ASSESSING THE ROLE OF DELAYS IN PULMONARY TB DETECTION AND CONTROL STRATEGIES ON TRANSMISSION DYNAMICS: A SYNTHETIC POPULATION APPROACH IN THIRUVANANTHAPURAM DISTRICT, KERALA

Dr Razi N

Dissertation submitted in partial fulfilment of the requirements for the award of

Master of Public Health

ACHUTHA MENON CENTRE FOR HEALTH SCIENCE STUDIES

SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND TECHNOLOGY

Thiruvananthapuram, Kerala, India – 695011

June 2025

Acknowledgements

First, I am deeply thankful to my guide, Dr. Biju Soman, for his constant support, encouragement, and guidance throughout this work.

I would also like to thank the faculty members of Achutha Menon Centre for Health Science Studies Dr. Rakhal Gaitonde, Dr. Srikant A, Dr. Jissa V. T, Dr. Jeemon P, Dr. Srinivasan Kannan, Dr. Mala Ramanathan, Dr. Ravi Prasad Varma and Dr. Manju Nair R for their valuable teaching and support during my MPH course.

Special thanks to Karan Kumar, Bhavesh Neekhra, and Philip Cherian from the Department of Computer Science, Ashoka University, for helping me with the development of the agent-based model and providing technical guidance.

I am grateful to Arun Jose, PhD scholar, for his thoughtful inputs and support throughout the thesis process.

I thank the State TB Cell for granting permission and providing access to the data used in this study.

A heartfelt thank you to Sangeetha Rachel John for encouraging me to pursue MPH and for being a steady source of support throughout.

Finally, I would like to thank my friends and family for their love, patience, and encouragement, which helped me complete this journey.

Declaration

I hereby declare that this dissertation titled "**Assessing the Role of Delays in Pulmonary TB Detection and Control Strategies on Transmission Dynamics: A Synthetic Population Approach in Thiruvananthapuram District, Kerala**" is the bonafide record of my original research. It has not been submitted to any other university or institution for the award of any degree or diploma. Information derived from the published or unpublished work of others has been duly acknowledged in the text.



Dr Razi N

Achutha Menon Centre for Health Science Studies

Sree Chitra Tirunal Institute for Medical Sciences and Technology

Thiruvananthapuram, Kerala, India - 695011

June 2025

Certificate

Certified that the dissertation titled "**Assessing the Role of Delays in Pulmonary TB Detection and Control Strategies on Transmission Dynamics: A Synthetic Population Approach in Thiruvananthapuram District, Kerala**" is a record of the research work undertaken by Dr Razi N in partial fulfillment of the requirements for the award of the degree of "Master of Public Health" under my guidance and supervision.

BIJU SOMAN

Digitally signed by BIJU SOMAN
Date: 2025.06.27 09:13:43 +05'30'

Dr Biju Soman

Professor & Head

Achutha Menon Centre for Health Science Studies

Sree Chitra Tirunal Institute for Medical Sciences and Technology

Thiruvananthapuram, Kerala, India - 695011

June 2025

Table of contents

Acknowledgements i

Declaration ii

Certificate iii

Abstract x

1 Introduction 1
- 1.1 Background . . . 1
- 1.2 Rationale for the study . . . 2

2 Literature Review 4
- 2.1 History of Infectious Disease Modeling . . . 4
- 2.2 Compartmental Models . . . 5
- 2.3 Agent-Based Models . . . 6
- 2.4 Synthetic Population . . . 7
- 2.5 Tuberculosis Transmission Dynamics . . . 7
- 2.6 Global and National Efforts for TB Elimination . . . 8
- 2.7 TB Control Interventions and Their Effectiveness . . . 9
- 2.8 Objectives . . . 10
  - 2.8.1 Major Objective . . . 10
  - 2.8.2 Minor Objectives . . . 10

3 Methods 11
- 3.1 Study Design . . . 11
- 3.2 Study Setting . . . 12
- 3.3 Nikshay Data . . . 13
  - 3.3.1 Variables Extracted . . . 13
  - 3.3.2 Analysis of Data . . . 13
  - 3.3.3 Integration into the Model . . . 14
- 3.4 Synthetic Population Generation . . . 15
  - 3.4.1 Data Sources . . . 15
  - 3.4.2 Population Attributes . . . 16
  - 3.4.3 Generation Technique . . . 17
  - 3.4.4 Making Synthetic Population Dynamic . . . 17
- 3.5 Input Parameters . . . 19
- 3.6 Modeling (Agent-Based Modeling) . . . 19
  - 3.6.1 Agents and Their Attributes . . . 19

3.6.2 Agent Schedule . . . 21
3.6.3 Disease States and Transitions . . . 21
3.6.4 TBI Screening and Intervention . . . 22
3.6.5 Time Steps and Simulation Period . . . 23
3.6.6 Model Calibration . . . 23
3.6.7 Scenarios Analysed . . . 23
3.6.8 Implementation Details . . . 24

4 Results **25**
4.1 Nikshay Data Analysis . . . 26
4.1.1 Notification Trends . . . 26
4.1.2 Treatment Outcome . . . 29
4.2 Synthetic Population Generation . . . 30
4.2.1 Comparisons with Census . . . 30
4.3 Demographic Projections . . . 32
4.3.1 Age-Specific Fertility Rates (ASFR) . . . 32
4.3.2 Age-Specific Mortality Rates (ASMR) . . . 33
4.3.3 Sex Ratio at Birth . . . 34
4.4 Dynamic Synthetic Population Characteristics . . . 35
4.4.1 Population Growth and Age Structure . . . 35
4.5 Model Calibration . . . 37
4.6 Baseline (Current Program) Scenario . . . 39
4.6.1 Pulmonary TB Incidence . . . 39
4.6.2 Prevalence of Pulmonary TB . . . 40
4.7 Intervention Scenario Results . . . 41
4.7.1 Delay in Diagnosing and Starting TB Treatment . . . 41
4.7.2 Household Contacts Screened for TBI and Started on TPT . . . 45

5 Discussion **49**
5.1 Nikshay and Secondary Data Analysis . . . 49
5.2 Synthetic Population Generation . . . 51
5.3 Modeling Delay and TPT Screening . . . 53
5.4 Comparisons Similar Studies . . . 55
5.5 Strengths and Limitations . . . 58
5.6 Conclusion . . . 60
5.7 Recommendations . . . 61

References **63**

Annexure **69**

List of Tables

3.1 Model parameters used for scenario simulations. 19

4.1 Age group-wise comparison of census data and synthetic population. . . . 31
4.2 Root-mean-square error comparing simulated weekly TB notifications to Nikshay data for candidate values of the per-contact infection probability (β). 38
4.3 Annual pulmonary TB incidence per 100,000 population under the baseline scenario, with percentage decreases from the previous year and from the base year (2025). 40
4.4 Annual pulmonary TB incidence per 100,000 population under differing diagnosis-delay scenarios, with percentage decreases from the previous year and from the base year (2025). 42
4.5 Annual pulmonary TB incidence per 100,000 population under 0 percent and 50 percent household-contact screening scenarios, with percentage changes from the previous year and from the base year (2025). 46

List of Figures

3.1 Map of Kerala highlighting Thiruvananthapuram district, the location of the present study. 12
3.2 Diagrammatic representation of steps in generating synthetic population . 16
3.3 Diagrammatic representation of the model for TB transmission dynamics and treatment. 21

4.1 Annual pulmonary TB notification rate 26
4.2 Gender trends in annual pulmonary TB notification rate 27
4.3 Age trends in annual pulmonary TB notification rate 28
4.4 Distribution of treatment outcome (2021-2024) 29
4.5 Comparison of sex distribution between census data and the synthetic population (2011) 30
4.6 Age distribution: census data vs synthetic population 31
4.7 Age-specific fertility rates by age groups across years (log scale) 32
4.8 Female age-specific mortality rates by age groups across years (log scale) 33
4.9 Male age-specific mortality rates by age groups across years (log scale) . 33
4.10 Sex ratio at birth 34
4.11 Population pyramid for the year 2016 generated using the dynamic synthetic population model 35
4.12 Population pyramid for the year 2021 generated using the dynamic synthetic population model 36
4.13 Population pyramid for the year 2026 generated using the dynamic synthetic population model 37
4.14 Projected annual pulmonary TB incidence per 100,000 population under the baseline scenario, 2025–2030. 39
4.15 Projected temporal trend in the prevalence of active pulmonary TB per 100,000 population under the baseline diagnostic-delay (78 days) and household-contact screening (93%) scenario, 2025–2030. 40
4.16 Projected annual pulmonary TB incidence per 100,000 population under scenarios of 4, 6 and 8-week diagnostic delays, 2025–2030. 41
4.17 Projected prevalence of pulmonary TB per 100,000 population under baseline and reduced diagnostic-delay scenarios, 2025–2030. 43
4.18 Projected proportion of the population with tuberculosis infection under baseline and reduced diagnostic-delay scenarios, 2025–2030. 44
4.19 Projected annual pulmonary TB incidence per 100,000 population under scenarios of 0 percent and 50 percent household-contact screening, 2025–2030 (baseline scenario for comparison). 45

4.20 Projected prevalence of pulmonary TB per 100,000 population under scenarios of 0 percent, 50 percent, and 93 percent household-contact screening, 2025–2030. 47
4.21 Projected proportion of the population with tuberculosis infection under scenarios of 0 percent, 50 percent, and 93 percent (baseline) household-contact screening, 2025–2030. 48

Glossary of Abbreviations

Abbreviation	Full Form
TB	Tuberculosis
TBI	Tuberculosis Infection
MDR TB	Multi-Drug-Resistant Tuberculosis
RR TB	Rifampicin-Resistant Tuberculosis
ABM	Agent-Based Model
TPT	Tuberculosis Preventive Therapy
HIV	Human Immunodeficiency Virus
TST	Tuberculin Skin Test
IGRA	Interferon-Gamma Release Assay
BCG	Bacillus Calmette–Guérin
IHDS II	India Human Development Survey II
IPU	Iterative Proportional Updating

Abstract

Introduction: India bears one of the highest global burdens of tuberculosis (TB). Traditional models often miss the complexity of TB transmission in diverse settings. This study presents an agent-based model (ABM) using a synthetic population for TB in India, simulating individual-level transmission in Thiruvananthapuram. The model explores how diagnostic delays and preventive treatment coverage affect TB trends, providing a foundation for more detailed, data-driven planning.

