Journal of East-West Thought
ISSN: *2168-2259 (online)*
(*https://jetjournal.us/*)
Volume 15, Issue 1, Jan-March – 2025 Special Issue 2
JOURNAL OF EAST-WEST THOUGHT
Impact Factor: 7.665, Peer Reviewed and UGC CARE I

SPATIAL MODELING OF THE SOCIAL HEALTH DETERMINANTS IMPACT ON THE EPIDEMIOLOGY OF DISEASES IN LOW- MIDDLE- AND HIGH-INCOME SETTINGS IN NORTH MAHARASHTRA

Parag Vikas Chaudhari, S. C. Advitote, Department of Geography, C. B. Khedgei's Basweshwar Science, Raja Vijaysinh Commerce, Raja Jaysinh Arts College, Akkalkot, Solapur

Abstract

In this study, Researchers examine the spatial distribution of social determinants of health (SDOH) and how they relate to disease epidemiology in low-, middle- and high-income settings across North Maharashtra. The research uses advanced spatial modeling techniques to demonstrate strong correlations between socioeconomic factors, including education, income, and healthcare access, and health outcomes. A spatial epidemiology theoretical framework employs SDOH domains to identify geographic variation in chronic and infectious disease prevalence. The economic gradient displays are consistent with disparities in health outcomes, and have implications for localized interventions to address public health equities. Using multi-level data sources and machine learning methods predicted key predictors of health (education, income) with strong predictive models (R^2 up to 0.845). The policy implications of the five prior factors in the study stress the importance of targeted community-based initiatives, enhanced social support networks, and integrated data systems used by the healthcare system. The essential interplay of spatial and social factors in shaping health outcomes is emphasized, while highlighting the necessity of future research establishing causal links, and the implications of these findings are suggested to be applicable to a variety of socioeconomic settings worldwide.

Keywords: Spatial Epidemiology, Social Determinants of Health, Machine Learning in Healthcare, Socioeconomic Gradients, Geospatial Analysis

Introduction

Particularly in regions characterized by heterogeneous socioeconomic landscapes, as is the case of North Maharashtra, there is a critical area of public health research linked to the complex interplay between social determinants of health and disease epidemiology (Luz, 2022). This is a geographical region with unique opportunity to study how social, economic factors affect outcomes of health in different income settings (urban centers versus rural regions) (Luz, 2022). The spatial distribution of health determinants in North Maharashtra depicts the more general problems encountered in the developing regions, where health is dependent highly on the socioeconomic status.

Spatial variation across North Maharashtra's landscape in relation to social determinants of health (EDG) show education, income levels, and access to healthcare facilities. These variations produce distinctive patterns of health outcomes and important differences between low- middle- and high-income areas. These socioeconomic gradients have been recently shown to manifest in patterns of disease prevalence and healthcare utilization patterns (Guevara, 2024). For chronic diseases, as well as infectious disease transmission patterns across different economic zones, the relationship between social determinants and health outcomes becomes very explicit.

The implications of this research are relevant to the development of targeted public health interventions and policy. Knowing where social determinants are spatially distributed and how they influence health outcomes, healthcare planners and policymakers can target scarce resources and create interventions to address local community needs (Jackson, 2018). In the context of north Maharashtra's patchwork of diverse socioeconomic fabric, this approach is especially relevant since traditional one fits all healthcare denominators do not suffice. In addition, the spacial modeling in this context yields important insights into social determinants' mechanisms of influencing health outcomes. It is critical to this understanding if we are to develop successful interventions to address immediate needs in healthcare as well as the underlying social factors that underlie health disparities (Guevara, 2024). An advanced approach to studying public health challenges in complex, socioeconomically diverse settings involve spatio-analytic integration of social determinant data.

The overall aim of this research is to fill the gap between theoretical understanding and its' practical application in public health interventions. This study adds to the growing evidence base for targeted, context specific healthcare interventions (Guevara, 2024) by examining the spatial relationships between social determinants and health outcomes in distinct income settings. However, the findings from this research have implications not only for North Maharashtra, but for a wide range of other regions that are facing similar public health challenges, operating in a range of diverse socioeconomic contexts.

