Journal of Medicine, Nursing & Public Health

Implications for Social Health Insurance Fund (SHIF) Implementation from Determinants of Health Insurance Uptake in Kenya Before and During the COVID-19 Pandemic

Namukoa Isaac Wekesa, Nicholas Mutothya, Ngesa Oscar & Boniface Malenje

ISSN: 2706-6606

Stratford Peer Reviewed Journals and Book Publishing
Journal of Medicine, Nursing & Public Health
Volume 9||Issue 1||Page 107-120||July||2026|
Email: info@stratfordjournals.org ISSN: 2706-6606

Implications for Social Health Insurance Fund (SHIF) Implementation from Determinants of Health Insurance Uptake in Kenya Before and During the COVID-19 Pandemic

***[1]Namukoa Isaac Wekesa, [2]Nicholas Mutothya, [3]Ngesa Oscar & [4]Boniface Malenje**

[1, 2,3]Department of Mathematics, Statistics and Physical Sciences, Taita Taveta University, Voi, Kenya

[4]Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

*Email of corresponding author: wek2nam@gmail.com

How to cite this article: **Wekesa, N. I., Mutothya, N., Ngesa, O. & Malenje, B. (2026). Implications for Social Health Insurance Fund (SHIF) Implementation from Determinants of Health Insurance Uptake in Kenya Before and During the COVID-19 Pandemic. *Journal of Medicine, Nursing & Public Health,* 9(1), 107-120.** https://doi.org/10.53819/81018102t4399

Abstract

Kenya is transitioning from the National Hospital Insurance Fund (NHIF) to the Social Health Insurance Fund (SHIF) to advance Universal Health Coverage (UHC). Effective SHIF implementation requires understanding the factors influencing health insurance uptake and how these changed during the COVID-19 pandemic. This study analyzed nationally representative data from two sources: the Kenya Integrated Household Budget Survey (KIHBS 2015/16), comprising 21,536 households, and the Kenya Demographic and Health Survey (KDHS 2022), with 19,582 households. Both datasets included complete insurance-status information. The analysis used descriptive statistics, logistic regression, Random Forest, Artificial Neural Networks, SHAP-based feature interpretation, and spatial analysis to identify and compare determinants of health insurance uptake. These determinants were examined both before and during the COVID-19 pandemic, clearly establishing the timeframes for comparison. Health insurance coverage was linked with educational attainment, household wealth, digital inclusion, and geographic location in both survey periods. Higher educational attainment and greater household wealth were consistently associated with higher probabilities of insurance coverage. Digital-access indicators, like mobile phone ownership, mobile money registration, mobile banking use, and television ownership, were also strong predictors. During the pandemic, wealth-related disparities grew. Healthcare utilization variables, such as hospitalizations and healthcare spending, became more predictive. County-level effects were among the strongest determinants in both statistical and machine-learning models, indicating persistent geographic inequalities in insurance uptake. Spatial analysis found lower predicted coverage in the Northeastern and some coastal counties compared to Central, Western, and parts of the Rift Valley regions. Health insurance uptake in Kenya is influenced by interconnected socioeconomic, digital, healthcare-utilization, and geographic factors. The findings suggest that for SHIF to succeed, strategies must target socioeconomic vulnerability, digital exclusion, and ongoing regional disparities. This will help achieve more equitable progress toward Universal Health Coverage.

Keywords: *Health Insurance Uptake, SHIF, Universal Health Coverage, Kenya, Machine Learning, Digital Inclusion, Health Financing, COVID-19*

1.0 Introduction

Universal Health Coverage (UHC) remains a central global health priority and is a key target under Sustainable Development Goal 3. UHC aims to ensure that all individuals have access to needed health services without experiencing financial hardship associated with healthcare utilization. Despite substantial progress in expanding healthcare access globally, many low- and middle-income countries (LMICs) continue to face significant challenges related to healthcare financing, inequitable access to services, and high levels of out-of-pocket(OOP) expenditure (World Health Organization, 2023).

In Kenya, health insurance is a critical mechanism for advancing UHC by improving financial protection and facilitating access to healthcare services. However, insurance coverage remains relatively low and unevenly distributed across population groups. Previous studies have consistently demonstrated that socioeconomic status, educational attainment, employment characteristics, geographic location, and healthcare-seeking behaviour influence participation in health insurance schemes (Barasa et al., 2018; Kazungu & Barasa, 2017). Consequently, substantial disparities persist between wealthier and poorer households, urban and rural populations, and different regions of the country.

To accelerate progress toward UHC, Kenya has embarked on major health financing reforms, including the transition from the National Hospital Insurance Fund (NHIF) to the Social Health Insurance Fund (SHIF) under the Social Health Insurance Act, 2023. The SHIF seeks to expand population coverage, strengthen financial protection, and promote equitable access to healthcare services across all population segments. Achieving these objectives requires a comprehensive understanding of the factors associated with health insurance uptake and how these factors vary across socioeconomic, demographic, technological, and geographic contexts (Ministry of Health, 2021; Parliament of Kenya, 2023).

The COVID-19 pandemic introduced additional complexity into health insurance participation patterns. Beyond its direct health effects, the pandemic disrupted household incomes, employment opportunities, healthcare utilization patterns, and financial security, potentially altering the determinants of insurance uptake. Emerging evidence suggests that the pandemic disproportionately affected vulnerable populations and may have amplified existing inequalities in healthcare access and financial protection (Demirgüç-Kunt et al., 2022; World Bank, 2020) . Understanding whether the drivers of insurance coverage changed during this period is therefore important for designing effective and equitable SHIF implementation strategies.

