and Addis Ababa (49.0%). In rural areas, the most vulnerable segment (R4) was heavily concentrated in Somali (79.0%) and Afar (59.5%), with limited presence elsewhere. Segment R3.2 was most common in Dire Dawa (44.4%), CERS (43.6%), Sidama (37.3%), Gambella (31.6%), and Southwest

Ethiopia (26.7%), while R3.1 predominated in Tigray (36.6%), Amhara (32.5%), and Benishangul-Gumuz (31.8%). Among the less vulnerable groups, R2.2 was most prevalent in Amhara (42.9%) and Oromia (40.7%), whereas R2.1 was particularly concentrated in rural Harari (80.0%) (Figure 10).



Figure 10: Distribution of urban and rural segments across regions of Ethiopia, 2025.

Geographic interpolation revealed distinct spatial clustering of the rural segments (Figure 10). Among them, R4 showed the strongest geographic concentration, being clustered in the Somali and Afar regions, eastern Oromia, and southern parts of SERS. Segment R3.1 was concentrated in southern Tigray and northeastern Amhara, while R3.2 clustered mainly in CERS and Sidama. In contrast, the less vulnerable segments, R2.1 and R2.2, were more widely dispersed across the country, with multiple localized hotspots (Figure 11).

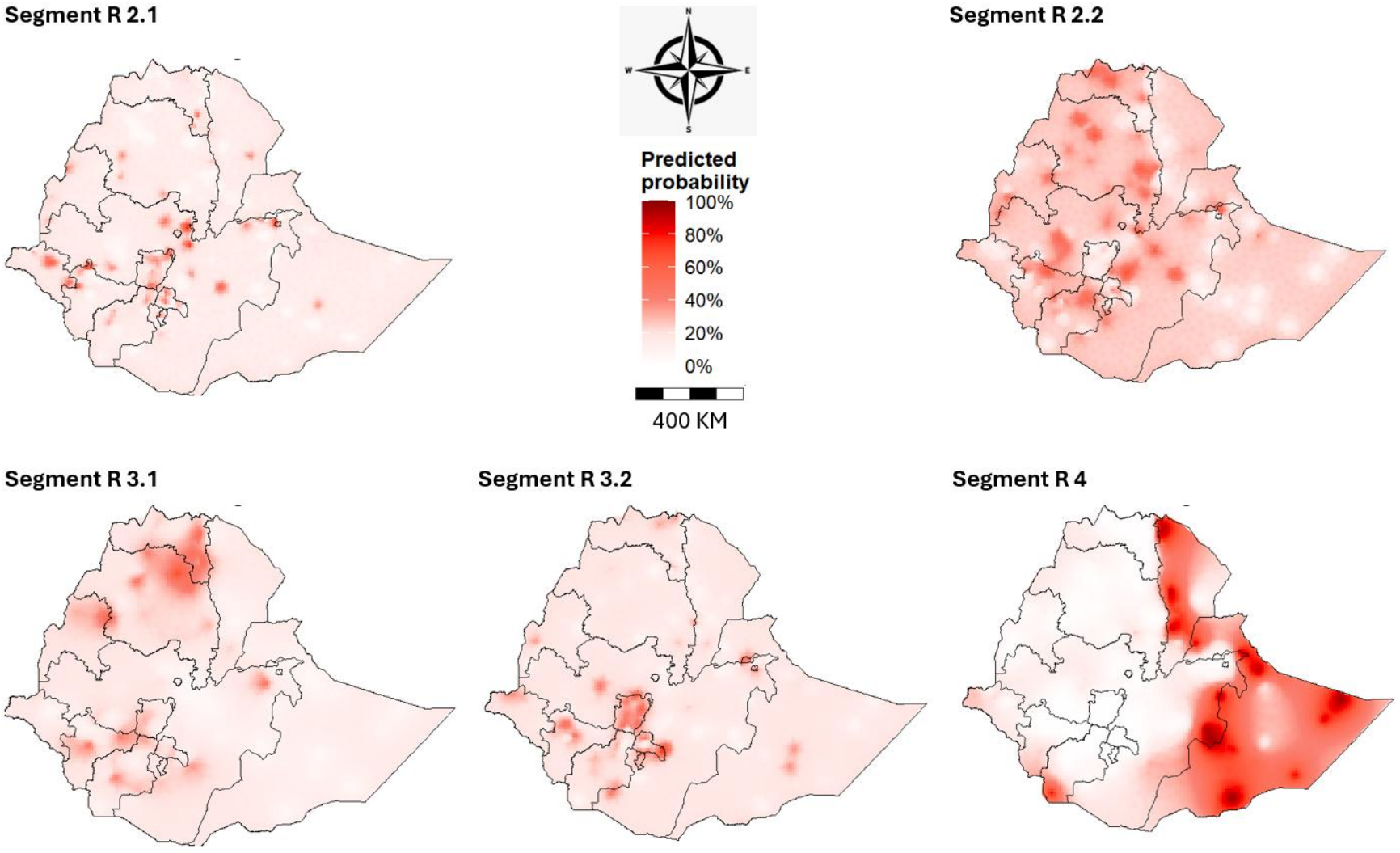


Figure 11: Geographic interpolation for rural segment membership among WRA in Ethiopia.

3.11. Development of Typing tool

The final TT included 17 vulnerability factors for the urban stratum and 16 for the rural stratum. As shown in Table 6, sensitivity, specificity, and balanced accuracy exceeded 80% for all segments except U.1.3, which achieved a sensitivity of 75%. Misclassification analysis indicated that most errors occurred between segments U.1.2 and U.1.3, with 12.0% of women in U.1.3 classified as U.1.2 and 11.4% of women in U.1.2 classified as U.1.3.

Table 6: Accuracy of the urban and rural typing tools

Typing tool type	Segment	Sensitivity	Specificity	PPV	NPV	Balanced accuracy
Urban typing tool	U 1.1	0.835	0.931	0.858	0.919	0.883
	U 1.2	0.792	0.899	0.742	0.921	0.845
	U 3	0.814	0.986	0.915	0.966	0.900
	U 1.3	0.746	0.909	0.720	0.919	0.827
Rural typing tool	R 2.1	0.807	0.961	0.846	0.949	0.884
	R 2.2	0.802	0.920	0.770	0.933	0.861
	R 3.1	0.804	0.958	0.818	0.955	0.881
	R 3.2	0.901	0.956	0.858	0.970	0.928
	R 4	0.820	0.985	0.883	0.975	0.903

4. **Discussion**

Using nationally representative data, this study segmented WRA in Ethiopia based on social determinants of health and demonstrated that these segments differ markedly in their risk profiles across multiple RMNCH+N outcomes. The findings indicate that the population segments represent distinct strata of health risk, creating an opportunity to design and deliver services that are tailored to their specific needs. Furthermore, the TT demonstrated acceptable levels of sensitivity and specificity in classifying women into their corresponding segments using a smaller set of indicators. This suggests that FHWs could feasibly identify segment membership and use this information to deliver targeted services.