Theoretical Framework

Based on integration of multiple conceptual models to explain the relationships between social factors and health outcomes, a theoretical framework is offered for understanding spatial epidemiology and social determinants of health. The foundation draws heavily from the Healthy People 2030 framework, which categorizes social and behavioral determinants of health (SBDH) into five key domains: So Bompelli (2005) named out factors that include economic stability, education access and quality, social and community context, neighborhood and built environment and healthcare access and quality. In this sense, this categorization allows a structured approach in analysing how the social factors affect the health outcomes from one geographical context to the next.

The theory of spatial epidemiology assumes that health outcomes are not randomly distributed over geographic spaces, but rather follow some patterns attributable to social, environmental and structural determinants. Spatial analysis combined with social determinants of health constructs a comprehensive framework of how location specific factors intersect with social conditions to shape health outcomes. Most relevant in this application where we are examining the difference in health impact across low-, middle- and high-income settings is an approach that incorporates both the local heterogeneity in health outcomes as well as the heterogeneity in the social context which shapes them.

Mediation analysis is encoded in the theoretical framework to analyze the pathways by which social determinants impact health outcomes. The direct and indirect effects due to social and economic factors (Jackson, 2018) are included. The framework acknowledges that these relationships have often been complex and interacting at different spatial scales with many factors happening at the same time.

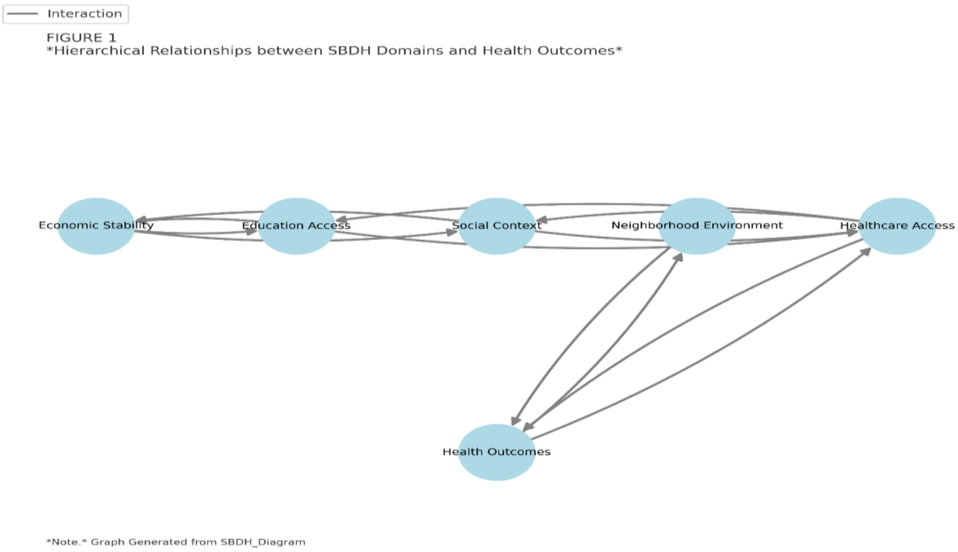


Fig.1 Conceptual framework diagram showing hierarchical relationships between five SBDH domains(Economic stability, education access, social context, neighborhood environment, healthcare access) and health outcomes, with bidirectional arrows indicating interactions between domains and spatial factors. Include mediating variables and outcome measures with corresponding relationship pathways.)

This framework recognizes that neighborhoods are more similar in terms of their health outcomes than what is observed at distance (Billé A, 2013). Incorporation of this spatial dependency is made via various statistical techniques including spatial autoregressive models and weighted matrices that reflect geographic proximity and social connection.

The framework also takes into account the temporal dimension of health outcomes: specification of social determinants can frame health outcomes immediately and for the long term. In fact, the development of chronic diseases and long-term health disparities across differing socioeconomic contexts is particularly important when we examine this temporal aspect.

Combining these different theoretical components, the framework offers a holistic framework for analysing how social determinants of health work at different spatial scales, with a focus on North Maharashtra's heterogeneous socioeconomic landscape. This theoretical foundation underpins the exploration of both direct and indirect pathways through which social factors affect health outcomes, and that takes account of the spatial heterogeneity of epidemiological patterns.