Methods: As a first step, we created a synthetic population for Thiruvananthapuram district covering 3.3 million persons. The 2011 census figures were updated yearly with projected births and deaths, using established demographic methods. Each person was considered as an agent and he/she followed daily household, school or workplace schedules. Disease states (susceptible, TB infection, active TB, treatment) and transitions were parameterised with available data. The per-contact transmission probability was calibrated to 2021-2024 notification counts (best fit $\beta = 0.2$). Scenarios compared the current program (mean diagnostic delay = 78 days, 93 percent household-contact screening) with delays of 4, 6, 8 weeks and screening coverage of 0 percent or 50 percent. Outcomes were the incidence and prevalence of active pulmonary TB and the proportion of the population with TB infection for 2025–2030.

Results: Under current conditions, annual pulmonary TB incidence fell 27 percent (41.9 to 30.4 per 100,000) and prevalence dropped to 21.7 per 100,000 by 2030. Cutting

the delay to four weeks accelerated decline: incidence fell 64 percent and prevalence to 7.5 per 100,000. Halving the screening coverage of household contacts of active pulmonary TB cases resulted in only a 10 percent decline in incidence compared to the base year, while stopping it reversed gains, with incidence 14 percent above baseline by 2030 and latent infection rising to 21.2 percent.

Conclusions: In this high-coverage setting, rapid diagnosis is the single most powerful lever, but sustained, near-universal household TPT is essential to prevent resurgence. This ABM is a preliminary framework that can be expanded with richer social-network data, drug-resistance dynamics and cost modules to refine predictions and guide future TB-free Kerala strategies.

Chapter 1

Introduction

1.1 Background

Tuberculosis (TB) has affected humans since ancient times. Evidence from human skeletons indicates that the disease has affected humans for thousands of years (Hershkovitz et al., 2015). Throughout history, TB has been known by various names, including consumption, phthisis, and the white plague, reflecting its devastating impact on societies across different eras (CDC, 2025). The etiology of TB remained unknown until 1882, when Dr. Robert Koch identified the causative agent, later named Mycobacterium tuberculosis (Sakula, 1982). Despite advancements in medical science and public health measures, TB, caused by Mycobacterium tuberculosis, remains a major global health challenge and continues to be one of the leading causes of death worldwide (Global Tuberculosis Report, 2024).

According to the Global Tuberculosis Report 2024 published by the World Health Organization, an estimated 10.8 million people developed TB in 2023 and caused an estimated 1.25 million deaths globally. The estimated incidence rate was 134 per 100,000 people per year. In 2023, 55 percent of the people who developed TB were men, 33 per-

cent were women, and 12 percent were children (aged 0–14 years). An estimated 400,000 people developed multidrug-resistant or rifampicin-resistant TB (MDR/RR-TB) in 2023. Worldwide, 175,923 people with MDR/RR-TB were enrolled in the treatment program in 2023. This level of enrolment is equivalent to approximately 44 percent of the estimated number of people who develop MDR/RR-TB each year (Global Tuberculosis Report, 2024). Approximately a quarter of the global population is estimated to be infected with TB, with the WHO South-East Asia, Western Pacific, and Africa regions comprising nearly 80 percent of TB infections (TBI) (Houben and Dodd, 2016).

In 2023, India reported 2.55 million TB cases, with a notification rate of 179 per 100,000 people. A total of 63,939 cases of MDR/RR-TB were diagnosed in the same year. The national TB prevalence was estimated at 312 per 100,000 (*National TB Prevalence Survey in India 2019-2021*, 2023). In Kerala, 20,846 TB cases were notified in 2024, of which 2,636 cases (12.6 percent of the state's total) were reported from Thiruvananthapuram district. Between 2021 and 2024, around two-thirds of the notified TB cases were pulmonary TB, the form responsible for community transmission of the disease (Ni-kshay Reports, 2025).

1.2 Rationale for the study

Tuberculosis remains a major public health issue in India, which carries one of the world's highest TB burdens. The country's National Strategic Plan (2017–2025) has set ambitious goals for reducing TB incidence, morbidity, and mortality, with the vision of achieving TB elimination by 2025. However, current data indicate that TB incidence and mortality rates are not declining as expected, suggesting that existing approaches may be insufficient to meet these objectives. Despite extensive control efforts, diagnostic delays

undermining efforts to reduce transmission and leading to continued TB morbidity and mortality.

While traditional compartmental models provide useful projections, they often lack the capacity to capture the complexity of TB transmission in diverse populations (Trauer et al., 2019). In contrast, agent-based models (ABMs) simulate interactions at the individual level, accounting for population heterogeneity and realistic social networks, which can reveal critical insights into disease spread.

Synthetic population data in ABMs is increasingly used in infectious disease modeling to enhance representation of demographic diversity (Willem et al., 2017). However, no ABMs using synthetic populations have yet been developed for TB in India, revealing a significant gap.

This study aims to address this gap by developing an ABM of TB transmission using a synthetic population for Thiruvananthapuram district. The model will simulate individual-level behaviors, diagnostic delays, and intervention coverage, offering insights into how these factors influence TB transmission over time. By developing this preliminary model, the study aims to set a foundation for more sophisticated, data-driven TB models that can better inform TB control strategies in India.

Chapter 2

Literature Review

2.1 History of Infectious Disease Modeling

Mathematical modeling and analysis are fundamental to the study of infectious disease epidemiology (Cuddington et al., 2005). In 1760, Bernoulli developed the first mathematical model of infectious disease transmission to evaluate the effectiveness of variolation, an early smallpox vaccination method (White and Enright, 2012). Subsequent contributions to infectious disease modeling were made by Hamer, Ross, and Hudson (Hamer, 1928; Ross and Hudson, 1916).

In 1927, Kermack and McKendrick introduced the SIR model, a framework using coupled non-linear differential equations to track changes in susceptible (S), infectious (I), and recovered (R) populations over time (Kermack et al., 1927). This framework, dividing individuals into S, I, and R compartments, became a cornerstone in epidemiological modeling, forming the foundation for analyzing infectious disease dynamics.

As the field evolved, extensions of the original SIR model emerged. Models such as SEIR, which includes an exposed (E) category of individuals, and SIRV, which accounts for vaccination (V), were developed to address the complexities of different diseases and

intervention strategies (Keeling et al., 2008).

Initially, most models were deterministic, considering the population as homogenous and offering a fixed output given certain parameters. However, with advancements in statistical methods and our understanding of infectious diseases (Kendall, 1956), there has been a shift towards stochastic compartmental models, which incorporate random variability to reflect the inherent uncertainty and randomness in disease transmission (Grassly and Fraser, 2008).

The rise of computational power has enabled development of advanced simulation-based approaches. ABMs, for example, simulate individual-level interactions within a population, capturing the detailed complexity of behaviors and social dynamics that drive disease spread (Richard and Lipsitch, 2024). This shift to simulation-based models marks a significant advancement, allowing epidemiologists to explore more complex disease patterns and intervention scenarios (Halloran et al., 2017).

2.2 Compartmental Models

Compartmental models provide a structured framework for understanding the dynamics of infectious disease transmission (Brauer, 2008). These models categorize a population into distinct compartments based on their infection status, allowing researchers to analyze the flow of individuals between these states over time. The basic compartments typically include Susceptible (S), Infected (I), and Recovered (R) individuals, commonly referred to as the SIR model. Individuals transition between these compartments at rates defined by specific parameters. Mathematically, these transitions are expressed through differential equations that describe how the number of individuals in each compartment changes over time (Tolles and Luong, 2020).

Traditional compartmental models, while widely used in infectious disease epidemiology, suffer from several limitations. These models often oversimplify the complexities of real-world disease transmission by homogeneity among individuals within compartments and they do not account for migration or changes in population size due to births and deaths. Also, these models struggle to capture the spatiotemporal dynamics of disease spread (Iranzo and Pérez-González, 2021).

2.3 Agent-Based Models

Agent-based models (ABMs) are computational modeling techniques that simulate the actions and interactions of individual agents to understand complex system(Bandini et al., 2009) . These models are being used across various fields, including social sciences, ecology, economics, and engineering, due to their ability to capture individual-level behaviors and emergent phenomena (Hammond, 2015).

Recently, ABMs has gained popularity in infectious disease modeling for its ability to simulate individual interactions and behaviors that drive disease spread (Willem et al., 2017). ABMs allow researchers to simulate individual level behaviours and interactions, offering detailed insights into disease transmission patterns over time. ABMs have been applied to model diseases like seasonal influenza, malaria, Ebola, and COVID-19 (Amadi et al., 2021; Ferguson et al., 2006; Halloran et al., 2008; Kerr et al., 2021; Siettos et al., 2015).

While traditional compartmental models are generally simpler, faster, and require fewer computational resources, ABMs offer a more nuanced, detailed approach at the cost of increased complexity and computational expense.