The application of population segmentation in healthcare and public health has grown considerably in recent years. Studies have demonstrated the utility of segmentation approaches in identifying population subgroups with distinct health needs. A recent study from the United States identified four unique clusters of African American adults based on social determinants of health and found substantial heterogeneity in unmet health needs across segments [30]. Similar approaches have been used to stratify people living with chronic illnesses and older adults receiving primary care services [31,32]. The present study extends this growing body of evidence by applying segmentation within the RMNCH+N context.

The Pathways segmentation framework has recently been applied in several countries, including Indonesia, Kenya, and Nigeria, using nationally representative DHS [17]. Across these settings, the analyses consistently identified distinct and internally homogeneous population segments characterized by different vulnerability profiles and health risks. For example, the segmentation conducted in northern Nigeria identified a highly disadvantaged group of women with extremely low CPR and very high U5MR, illustrating how social vulnerability can translate into poor health outcomes [17]. The consistency of findings across countries suggests that segmentation based on social determinants of health is a robust approach for identifying population subgroups with unique needs, allowing for tailoring of interventions.

The findings also highlight the potential of population segmentation as a novel framework for measuring and addressing health inequalities. Traditionally, inequalities in RMNCH+N outcomes have been assessed using individual dimensions such as wealth, education, area of residence, geographic location, gender empowerment, ethnicity, or disability status. While these indicators are valuable, they often capture only a single dimension of health inequality. In reality, vulnerabilities frequently intersect and accumulate within the same individuals and households. By integrating multiple social determinants into a unified segmentation framework, population segmentation provides a more holistic measure of social disadvantage and allows policymakers to identify groups experiencing overlapping vulnerabilities.

The findings also have implications for the growing field of precision public health. Building on the concepts of precision medicine, precision public health emphasizes delivering the right intervention to the right population at the right time [12]. However, practical approaches for operationalizing precision public health through social determinants of health remain limited. The present study demonstrates that segmentation of women based on social vulnerabilities can generate relatively homogeneous groups with distinct health profiles and service needs. Such information can support the design of targeted, context-specific interventions that are more responsive to the realities faced by different groups of women.