Recent advances in machine learning (ML) offer new opportunities to examine health insurance uptake using large, complex population datasets. Unlike traditional statistical approaches that primarily focus on estimating associations, ML methods can identify complex nonlinear relationships, evaluate predictive performance, and provide insights into the relative importance of different determinants. When combined with conventional regression techniques and spatial analysis, these approaches offer a more comprehensive understanding of the factors influencing insurance participation and the extent to which their importance changes over time (Géron, 2019; Goodfellow et al., 2016; Lundberg & Lee, 2017; Russell & Norvig, 2021).

Using nationally representative data from the Kenya Integrated Household Budget Survey (KIHBS 2015/16) and the Kenya Demographic and Health Survey (KDHS 2022), this study examines the socioeconomic, digital, healthcare-related, and geographic determinants of health insurance uptake before and during the COVID-19 pandemic. The study further draws policy-relevant lessons for SHIF implementation by identifying population groups and geographic areas that may require targeted interventions to improve insurance enrolment and advance progress toward Universal Health Coverage in Kenya.

2.0 Methods

2.1 Study Design and Data Sources

This study employed a comparative cross-sectional design using nationally representative household survey datasets collected before and during the COVID-19 pandemic in Kenya. The pre-pandemic dataset was obtained from the Kenya Integrated Household Budget Survey (KIHBS) 2015/16 conducted by the Kenya National Bureau of Statistics (KNBS), while the during-pandemic dataset was obtained from the Kenya Demographic and Health Survey (KDHS) 2022 implemented by KNBS in collaboration with the Ministry of Health and development partners.

The KIHBS 2015/16 dataset contained information on 21,536 households with valid responses regarding health insurance status, whereas the KDHS 2022 dataset contained 19,582 households that met the same inclusion criteria. Both surveys employed multistage stratified sampling designs and are nationally representative of Kenyan households. The surveys provide comprehensive information on demographic characteristics, socioeconomic conditions, healthcare utilization, digital access, and geographic location, making them suitable for examining determinants of health insurance uptake.

2.2 Study Variables

The outcome variable was health insurance coverage, coded as a binary indicator of whether a household or respondent reported having health insurance (insured = 1; uninsured = 0). Independent variables were selected based on previous literature on health insurance participation and data availability across both surveys. These variables were grouped into five broad domains: Demographic characteristics (age, gender, marital status, household size, and locality); Socioeconomic characteristics (education level, wealth status, and social assistance indicators); Healthcare utilization and need indicators (hospitalization, outpatient care utilization, healthcare expenditure, and pregnancy status); Digital inclusion indicators (mobile phone ownership, mobile money registration, mobile banking use, computer ownership, and television ownership); and Geographic variables (county of residence). To facilitate comparability across datasets, variables were harmonized and recoded into consistent categories prior to analysis.

2.3 Analytical Approach

The analysis was conducted in four stages. First, descriptive statistics were computed to summarize the characteristics of the study populations and compare health insurance coverage patterns between the pre-pandemic and during-pandemic periods. Continuous variables were summarized using means and standard deviations, while categorical variables were summarized using frequencies and percentages.

Second, bivariate analyses were performed to examine associations between study variables and health insurance coverage. Chi-square tests were used for categorical variables, while independent sample t-tests were used for continuous variables. Univariate logistic regression (LR) coefficients were additionally estimated to assess the direction and magnitude of associations.

Third, multivariable LR models were fitted separately for the KIHBS 2015/16 and KDHS 2022 datasets to identify independent determinants of health insurance uptake. Adjusted odds ratios (AORs), 95% confidence intervals (CIs), and p-values were used to quantify associations. A pooled interaction model was subsequently estimated to evaluate whether the relationship between healthcare need indicators and health insurance coverage changed significantly during the COVID-19 period.

Fourth, machine learning (ML) models were developed to assess predictive performance and identify influential determinants of insurance uptake. Three supervised classification algorithms were evaluated: LR, Random Forest (RF), and Artificial Neural Networks (ANN). Each dataset was partitioned into training (70%) and testing (30%) subsets. Model performance was evaluated using accuracy, area under the receiver operating characteristic curve (AUC-ROC), F1-score, precision, recall, and Brier score.

To improve model interpretability, RF feature importance measures and Shapley Additive Explanations (SHAP) were computed to quantify the relative contribution of predictor variables. Predictors were further grouped into conceptual constructs to facilitate substantive interpretation of feature importance patterns across survey periods.

2.4 Spatial Analysis

To examine geographic variation in health insurance uptake, county-level predicted probabilities generated from the RF models were aggregated and mapped using Geographic Information System (GIS) techniques. Choropleth maps were produced to visualize spatial disparities in predicted insurance coverage before and during the COVID-19 pandemic. The spatial analysis enabled the identification of geographic clusters with persistently high or low predicted insurance uptake.

2.5 Ethical Considerations

The study utilized publicly available, anonymized secondary data obtained from the Kenya National Bureau of Statistics and the Demographic and Health Surveys (DHS) Program. No personally identifiable information was accessed. Ethical approval for the original surveys was obtained by the respective implementing agencies. Permission to access and analyze the datasets was granted by the relevant data custodians. The present analysis involved secondary use of anonymized data and therefore posed minimal risk to participants.