Synthetic populations are crucial for building effective ABMs in infectious disease

modeling. They provide realistic demographic representations, address data limitations, and allow for individual heterogeneity—all of which enhance the accuracy and reliability of ABMs (Huynh et al., 2013).

2.4 Synthetic Population

Synthetic populations are artificially created datasets that replicate the demographic and socio-economic profiles of actual populations in a defined region. Constructed using data from sources such as census records, surveys, and other statistical information, synthetic populations are designed to closely reflect the attributes of the target population, maintaining the joint distribution of these attributes for more accurate representation (Zhu et al., 2024). They are simplified microscopic representation of the real target population, capturing only the specific variables of interest for the purpose (Chapuis et al., 2022). It's especially valuable in fields like public health and urban planning, where collecting actual data may be difficult due to privacy, cost, or logistical constraints (Geard et al., 2012).

2.5 Tuberculosis Transmission Dynamics

Tuberculosis primarily spreads through airborne droplets released when an infected person coughs, speaks, or sneezes, with coughing being the most effective method of producing infectious aerosols (Turner and Bothamley, 2015). When individuals with infectious TB receive appropriate treatment, their infectiousness decreases rapidly, significantly reducing transmission risk (Dharmadhikari et al., 2014).

Close contacts of active TB cases are susceptible to becoming TB infected (TBI) and, if infected, to progressing to active TB, particularly within the first year after ex-

posure (Fox et al., 2013). Numerous studies now support the transmission of both drug-susceptible and drug-resistant TB within household settings (Shah et al., 2014; Zelner et al., 2014). Transmission is especially common in households where the index case is smear-positive and contacts are younger than 15 years (Zelner et al., 2014).

Demographic factors, such as age structure, play a crucial role in TB transmission dynamics. Mathematical models have shown that age-specific patterns of disease progression can impact transmission dynamics, with adult males often posing a higher transmission risk due to more extensive social interactions compared to other age groups (Dong et al., 2022; Miller et al., 2021).

To reduce the TB incidence, interrupting transmission is essential. Targeted interventions that reduce the infectiousness or duration of infectiousness in active TB cases, limit contact rates, and decrease the susceptibility of close contacts are possible strategies for transmission control (Liyew et al., 2023).

2.6 Global and National Efforts for TB Elimination

The World Health Organisation End TB Strategy (2015-2035), adopted by the World Health Assembly in 2014, aims to end the global TB epidemic. End TB 2035 targets are 95 percent reduction in TB deaths (compared with 2015), 90 percent reduction in TB incidence rate (compared with 2015) and no affected families facing catastrophic costs due to TB.

The goal of India's National Strategic Plan (2017-2025) is to achieve a rapid decline in burden of TB, morbidity and mortality while working towards elimination of TB in India by 2025. The plan outlines ambitious strategies and resource commitments to drive this decline, aligning with global End TB targets to achieve a TB-free India by 2030. However,

TB incidence and mortality rates are not on track with these targets. To make significant progress, enhanced coverage of drug-resistance testing and improved access to treatment are essential (Global Tuberculosis Report, 2024). Also, India has a high prevalence of TB infection (31.4% in adults), requiring targeted interventions to prevent progression to active TB and meet elimination goals (Panda, 2023).

2.7 TB Control Interventions and Their Effectiveness

Eliminating the tuberculosis infection (TBI) reservoir is now widely recognized by the scientific community as a pivotal intervention for achieving TB elimination targets (*WHO Consolidated Guidelines on Tuberculosis: Tuberculosis Preventive Treatment*, 2020). A main strategy in this effort is targeted screening for TBI, enabling early identification and treatment of individuals at high risk, which significantly reduces the progression to active TB (Matteelli et al., 2023). In India, tuberculosis preventive therapy (TPT) is provided to high-risk groups, such as individuals living with HIV, household contacts of TB patients, and other vulnerable populations, after ruling out active TB disease (India TB Report, 2024). Tuberculin skin tests (TST) and interferon-gamma release assays (IGRA) are widely used to identify TBI in these populations.

Complementing TBI screening, early identification of presumptive TB cases through active case-finding is a crucial strategy to accelerate case detection to achieve the goal of End TB strategy (Prathiksha et al., 2019).

Additional measures, including BCG vaccination, offer moderate, short-term protection against severe TB in children (Borgdorff et al., 2002). Nutritional support, especially in undernourished communities, plays an essential role in reducing TB incidence and transmission (Bhargava et al., 2023).

2.8 Objectives

2.8.1 Major Objective

- To examine how delays in identifying active TB affect pulmonary TB transmission in the synthetic population.

2.8.2 Minor Objectives

- To create a dynamic synthetic population of Thiruvananthapuram district that represents the real population.
- To evaluate how specific TB control measures, like active case finding and TB preventive treatment, reduce pulmonary TB incidence in the synthetic population.
- To analyze how changes in the coverage of TB control measures influence the progress of TB elimination in Thiruvananthapuram district using the synthetic population.

Chapter 3

Methods

3.1 Study Design

This study used a computational modeling design using agent-based modeling (ABM) to simulate TB transmission dynamics within a synthetic population.

An agent-based modeling approach was selected instead of a traditional compartmental model to capture the individual-level heterogeneity and complex contact patterns that drove TB transmission. Each "agent" was assigned with its own attributes (for example age, household) and was made to interact according to explicit daily schedules. This individual-based perspective was essential for the evaluation of interventions such as household screening and preventive therapy that operated at the person and household levels and depended on the timing and contacts.

TB transmission dynamics were simulated for the period 2025 – 2030. Model parameters were calibrated with TB notification data from 2021 – 2024. The primary outcome, TB incidence, was estimated under several intervention scenarios: the prevailing programme conditions; treatment initiation delay by four, six, or eight weeks after progression to active disease; and varying levels of TPT coverage among household contacts

of index cases.

3.2 Study Setting

The study focuses on Thiruvananthapuram district of Kerala, located on India's south-western coast (Figure 3.1). According to 2011 census, Thiruvananthapuram district had a population of 33,01,427, out of which 15,81,678 are males and 17,19,749 are females. There were 20,846 notified TB cases in the state of Kerala in the year 2024 and 2,636 of the notified cases were from Thiruvananthapuram district, which accounts for nearly 12.64 percent of the state's total (Ni-kshay Reports, 2025).

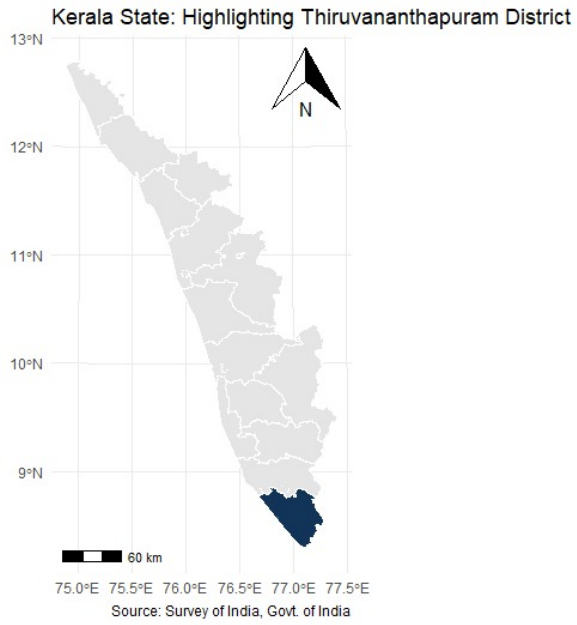


Figure 3.1: Map of Kerala highlighting Thiruvananthapuram district, the location of the present study.

3.3 Nikshay Data

3.3.1 Variables Extracted

From the Nikshay database, the following patient-level variables for all registered TB cases were extracted:

- **Diagnosis date**: date on which TB was bacteriologically or clinically diagnosed.
- **Treatment initiation date**: date on which anti-TB therapy began.
- **Age**: patient's age at the time of diagnosis.
- **Gender**: recorded as male, female or transgender.
- **Type of case**: categorized as pulmonary or extrapulmonary TB.
- **Treatment outcome**: final recorded outcome (for example, cured, treatment complete, treatment failure, regimen change, died).
- **Date of treatment outcome**: date on which treatment outcome was declared.

Only pulmonary TB cases were retained for further analysis, as they drive community transmission, which is the focus of this study.

3.3.2 Analysis of Data

To prepare the data for incidence estimation and age-structured modeling, the following steps were performed:

1. **Age grouping**: Patient ages were binned into five-year intervals (0–4, 5–9, 10–14, …, ≥80).
2. **Treatment outcome grouping**: Treatment outcomes were combined into three categories:
    - **Success**: "cured" or "treatment complete."

- **Failure**: "treatment failure" or "treatment regimen changed."
- **Died**: all records marked as "died."

3. **Notification rate calculation**: Using the diagnosis dates, both weekly and monthly pulmonary TB notification rates (cases per 100,000 population) were calculated, with population denominators from census.
4. **Disaggregated rates**: Notification rates were stratified by gender and by age group to identify differential patterns across these subpopulations.
5. **Initial prevalence estimates**: The baseline prevalence of active TB (both undetected and those already under treatment) per 100,000 population were estimated. This was used to assign initial states for the agents in the model.

One transgender individual was recorded in the Nikshay data for the year 2021; however, no further disaggregated analysis or simulation was conducted due to the small count.