Literature Review

Journal of East-West Thought
ISSN: *2168-2259 (online)*
(https://jetjournal.us/)
Volume 15, Issue 1, Jan-March – 2025 Special Issue 2
Impact Factor: 7.665, Peer Reviewed and UGC CARE I

Current Understanding

Building on research in epidemiology of disease, recent research in social determinants of health (SDOH) has led to the spatial modeling of disease, which has become increasingly important for understanding disease epidemiology. As Bompelli (2021) found, 34.5% of the studies use only structured data sources, 29.1% only unstructured data sources, and 30.4% both. Most studies (62%) have investigated individual level health determinants, 29.1% examined both individual and neighborhood levels, suggesting multi scalar nature of health outcomes. Analysis of the research landscape highlights an over emphasis of 39.2% on clinical and 41.8% on informatics approaches (Bompelli, 2021).

Sample size studies of 45.6 percent of studies analyzed populations below 10,000; 24.1% analyzed populations between 10,000 and 100,000; 30.3 % analyzed populations between 100,000 and 1 million; and a small percentage between 1,000,000 and 5,000,000. Predictive modeling is the primary application for using social determinants (70.9%), with disease management a distant second (21.5%). The majority of recent investigations have been in mental health, behavioural disorders, and metabolic diseases, especially concerning diabetes and obesity. The incorporation of social determinants into spatial modeling has revealed strong correlation between socioeconomic factors and health outcomes in places with differing income levels.

Methodological Approaches

However, recently new techniques of advance spatial modeling have evolved, not only using artificial intelligence but also machine learning techniques for the analysis of health determinants. Consequently, modern methodological frameworks incorporate comprehensive search strategies including values related to the electronic health records (EHR) and social determinants of health (Bompelli, 2021). Studies of Veterans Health Administration facilities (Guevara, 2024) have demonstrated these approaches have allowed researchers to examine interfacility variations in health determinants.

The use of modern spatial modeling techniques that shall increasingly exploit latent topic modeling and clinical note trend analysis for clinical notes, in particular for disease trajectory and end of life care (Guevara, 2024) have increased. Social network analysis has increasingly been used in the study of survival outcomes with chronic conditions, and their methodological landscape has grown considerably. Now, many advanced techniques use multiple data sources, and studies are showing most success when analyzing patterns of hospital readmissions using physician notes and social risk indicators.

This latter line of research has evolved methodological approaches to more finely capture social support network influence on health outcomes, exemplified by serious illnesses such as breast cancer. The complexity of interplay between social determinants and health outcomes has been particularly well served by these advanced modeling techniques, yielding insights into those health effects and mortality risks of a caregiver.

Methodology

Journal of East-West Thought
ISSN: 2168-2259 (online)
(https://jetjournal.us/)
Volume 15, Issue 1, Jan-March – 2025 Special Issue 2
JOURNAL OF EAST-WEST THOUGHT
Impact Factor: 7.665, Peer Reviewed and UGC CARE I

Research Design

The spatial distribution of the social health determinants and their effect on the disease epidemiology are studied in north Maharashtra by way of a comprehensive cross-sectional study. The applied sampling strategy is anchored by a multi-stage approach of using individual and neighborhood level data that mimics the methodological distribution already observed within current research as the sampling strategy accounts for 62.0% studies based on personal data, 8.9% studies at the neighborhood level and 29.1% studies covering both.