3.0 Findings and Discussion

3.1 Socioeconomic Determinants of Insurance Uptake

The multivariable analyses (Table 1) revealed that socioeconomic characteristics were among the strongest determinants of health insurance uptake in both the pre-pandemic and pandemic periods. Educational attainment, wealth status, household size, marital status, and locality consistently demonstrated significant associations with insurance coverage after adjustment for other demographic, healthcare-related, digital, and geographic factors.

https://doi.org/10.53819/81018102t4399

Stratford Peer Reviewed Journals and Book Publishing
Journal of Medicine, Nursing & Public Health
Volume 9||Issue 1||Page 107-120||July||2026|
Email: info@stratfordjournals.org ISSN: 2706-6606

Table 1: Adjusted odds ratios for determinants of health insurance coverage in Kenya

		Pre-pandemic		During-pandemic	
Predictor	**Category (vs Ref)**	**Adjusted odds ratio (95% CI)**	**p-value**	**Adjusted odds ratio (95% CI)**	**p-value**
Age (per year increase)	—	1.01 (1.01–1.02)	<0.001	1.02 (1.02–1.02)	<0.001
Computer ownership	Yes (vs No)	1.81 (1.52–2.14)	<0.001	1.43 (1.23–1.66)	<0.001
County	Include	Include	<0.001	Include	0.002
Gender	Male (vs Female)	1.00 (0.85–1.17)	0.970	0.95 (0.87–1.04)	0.274
Highest education level	Post-secondary (vs No/Pre-primary)	8.08 (5.58–11.71)	<0.001	7.11 (5.73–8.81)	<0.001
	Primary (vs No/Pre-primary)	1.58 (1.10–2.26)	0.013	1.69 (1.45–1.97)	<0.001
	Secondary (vs No/Pre-primary)	3.30 (2.30–4.75)	<0.001	3.49 (2.95–4.12)	<0.001
	University (vs No/Pre-primary)	11.37 (7.61–16.99)	<0.001	11.16 (9.20–13.54)	<0.001
Hospitalization in last 12 months	Yes (vs No)	1.26 (0.87–1.81)	0.219	1.72 (1.46–2.02)	<0.001
Household size (per additional member)	—	0.90 (0.86–0.95)	<0.001	0.98 (0.96–1.00)	0.034
Inpatient spending (log[KES+1])	—	1.06 (1.01–1.11)	0.012	1.03 (1.01–1.04)	<0.001
Marital status	Never married (vs Married/Living together)	0.40 (0.34–0.47)	<0.001	0.57 (0.50–0.65)	<0.001
	Separated/Divorced (vs Married/Living together)	0.38 (0.31–0.46)	<0.001	0.57 (0.42–0.76)	<0.001
	Widowed (vs Married/Living together)	0.50 (0.42–0.59)	<0.001	0.83 (0.72–0.95)	0.009
Outpatient spending (log[KES+1])	—	1.04 (1.01–1.06)	0.003	1.03 (1.02–1.04)	<0.001
Owns mobile phone	Yes (vs No)	1.98 (1.56–2.53)	<0.001	1.29 (1.04–1.60)	0.023
Place of residence	Urban (vs Rural)	1.14 (1.04–1.25)	0.004	0.66 (0.60–0.74)	<0.001
Pregnancy in household	Yes (vs No)	1.47 (1.24–1.75)	<0.001	0.95 (0.79–1.14)	0.592
Received NGO social assistance	Yes (vs No)	0.92 (0.73–1.16)	0.468	1.55 (1.18–2.04)	0.002
Received assistance from friends/relatives	Yes (vs No)	1.46 (0.64–3.33)	0.370	1.10 (0.94–1.29)	0.255
Received county social assistance	Yes (vs No)	1.45 (0.97–2.16)	0.067	1.32 (1.10–1.57)	0.002
Received national social assistance	Yes (vs No)	0.93 (0.84–1.03)	0.158	1.13 (0.99–1.30)	0.076
Received religious social assistance	Yes (vs No)	0.81 (0.56–1.16)	0.242	0.91 (0.68–1.22)	0.529
Registered for mobile money	Yes (vs No)	1.92 (1.54–2.40)	<0.001	1.88 (1.44–2.46)	<0.001
Television ownership	Yes (vs No)	2.40 (2.19–2.64)	<0.001	1.40 (1.27–1.54)	<0.001
Uses mobile banking	Yes (vs No)	2.24 (2.03–2.47)	<0.001	0.83 (0.67–1.04)	0.101
Wealth quintile	Poorer (vs Middle)	0.73 (0.64–0.84)	<0.001	0.60 (0.53–0.68)	<0.001
	Poorest (vs Middle)	0.96 (0.80–1.16)	0.670	0.32 (0.28–0.38)	<0.001
	Richer (vs Middle)	0.99 (0.85–1.14)	0.844	1.67 (1.49–1.87)	<0.001
	Richest (vs Middle)	1.20 (0.97–1.48)	0.097	3.38 (2.89–3.96)	<0.001

NOTE: Educational attainment emerged as one of the strongest and most consistent determinants of insurance coverage across both survey periods. University-level education was associated with substantially higher odds of coverage in both the pre-pandemic (AOR = 11.37, 95% CI: 7.61–16.99) and during-pandemic (AOR = 11.16, 95% CI: 9.20–13.54) models.

Educational attainment emerged as one of the most influential predictors across both survey periods. Individuals with secondary, postsecondary, and university education had substantially higher odds of health insurance coverage than those with little or no formal education. The positive relationship between education and insurance participation remained consistent before and during the COVID-19 pandemic, indicating that educational attainment continues to play a critical role in facilitating access to health insurance. Higher education levels may enhance awareness of insurance benefits, improve understanding of enrolment processes, and increase access to formal employment opportunities that are often linked to insurance coverage.

Wealth status was also strongly associated with health insurance uptake. Households in higher wealth quintiles consistently had a greater likelihood of having insurance coverage than those in lower wealth quintiles. The observed wealth gradient became more pronounced during the pandemic period, suggesting that economic vulnerability may have become an increasingly

https://doi.org/10.53819/81018102t4399

important barrier to insurance participation. This finding indicates that financial constraints remain a significant obstacle to achieving equitable health insurance coverage in Kenya.