3.3.3 Integration into the Model

The processed Nikshay data were incorporated into the ABM in the following ways:

- **Initial conditions**: The estimated numbers of undetected active cases and cases already on treatment were used to assign disease states to the model population at time zero, scaled to match the synthetic population of Thiruvananthapuram district.
- **Model Calibration**: Weekly notification time series from Nikshay were used to adjust the transmission probability parameter of the model. Iterative model simulations were ran, comparing simulated incidence counts against notification counts and finding the per-contact infection risk until model outputs closely tracked the reported weekly notification patterns.
- **Deriving Input Parameters for Treatment Outcomes**: Treatment outcome from

the Nikshay database was used to parametrize the model outcomes.

3.4 Synthetic Population Generation

3.4.1 Data Sources

- **Census Tables**: Age, sex, household size marginal data from the 2011 Census (Census tables | Government of India, 2011).
- **India Human Development Survey-II (IHDS-II) Microdata**: Individual level survey data for generating the synthetic population using iterative proportional updating (IPU) (Desai and Vanneman, 2018).
- **Gridded Population of the World**: Data for spatially distributing households based on population density across Thiruvananthapuram (Gridded Population of the World (GPW), v4 | SEDAC, 2016).
- **Sample Registration System Statistical Report**: Age specific fertility and mortality rates data to consider dynamic population adjustments over time (Government of India | Office of the Registrar General & Census Commissioner, India, 2025) .
- **Census Projections**: Sex ratio at birth data to assign sex of newborn individuals in the dynamic population (Population Projections for India and States, 2011-2036, 2019).

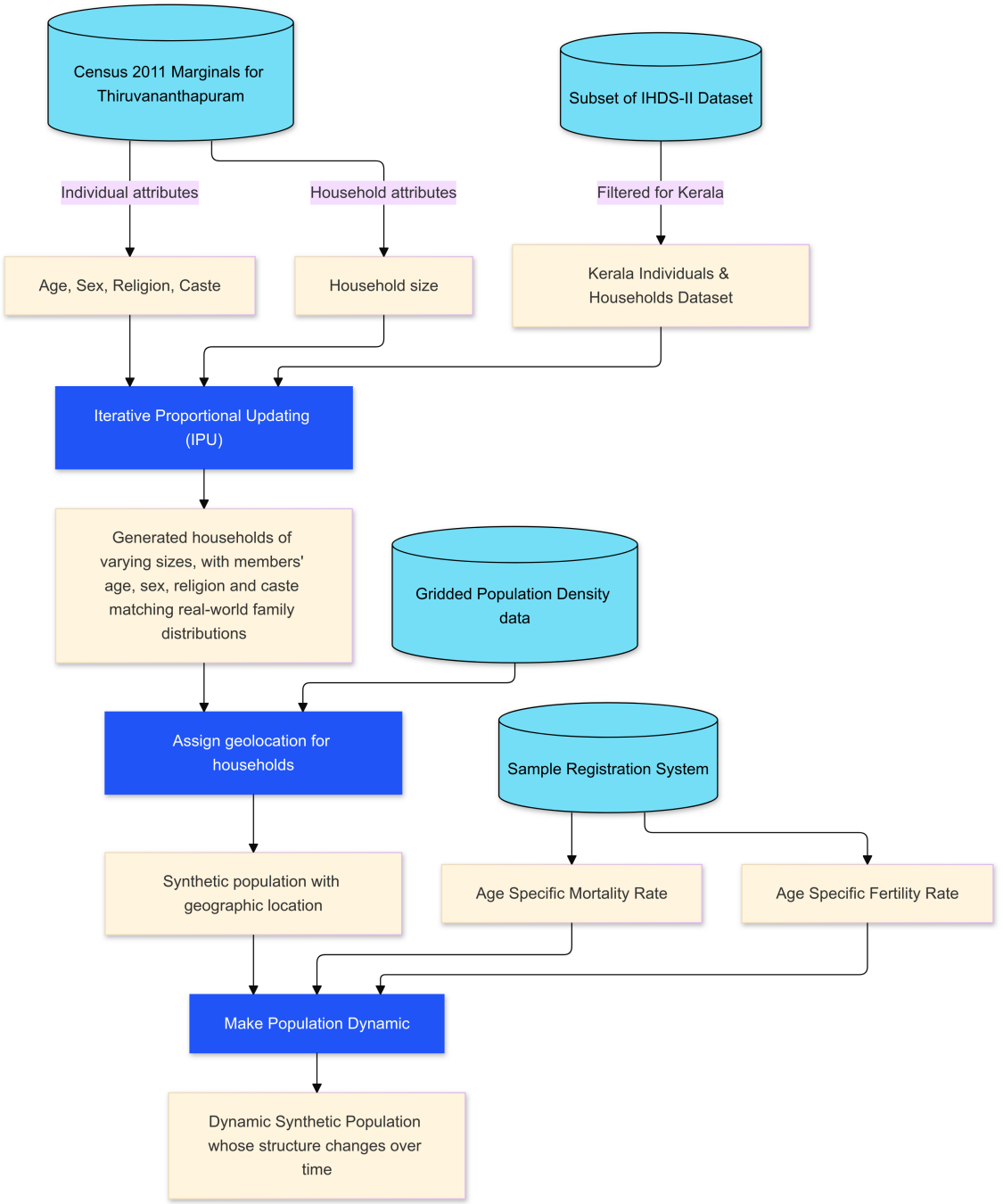


Figure 3.2: Diagrammatic representation of steps in generating synthetic population

3.4.2 Population Attributes

- Unique ID for individuals
- Age
- Sex

- Household ID
- School ID
- Work ID

3.4.3 Generation Technique

The baseline synthetic population for Thiruvananthapuram district was generated by adapting the framework originally developed by Neekhra et al. for Pune district (Neekhra et al., 2023). The Python-based program, was designed to use 2011 Indian Census marginals and IHDS II microdata to construct a scalable synthetic population across India (Figure 3.2). It was modified to reflect the demographic structure of Thiruvananthapuram district. Census-derived household compositions, age distributions, and sex ratios were used to assign agents to households, thereby producing a static population that reflected the 2011 district-level demographics.

3.4.4 Making Synthetic Population Dynamic

The synthetic population used in this study was constructed based on the 2011 Census of India and 2011-12 micro survey data. However, since the modeling period extends from 2021 to 2030, it was essential to account for demographic changes over time to ensure realism and temporal validity. To achieve this, the synthetic population was made dynamic by simulating births and deaths on an ongoing basis.

New individuals were added to the population through simulated births using age-specific fertility rates (ASFR), while existing individuals were removed based on age-specific mortality rates (ASMR). The sex of newborns was determined probabilistically according to the sex ratio at birth for each year (Figure 3.2).

ASFR and ASMR values up to the year 2020 were sourced from the Sample Registration System (SRS). For the period 2021–2030, future ASFR and ASMR values were forecasted using the Lee-Carter method, a standard demographic modeling approach that captures trends and age-specific changes in fertility and mortality. The forecasted values were visualized and found to be consistent with trends observed in similar epidemiological and demographic studies, providing confidence in their use for dynamic simulation.

Sex ratio values were obtained from Census Bureau projections for the years 2016, 2021, 2026, and 2031. For intervening years not directly available, values were linearly interpolated to maintain continuity.

Net migration was assumed zero over the modeling period. This assumption was made to simplify population dynamics in the absence of reliable annual migration data at the district level and to focus the model on internal demographic changes and disease transmission patterns.

By implementing these demographic dynamics, the ABM reflected realistic population growth, aging, and mortality patterns, thereby improving the accuracy of TB transmission projections over the simulation horizon.

For record-keeping and future analysis, a CSV file containing the population structure for each simulation year was saved separately.

3.5 Input Parameters

Table 3.1: Model parameters used for scenario simulations.

Description	Value	Source
Initial TB infection (TBI) percentage	20.5%	Mohan et al. (2022)
Initial active TB per 100,000	8.5	Nikshay data (2021 -2024)
Active TB treatment duration (days)	180	NTEP | Knowledge Base (2025)
Treatment success rate	84%	Nikshay data (2021 -2024)
TB-related mortality rate	12%	Nikshay data (2021 -2024)
Treatment failure rate	4%	Nikshay data (2021 -2024)
Household screening coverage	93.0%	India TB Report (2023)
Household screening delay after TB diagnosis (days)	14	Expert opinion
Screening test sensitivity (IGRA)	81%	Jonas et al. (2023)
TPT (TB preventive therapy) duration (days)	90	NTEP | Knowledge Base (2025)
TPT completion rate	81%	Global Tuberculosis Report (2024)
TPT efficacy	99.9%	Sterling et al. (2011)
Mean TB notification delay (days)	78	Sebastian et al. (2021)
Men working percentage	76.1%	Annual Report, Periodic Labour Force Survey (2024)
Women working percentage	39.6%	Annual Report, Periodic Labour Force Survey (2024)

3.6 Modeling (Agent-Based Modeling)

3.6.1 Agents and Their Attributes

The model used the synthetic population generated as mentioned above. Each synthetic individual (“agent”) was represented by a unique identifier and the following char-

acteristics:

3.6.1.1 Demographic Attributes

- Age (in years)
- Sex
- Household ID
- Workplace ID
- School ID

3.6.1.2 Epidemiological Attributes

- Disease state: Susceptible, TBI, Active TB (undetected), Treatment
- Time since TBI: in days
- Whether progress to active TB from TBI or not: TRUE or FALSE
- Active TB time: in days
- Treatment time: in days
- Treatment notification time
- Will infect others or not: TRUE or FALSE
- TPT status: Not on TPT, On TPT, Completed TPT
- TPT time: in days
- Completed TPT: TRUE or FALSE

Demographic processes were driven by age-specific fertility rates (ASFR) to add new agents, age-specific mortality rates (ASMR) to remove agents, and the sex ratio at birth to assign sex of newborns.