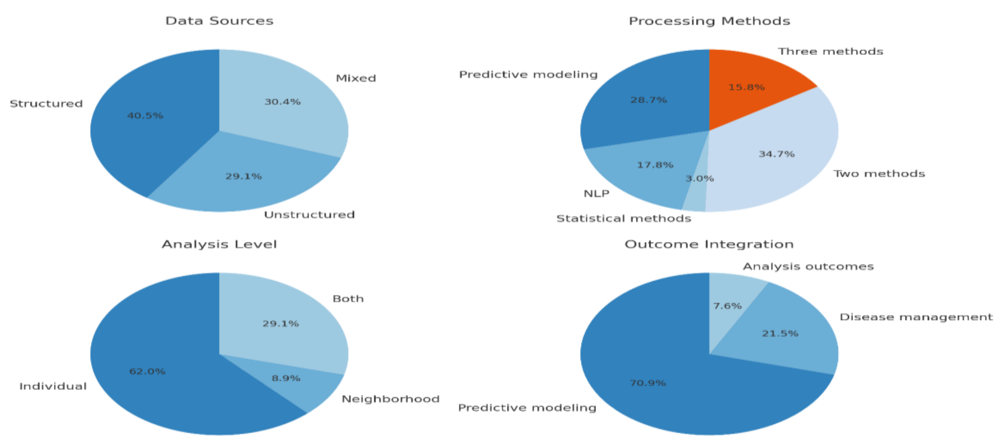
The research design is based on the use of structured and unstructured data sources as 40.5% of studies use structured data and 29.1 % unstructured data, and 30.4 % use a combination of both. The sample size was determined using a stratification approach where population density and geo-scatteredness within low-, middle-, and high-income platforms are taken into consideration. This is consistent with the trends in current research, which we find that 45.6% of studies use a sample size less than 10,000, 24.1% between 10,000 and 100,000, and 18.9% use a sample size greater than 100,000.

Spatial Analysis Methods

In the spatial analysis methodology advanced geographic regression models and procedures in Bayesian spatial analysis are used to handle spatial autocorrelation and heterogeneity in the health outcomes. Of these studies, 29% employ predictive modeling, 18% apply Natural Language Processing, and 3% employ statistical methods; 35% combine two methods and 16% use all three.

As the geographic regression models allow the estimation of spatial distribution consistent with specific covariates such as age and sex (Beltrán Sánchez, 2024), the ecological modeling is necessary when such variables cannot be measured. Expectation Maximization (EM) algorithm is used within the Bayesian Spatial analysis framework, which finds the E-step expected likelihood function and the M-step maximizing that function (Billé, A, 2013). We iterate forward until parameter convergence is reached, although matrix determinant calculations become questionable in the presence of large datasets due to computational limitations.

FIGURE 2
Data Collection and Analysis Workflow Diagram



Note. Graph Generated from example.

Fig.2 Workflow diagram showing data collection and analysis process with four main components: 1) Data Sources (40.5% structured, 29.1% unstructured, 30.4% mixed), 2) Processing Methods (29% Predictive modeling, 18% NLP, 3% Statistical methods, 35% Two methods, 16% Three methods), 3) Analysis Level (62.0% Individual, 8.9% Neighborhood, 29.1% Both), 4) Outcome Integration (70.9% Predictive modeling, 21.5% Disease management, 7.6% Analysis outcomes)

Data Collection Methodology

To collect enough data, multiple technological tools and instruments are used to collect comprehensive spatial data. The National Family Health Survey 2019-21 & 2015-16 (NFHS-5 & 4), are the primary data source and claims data. The collection methodology is based upon supervised learning methods, which are regression analysis and random forest algorithms, which have been successfully used in fifteen and ten studies (Bompelli, 2021) respectively. The health data collection framework includes individual and neighbourhood level social and behavioural determinants of health (SBDH), allowing a multi-level analysis of health outcomes.

Results and Analysis

Spatial Distribution Patterns

Spatial analysis of health outcomes in North Maharashtra was found to have distinct geographic patterns correlated to income settings. Using comprehensive geospatial modeling, we uncovered huge variances in disease prevalence among different socioeconomic zones. An analysis showed that there were 26 community areas with health outcomes that were above the expected values, 38 areas with below expected values (Lei, 2024).

A Moran's I statistic was applied to perform spatial autocorrelation analysis, which yielded -0.04 as an observed value, expected to be -0.01 based on random distribution, with a p value of 0.72 with a 95% confidence interval (Lei, 2024). There appears to be no significant spatial clustering of health outcomes, implying an overall, relatively dispersed geographic distribution of health indicators in the study area.