Household characteristics also influenced insurance participation. Smaller households generally exhibited higher probabilities of insurance coverage than larger households, while married individuals were more likely to be insured than respondents who were never married, separated, divorced, or widowed. Age similarly showed a positive association with coverage, with older individuals exhibiting higher levels of insurance participation than younger individuals.

Locality remained an important determinant across both periods. Urban residents consistently exhibited higher insurance coverage than rural residents, although the magnitude of this difference varied between survey periods. The findings suggest that disparities in economic opportunities, access to information, healthcare infrastructure, and insurance services may continue to contribute to differences in coverage between urban and rural populations.

Overall, the results indicate that health insurance uptake in Kenya remains strongly patterned by socioeconomic advantage. Educational attainment, wealth status, household structure, and locality collectively influence participation in health insurance schemes and continue to shape inequalities in coverage before and during the COVID-19 pandemic. These findings have important implications for SHIF implementation, particularly regarding improving enrolment among economically disadvantaged and socially vulnerable populations.

3.2 Digital Inclusion and Insurance Uptake

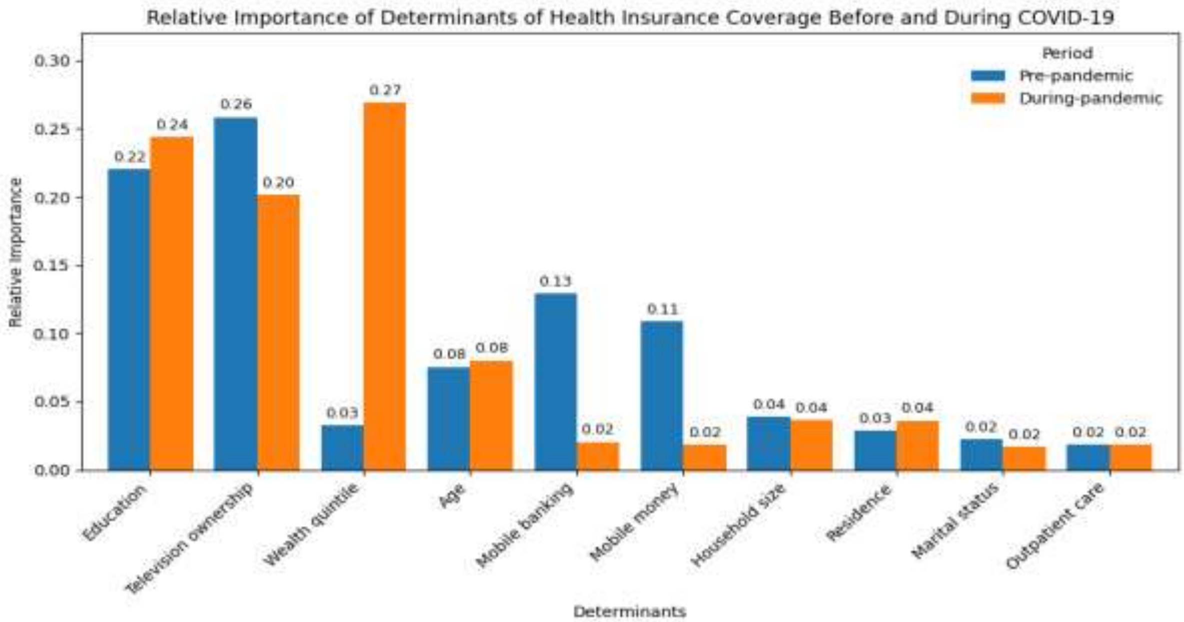

Figure 1: Relative Importance of Major Determinants of Health Insurance Uptake

NOTE: Prior to the pandemic, mobile banking and mobile money were among the most influential predictors of insurance uptake, whereas during the pandemic the relative importance of wealth increased substantially while the influence of digital-financial variables declined.

Digital inclusion emerged as an important determinant of health insurance uptake across both survey periods. Indicators of access to digital technologies and digital financial services were consistently associated with higher probabilities of insurance coverage in both the regression and machine learning analyses. Mobile phone ownership, mobile money registration, mobile banking usage, television ownership, and computer ownership were among the variables that demonstrated significant associations with insurance participation.

The multivariable analyses indicated that individuals with access to digital technologies generally exhibited higher levels of insurance coverage than those without such access. Mobile phone ownership and mobile banking usage were particularly important predictors before the pandemic, reflecting the growing role of digital financial systems in facilitating access to information, financial transactions, and formal service delivery. Television ownership and computer ownership also demonstrated positive associations with insurance coverage, suggesting that access to information and communication technologies may influence awareness and understanding of health insurance products.

The ML analyses (Figure 1) further reinforced the importance of digital inclusion. RF feature-importance rankings and SHAP analyses consistently identified several digital-access indicators among the most influential predictors of insurance coverage. Prior to the pandemic, mobile banking, mobile phone ownership, and television ownership ranked among the strongest predictors in the model. During the pandemic, although the relative importance of some digital indicators declined, they continued to contribute meaningfully to predicting insurance uptake.

The findings suggest that digital inclusion may influence health insurance participation through several pathways. Access to digital technologies can increase exposure to information about insurance products, facilitate communication with service providers, and enable access to mobile financial services that simplify premium payment and enrolment processes. In addition, digital-access indicators often reflect broader socioeconomic advantages that may facilitate participation in formal health financing arrangements (Demirgüç-Kunt et al., 2022).

From a SHIF implementation perspective, these findings highlight the potential value of leveraging Kenya's well-developed digital ecosystem to expand insurance coverage. Mobile-based enrolment systems, digital payment platforms, SMS communication channels, and other technology-enabled services may enhance accessibility and convenience for beneficiaries. At the same time, the results underscore the importance of addressing digital exclusion among populations with limited access to technology, particularly in underserved and economically disadvantaged communities. Overall, the findings demonstrate that digital inclusion is closely linked to health insurance uptake in Kenya and is an important consideration for efforts to expand coverage and improve participation under SHIF.

