Artificial Intelligence-Enabled Clinical Decision Support for Non-Communicable Disease Prevention among Tribal Communities Using Lifestyle Behaviour: Integrating Risk Prediction and Healthcare

Sravanthi. V [1], Prof. T. Sobha Rani [2]

[1]Research Scholar, Department of Business Management, Sri Padmavati Mahila Visvavidyalayam, Tirupati

[2] Professor, Department of Communication & Journalism, Sri Padmavati Mahila Visvavidyalayam, Tirupati

Abstract

Non-communicable diseases (NCDs) are emerging as a significant health problem for tribal populations, aside from the established problems of infectious diseases, undernutrition, and healthcare access. The epidemiological transition, dietary changes, increased tobacco and alcohol use, occupational changes, decreased physical activity, obesity, and inadequate screening have contributed to the rising burden of NCDs in tribal populations. Evidence from India shows that non-communicable diseases in tribal areas are significant causes of mortality, and many cases of hypertension and diabetes mellitus go undiagnosed. In a multi-centric community-based study conducted in 12 tribal districts, 66% of deaths in individuals aged 15 years and older were due to NCDs, among which cardiovascular diseases were the leading cause of mortality in most of the study sites. A systematic review and meta-analysis of 42 studies also reported a significant prevalence of hypertension in the tribal population of India. Recent multi-centric studies have highlighted the burden of type 2 diabetes and prediabetes in tribal populations, with hypertension, obesity, and smokeless tobacco use as significant risk factors. Artificial intelligence (AI) and machine learning (ML) can enable multidimensional risk assessment and personalized healthcare by analysing demographic, behavioural, anthropometric, clinical, and longitudinal data using algorithms such as the Random Forest, XGBoost, and Long Short-Term Memory

(LSTM). Most current AI-based clinical models, however, were not designed with tribal populations in mind, raising questions about representation, population-specificity, biases, privacy, and interpretability. This integrative narrative review highlights the relationships between the increased burden of NCD, lifestyle and behavioural risk factors, AI-based risk prediction models and decision support systems, and healthcare for tribal populations. Based on the evidence, this review introduces a conceptual framework for a culturally sensitive Clinical Decision Support System (CDSS) designed to leverage tribal community-specific factors, such as lifestyle and behavioural data, along with clinical data, an AI-based predictive analytics module, and community-specific clinical guidelines. The review argues that CDSS should be designed as an augmentation to clinical decision-making rather than a substitute for human expertise. Community health workers can play a critical role in this framework by bridging the gap between AI-assisted clinical decision-making and community health. The review suggests that AI-enabled tribal healthcare can evolve from predictive analytics to a closed-loop system that includes promoting appropriate lifestyle changes, improving screening, facilitating referrals, and monitoring. The framework can guide future research directions, including population-specific AI model development and empirical studies of AI-based healthcare delivery in tribal populations.

Keywords: Artificial intelligence; Clinical decision support system; Tribal communities; non-communicable diseases; Machine learning; Random Forest; XGBoost; LSTM; Lifestyle; Preventive healthcare; Equity; India.

1. Introduction

Non-communicable diseases (NCDs) are a significant constituent of the global disease burden. Cardiovascular diseases, cancer, chronic respiratory diseases, and diabetes contribute significantly to the overall morbidity and mortality as well as pose several modifiable behavioural risk factors for humans. According to the World Health Organization, tobacco use, harmful alcohol consumption, physical inactivity, and unhealthy diets are the four key behavioural risk factors for NCDs. The aforementioned modifiable behaviours have been associated with increased blood pressure, blood glucose levels, lipid levels, overweight, and obesity. [1]

The importance of NCDs in tribal populations has been increasingly recognized, particularly with the growing emphasis on reducing health inequalities. Tribal populations are highly diverse groups of people who differ significantly in their geographical distribution, cultural practices, dietary patterns, occupations, socio-economic status, and access to healthcare services. Despite the apparent heterogeneity, the health status of indigenous peoples has consistently shown worse health outcomes as compared to the general population. Anderson et al. emphasized the importance of high-quality indigenous health data for evidence-based health policy and service planning. [2]

Historically, tribal populations in India have been largely addressed in terms of the greater importance of infectious diseases and poor nutrition. However, the recent paradigm shift towards the growing relevance of NCDs appears to apply to tribal populations as well. Sathiyanarayanan et al. noted the rising prevalence of lifestyle-related diseases among tribal populations and argued for the need to address NCDs in tribal health programs. [3]

More recent studies highlight the significance of this issue even further. Kaur et al. carried out a community-based mortality survey among 12 highly tribal districts (with more than 50% of the population identifying as tribal). Out of 5,292 deaths of the respondents aged 15 or older, 66% were linked to NCDs, 15% to infectious diseases, and 11% to injuries. Moreover, the researchers found cardiovascular diseases to be the chief cause of death in 10 out of 12 study sites. [4]

The researchers also highlight the importance of addressing NCDs in tribal populations on an individual level. Hazarika and Babu carried out a systematic review and meta-analysis of 42 studies regarding hypertension prevalence in tribal populations of India. The researchers report a pooled estimate of 23.66% of hypertension prevalence in males, 23.37% in females, and 16.68% in combined pooled populations with large variability between the studies. [5]