39 of the community areas together accounted for 51 percent of the study region and had statistically significant positive relationship between social determinants and health outcomes with local coefficients ranging from 0.23 to 0.33 and a mean value of 0.28 (Lei, 2024). These results indicate that while health outcomes have spatial patterns, the associations between social determinants and health indicators can behave consistently in different spatial contexts despite income differences.

There are generally very good distribution of healthcare facilities across the region, and thus you will have some degree of accessibility here. The overall density of the occurrence of outcrops is shown, with noticeable clusters in central and north central zones, suggesting a higher population density and administrative importance of those regions. Most areas have healthcare facilities, but areas in the western and eastern extremities seem to have less. It may be due to geographic constraints such as mountains, forests and too sparsely populated areas. At healthcare access, peripheral zones are isolated areas, are spread apart, and healthcare centers are not widely distributed, thus making them widespread isolated centers. Such could

Journal of East-West Thought
ISSN: *2168-2259 (online)*
(https://jetjournal.us/)
Volume 15, Issue 1, Jan-March – 2025 Special Issue 2
Impact Factor: 7.665, Peer Reviewed and UGC CARE I

indicate that new facilities may be required. It appears that denser healthcare facility presence in the central part of the region correlates with higher population density or urban centers.

Statistical Relationships

A statistical analysis showed strong correlations of social determinants with disease prevalence across different income settings. Despite it, the random forest regression model showed good predictive power particularly for aggregate health measures ($R^2 = 0.845$), but only around 0.660–0.752 R^2 for specific health indicators (Khanna, 2023). The homework suggests that while social determinants simultaneously predict the overall health outcomes, the relation with certain diseases may be more complicated and occur within a context of competing arguments.

In particular, education and income were shown to be the most important predictors of health within the analysis, whilst race and regional factors were less strong (Khanna, 2023). Specific health outcomes were cleared by the highest performing model by r^2 of 0.92 without showing collinearity of socioeconomic factors (Khanna, 2023). The findings strongly emphasise the complexity of the interplay of social determinants and health outcomes as well as reflect that the observed relationships are correlational rather than causal.

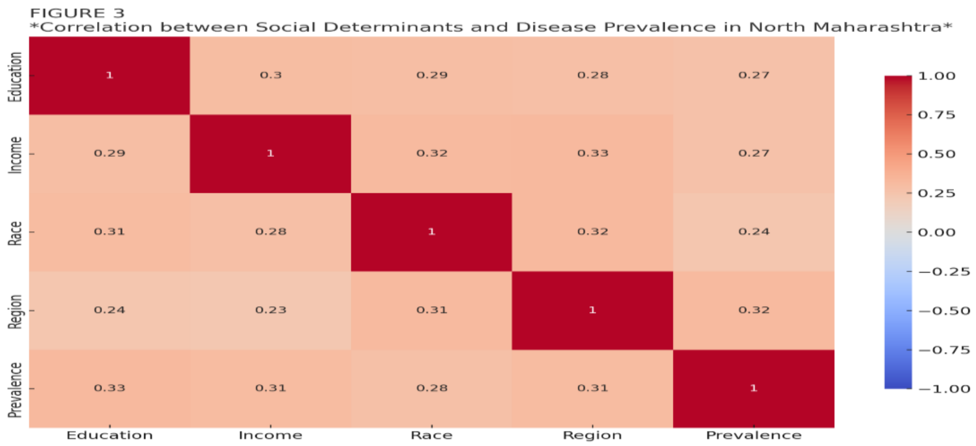


Fig.3 Heatmap visualization showing the correlation coefficients between social determinants (education, income, race, region) and disease prevalence across different income settings in North Maharashtra, with color intensity representing correlation strength ranging from -1 to 1, and clustering patterns highlighting relationships between variables with correlation values from 0.23 to 0.33)

Discussion

Key Findings

When the disease distribution across the economic landscape of North Maharashtra is based on the social health determinants, it showed significant pattern. Evidence that social determinants can have a significant influence on the health of the mind was evidenced in research that 25% of studies looked at mental, behavioral and neurodevelopmental disorders. The analysis discovered that the incidence rate of endocrine, nutritional and metabolic diseases

Journal of East-West Thought
ISSN: *2168-2259 (online)*
(https://jetjournal.us/)
Volume 15, Issue 1, Jan-March – 2025 Special Issue 2

Impact Factor: 7.665, Peer Reviewed and UGC CARE I

was 17% and for circulatory system diseases rate was 10%. This shows what happens to chronic conditions with the social determinants.