3.3 Healthcare Need and Utilization

Unlike Table 1, which presents the fully adjusted multivariable models incorporating socioeconomic, healthcare, digital, and geographic variables, Table 2 reports estimates from models developed specifically to evaluate healthcare need indicators and their interaction with the COVID-19 period. Consequently, adjusted odds ratios differ because the underlying model specifications address different analytical objectives.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Medicine, Nursing & Public Health
Volume 9||Issue 1||Page 107-120||July||2026|
Email: info@stratfordjournals.org ISSN: 2706-6606

Table 2: Multivariable Logistic Regression of Factors Associated with Health Insurance Coverage

	Pre-Pandemic		During-Pandemic		Pooled interaction model	
Variable	**AOR (95% CI)**	**p-value**	**AOR (95% CI)**	**p-value**	**AOR (95% CI)**	**p-value**
Healthcare need indicator	1.00 (0.90–1.11)	0.985	0.89 (0.81–0.97)	0.011	0.96 (0.88–1.05)	0.389
Male	0.82 (0.74–0.91)	<0.001	0.95 (0.87–1.04)	0.266	0.91 (0.85–0.97)	0.004
Urban residence	1.01 (0.93–1.10)	0.764	0.63 (0.57–0.70)	<0.001	1.03 (0.97–1.10)	0.296
Never married	0.43 (0.37–0.51)	<0.001	0.57 (0.50–0.66)	<0.001	0.55 (0.50–0.61)	<0.001
Separated/Divorced	0.42 (0.35–0.51)	<0.001	0.58 (0.44–0.77)	<0.001	0.52 (0.44–0.60)	<0.001
Widowed	0.49 (0.42–0.58)	<0.001	0.79 (0.69–0.90)	<0.001	0.65 (0.59–0.73)	<0.001
Primary education	2.56 (1.82–3.62)	<0.001	1.54 (1.34–1.76)	<0.001	1.58 (1.40–1.77)	<0.001
Secondary education	5.32 (3.76–7.54)	<0.001	3.10 (2.66–3.61)	<0.001	3.46 (3.05–3.93)	<0.001
Post-secondary education	11.23 (7.87–16.02)	<0.001	6.21 (5.08–7.60)	<0.001	7.44 (6.45–8.59)	<0.001
University education	14.43 (9.83–21.17)	<0.001	9.71 (8.12–11.61)	<0.001	10.95 (9.42–12.73)	<0.001
Poorer wealth quintile	0.69 (0.60–0.78)	<0.001	0.57 (0.51–0.65)	<0.001	0.78 (0.71–0.84)	<0.001
Poorest wealth quintile	0.80 (0.67–0.96)	0.019	0.33 (0.28–0.38)	<0.001	0.80 (0.73–0.87)	<0.001
Richer wealth quintile	1.03 (0.89–1.18)	0.729	1.64 (1.47–1.83)	<0.001	1.00 (0.93–1.08)	0.941
Richest wealth quintile	1.27 (1.06–1.58)	0.035	3.30 (2.84–3.84)	<0.001	1.20 (1.10–1.31)	<0.001
Hospitalized	1.31 (0.92–1.86)	0.142	1.57 (1.34–1.84)	<0.001	1.50 (1.31–1.72)	<0.001
Outpatient care	1.17 (1.03–1.32)	0.013	1.09 (0.99–1.20)	0.094	1.08 (1.01–1.17)	0.033
Mobile money registration	2.36 (1.91–2.93)	<0.001	1.56 (1.21–2.02)	<0.001	1.90 (1.63–2.21)	<0.001
Mobile banking usage	2.13 (1.94–2.34)	<0.001	0.84 (0.68–1.04)	0.101	1.85 (1.70–2.01)	<0.001
Pandemic period	—	—	—	—	0.86 (0.78–0.95)	0.003
Healthcare need indicator × Pandemic	—	—	—	—	0.97 (0.87–1.08)	0.605

Multivariable LR analysis of factors associated with health insurance coverage in before and during the COVID-19 pandemic.

Healthcare need and healthcare utilization variables demonstrated important but comparatively modest associations with health insurance uptake (Table 2). Across the regression and ML analyses, indicators such as hospitalization, outpatient-care utilization, healthcare expenditure, and pregnancy status were associated with insurance coverage, particularly during the pandemic period. However, their overall contribution to predicting insurance uptake remained lower than that of socioeconomic and digital-inclusion factors.

The pooled interaction model was used to assess whether the relationship between healthcare need and health insurance coverage changed significantly during the COVID-19 pandemic.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Medicine, Nursing & Public Health
Volume 9||Issue 1||Page 107-120||July||2026|
Email: info@stratfordjournals.org ISSN: 2706-6606