3.6.2 Agent Schedule

Each agent was assigned a daily activity pattern determined by age. Children and adolescents aged 4–18 years were assumed to attend school during daytime and to return to their households at night. Adults aged 18–65 years were considered eligible for employment; the proportion employed was sampled from the labour-force participation proportions reported in the Periodic Labour Force Survey. Employed adults were modelled as spending the working day at a workplace and the remainder of the time at home, whereas unemployed adults stayed at home throughout the day. Infants younger than 4 years and adults older than 65 years were likewise assumed to remain in the household at all times. Within each location (household, school, workplace), a well-mixed contacts were assumed such that every susceptible agent present shares risk from all infectious agents present.

3.6.3 Disease States and Transitions

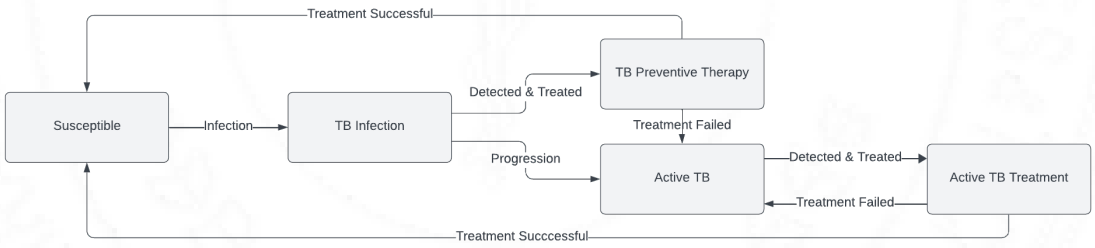


Figure 3.3: Diagrammatic representation of the model for TB transmission dynamics and treatment.

Four disease states were represented in the model: Susceptible, Tuberculosis Infection (TBI), Active TB, and Treatment (Figure 3.3). Agents in the Susceptible state acquired infection through contact with infectious individuals. Once infected, agents entered the non-infectious TBI state. While in TBI, two alternative pathways were possible. First,

agents could initiate a 90-day course of TPT, for which the probability of completing treatment was specified in the model parameters; successful completion returned the agent to the Susceptible state, whereas non-completion left the agent in TBI. Second, agents who did not receive or who failed to complete preventive therapy faced a risk of progression to active disease.

Progression from TBI to Active TB was governed by age-independent cumulative risks: 10 percent of all latent infections were assumed to progress over the lifetime, with 45 percent of those progressions occurring during the first year after infection, 62 percent by the end of the second year, 83 percent by year five, 99 percent by year twelve, and the remainder thereafter. These cumulative probabilities were converted into annual hazard rates and subsequently into per-time step transition probabilities.

Upon transition to the Active TB state, agents were assumed to be infectious until diagnosis. After a modeled diagnostic delay, treatment was initiated and agents entered the Treatment state, where a 180-day regimen was simulated. Individuals remained infectious for the first two weeks of treatment and were then considered non-infectious. Treatment outcomes were modeled as follows: cure returned the agent to the Susceptible state; treatment failure resulted in a return to Active TB with all relevant timers reset; and tuberculosis-related death permanently removed the agent from the simulated population.

3.6.4 TBI Screening and Intervention

In the baseline scenario, household contacts of each newly notified pulmonary TB case were screened 14 days after the index patient's diagnosis, with a coverage of 93 percent as reported in the India TB Report. Screening included an interferon-gamma release assay, the sensitivity of which was set at 81 percent. Contacts who tested positive for latent

infection were started on tuberculosis-preventive therapy, whereas those in whom active disease was detected were placed immediately on standard TB treatment.

3.6.5 Time Steps and Simulation Period

The model advanced in half-daily time steps. At each time step, agents were relocated according to their schedule (home, school, workplace), and all disease-state transitions are recorded. The simulation ran from 1st January 2025 to 31st December 2030.

3.6.6 Model Calibration

The model was calibrated over the period 2021–2024 using 2024 TB notification data from the Nikshay database. The per-contact transmission probability was determined by minimizing the root mean squared error (RMSE) between the simulated weekly incidence and the observed weekly notifications from the Nikshay database for 2021–2024.

3.6.7 Scenarios Analysed

- Current Scenario
  - Delay in diagnosis of 78 days
  - 93 percent household contacts (HHC) screened for TBI and started on TPT
  - 14 days to screen for TBI in HHC after a person diagnosed with active TB
- Delay in diagnosing and starting ATT
  - 8 weeks
  - 6 weeks
  - 4 weeks
- Coverage in percentage of HHC screened for TBI and started on TPT

- 50 percent
- 0 percent

3.6.8 Implementation Details

The model was implemented in Julia v1.11.4 programming language in order to benefit from that language's high-performance capabilities for large-scale, individual-based simulations. At the end of each simulated day, a comma-separated values (CSV) file was written that recorded the daily counts of agents in each disease state—Susceptible, TBI, Active TB, and Treatment—together with the numbers of new TBI and new active TB cases arisen during that day.

The codebase was built for maximum flexibility and reusability, enabling researchers to simulate different geographic settings or intervention scenarios without altering the core code. By simply updating input files—such as alternative synthetic populations, Nikshay-derived TB notification data, or new intervention strategies—and adjusting parameters, users can adapt the model to diverse contexts with ease. All file paths and model parameters were written in a single, human-readable configuration file for easy scenario setup. This plug-and-play design allows for both the replication of published analyses and the rapid exploration of new hypotheses, all without touching the underlying source code.

Chapter 4

Results

This study aimed to investigate how delays in tuberculosis diagnosis and preventive therapy coverage among household contacts influence pulmonary TB transmission dynamics within a synthetic population of Thiruvananthapuram district, Kerala. In this chapter, the findings are presented in relation to the study objectives, with interpretations grounded in the model framework and contextualized against existing literature.

The chapter begins by summarizing the key characteristics of the input data, particularly the Nikshay notifications that informed the calibration process and the demographic projections used for synthetic population generation. Subsequently, the calibration performance of the agent-based model is described. The baseline simulation results under current program conditions are then detailed, followed by the results from various intervention scenarios targeting diagnostic delay and household contact management(Global Tuberculosis Report, 2024).

4.1 Nikshay Data Analysis

4.1.1 Notification Trends

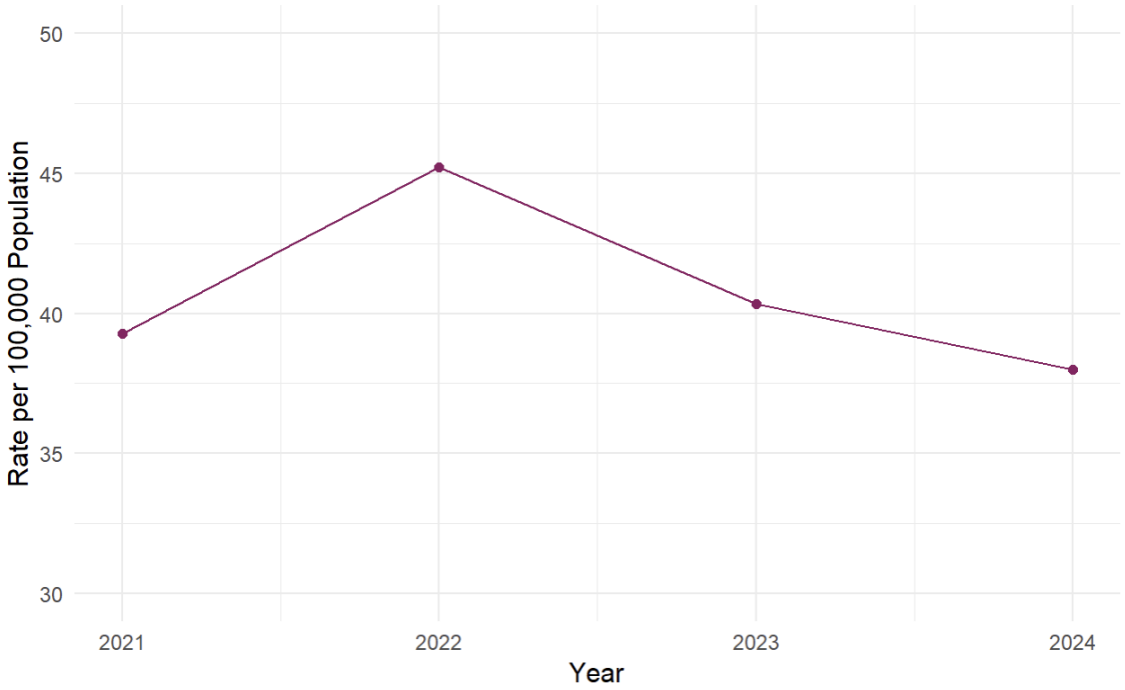


Figure 4.1: Annual pulmonary TB notification rate

Figure 4.1 shows the annual pulmonary tuberculosis (TB) notification trend extracted from the Nikshay database which was used for model calibration. An initial rate of 39.3 per 100,000 population was recorded in 2021. This figure rose to a peak of 45.2 per 100,000 in 2022, representing a 15.1 percent increase relative to 2021. Thereafter, the rate declined to 40.3 per 100,000 in 2023 (a 10.8% reduction from the 2022 peak) and fell further to 38.0 per 100,000 in 2024, yielding an overall decrease of 1.3 per 100,000 (3.3%) compared with the baseline year. In absolute terms, 1,297 pulmonary TB cases were notified in 2021, the count increased to 1,493 in 2022, then decreased to 1,332 in 2023 and to 1,254 in 2024.