A multicentric study on tribal populations from six districts evaluated the prevalence of type 2 diabetes. The researchers have used random and fasting blood glucose tests to analyse 8,486 adults and 3,131 adults, respectively. They report the prevalence of diabetes of 4.77% and 6.80% among the tested population using random and fasting blood glucose tests, respectively. At the same time, 8.69% of the fasting tests population were diagnosed with prediabetes. Interestingly, diabetes was found to be associated with age, smokeless tobacco use, hypertension, and obesity. [6]

Finally, a similar multicentric study on tribal populations from six districts analysed the state of hypertension in tribal populations. It evaluated prevalence, awareness, treatment, control, and risk factors associated with hypertension. It appears that tribal populations are in great need of screening and management for hypertension, especially considering several key behavioural risk factors for NCDs. [7]

The reviewed body of literature highlights the great importance of NCD prevention, particularly among tribal populations. While treatment obviously plays a critical role in managing NCDs, screening, and behavioural risk factor reduction are of particular importance. Early identification of NCDs, as well as addressing individual risk factors for the development of NCDs, improved treatment adherence, supportive communication strategies, and timely referrals are of particular importance for preventing NCDs in tribal populations.

Artificial intelligence (AI) has become an increasingly important aspect of modern healthcare. Its ability to analyse hundreds or thousands of variables has allowed AI to be successfully applied in identifying chronic disease risk factors. A systematic review of 42 studies with AI applications in healthcare found multiple examples of machine learning (ML) identification of chronic disease occurrence, progression, and determinants. In particular, the review notes the use of Random Forest methodology and similar approaches. Overall, the review underlines ML's ability to process numerous independent variables and recognize patterns associated with chronic disease development. [8]

AI has also been successfully applied to developing clinical decision support systems (CDSS). However, the evidence on the effectiveness of CDSS is mixed, which has prompted Susanto et al. to undertake a systematic review of 32 studies on the effectiveness of machine learning in clinical decision support systems. Notably, all of the reviewed studies used ML in either secondary or tertiary care in developed countries. The authors conclude by stating the need to further explore these interventions in low- and middle-income countries. [9]

This systematic review highlights the importance of CDSS in healthcare, with a particular focus on the lack of research opportunities in low- and middle-income countries. However, the research question that emerges from the systematic review undertaken is not only:

How can artificial intelligence be leveraged to improve health outcomes?

but rather:

How can artificial intelligence be leveraged together with an individual's lifestyle and cultural context for disease prevention and management?

Particularly in the context of NCD prevention and management, this paper aims to synthesize the current understanding of tribal lifestyles, the application of AI in chronic disease prediction, and AI-enabled behaviour change and clinical decision support.

2. Aim of the Review

The main aim of this review is to assess the role of artificial intelligence-enabled clinical decision support in preventing non-communicable diseases (NCD) by analysing lifestyle behaviour, clinical risk factors, predictive modelling and culturally responsive preventive healthcare in tribal communities.

3. Objectives

This review has five objectives:

1. To examine the emerging burden of NCD among tribal populations with a focus on hypertension, diabetes and cardiovascular risk
2. To examine lifestyle and behavioural factors associated with NCD risk among tribal populations
3. To review the role of AI and machine-learning in NCD risk prediction and clinical decision support
4. To examine the role of digital technologies in promoting NCD preventive behaviours
5. To propose a culturally responsive AI-enabled clinical decision support framework for NCD prevention among tribal populations

4. Review Approach

4.1 Type of Review

This paper follows an integrative narrative review and conceptual framework approach. The reason for this choice is that this paper synthesizes published evidence from several areas. This review is not a formally registered systematic review. The literature was searched across four areas:

- tribal and indigenous health
- NCD epidemiology and behavioural risk
- AI and machine-learning prediction

- digital health and clinical decision support

Published studies were prioritized, including those in peer-reviewed journals and reputable health organizations. Original epidemiological studies, systematic reviews, meta-analysis and methodological publications were prioritized. This review can evolve into a formal systematic or scoping review, following the PRISMA framework, with a clearly defined search strategy, screening process and quality assessment in the future. A future version of this review can also be a systematic review if the search strategy and screening process are formalized and if this review follows the PRISMA framework. Alternatively, this review can become a scoping review if it focuses on mapping the literature and key concepts rather than on synthesizing quantitative data. However, at this stage, this paper focuses on conceptual integration and framework development.

5. NCD Burden among Tribal Communities

5.1 Epidemiological Transition

The emerging burden of NCD among tribal populations should be interpreted within the context of epidemiological transition. Tribal populations have traditionally been prioritized in terms of health: health programs have focused on treating tribal populations for infectious diseases, addressing micronutrient deficiencies and providing maternal and child health services. However, the epidemiological transition is now affecting tribal populations: hypertension, diabetes and cardiovascular diseases are also becoming serious health concerns for tribal populations. Sathiyanarayanan et al. examine diabetes and hypertension among tribal adults and conclude that non-communicable diseases are an emerging health priority for tribal populations. [3] The mortality study by Kaur et al. provides more population-level evidence for the changing health priorities in tribal populations. In 12 tribal districts, two-thirds of all deaths among people aged 15 years and older were attributed to NCDs. [4] This means that the health system needs to address not only communicable diseases and maternal and child health but also NCDs. A dual health strategy is needed for tribal populations: the prevention and management of communicable diseases and NCDs.