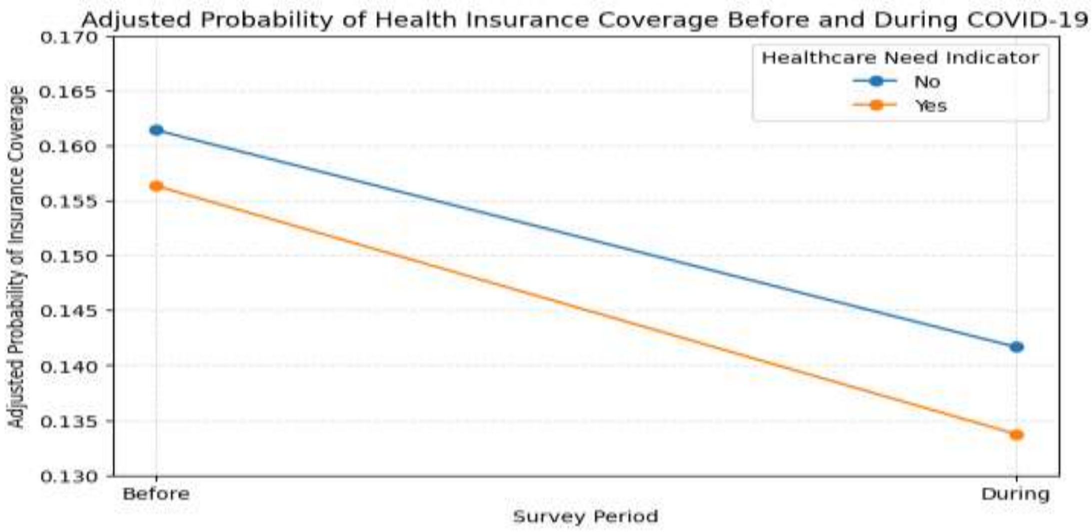

Figure 2: Adjusted Probability of Health Insurance Coverage

NOTE: Adjusted predicted probabilities were estimated from the multivariable LR model fitted to the KIHBS 2015/16 and KDHS 2022 datasets. Higher values indicate a greater predicted probability of coverage. The results compare households with and without a healthcare need indicator in Kenya.

The results indicate that adjusted probabilities of insurance coverage declined during the pandemic period for both groups. Prior to the pandemic, households without healthcare needs exhibited an adjusted probability of insurance coverage of approximately 16.2%, compared with 15.6% among households with healthcare needs. During the pandemic, these probabilities declined to approximately 14.2% and 13.4%, respectively.

Although households with healthcare needs had slightly lower predicted probabilities of insurance coverage than those without, the magnitude of the difference remained relatively small. Furthermore, the interaction between healthcare need and pandemic period was not statistically significant in the pooled regression model, indicating that the association between healthcare need and insurance coverage did not change significantly between the pre-pandemic and pandemic periods.

Additional analyses of healthcare utilization variables revealed that hospitalization and healthcare expenditures became more influential during the pandemic period. In the machine learning models, outpatient and inpatient expenditure variables increased substantially in relative importance, while hospitalization emerged as a stronger predictor of insurance coverage than in the pre-pandemic models. These findings suggest that direct interaction with the healthcare system became increasingly important during the pandemic, likely reflecting heightened health risks, increased healthcare utilization, and greater exposure to healthcare costs.

Overall, the results indicate that healthcare needs and utilization were associated with health insurance coverage, particularly during the pandemic. However, their influence remained secondary to socioeconomic position, educational attainment, and digital inclusion, which consistently emerged as the dominant determinants of insurance uptake across analytical approaches.

Stratford Peer Reviewed Journals and Book Publishing
Journal of Medicine, Nursing & Public Health
Volume 9||Issue 1||Page 107-120||July||2026|
Email: info@stratfordjournals.org ISSN: 2706-6606

3.4 Geographic Inequalities in Insurance Coverage

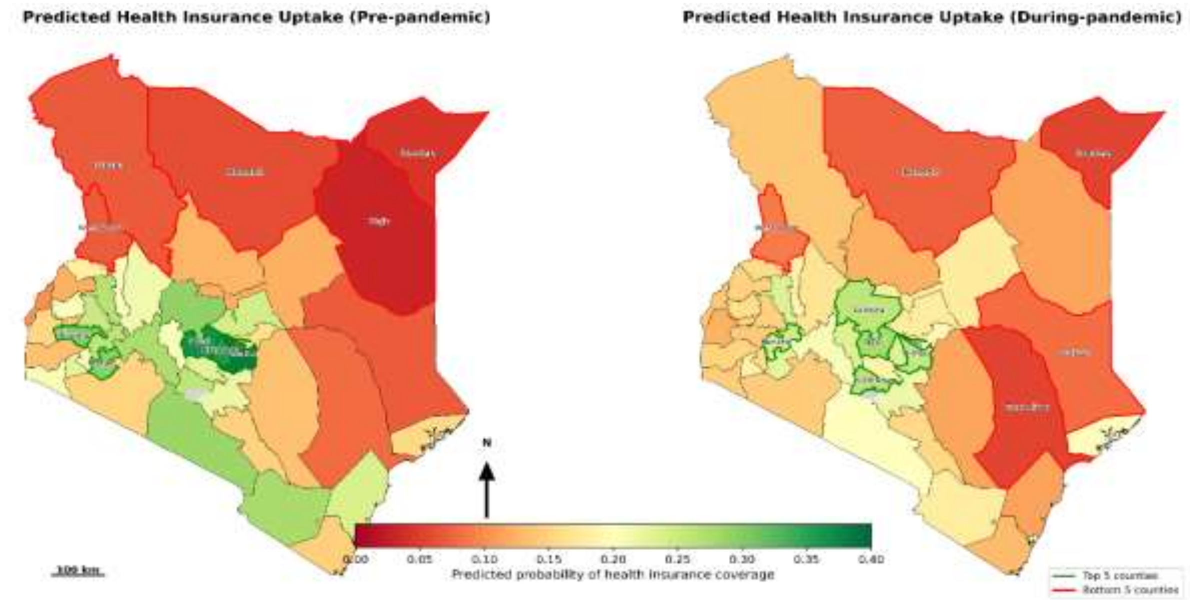

Figure 3: Predicted County-Level Health Insurance Uptake Before and During the COVID-19 Pandemic in Kenya

NOTE: Predicted probabilities were derived from Random Forest models fitted to the pre-pandemic (KIHBS 2015/16) and during-pandemic (KDHS 2022) datasets. Darker green shading indicates higher predicted uptake, while darker red shading indicates lower predicted uptake.