The performance of the TT further supports the practical applicability of the segmentation framework. Sensitivity and specificity exceeded the predefined target of 80% for eight of the nine segments, indicating that the tool can reliably classify women into their respective groups using a relatively small set of variables. The only exception was segment U1.3, for which sensitivity was slightly below the target threshold. These findings suggest that while the TT is suitable for programmatic use, caution may be warranted when classifying women into segment U1.3, particularly because misclassifications occurred most frequently between segments U1.2 and U1.3.

Several strengths of this study warrant consideration. First, the segmentation was informed by a qualitative investigation, ensuring that the factors are contextually relevant. Second, these vulnerabilities were measured through a primary survey, enabling the inclusion of a comprehensive set of social determinants that may not be readily available in secondary data analysis. Third, unlike conventional risk stratification approaches that classify individuals based on health outcomes, this study segmented women exclusively based on vulnerabilities, creating opportunities to support women at increased risk before adverse outcomes occur. Finally, by simultaneously incorporating a broad range of RMNCH+N outcomes during segment ranking, the framework supports integrated programming and provides a comprehensive understanding of intersecting vulnerabilities. Nevertheless, because segments may exhibit distinct patterns of vulnerability for specific outcomes, intervention design should consider both the overall vulnerability profile of each segment and its relationship with particular health outcomes of interest.

The findings should also be interpreted taking into consideration the following limitations. The segmentation model was developed using data from women who had given birth within the previous ten years; therefore, its applicability to women without pregnancy experience remains uncertain. In addition, some characteristics used to define the segments may change over time, and therefore segmentation solutions should not be viewed as static and require periodic updating.

Overall, the study demonstrates that population segmentation based on social determinants of health can identify meaningful and actionable subgroups of Ethiopian WRA with distinct vulnerability and RMNCH+N risk profiles. The approach offers a promising approach for advancing precision public health and strengthening equity-oriented health programming. Future research should focus on assessing the implementation feasibility of this approach through an implementation science lens, as well as evaluating the effectiveness and cost-effectiveness of segment-tailored interventions in improving health outcomes.

5. Conclusion

This study demonstrates that population segmentation based on social determinants of health can identify practically meaningful subgroups of Ethiopian WRA with distinct vulnerability profiles and markedly different RMNCH+N outcomes. By moving beyond conventional measures of inequality, the segmentation approach provides a more comprehensive understanding of intersecting vulnerabilities and offers a practical framework for precision public health. The accompanying TT further enhances the programmatic utility of the approach by enabling the identification of women at greatest risk and supporting the delivery of tailored interventions. These findings provide a strong foundation for the design, implementation, and evaluation of segment-specific strategies aimed at improving RMNCH+N outcomes and reducing health inequities in Ethiopia.