6. Hypertension among Tribal Populations

6.1 Silent Killer

Hypertension is a particularly insidious NCD since it often does not present any symptoms and is often diagnosed too late, after serious complications have already occurred. The systematic review by Hazarika and Babu analysed 42 studies on hypertension in Indian tribal populations. The review found that pooled prevalence estimates varied widely across tribal populations but were consistently high across all populations. [5] Another multicentric study by Babu et al. found that hypertension prevalence was high, but tribal populations also had poor awareness, treatment and control of hypertension. [7] In other words, there is a continuum of care:

- risk factors for hypertension
- hypertension
- undiagnosed hypertension
- hypertension complications

Artificial intelligence can play a role in the first step: by stratifying populations and individuals according to risk factors. However, it is important to note that AI should not be used for diagnosing hypertension: diagnosis requires proper measurements and clinical evaluation. Artificial intelligence can support clinical decision-making but cannot replace it.

7. Diabetes among Tribal Populations

7.1 Hidden Risk Factors

Diabetes is another serious NCD that is an emerging health concern for tribal populations. The multicentric study by Babu et al. analysed type 2 diabetes prevalence in India and found that it ranged from 4.77% (random blood glucose) to 6.80% (fasting blood glucose). The study also found that 8.69% of the population had prediabetes (according to fasting blood glucose). The study found that diabetes was associated with age, smokeless tobacco use, hypertension and obesity. [6] These findings are useful for developing an AI model since they suggest that diabetes is a multifactorial NCD. In other words, there is not one risk factor but several:

- age
- obesity
- hypertension
- tobacco use

An AI model can combine these risk factors and estimate the likelihood of developing diabetes.

8. Lifestyle and Behavioural Determinants of NCDs

8.1 Tobacco Use

Tobacco is one of the main modifiable NCD risk factors, according to the World Health Organization. [1] Tribal populations are particularly vulnerable to the health effects of tobacco. In addition, tribal populations may use both smoked and smokeless tobacco. The Indian multicentric diabetes study found that smokeless tobacco was a risk factor for diabetes. [6] Similarly, the behavioural study on tribal populations found that tobacco use was associated with hypertension. [7] In other words, tobacco use is a risk factor for several NCDs and should be considered in an NCD prevention model. The variables to consider are numerous:

- current tobacco use
- type of tobacco

- frequency
- duration
- smokeless tobacco use
- smoking history
- household exposure

The goal of including these variables in an AI model is to estimate the likelihood of NCDs and support prevention, not to pathologize individuals.

9. Alcohol Consumption

9.1 Harmful Use of Alcohol

Harmful alcohol use is another key NCD risk factor. [1] There is evidence that alcohol consumption is a risk factor for hypertension among tribal populations. [7] However, it is important to recognize that alcohol use is a complex behaviour that varies from community to community. In other words, it is not enough to simply look at whether a person consumes alcohol or not. An AI model for preventing harmful alcohol use should consider several variables:

- frequency
- quantity
- duration
- age
- other risk factors

It should also consider how alcohol interacts with other risk factors. The output of such a model can be used to counsel individuals and refer them to medical services if needed.

10. Dietary Behaviour

10.1 Healthy Eating

Diet is a key behavioural determinant of NCDs. It is also one of the most culturally sensitive aspects of health. Tribal populations are nutritionally diverse: they are accustomed to different foods and diets depending on their environment and economic status. In other words, it is not enough to simply recommend a healthy diet to a tribal population. A culturally responsive AI model for diet should include the following variables:

- available foods
- traditional dietary patterns
- household income
- access to food
- seasonality
- preferences
- metabolic risk

The aim is to combine traditional dietary knowledge with nutritional evidence and individual risk factors. The recent review on NCDs among tribal populations in India specifically notes the importance of addressing the erosion of traditional food practices, behavioural risk factors and social inequalities. [10]

11. Physical Activity and Occupational Behaviour

Physical activity is another important modifiable determinant of NCD risk. However, traditional health-related questionnaires may misrepresent the type of physical activity undertaken by members of the tribe. For instance, someone may not report 'exercises' but may engage in agricultural, forest, walking, manual home activities or occupational activity. The following variables should be considered to better capture the nuances of physical activity:

- occupation-specific activity
- time spent walking
- agricultural activity
- manual home tasks
- recreation time
- seated time

This may better represent a more culturally relevant description of physical activity.

12. Healthcare Access and Health-Seeking Behaviour

Lifestyle behaviours are intrinsically linked to healthcare access and health-seeking behaviours. Tribes may have issues with remoteness, transport and the availability of health services and specific, specialist care. As such, a member of the tribe with high NCD risk may not have timely access to screening services. Such a consideration may suggest that an AI model may also be useful for screening prioritization and healthcare resource allocation, for instance: High predicted NCD risk but low healthcare access to community screening prioritization However, such considerations must be undertaken with caution to avoid inadvertent discrimination.