Geographic variation in health insurance uptake remained evident across Kenya in both the pre-pandemic and pandemic periods. Figure 3 presents county-level predicted probabilities of health insurance coverage derived from the Random Forest models, illustrating the spatial distribution of insurance uptake before and during the COVID-19 pandemic.

The maps reveal substantial geographic heterogeneity in predicted insurance coverage across counties. Prior to the pandemic, counties located in the Central, Western, and parts of the Rift Valley regions generally exhibited the highest predicted probabilities of insurance uptake. Several counties, including Kiambu, Nyeri, and Kericho, as well as neighbouring counties in the central highlands, consistently appeared in the highest-probability categories. In contrast, counties in the Northeastern region, particularly Wajir, Mandera, and Garissa, exhibited among the lowest predicted coverage probabilities. Lower uptake was also observed in selected counties within the Coast region.

A broadly similar spatial pattern persisted during the pandemic period. Counties in Central Kenya and parts of the Rift Valley continued to exhibit relatively high predicted probabilities of insurance coverage, while counties in the Northeastern region remained concentrated within the lowest probability categories. Although some changes in county rankings were observed, the overall geographic distribution of higher- and lower-coverage counties remained largely stable across survey periods.

These spatial patterns align with the county effects observed in the multivariable regression models, in which county of residence remained a significant predictor of insurance coverage after adjustment for demographic, socioeconomic, healthcare utilization, and digital inclusion factors. The persistence of county-level variation suggests that contextual influences operating beyond individual household characteristics continue to shape insurance participation.

The observed geographic disparities may reflect differences in economic development, healthcare infrastructure, availability of health facilities, population density, Labor-market structure, administrative capacity, and access to insurance enrolment mechanisms across counties. Counties with higher predicted uptake tend to coincide with regions characterized by greater economic activity and stronger health-service availability, whereas counties with lower uptake are concentrated in historically underserved and geographically remote regions.

Overall, the findings indicate that health insurance uptake in Kenya is not distributed uniformly across space. Despite changes in socioeconomic conditions and healthcare utilization patterns during the COVID-19 pandemic, substantial geographic inequalities persisted. The stability of these spatial disparities across survey periods suggests that county-level contextual factors remain important determinants of insurance participation and should be considered in efforts to expand equitable coverage under SHIF.

3.5 Implications for SHIF Implementation

The findings of this study provide important insights for the implementation of Kenya's SHIF and ongoing efforts to advance UHC. Across descriptive, regression, ML, and spatial analyses, health insurance uptake was associated with educational attainment, wealth status, digital inclusion, healthcare utilization, and geographic location. These patterns suggest that expanding insurance coverage will require targeted strategies that address both financial and non-financial barriers to enrolment.

Educational attainment emerged as one of the strongest determinants of insurance coverage across all analytical approaches. Individuals with secondary, post-secondary, and university education were substantially more likely to possess health insurance than those with limited formal education. This finding suggests that awareness, understanding of insurance benefits, and ability to navigate enrolment processes remain important factors influencing participation. Effective SHIF implementation may therefore require communication and enrolment approaches tailored to populations with lower educational attainment, particularly in rural and underserved areas.

The strong and persistent influence of wealth-related variables indicates that economic barriers continue to play a major role in insurance participation. Although SHIF aims to improve equity in health financing, the findings suggest that economically vulnerable households remain at risk of exclusion. The increasing importance of wealth-related indicators during the pandemic period further highlights the sensitivity of insurance participation to household financial circumstances. These results underscore the importance of mechanisms that support enrolment among low-income and financially vulnerable populations.

Digital inclusion also emerged as an important dimension of insurance uptake. Mobile phone ownership, mobile money registration, mobile banking usage, and access to information technologies were consistently associated with insurance coverage. The findings suggest that Kenya's digital ecosystem provides an important platform for expanding enrolment, premium collection, beneficiary communication, and access to services. However, they also indicate that populations with limited digital access may face additional barriers to participation, emphasizing the need for complementary non-digital enrolment pathways.

The analyses further demonstrated that healthcare utilization variables, including hospitalization and healthcare expenditure, became more influential during the pandemic period. Although healthcare need alone was not a significant predictor in the pooled interaction model, increased utilization of healthcare services was associated with higher probabilities of insurance coverage. These findings suggest that engagement with the healthcare system may create opportunities for insurance enrolment and retention, particularly among households experiencing greater exposure to healthcare costs.

Geographic disparities remained evident throughout the analyses. County-level effects remained significant after adjustment for individual and household characteristics, while spatial modeling identified persistent clusters of high and low predicted insurance uptake. The concentration of lower coverage in parts of the North Eastern and coastal regions indicates that structural and contextual factors continue to influence participation in health insurance. These findings suggest that nationally uniform implementation approaches may be insufficient and that county-specific strategies may be necessary to address local barriers to enrolment.

Taken together, the results indicate that successful SHIF implementation will depend on addressing multiple and interconnected determinants of insurance participation. Educational inequalities, economic vulnerability, digital exclusion, healthcare utilization patterns, and geographic disparities all contribute to variation in insurance coverage. The consistency of these findings across statistical, machine learning, and spatial analytical approaches strengthens confidence that these factors represent important considerations for achieving equitable and sustainable expansion of health insurance coverage in Kenya.

4.0 Conclusion

This study examined the determinants of health insurance uptake in Kenya before and during the COVID-19 pandemic using nationally representative data from the KIHBS 2015/16 and KDHS 2022 surveys. By integrating logistic regression, machine learning techniques, and spatial analysis, the study provided a comprehensive assessment of how demographic, socioeconomic, digital, healthcare-related, and geographic factors influenced insurance participation across the two periods.