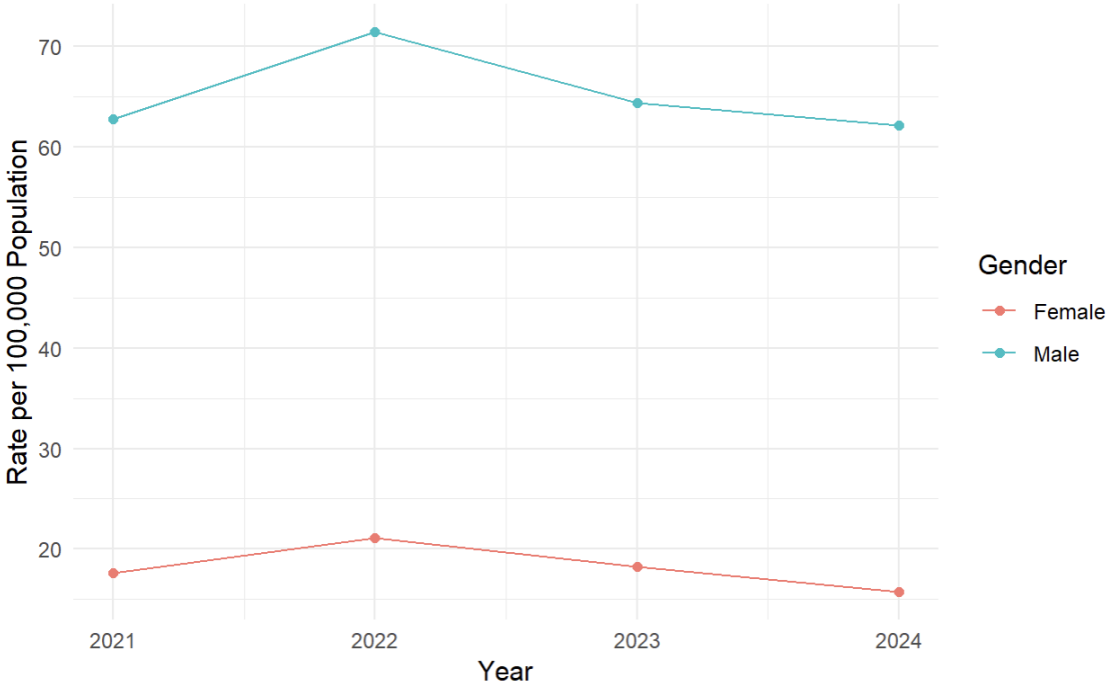


Figure 4.2: Gender trends in annual pulmonary TB notification rate

When notification rates were disaggregated by gender, markedly higher burdens were observed among men throughout the study period (Figure 4.2). Among males, the pulmonary tuberculosis notification rate increased from 62.8 per 100,000 population in 2021, to a peak of 71.4 per 100,000 in 2022, before declining to 64.4 per 100,000 in 2023 and 62.1 per 100,000 in 2024. Corresponding case counts were 993 in 2021, 1,130 in 2022, 1,018 in 2023, and 983 in 2024.

Female notification rates were consistently lower. An initial rate of 17.6 per 100,000 in 2021 rose to 21.1 per 100,000 in 2022, then fell to 18.3 per 100,000 in 2023 and 15.8 per 100,000 in 2024. The number of female cases notified were 303 in 2021, 363 in 2022, 314 in 2023, and 271 in 2024.

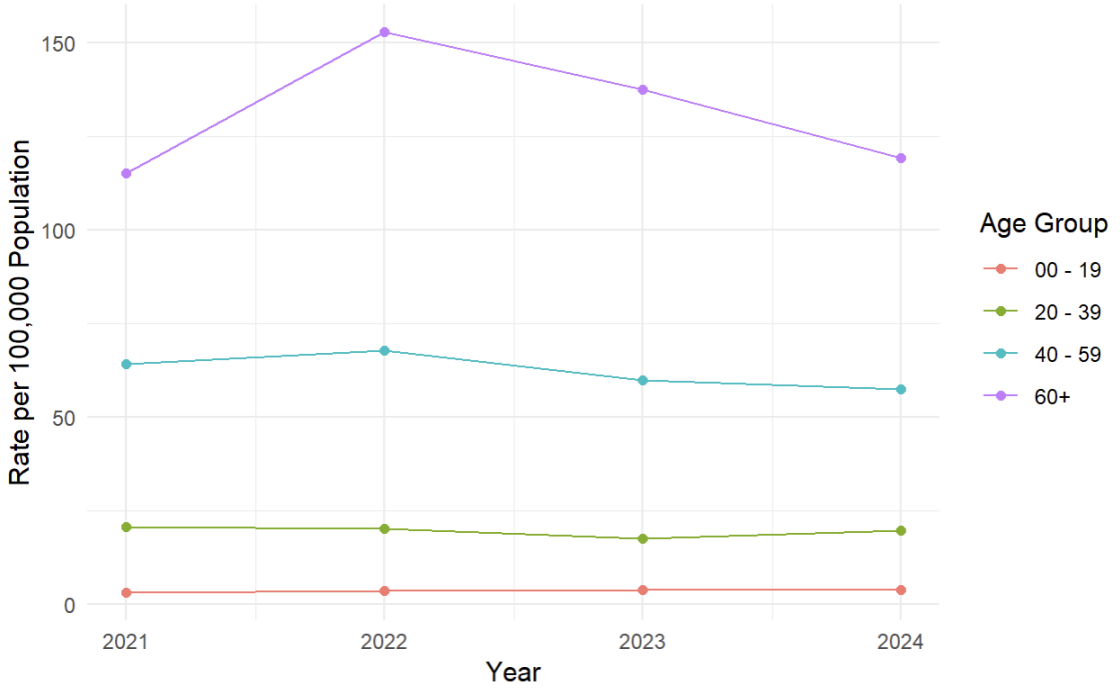


Figure 4.3: Age trends in annual pulmonary TB notification rate

Figure 4.3 shows the age-stratified trends in pulmonary tuberculosis notifications derived from the Nikshay database. Throughout the period, the ≥ 60-year group consistently recorded the highest notification rates, whereas the 0–19-year group registered the lowest. In 2021, rates of 115.0, 64.1, 20.5, and 3.1 cases per 100,000 population were documented in the ≥ 60, 40–59, 20–39, and 0–19-year groups, respectively. All groups except the 0–19-year reached their maximum in 2022: the ≥ 60-year rate increased to 153.0 per 100,000, the 40–59-year group to 67.8 per 100,000, and the 20–39-year group to 20.2 per 100,000. The 0–19-year group continued to rise, attaining its peak of 3.8 per 100,000 in 2023. By 2024, notification rates had declined across all age groups. Notification rate fell to 119.0, 57.3, 19.7, and 3.7 per 100,000 in the ≥ 60, 40–59, 20–39, and 0–19 year groups, respectively.

4.1.2 Treatment Outcome

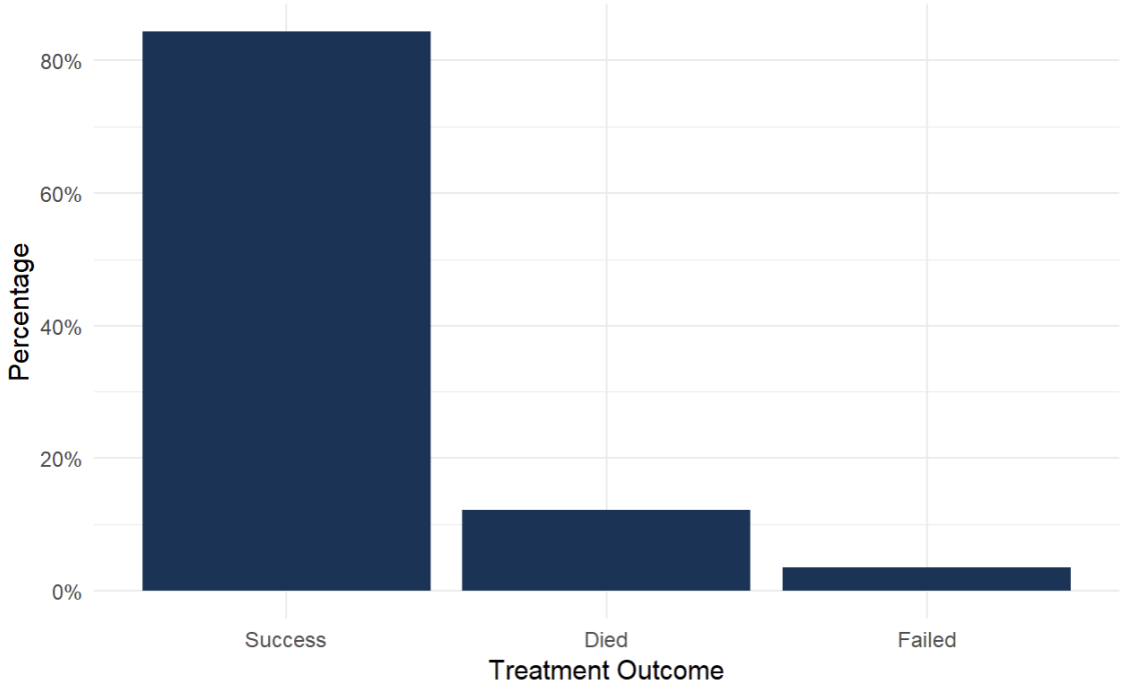


Figure 4.4: Distribution of treatment outcome (2021-2024)

A total of 5,376 treatment records were retrieved from the Nikshay database for the study period. Outcome information was missing for 761 patients, representing 14.2 percent of all entries. Among the 4,615 records with documented outcomes (Figure 4.4), treatment was reported as successful in 3,893 patients (84%), death occurred in 560 patients (12%), and treatment failure was recorded in 162 cases (4%).