13. Artificial Intelligence and NCD Risk Prediction

AI-driven prediction has gained much interest in chronic disease management. Delpino et al. undertook a systematic review of 42 studies of chronic disease prediction using machine learning. They found that machine learning can predict chronic disease onset, progression and determinants. Random Forest was highlighted as one of the algorithms that performed best for prediction. [8] Notably, the ability of machine learning algorithms to undertake complex associations between variables and prediction makes them useful candidates for determining NCD risk. For instance,

- Age + BMI + BP + glucose + tobacco + alcohol + activity + family history can be related to NCD onset. However, despite high predictive performance, a machine learning model must also demonstrate calibration, external validity, interpretability, fairness, clinical utility and feasibility.

14. Random Forest in the Proposed Framework

Random Forest is an ensemble learning method developed by Breiman which combines trees using randomized feature selection and aggregation. [11] The application of random forest in the analysis of tribal NCD data may be useful due to the potential ability to capture complex associations and interactions between variables. Within the proposed framework, random forests might be used for data which contain the following:

- demographic variables
- behavioural variables
- anthropometric measurements
- clinical measurements
- socioeconomic variables

However, the efficacy of random forests must be demonstrated rather than assumed.

15. XGBoost in the Proposed Framework

XGBoost is a scalable gradient-boosting algorithm developed by Chen and Guestrin. [12] It is often useful for structured prediction tasks. Within the context of a tribal NCD dataset, XGBoost can be applied to undertake the following associations:

- BP + BMI + glucose + tobacco + alcohol + diet + activity + age + family history

However, similar to random forests, the utility of XGBoost for such an application must be demonstrated.

16. LSTM for Longitudinal NCD Monitoring

LSTM networks are a type of recurrent neural network developed by Hochreiter and Schmidhuber which allow for the learning of long-term dependencies. [13] The relevance of LSTM networks to the current proposal mostly depends on available data. If longitudinal data is available (e.g., 6 months of monitoring including BP, blood glucose, weight, activity, diet, medication adherence), then LSTMs can potentially be more informative than standard models which use baseline only. Within the proposed framework,

- Random Forest/XGBoost → structured data
- LSTM → longitudinal data

17. Clinical Decision Support Systems

A clinical decision support system (CDSS) generally involves information about patients, clinicians or related matters which is used to aid in health-related decisions. As such, the AI model described in this paper is only a part of a larger CDSS. Ideally, a comprehensive CDSS would also involve:

- Community/health data
- Data preparation
- Knowledge base
- AI prediction
- Risk stratification
- Recommendation
- Human decision-making
- Follow-up

Susanto et al. undertook a review of 32 clinical decision-support studies which applied ML. They found that most CDSSs were actually "assistance systems." Most notably, all clinicians had to review and approve the suggestions, and all studies were undertaken in developed countries, with significant opportunities for research in low-and middle-income countries. [9] This observation highlights the need to consider CDSSs for use in tribal and other disadvantaged populations.

18. Proposed AI-driven CDSS Framework The proposed framework.

Tribal Community & Patient Data

↓

Data Preparation & Preprocessing

↓

AI Prediction Layer

(Random Forest / XGBoost / LSTM)

↓

Risk Stratification

(Low / Moderate / High)

↓

Clinical Knowledge + Cultural Context

↓

Personalized Preventive Recommendations

↓

Community Health Worker + Clinician Review

↓

Actionable Health Intervention

(Screening / Counselling / Referral / Lifestyle Modification)

↓

Follow-up & Monitoring

↺

New Health Data → Updated Risk Assessment

AI-enabled Clinical Decision Support Framework for NCD Prevention among Tribal Communities.

The framework generally starts with obtaining community data followed by data preparation and preprocessing. Once data are ready, they can be used for AI prediction which informs stratification and recommendation. The recommendation ideally involves discussion with clinicians and/or community members followed by actionable health steps. Notably, the AI model is only a part of the framework; as such, the figure shows it as the prediction layer.

19. Components of the Proposed CDSS 19.1 Community and Patient Data

The first layer generally comprises multidimensional information. The following variables might be considered:

- Demographic
- Age
- Sex
- Education
- Occupation
- Behavioural
- Tobacco
- Alcohol
- Diet
- Physical activity
- Sedentary time
- Anthropometric
- Height
- Weight

- BMI
- Waist circumference
- Clinical
- BP
- Blood glucose
- HbA1c if available
- Lipid profile
- Family history
- Contextual
- Healthcare access
- Geographic location
- Socioeconomic status
- Language
- Community-specific characteristics

Raw data cannot be used to train a prediction model. As such, there is a need for a data preprocessing layer. The following steps may be considered:

- data cleaning
- missing data evaluation
- encoding
- normalization/scaling when appropriate
- assessment of outliers
- feature selection
- class imbalance consideration if applicable

Generally, the exact steps will depend on the data and the prediction model. Notably, preprocessing steps should also be reported since the omission or addition of specific steps can introduce biases or other issues, such as data leakage.

21. AI Prediction Layer

- The proposed framework includes several AI options:
- Random Forest: for structured nonlinear prediction,
- XGBoost: for structured prediction,
- LSTM: for longitudinal data.