The findings demonstrate that health insurance uptake in Kenya is shaped by a multidimensional set of determinants. Educational attainment consistently emerged as one of the strongest predictors of insurance coverage, with individuals possessing higher levels of education exhibiting a substantially greater likelihood of enrolment. Wealth status also remained a dominant determinant, with socioeconomic inequalities becoming more pronounced during the pandemic period. Households in higher wealth categories consistently demonstrated higher coverage levels than those in lower wealth groups.

Digital inclusion constituted another important dimension of insurance participation. Mobile phone ownership, mobile money registration, mobile banking usage, and access to information technologies were positively associated with health insurance coverage and featured prominently in both regression and ML analyses. These findings highlight the growing importance of digital connectivity in facilitating access to health insurance and related financial services.

Healthcare utilization variables, including hospitalization and healthcare expenditure, became increasingly important during the pandemic period, suggesting heightened sensitivity to healthcare needs and financial risk exposure. However, healthcare-need indicators alone were not sufficient to explain insurance participation, indicating that enabling socioeconomic resources remain more influential determinants of coverage than healthcare need itself.

The study also identified persistent geographic inequalities in insurance uptake. County-level effects remained significant after adjustment for other determinants, and spatial analyses revealed stable clusters of higher and lower predicted coverage across regions. Counties in Central, Western, and parts of the Rift Valley generally exhibited higher predicted uptake, whereas counties in the North Eastern region and selected coastal counties consistently demonstrated lower coverage probabilities.

Across analytical approaches, the findings indicate that the COVID-19 pandemic did not fundamentally alter the core determinants of health insurance uptake. Rather, it intensified the influence of wealth, healthcare expenditure, and financial vulnerability while reinforcing existing geographic disparities. The convergence of evidence from regression, machine learning, and spatial analyses strengthens confidence in the robustness of these findings and demonstrates the value of combining traditional statistical methods with modern machine learning approaches to understand health insurance participation. Overall, the study concludes that achieving equitable expansion of health insurance coverage in Kenya will require attention not only to healthcare needs but also to broader socioeconomic, digital, and geographic inequalities that continue to shape participation in health insurance schemes. These findings provide evidence that can support the implementation of SHIF and contribute to ongoing efforts to advance Universal Health Coverage in Kenya.

5.0 Recommendations

Based on the findings of this study, the following recommendations can support the effective implementation of the SHIF and broader efforts to advance UHC in Kenya.

Educational attainment and household wealth consistently emerged as the strongest determinants of health insurance uptake. SHIF implementation should therefore prioritize enrolment strategies targeting low-income households, individuals with limited formal education, and other socially vulnerable populations. Community-based outreach programs, simplified enrolment procedures, and targeted awareness campaigns may help reduce barriers to participation among underserved groups.

Digital inclusion indicators, including mobile phone ownership, mobile money registration, and mobile banking usage, were strongly associated with insurance coverage. SHIF should continue to leverage Kenya's digital ecosystem to facilitate enrolment, premium contributions, beneficiary communication, and service access. However, alternative non-digital enrolment channels should remain available to ensure that populations with limited digital access are not excluded from coverage.

Healthcare utilization variables became more important during the pandemic period, suggesting that contact with the healthcare system creates opportunities for insurance enrolment. Health facilities should therefore be strengthened as enrolment and information points where eligible individuals can receive guidance on registration, benefits, and contribution requirements.

Persistent geographic disparities indicate that a uniform national implementation approach may not adequately address local barriers to insurance participation. Counties with consistently low predicted coverage, particularly in parts of the North Eastern and coastal regions, may require tailored enrolment approaches that consider local socioeconomic conditions, service accessibility, and population characteristics.

The study demonstrates the value of combining conventional statistical methods, machine learning techniques, and spatial analysis to identify determinants of insurance uptake. Policymakers should strengthen the use of data analytics to monitor enrolment trends, identify underserved populations, and evaluate the effectiveness of SHIF implementation strategies across regions and population groups.

REFERENCES

Barasa, E., Rogo, K., Mwaura, N., & Chuma, J. (2018). Kenya National Hospital Insurance Fund Reforms: Implications and Lessons for Universal Health Coverage. *Health Systems & Reform*, *4*(4), 346–361. https://doi.org/10.1080/23288604.2018.1513267

Demirgüç-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2022). *The Global Findex Database 2021: Financial Inclusion, Digital Payments, and Resilience in the Age of COVID-19*. The World Bank. https://doi.org/10.1596/978-1-4648-1897-4

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (Second edition). O'Reilly.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. The MIT Press.

Kazungu, J. S., & Barasa, E. W. (2017). Examining levels, distribution and correlates of health insurance coverage in Kenya. *Tropical Medicine & International Health*, *22*(9), 1175–1185. https://doi.org/10.1111/tmi.12912

Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1705.07874

Ministry of Health. (2021). *Kenya Universal Health Coverage Policy 2020-2030*. http://guidelines.health.go.ke:8000/media/Kenya_Universal_Health_Coverage_Policy_2020__2030.pdf

Parliament of Kenya. (2023). *Social Health Insurance Act*. Parliament of Kenya. https://new.kenyalaw.org/akn/ke/act/2023/16/eng@2023-11-24

Russell, S. J., & Norvig, P. (with Chang, M., Devlin, J., Dragan, A., Forsyth, D., Goodfellow, I., Malik, J., Mansinghka, V., Pearl, J., & Wooldridge, M. J.). (2021). *Artificial intelligence: A modern approach* (Fourth Edition). Pearson.

World Bank. (2020). *Kenya Economic Update: Navigating the Pandemic*. World Bank. https://ourworldindata.org/covid-vaccinations?utm_source=chatgpt.com

World Health Organization. (2023). *Tracking Universal Health Coverage: 2023 Global Monitoring Report* (1st ed). World Health Organization.