4.2 Synthetic Population Generation

4.2.1 Comparisons with Census

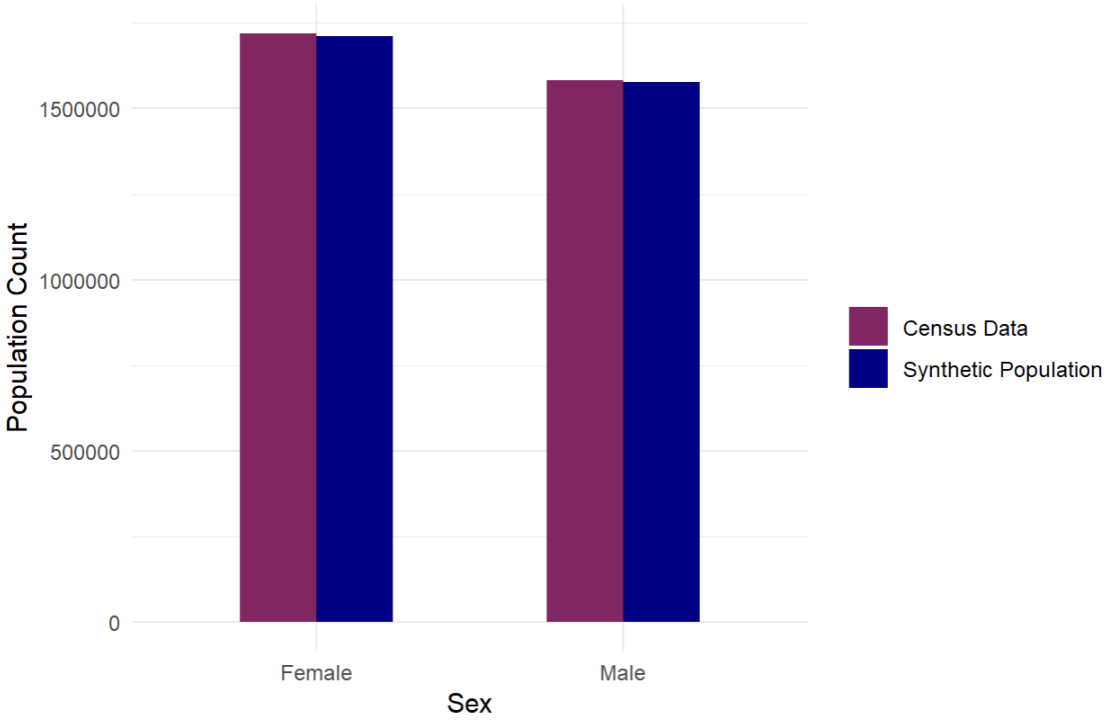


Figure 4.5: Comparison of sex distribution between census data and the synthetic population (2011)

The census recorded 1,719,749 females and 1,581,678 males, whereas the synthetic dataset contained 1,712,212 females and 1,576,798 males (Figure 4.5). The differences was only 0.44 percent for females and 0.31 percent for males. The female-to-male ratio was identical in both the population with 1,087 females per 1,000 males in the census versus 1,086 per 1,000 in the synthetic population.

The synthetic population replicated the census age structure very closely. For each five-year age groups, the counts in the synthetic dataset differed from the census by fewer than 2,300 individuals, representing deviations of less than one per cent (Table 4.1 and Figure 4.6). The greatest proportional departure was noted in the 20–24-year group, where the synthetic count (264,738) was 0.8 percent lower than the census figure (266,969).

Table 4.1: Age group-wise comparison of census data and synthetic population.

Age Group	Census Data	Synthetic Population	Absolute Difference	Percent Difference (%)
00 - 04	216,891	216,193	698	0.3
05 - 09	227,143	227,362	219	0.1
10 - 14	263,246	263,584	338	0.1
15 - 19	255,313	253,738	1575	0.6
20 - 24	266,969	264,738	2231	0.8
25 - 29	258,879	257,337	1542	0.6
30 - 34	245,395	244,954	441	0.2
35 - 39	257,169	256,246	923	0.4
40 - 44	241,071	240,877	194	0.1
45 - 49	245,603	244,798	805	0.3
50 - 54	202,305	203,057	752	0.4
55 - 59	188,288	187,734	554	0.3
60 - 64	147,631	146,998	633	0.4
65 - 69	108,793	108,676	117	0.1
70 - 74	74,670	74,126	544	0.7
75 - 79	48,531	48,441	90	0.2
80+	49,806	50,151	345	0.7

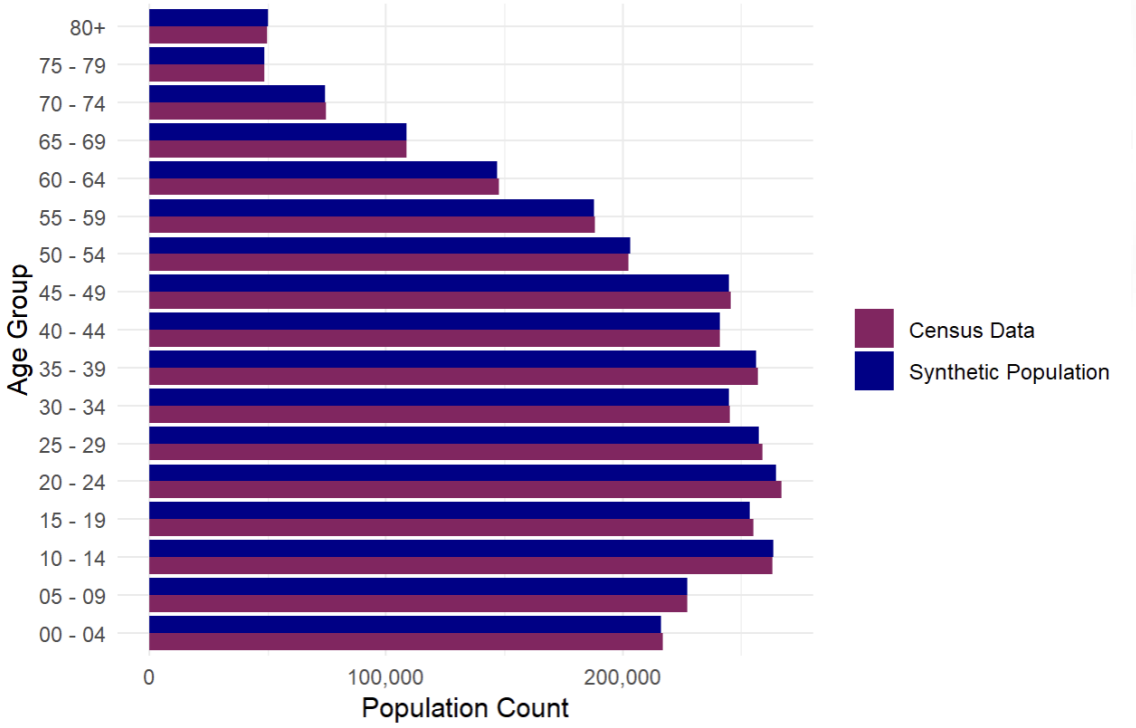


Figure 4.6: Age distribution: census data vs synthetic population

4.3 Demographic Projections

4.3.1 Age-Specific Fertility Rates (ASFR)

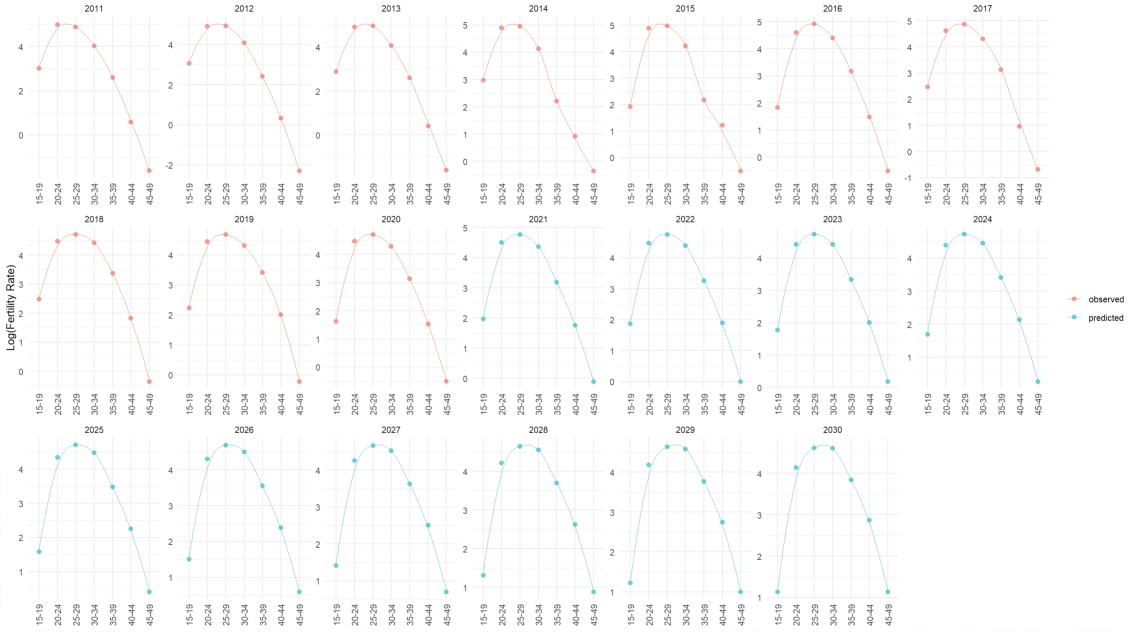


Figure 4.7: Age-specific fertility rates by age groups across years (log scale)

Age-specific fertility rates (ASFRs) were plotted on a logarithmic scale for each year from 2011 to 2030 which was used for the construction of the dynamic synthetic population. Observed data from the Sample Registration System (2011 – 2020) displayed the characteristic inverted-U pattern, with the highest fertility initially centred in the 20–24-year age group (Figure 4.7). Over the decade, the apex of the curve gradually shifted to the 25–29-year age group, indicating a postponement of child-bearing to slightly older ages.