A future empirical study will be required to compare the performance of these algorithms. For instance, a study can compare the performance of Random Forest, XGBoost and LSTM against standard statistical models, such as logistic regression. In turn, the best-fitting model should not be chosen based on predictive accuracy only but rather on a combination of the following criteria:

- Discrimination
- Calibration
- Sensitivity
- Specificit
- Interpretability
- Fairness
- Clinically usefulness

22. Risk Stratification

The model can be used to generate a risk score. The risk scores can then be stratified into categories, for instance:

- Low risk
- Moderate risk
- High risk

The exact cut-offs will need to be determined based on clinical relevance but a possible consideration is as follows:

Low risk: general preventive health advice and routine screening

Moderate risk: targeted prevention advice and monitoring

High risk: intensive screening, professional evaluation and referrals

23. Clinical Knowledge Base

The prediction result alone is not sufficient for informing health decisions. Ideally, predictions should be interpreted within the context of relevant clinical guidelines and screening/prevention recommendations, as well as community-specific considerations. By integrating such information, the prediction moves from being an 'ML model' to a 'CDSS.'

24. Culturally Responsive Recommendations

The proposed framework can be used to support culturally responsive recommendations, such as the following:

- Language
- The recommendation should be in a language understood by the recipient.
- Food
- Use of local food when developing dietary recommendations.
- Occupation
- Consideration of occupational activity when making exercise recommendations.
- Healthcare access
- Recommendations should consider the distance of healthcare facilities.
- Community structure
- Use of community health workers where appropriate.
- The 2025 review of NCDs among Indian tribes emphasized the importance of equity and community engagement. [10]

25. AI and Behaviour Change

Ultimately, the value of AI for NCD prevention will be determined by its ability to not only predict but also to facilitate behaviour change. One pathway by which AI can contribute is via the following sequence:

Risk prediction ↓ Risk explanation ↓ Personalized communication ↓ Behavioural change ↓ Follow-up ↓ Updated information ↓ Revised risk estimation

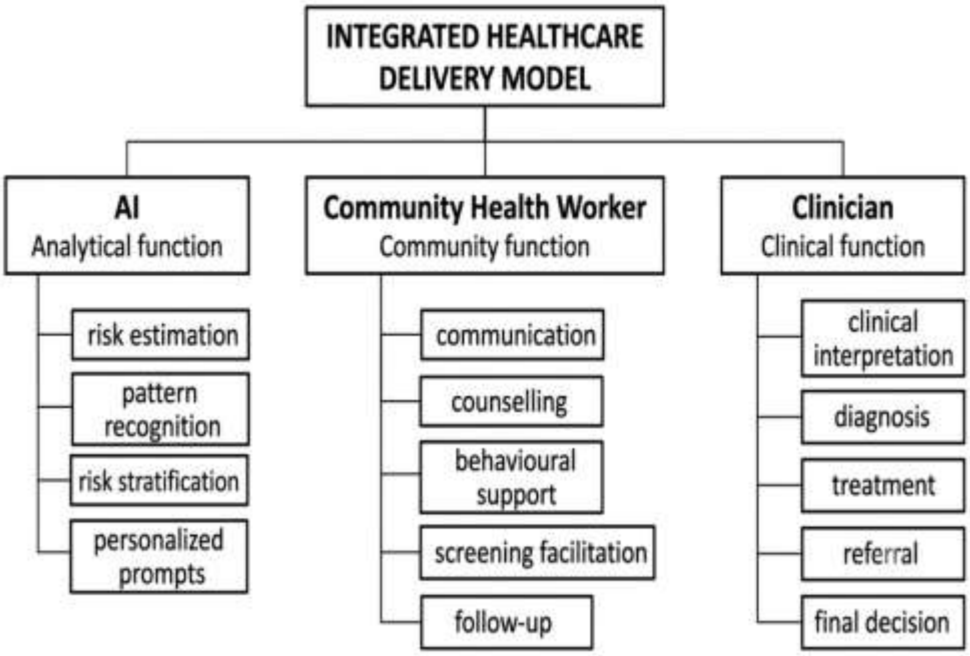
This is a continuous loop which involves an iterative change in health and risk.

26. Digital Health Evidence

Evidence from digital health research supports the general proposition that technology can contribute to health behaviour change. First, Palmer et al. conducted a systematic review of 71 randomized trials of mobile interventions for smoking cessation, physical activity, diet and alcohol reduction. Evidence was strongest for SMS-based smoking cessation interventions, with more mixed results for physical activity, diet and alcohol reduction. [14] Jahan et al. conducted a systematic review of 24 studies (2,378 participants) on smartphone applications for diabetes prevention, reporting that these were associated with weight loss and decreased BMI, but that socially disadvantaged populations were under-represented and that future studies should focus on personalization and diversity. [15] More recently, a systematic scoping review of digital interventions for NCD prevention found that messaging platforms, mobile applications and wearables were used to encourage healthy behaviours, with effectiveness depending on the fit between intervention functions and behavioural targets. [16] These findings do not demonstrate that AI-based interventions are effective in tribal populations, but rather that digital health interventions more broadly have potential. The distinction is important.

27. Role of Community Health Workers

The proposed system should not be viewed as a technology-centric intervention. Community health workers can act as the interface between the digital system and the community. The following diagram depicts the roles of AI, community health workers and clinicians in the proposed system.