References

1. Central Statistical Authority (CSA) and ORC Macro. *Ethiopia Demographic and Health Survey 2000*. Addis Ababa, Ethiopia and Calverton, Maryland, USA: CSA and ORC Macro; 2001.
2. Ethiopian Statistical Service, The DHS Program, ICF. *Ethiopia Demographic and Health Survey 2024–25.* Addis Ababa, Ethiopia, and Rockville, Maryland, USA: ESS and ICF; 2026.
3. World Health Organization (WHO), *Trends in maternal mortality 2000 to 2023: estimates by WHO, UNICEF, UNFPA, World Bank Group and UNDESA/Population Division*. Geneva: WHO; 2024.
4. United Nations Children's Fund (UNICEF). Under-five mortality. UNICEF Data. Available from: https://data.unicef.org/topic/child-survival/under-five-mortality/, Accessed on 24 May 2026.
5. Yigzaw M, Zakus D, Tadesse Y, Desalegn M, Fantahun M. Paving the way for universal family planning coverage in Ethiopia: an analysis of wealth related inequality. *Int J Equity Health*. 2015; 14: 77. doi: 10.1186/s12939-015-0214-7.
6. Asefa A, Gebremedhin S, Marthias T, Nababan H, Christou A, Semaan A, et al. Wealth-based inequality in the continuum of maternal health service utilisation in 16 sub-Saharan African countries. *Int J Equity Health*. 2023; 22(1): 203. doi: 10.1186/s12939-023-02015-0.
7. Shiferie F, Gebremedhin S, Andargie G, Tsegaye DA, Alemayehu WA, Fenta TG. Decomposition analysis of socioeconomic inequalities in vaccination dropout in remote and underserved settings in Ethiopia. *Am J Trop Med Hyg*. 2024;111(1):196–204. doi:10.4269/ajtmh.23-0816.
8. Arroyave L, Saad GE, Victora CG, Barros AJD. Inequalities in antenatal care coverage and quality: an analysis from 63 low and middle-income countries using the ANCq content-qualified coverage indicator. *Int J Equity Health*. 2021; 20(1): 102. doi: 10.1186/s12939-021-01440-3.
9. Denu Z, Defar A, Persson L, Lemma S, Berhanu D, Getachew T, et al. Socio-economic and geographic equity in maternal health services utilization in Ethiopia: a community-based cross-sectional study. *BMC Health Serv Res*. 2025;25(1):610. doi:10.1186/s12913-025-12639-3.
10. Tessema NS, Geda NR. Trends in maternal education-based inequalities in under-five mortality in Ethiopia: multilevel, decomposition and concentration index analyses. *Discov Soc Sci Health. 2024;* **4**: 59. doi:10.1007/s44155-024-00122-z.
11. Handebo S, Demie TG, Gessese GT, Woldeamanuel BT, Biratu TD. Effect of women's literacy status on maternal healthcare services utilisation in Ethiopia: a stratified analysis of the 2019 mini–Ethiopian Demographic and Health Survey. *BMJ Open*. 2023;13(11):e076869. doi:10.1136/bmjopen-2023-076869.
12. Khoury MJ, Iademarco MF, Riley WT. Precision Public Health for the Era of Precision Medicine. *Am J Prev Med.* 2016; 50(3): 398-401. doi: 10.1016/j.amepre.2015.08.031.
13. Frieden TR. Six components necessary for effective public health program implementation. *Am J Public Health*. 2014; 104(1): 17-22. doi: 10.2105/AJPH.2013.301608.
14. Yan S, Kwan YH, Tan CS, Thumboo J, Low LL. A systematic review of the clinical application of data-driven population segmentation analysis. *BMC Med Res Methodol.* 2018; 18 :121. doi:10.1186/s12874-018-0584-9.
15. Root ED, Mishra A, Lin Y, Cooper J, Kalnoky M, Gao L, et al. The Pathways segmentation approach: a novel application for identifying intersecting vulnerabilities in public health. *VeriXiv [preprint]*. 2025;2:143
16. Olstad DL, McIntyre L. Reconceptualising precision public health. *BMJ Open*. 2019;9(9):e030279. doi:10.1136/bmjopen-2019-030279.
17. Pathways. Population segmentations. Available from: https://withpathways.org/en/segmentations, accessed on Aug 19, 2026.

18. Johnson TP, Wendland M. A role for design in global health: Making the concept of vulnerability actionable. *She Ji*. 2022;8(4):486-503.
19. Ethiopian Public Health Institute (EPHI), ICF. *Ethiopia Mini Demographic and Health Survey 2019: Final Report*. Addis Ababa, Ethiopia, and Rockville, MD, USA: EPHI and ICF; 2021.
20. World Health Organization. The WHO-Five Well-Being Index (WHO-5). Available from: https://www.who.int/publications/m/item/WHO-UCN-MSD-MHE-2024.01, accessed on Aug 19, 2026.
21. Collins LM, Lanza ST. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Hoboken (NJ): John Wiley & Sons; 2010.
22. Croft TN, Aileen MJ, Courtney AK. *Guide to DHS Statistics*. Rockville, MD, USA: ICF; 2018.
23. World Bank. *Ethiopia regional poverty report: Promoting equitable growth for all regions*. Washington (DC): World Bank; 2020.
24. Food and Agriculture Organization of the United Nations (FAO). *Methods for estimating comparable prevalence rates of food insecurity experienced by adults throughout the world*. Rome: FAO; 2013.
25. Kocalevent RD, Berg L, Beutel ME, Hinz A, Zenger M, Härter M, et al. Social support in the general population: standardization of the Oslo Social Support Scale (OSSS-3). *BMC Psychol*. 2018;6(1):31.
26. European Social Survey European Research Infrastructure (ESS ERIC). *European Social Survey*. Bergen (Norway): ESS ERIC; 2023. Available from: https://www.europeansocialsurvey.org accessed on July 01, 2024.
27. Grootaert C, Narayan D, Jones VN, Woolcock M. *Measuring Social Capital: An Integrated Questionnaire*. Washington (DC): World Bank; 2004.
28. Sedlander E, Bingenheimer JB, Long MW, Swain M, Rimal RN. The G-NORM Scale: Development and validation of a theory-vased Gender Norms Scale. *Sex Roles*. 2022;87(5-6):350-363. doi:10.1007/s11199-022-01319-9.
29. World Health Organization, United Nations Children's Fund. *Progress on household drinking water, sanitation and hygiene 2000–2022: special focus on gender*. Geneva: WHO; 2023.
30. Suzuki S, Longcoy J, Isgor Z, Avery E, Johnson TJ, Yang E, et al. Clustering of social determinants of health as an indicator of meaningful subgroups within an African American population: application of latent class analysis. *Int J Environ Res Public Health*. 2024;21(6):676. doi:10.3390/ijerph21060676.
31. Smeets RGM, Elissen AMJ, Kroese MEAL, Hameleers N, Ruwaard D. Identifying subgroups of high-need, high-cost, chronically ill patients in primary care: a latent class analysis. *PLoS One*. 2020;15(1):e0228103. doi:10.1371/journal.pone.0228103.
32. Hutchins F, Thorpe J, Maciejewski ML, et al. Clinical outcome and utilization profiles among latent groups of high-risk patients: moving from segmentation towards intervention. *J Gen Intern Med*. 2022;37(10):2429-2437. doi:10.1007/s11606-021-07166-w.