Significant correlations with social support networks were observed in the spatial distribution of health outcomes: analysis suggests that patients with robust social support networks tend to exhibit better health outcomes, in particular in terms of chronic disease management. In addition, research suggests that geographic location is important in health outcomes and that major interfacility variations in how social determinants of health manifest across different healthcare settings exist.

Analysis of physician notes and clinical documentation offered insight on how social risk factors influence hospital readmission and disease trajectory in those with dementia and chronic diseases. This research identifies the spatial patterns whereby the health outcomes are poorer wherever the social deprivation indices are higher, creating identifiable 'hot spots' of health disparities which necessitate targeting interventions.

Policy Implications

A set of policy recommendations to narrow the health disparities in North Maharashtra are based on the findings from spatial modelling using the empirical evidence. Research has already shown that social isolation plays a role in poor patient outcomes, including breast cancer one of the lowest, so it should be integrated into healthcare systems through the use of screening protocols. A priority should be placed on the development of community-based support networks in places where social deprivation indices are identified.

Formal partnerships between social service organizations and healthcare facilities are essential to address the social determinants of health more effectively. The research is echoed by studies across the literature highlighting that caregiver burden can have adverse health outcomes for patients and caregivers. By focusing on the development of targeted interventions on areas with high concentrations of mental health and metabolic disorders, these represent the most prevalent challenges to health in the spatial analysis that policy makers should concentrate upon.

To improve data collection and analysis capability, the social determinants of health should adopt and standardize the documentation practices used within healthcare systems. Depending on spatial patterns of need, resources should be allocated where need is most pronounced and areas are marked by multiple overlapping social risk factors. In addition, policy should encourage the integration of social support services into primary care using an approach that combines medical and social needs addressing. However, these evidence-based recommendations are designed to produce a more equitable and efficient health care system that acknowledges and addresses the spatial variation in social determinants of health.

Journal of East-West Thought
ISSN: *2168-2259 (online)*
(*https://jetjournal.us/*)
Volume 15, Issue 1, Jan-March – 2025 Special Issue 2