Thus, the proposed model involves AI + Community Health Worker + Clinician rather than AI-only solutions. This is consistent with the World Health Organization's general principle that humans should remain in control of healthcare systems and clinical decision-making. [17]

28. Explainable AI

Interpretability is a critical issue in healthcare, since clinicians need to understand the rationale behind computational predictions. Xu et al. conducted a systematic review of the interpretability of AI-based clinical decision support systems, concluding that explainability was a major barrier to clinical adoption. [18] For tribal health, this is a critical consideration. Instead of simply presenting a risk score such as "Risk = 0.84", the system could potentially explain the rationale behind the prediction, for example: "The estimated risk is elevated, with blood pressure, tobacco use, BMI and age contributing to the model's prediction". These should be viewed as approximations of causation, not proof of it.

29. Bias and Health Equity

AI can amplify biases present in training data. Colacci et al. conducted a systematic review of 6,448 articles on sociodemographic bias in clinical ML models, identifying 91 studies that met their inclusion criteria. Of the 91 studies, 74.7% reported evidence of bias, and of those, 87% reported evidence of bias against socially disadvantaged groups. [19] This is directly relevant to the tribal health context, since if tribal populations are under-represented in the training dataset, the model will perform poorly for those populations. Future tribal health AI applications should therefore incorporate population-specific datasets, external validation, subgroup analysis, fairness audits, calibration, transparency and community feedback.

30. Ethical Considerations

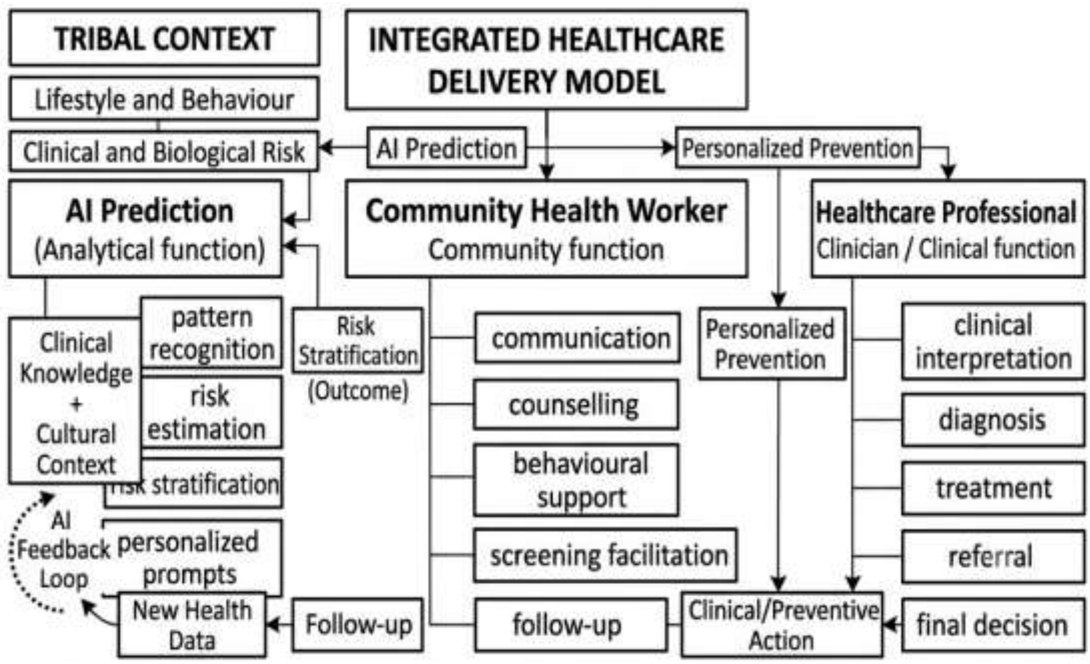
The World Health Organization's guidance on ethics and governance of AI for health identifies human autonomy, privacy, confidentiality, informed consent, transparency, accountability and equity as key principles. [17] In addition, special consideration should be given to the following issues in the tribal health context:

- **Data ownership:** Who controls community health data?
- **Consent:** How is informed consent obtained?
- **Privacy:** How are personal health records protected?
- **Representation:** Are the community's health needs adequately represented in the dataset?
- **Cultural sensitivity:** Are recommendations culturally appropriate?
- **Human oversight:** Can clinicians override AI recommendations?

These issues should be considered prior to implementation.

31. Proposed Research Model

The conceptual model can be summarized in the following diagram:



This creates a closed-loop preventive healthcare model.

32. Why the Proposed Model Is Different

Much of the literature falls into one or more of the following categories:

Category 1: Studies describing the epidemiology of NCDs in tribal populations.

Category 2: Studies investigating lifestyle/behavioural risk factors.

Category 3: AI studies predicting NCDs in general populations.

Category 4: Digital health interventions promoting lifestyle change.

Category 5: AI-enabled clinical decision-support research.

The proposed framework synthesizes these areas of research, linking tribal NCD epidemiology + lifestyle/behaviour + AI prediction + culturally appropriate communication + clinical decision-support + human oversight. This is the conceptual innovation of this paper.

33. Research Gaps

The review identifies the following major gaps.