Abbreviations

ANC	Antenatal Care
BIC	Bayesian Information Criterion
BMI	Body Mass Index
CART	Classification and Regression Tree
CERS	Central Ethiopia Regional State
CP	Complexity Parameter
CPR	Contraceptive Prevalence Rate
DHS	Demographic and Health Survey
EA	Enumeration Area
EDA	Exploratory Data Analysis
ESS	Ethiopian Statistical Service
FP	Family Planning
G-NORM	Gender Norms Scale
IDM	Institute for Disease Modeling
IPV	Intimate Partner Violence
KMC	Kangaroo Mother Care
LCA	Latent Class Analysis
LMIC	Low- and Middle-Income Country
MCV	Measles-Containing Vaccine
MUAC	Mid-Upper Arm Circumference
NPV	Negative Predictive Value
OSSS-3	Oslo Social Support Scale, 3-item Version
PCA	Principal Component Analysis
PC	Principal Component
PoE	Panel of Experts
PPV	Positive Predictive Value
RMNCH+N	Reproductive, Maternal, Newborn, Child Health and Nutrition
SD	Standard Deviation
SERS	South Ethiopia Regional State
TT	Typing Tool
U5MR	Under-Five Mortality Rate
WHO	World Health Organization
UNICEF	United Nations Children's Fund
WRA	Women of Reproductive Age

Supplementary Information

- Supplementary File 1: Technical report from which the manuscript emanated

Acknowledgements

The authors would like to acknowledge Dr. Elizabeth Omoluabi who led the design of the questionnaire.

Authors Contributions

SG refined the data collection tools, led survey implementation, refined the data analysis scripts, conducted the data analysis, and drafted the manuscript. **AA** refined the data collection tools, participated in survey implementation, and refined the data analysis scripts. **AAS, FS, AW, and GAB** participated in survey implementation and contributed to data analysis. **DA, WAA, FN, LP, and FD** provided technical support. **LG, MK, and JK** developed the data analysis scripts and provided technical support. **CR and GH** contributed to the development of the data collection tools and provided technical support. **ER and TJ** conceived the methodological approach and provided technical support. All authors reviewed and approved the final manuscript.

Funding

That project was funded by the Gates Foundation (INV-060889).

Data Availability

The data supporting the findings of this study shall be provided upon reasonable request.

Declarations

Ethics approval and consent to participate

Ethical approval was obtained from the Research Ethics Review Committee of the Ethiopian Public Health Institute (EPHI/OG/760/25). The study was conducted in accordance with national and international ethical guidelines, including the principles of the Declaration of Helsinki. Written informed consent was obtained from all adult participants. For participants under 18 years of age, written informed consent was obtained from a parent or legal guardian, along with assent from the participant.

Consent for publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

Artificial Intelligence Disclosure

Microsoft Copilot was used solely to assist with language editing, grammar correction, and improving readability. The authors independently reviewed and revised all suggested edits and retained full responsibility for the manuscript's content. The tool was not used to generate scientific content, conduct analyses, interpret results, or draw conclusions.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PathwaysWRAsegmentationreportV2.docx