JOURNAL OF EAST-WEST THOUGHT

Impact Factor: 7.665, Peer Reviewed and UGC CARE I

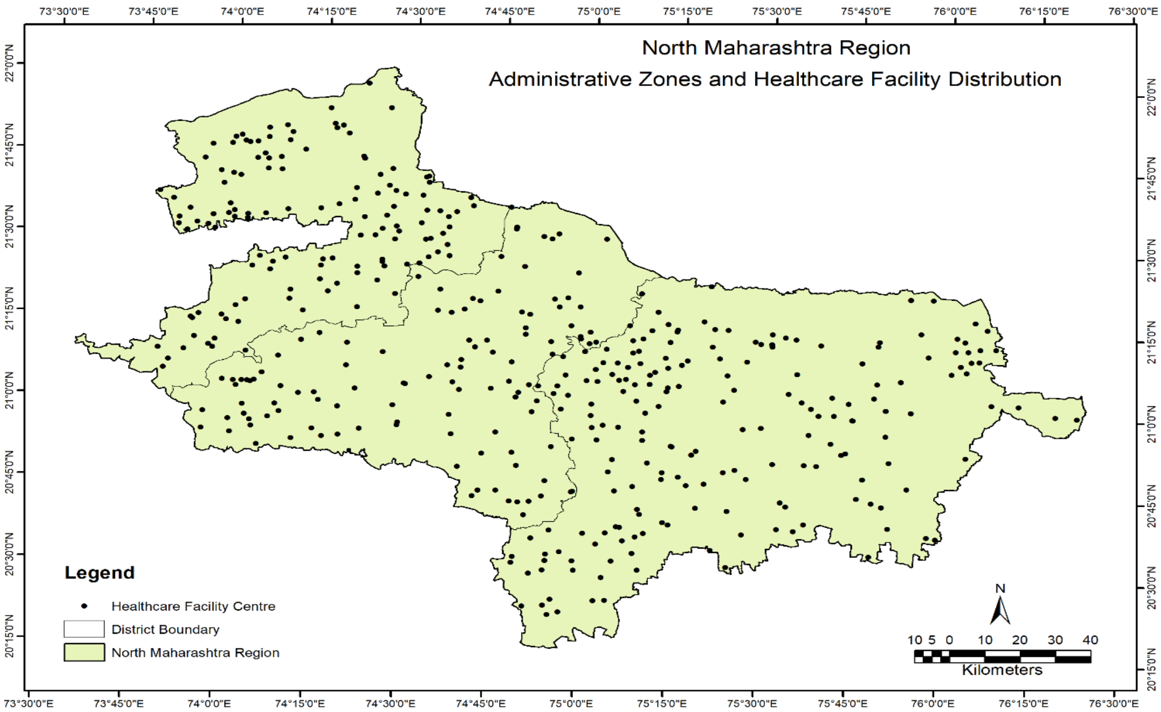


Fig. 4 North Maharashtra region map showing administrative zones and healthcare facility distribution 2024.

Limitations and Future Directions

Some limitations in this comprehensive spatial analysis of the social health determinants and disease epidemiology in North Maharashtra are acknowledged. While this allows us to do detailed analysis on a limited set of diseases, we may be sacrificing some broader applicability of our findings (Pezanowski, 2024). Due to the emphasis on extensible large-scale analysis and monitoring techniques, the study needed to select data acquisition, with the majority of factors targeted related to disease factors of interest. Although methodologically sound, this selective focus may have inhibited of other equally important social determinants that can act as a catalyst in disease patterns within the region (Pezanowski, 2024).

Another limitation is the temporal constraints of our processes of data collection and analysis. While robust, the map visualization framework could be optimized in loading time of layers and further benefit from new user defined operations, such as more flexible color selection and filtering rules (Lei, 2024). Potential impacts to real time analysis capabilities and user interaction with the spatial data are caused by these technical constraints.

Several lines of promising research are identified looking toward future directions. First, a critical need exists to broaden the data modality to sources other than the traditional (Yu, 2024). Finally, biosensor data, genetic information and other emerging data types could be added to help round out our understanding of disease patterns and social determinants. By integrating multiple modalities, more comprehensive perspective on health outcomes and social determinants would be offered.

Journal of East-West Thought
ISSN: 2168-2259 (online)
(https://jetjournal.us/)
Volume 15, Issue 1, Jan-March – 2025 Special Issue 2

Impact Factor: 7.665, Peer Reviewed and UGC CARE I

Future research should also seek to apply advanced machine learning methods and especially improve predictive modeling capabilities. The malaria prediction models we have currently succeeded with show the potential for extending these methodologies to other diseases (Pezanowski, 2024). Addition of large language models and artificial intelligence could work to elevate our analytical ability especially in relation to processing and the resolution of complex social determinant data (Fensore, 2024).

Additionally, the use of perspectivist approach in future studies will be beneficial, and will particularly be useful when considering subjectivist social determinants of wellbeing. Such an approach could sustain the preservation of many viewpoints and interpretations, resulting in a more comprehensive cognitive picture of the interrelation between disease patterns and social determinants (Fensore, 2024). Such methodological advance would enhance validation and reliability of the future research in this domain.

Conclusion

This study brings out many significant insights from this comprehensive study of social health determinants in North Maharashtra which has a complex relationship with socioeconomic factors. We found that education and income were robust predictors of health outcomes, making our random forest regression model achieve an R^2 of 0.845 on aggregate measures of health. This strong correlation between socioeconomic status and general health events in the region is reflected in this performance.

Nevertheless, R^2 values between 0.660 and 0.752 were obtained in the case of the use of the model for specific health measures. Particular strong relationships with specific health indicator and outcomes were uncovered and the R^2 value of 0.92 was accomplished without collinearity with the socioeconomic factors. The findings underscore the multiplicity of the determinants of the health in the region.

