TUPED353

Mapping the Future: Leveraging Lasso Regression and GWR to Guide HIV Testing and Prevention Strategies in Tanzania

Amobi Onovo[1,2], Joseph d. Nondi[3], Raymond I. Majengo[3], Emmanuel Mtui[4], Rosemary Mrina[1,2], Eniko Akom[1,2], Patricia Agaba[1,2]

[1]U.S Military HIV Research Program, CIDR, Walter Reed Army Institute of Research, Silver Spring, MD, USA
[2]The Henry M. Jackson Foundation for the Advancement of Military Medicine Inc., Bethesda, MD, USA
[3]U.S Military HIV Research Program, Walter Reed Army Institute of Research, U.S Army Medical Research Directorate-Africa/Tanzania
[4]HJF Medical Research International, Tanzania

BACKGROUND

- A comprehensive understanding of the local sociodemographic, cultural, biological, and behavioral factors influencing HIV transmission in Tanzania is crucial for developing effective public health interventions aimed at reducing the burden of HIV.
- Previous studies have often adopted a generalized approach to analyzing risk factors influencing HIV acquisition, potentially overlooking regional variations and unique risk factors specific to different communities. Our study employs a combination of techniques to highlight these regional variations and community-specific risk factors.

MAIN OBJECTIVE

- To identify key regional drivers of HIV acquisition in Tanzania's culturally diverse regions, using advanced statistical methods, to inform targeted public health interventions.

METHODS

- A nuanced dual analytical design was employed across Tanzania's thirty-one regions.
- Lasso Regression, implemented in three distinct models—Cross-Validation (CV) for optimal regularization, Minimum Bayesian Information Criterion (minBIC) for balanced model complexity, and Adaptive Lasso for refined variable weighting—guided the selection of key variables.
- Concurrently, Geographical Weighted Regression (GWR) examined the spatial dynamics in these variables' relationship with HIV acquisition.
- We analyzed comprehensive data of persons above 15 years from the 2022-2023 Tanzania HIV Impact Survey and the 2022 Tanzania Demographic and Health Survey.
- The dataset, comprising 49 explanatory variables, was divided into training and testing groups for Lasso model selection and validation, respectively

RESULTS

- Both CV and minBIC Lasso models consistently selected 4 key variables as most influential in determining HIV acquisition.
- The variables are prioritized based on the absolute values of their coefficients. Those with the largest magnitudes are listed first, indicating their relative influence, and are ranked by importance in the model.
- These were, in order of ranking: men's alcohol consumption (14.8%, IQR: 9.7%-24.5%), women's primary education completion (44.9%, IQR: 38.3%-48.7%), age at first sexual intercourse among women (16.9%, IQR: 16.6%-18.1%), and women tested for HIV and received their test results (35.5%, IQR: 30.4%-40.4%).
- Spatial analysis revealed distinct regional patterns: men's alcohol consumption was significantly related to HIV acquisition in the northeast, women's completion of primary education in the southeast, and the awareness of HIV test results among women in the southwest region. Multicollinearity led to the exclusion of age at first sexual intercourse from the GWR model.

Comparison of Variable Selection for CV, minBIC and Adaptive Lasso

Variables	CV	minBIC	Adaptive
Percentage of men aged 15–49 who have consumed at least one alcoholic drink in the last 1 month	X	X	X
Percent distribution of women aged 15–49 who completed primary education	X	X	X
Median age at first sexual intercourse among women aged 25–49	X	X	
Percentage of women aged 15-49 who were tested for HIV and received their test results	X	X	X

Validation of Lasso Regression Models

Lasso models	Sample	MSE	R-squared
CV	Training dataset	0.1124	0.840
	Testing dataset	0.1799	0.832
minBIC	Training dataset	0.1124	0.840
	Testing dataset	0.1797	0.831
Adaptive	Training dataset	0.1457	0.793
	Testing dataset	0.2316	0.780

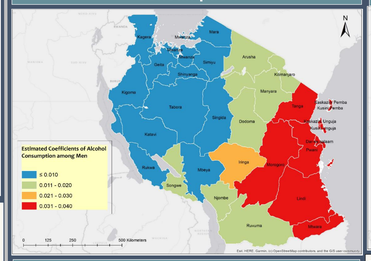
GWR Model Performance

78%

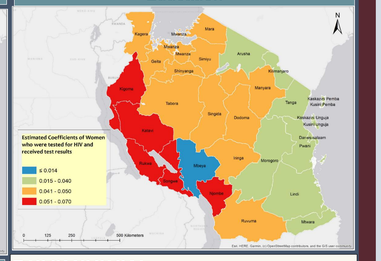
The GWR model demonstrated strong fit, with an adjusted R-squared of 0.784, indicating its reliability in explaining spatial trends in HIV acquisition.

Utilizing GWR to Model Spatial Relationships Between Key Variables and HIV Acquisition

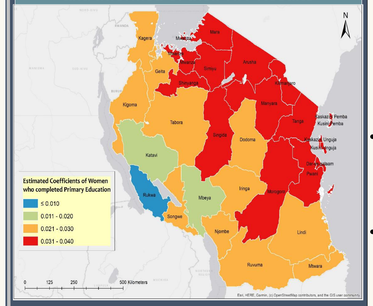
HIV Prevalence and Men's Alcohol Consumption



HIV prevalence and Women's Primary Education



HIV Prevalence and Women tested for HIV and Informed of their HIV Results



DISCUSSION

- This study highlights the need for regionally-specific HIV prevention strategies. Areas with high male alcohol consumption (northeast) could benefit from programs addressing alcohol use.
- Conversely, areas with low female education rates (southeast) might see success with interventions focused on female empowerment and education.
- The study also found that wider access to HIV testing for women and clear communication of results played a positive role in reducing acquisition. rates.

CONCLUSION

This study's insights into spatial dynamics of HIV acquisition provide a strategic foundation for implementing targeted, data-driven HIV prevention and management programs, particularly in high-impact areas identified through our analysis. Targeted HIV case finding efforts should prioritize those at higher risk of acquisition in each region.

This work has been supported by the President's Emergency Plan for AIDS Relief (PEPFAR) through the U.S. Department of Defense under the cooperative agreement (W81XWH-18-2-0040) between the Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., and the U.S. Department of Defense (DOD).

Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the author and should not be construed as official, or as reflecting true views of the Department of the Army, the Department of Defense or HJF.