33.1 Population Representation Gap

Tribal populations are under-represented in many AI datasets.

33.2 Behavioural Data Gap

Clinical AI models tend to focus on clinical biomarkers, with less emphasis on culturally-specific lifestyle/behavioural data.

33.3 Cultural Adaptation Gap

Generic AI recommendations may not be appropriate for diverse tribal populations.

33.4 Prediction-to-Action Gap

Many AI studies focus on prediction accuracy rather than behavioural or clinical impact.

33.5 Longitudinal Data Gap

LSTM-based approaches require longitudinal data, which are limited for tribal populations.

33.6 Explainability Gap

Many high-performing AI models are not interpretable to clinicians or communities.

33.7 Equity Gap

AI models require rigorous evaluation for subgroup performance and algorithmic bias.

33.8 Implementation Gap

There is limited literature on the implementation of AI-based clinical decision-support systems, particularly in low-resource and tribal settings. Susanto et al.'s review found that all 30 studies included in their review were from high-income countries, demonstrating the implementation gap. [9]

34. Proposed Future Research Design

 ISSN (Online): 2582-3078

The conceptual framework can be transformed into a research proposal as follows.

Phase 1: Community assessment

- Assess:
- lifestyle;
- diet;
- tobacco;
- alcohol;
- physical activity;
- healthcare access;
- health literacy.

Phase 2: Clinical assessment

- Collect:
- blood pressure;
- BMI;
- waist circumference;
- glucose;
- lipids;
- family history.

Phase 3: Dataset development

Develop a population-specific dataset.

Phase 4: Model development

- Compare:
- Logistic Regression;
- Random Forest;

- XGBoost;
- LSTM (if longitudinal data are available).

Phase 5: Model evaluation

- Assess:
- sensitivity;
- specificity;
- precision;
- recall;
- F1-score;
- ROC-AUC;
- calibration;
- clinical utility.

Phase 6: Explainability

Identify and communicate the key predictors.

Phase 7: Fairness evaluation

Compare performance across relevant subgroups.

Phase 8: CDSS development

Develop a clinical decision-support system incorporating the model and population-specific recommendations.

Phase 9: Community implementation

Implement the system through community health workers and appropriate facilities.

Phase 10: Outcome evaluation

- Measure:
- screening uptake;

- health knowledge;
- tobacco use;
- alcohol use;
- diet;
- physical activity;
- blood pressure;
- glucose;
- BMI;
- healthcare utilization.

This final step is critical, since a technically sophisticated model does not necessarily lead to improved health outcomes.

35. Implications for Public Health

- The proposed framework has several public-health level implications.
- Early identification
- High-risk individuals can be identified prior to the development of severe disease.
- Targeted screening
- Healthcare resources can be focused on individuals with elevated risk.
- Personalized communication
- Health messages can be tailored to individuals' risk profiles and cultural context.
- Community participation
- Community health workers can help bridge the gap between AI and tribal populations.
- Continuous monitoring
- Serial measurements can be used to estimate changes in risk.
- Health-system planning
- Aggregated, de-identified data can potentially be used for health-system planning at scale.

These are potential implications of the proposed framework, but they have not been demonstrated by this review.

36. Limitations

This review has several limitations. First, tribal populations are diverse, and findings from one tribal population cannot be generalized to other tribal populations. Second, much of the evidence on tribal NCD is cross-sectional, meaning that causal inferences and behavioural change cannot be inferred. Third, there is limited direct evidence on the role of AI in tribal NCD prevention, meaning that the proposed CDSS is conceptual. Fourth, digital health interventions cannot be assumed to have equivalent impact to AI-based interventions. Fifth, the performance of any AI model depends on the representativeness and quality of its training and validation datasets. Finally, AI is not a panacea for socioeconomic inequities, healthcare access, geographic isolation and other structural determinants of health.

37. Conclusion

NCDs are an emerging health priority among tribal populations in India. Evidence from mortality studies, hypertension studies and diabetes studies suggests that cardiovascular diseases, hypertension and diabetes are already significant causes of morbidity and mortality. [4–7] At the same time, tobacco use, alcohol consumption, dietary transition, physical inactivity, obesity and socioeconomic factors can contribute to NCD risk. These factors should be understood in their cultural context rather than viewed as individual behaviours.

Artificial intelligence can help to address the complexity of NCD prevention by integrating diverse data types to identify individuals who may benefit from preventive interventions. Random Forest and XGBoost algorithms can be applied to structured data on clinical and behavioural risk factors, while Long Short-Term Memory (LSTM) networks can be used to analyze longitudinal data if available. [8,11–13] However, predictive models alone are not sufficient to support clinical decision-making. A useful clinical decision-support system would link patient/community data → predictive analytics → risk stratification → clinical knowledge → culturally appropriate recommendations → human intervention → monitoring and evaluation. Thus, the framework proposed in this review views AI as part of a larger healthcare system. Community health workers can help to interpret risk estimates and design appropriate prevention strategies, while clinicians can ensure that the recommendations are evidence-based and context-appropriate.