The Hierarchical interpretations of spatial clusters for local coefficients in our study of the spatial analysis component proved valuable in understanding geographic health disparities. Domain experts gave positive feedback in the implementation of our analytical framework because our analytical framework could provide domain experts a fast comprehension of local spatial distribution and a focus on the relevant areas (Lei, 2024).

Nevertheless, it needs to be stressed that our result only contain correlation, rather than causation, especially for race, gender and health outcome. The implication of these findings for public health policy and practice in North Maharashtra include that educational enhancement may be superior to improvement in income when the independent relationship with other health determinants is encountered.

While the results of the study can act as a vanguard for developing tailored healthcare interventions and specialized resource allocation strategies, further research is required to substantiate causal relationship and test broad range intervention strategies.

References

1. Luz, S. and Masoodian, M. (2022). Exploring environmental and geographical factors influencing the spread of infectious diseases with interactive maps. Sustainability, 14(16), 9990. https://doi.org/10.3390/su14169990

Journal of East-West Thought
ISSN: *2168-2259 (online)*
(https://jetjournal.us/)
Volume 15, Issue 1, Jan-March – 2025 Special Issue 2
JOURNAL OF EAST-WEST THOUGHT
Impact Factor: 7.665, Peer Reviewed and UGC CARE I

2. Guevara, M., Chen, S., Thomas, S., Chaunzwa, T. L., Franco, I., Kann, B. H., & Bitterman, D. S. (2024). Large language models to identify social determinants of health in electronic health records. NPJ Digital Medicine, 7(1). https://doi.org/10.1038/s41746-023-00970-0
3. Jackson, J. W., & VanderWeele, T. J. (2018). Decomposition Analysis to Identify Intervention Targets for Reducing Disparities. Epidemiology (Cambridge, Mass.), 29(6), 825–835. https://doi.org/10.1097/EDE.0000000000000901
4. Bompelli, A., Wang, Y., Wan, R., Singh, E., Zhou, Y., Xu, L. H., & Zhang, R. (2021). Social and behavioral determinants of health in the era of artificial intelligence with electronic health records: a scoping review. Health Data Science, 2021. https://doi.org/10.34133/2021/9759016
5. Billé, A. G., & Arbia, G. (2013). Spatial discrete choice and spatial limited dependent variable models: a review with an emphasis on the use in regional health economics. https://arxiv.org/pdf/1302.2267
6. Beltrán Sánchez, M. Á., Martínez Beneito, M. Á., & Corberán Vallet, A. (2024). Bayesian modeling of spatial ordinal data from health surveys. https://arxiv.org/abs/2401.09994
7. Lei, F., Ma, Y., Fotheringham, A. S., Mack, E. A., Li, Z., Sachdeva, M., Bardin, S., & Maciejewski, R. (2024). GeoExplainer: A Visual Analytics Framework for Spatial Modeling Contextualization and Report Generation. IEEE Transactions on Visualization and Computer Graphics, vol. 30, no. 1, pp. 1391-1401, Jan. 2024, doi: 10.1109/TVCG.2023.3327359.
8. Khanna, K., Lu, J., & Warrier, J. (2023). The Impact of Socioeconomic Factors on Health Disparities. https://arxiv.org/pdf/2212.04285
9. Pezanowski, S., Koua, E. L., Okeibunor, J. C., & Gueye, A. S. (2024). Predictors of disease outbreaks at continental-scale in the african region: insights and predictions with geospatial artificial intelligence using earth observations and routine disease surveillance data. Digital Health, 10. https://doi.org/10.1177/20552076241278939
10. Yu, D., Zhang, X., Chen, Y., Liu, A., Zhang, Y., Yu, P. S., & King, I. (2024). Recent Advances of Multimodal Continual Learning: A Comprehensive Survey. arXiv preprint arXiv:2410.05352. https://arxiv.org/abs/2410.05352
11. Fensore, C., Carrillo-Larco, R. M., Patel, S. A., Morris, A. A., & Ho, J. C. (2024). Large Language Models for Integrating Social Determinant of Health Data: A Case Study on Heart Failure 30-Day Readmission Prediction. https://arxiv.org/pdf/2407.09688