The most important challenge is to ensure that the AI system is representative, equitable, interpretable, clinically useful and culturally appropriate. Future research should focus on developing population-specific datasets, validating algorithms, assessing fairness, incorporating behavioural and cultural variables, and evaluating whether AI-based interventions can effectively change health behaviours and clinical outcomes. The framework proposed in this review can guide such research by linking tribal health, NCD prevention, lifestyle behaviours, artificial intelligence and clinical decision-support within a single conceptual model.

References

1. World Health Organization. *Noncommunicable diseases: Risk factors and global trends*. Geneva: World Health Organization. The WHO identifies tobacco use, harmful alcohol use, physical inactivity and unhealthy diet as major behavioural risk factors for NCDs.

2. Anderson I, Robson B, Connolly M, et al. Indigenous and tribal peoples' health (The Lancet-Lowitja Institute Global Collaboration): a population study. *Lancet*. 2016;388(10040):131–157. doi:10.1016/S0140-6736(16)00345-7.

3. Sathiyanarayanan S, Muthunarayanan L, Devaparthasarathy TA. Changing perspectives in tribal health: rising prevalence of lifestyle diseases among tribal population in India. *Indian J Community Med*. 2019;44:342–346. doi:10.4103/ijcm.IJCM_40_19.

4. Kaur P, Borah PK, Uike PV, et al. Non-communicable diseases as a major contributor to deaths in 12 tribal districts in India. *Indian J Med Res*. 2022;156(2):250–259. doi:10.4103/ijmr.ijmr_3332_21.

5. Hazarika CR, Babu BV. Prevalence of hypertension in Indian tribal population: a systematic review and meta-analysis. *J Racial Ethn Health Disparities*. 2024;11(1):451–467. doi:10.1007/s40615-023-01532-6.

6. Babu BV, Hazarika CR, Raina SK, et al. Prevalence of type 2 diabetes among tribal population of India: a multi-centric cross-sectional study. *J Natl Med Assoc*. 2024;116(2 Pt 1):153–164. doi:10.1016/j.jnma.2024.01.004.

7. Babu BV, Hazarika CR, Raina SK, et al. Hypertension prevalence, awareness, treatment, control and risk factors in tribal population of India: a multi-centric cross-sectional study. *J Racial Ethn Health Disparities*. 2024;11(6):3678–3689. doi:10.1007/s40615-023-01817-w.

8. Delpino FM, Costa ÂK, Farias SR, Chiavegatto Filho ADP, Arcêncio RA, Nunes BP. Machine learning for predicting chronic diseases: a systematic review. *Public Health*. 2022;205:14–25. doi:10.1016/j.puhe.2022.01.007.

9. Susanto AP, Lyell D, Widyantoro B, Berkovsky S, Magrabi F. Effects of machine learning-based clinical decision support systems on decision-making, care delivery, and patient outcomes: a scoping review. *J Am Med Inform Assoc*. 2023;30(12):2050–2063. doi:10.1093/jamia/ocad180.

10. Devi SR, Eashwar VMA, Pandian S, Sekhar MA, Manivannan P. Non-communicable diseases among tribal populations in India: epidemiology, social determinants, and tailored public health approaches. *Cureus*. 2025;17(11). doi:10.7759/cureus.96899.

11. Breiman L. Random forests. *Machine Learning*. 2001;45:5–32. doi:10.1023/A:1010933404324.

12. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016:785–794. doi:10.1145/2939672.2939785.

13. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*. 1997;9(8):1735–1780. doi:10.1162/neco.1997.9.8.1735.

14. Palmer M, Sutherland J, Barnard S, et al. The effectiveness of smoking cessation, physical activity/diet and alcohol reduction interventions delivered by mobile phones for the prevention of non-communicable diseases: a systematic review of

randomised controlled trials. *PLoS One*. 2018;13(1). doi:10.1371/journal.pone.0189801.

15. Jahan E, Almansour R, Ijaz K, et al. Smartphone applications to prevent type 2 diabetes: a systematic review and meta-analysis. *Am J Prev Med*. 2024;66(6):1060–1070. doi:10.1016/j.amepre.2024.01.008.

16. Shaluhiyah Z, Qatrannada SA, Kusumawati A, Miah MS. The effect of digital health intervention in promoting healthy behavior: a systematic scoping review on strategies to prevent non-communicable diseases. *Dialogues Health*. 2025;7:100258. doi:10.1016/j.dialog.2025.100258.

17. World Health Organization. *Ethics and governance of artificial intelligence for health: WHO guidance*. Geneva: World Health Organization; 2021.

18. Xu Q, Xie W, Liao B, et al. Interpretability of clinical decision support systems based on artificial intelligence from technological and medical perspective: a systematic review. *Comput Intell Neurosci*. 2023;2023:9919269. doi:10.1155/2023/9919269.

19. Colacci M, Huang YQ, Postill G, et al. Sociodemographic bias in clinical machine learning models: a scoping review of algorithmic bias instances and mechanisms. *J Clin Epidemiol*. 2025;178:111606. doi:10.1016/j.jclinepi.2024.111606.

20. Nojomi M, Babaee E, Rampisheh Z, et al. AI-powered clinical decision support systems in disease diagnosis, treatment planning, and prognosis: a systematic review. *Med J Islam Repub Iran*. 2025;39:81. doi:10.47176/mjiri.39.81.