The Lee–Carter projections for 2021 – 2030 preserved this overall shape while extending the shift. Throughout the projection period, the 25–29-year group remained the modal age of child-bearing, and the magnitude of fertility declined at every age. The 15–19-year cohort consistently exhibited the lowest rates and experienced the sharpest relative decline.

By 2030, fertility was markedly concentrated in women aged 25–34 years, with substantially reduced contributions from both adolescent (15–19 years) and older (40 years and above) age groups.

4.3.2 Age-Specific Mortality Rates (ASMR)

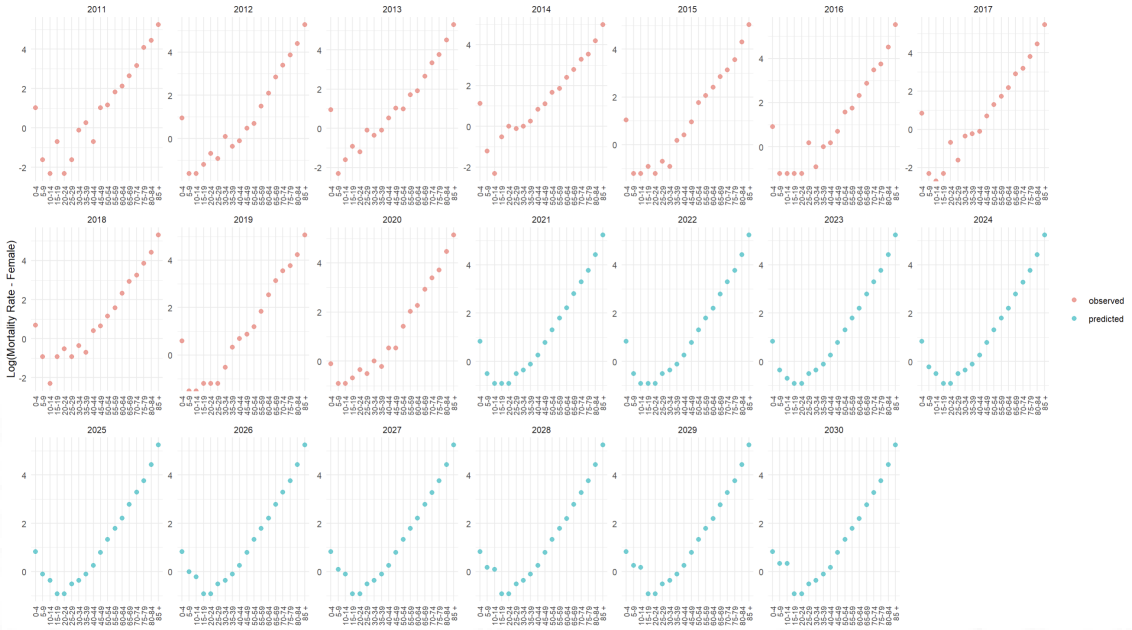


Figure 4.8: Female age-specific mortality rates by age groups across years (log scale)

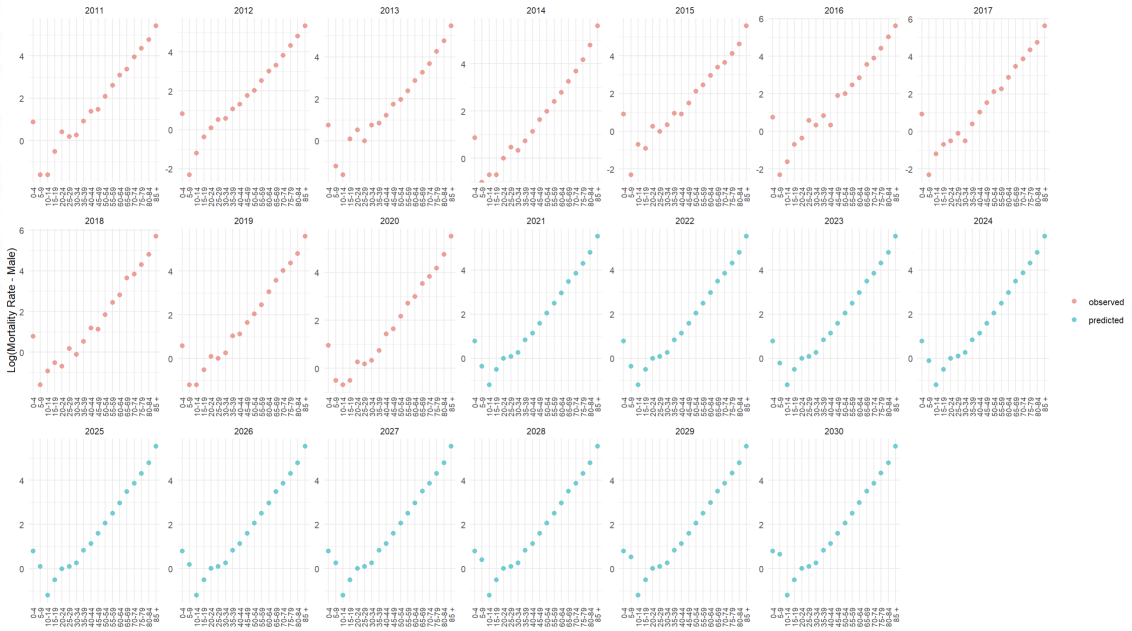


Figure 4.9: Male age-specific mortality rates by age groups across years (log scale)

Age-specific mortality rates (ASMRs) for males and females were plotted on a logarithmic scale for each year from 2011 to 2030 which was used for the construction of the dynamic synthetic population. The observed data from the Sample Registration System (2011 – 2020) exhibited the characteristic J-shaped pattern: mortality was highest in infancy, fell sharply during childhood and young adulthood, and then increased approximately log-linearly with advancing age (Figure 4.8 and Figure 4.9) . The Lee–Carter projections for 2021 – 2030 preserved the characteristic J-shaped age pattern.

4.3.3 Sex Ratio at Birth

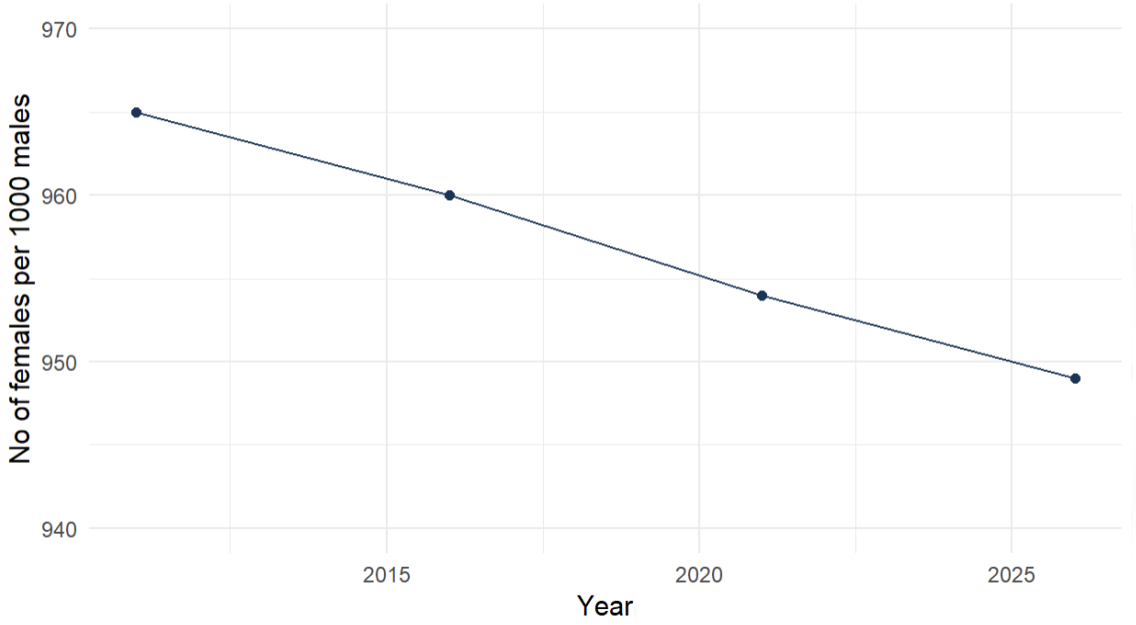


Figure 4.10: Sex ratio at birth

Across the period, the sex ratio at birth showed a steady decline, falling from 965 females per 1,000 males in 2011 to 960 in 2016, 954 in 2021, and 949 in 2026 (Figure 4.10).

4.4 Dynamic Synthetic Population Characteristics

4.4.1 Population Growth and Age Structure

The synthetic-population pyramids for 2016, 2021 and 2026 confirmed a progressive demographic ageing of the study population. In 2016 the structure still retained a modestly pyramidal form, with children and adolescents (0–14 years) representing a visibly broader share than the oldest groups. By 2021 the base had contracted further due to the continued decline in recent births.

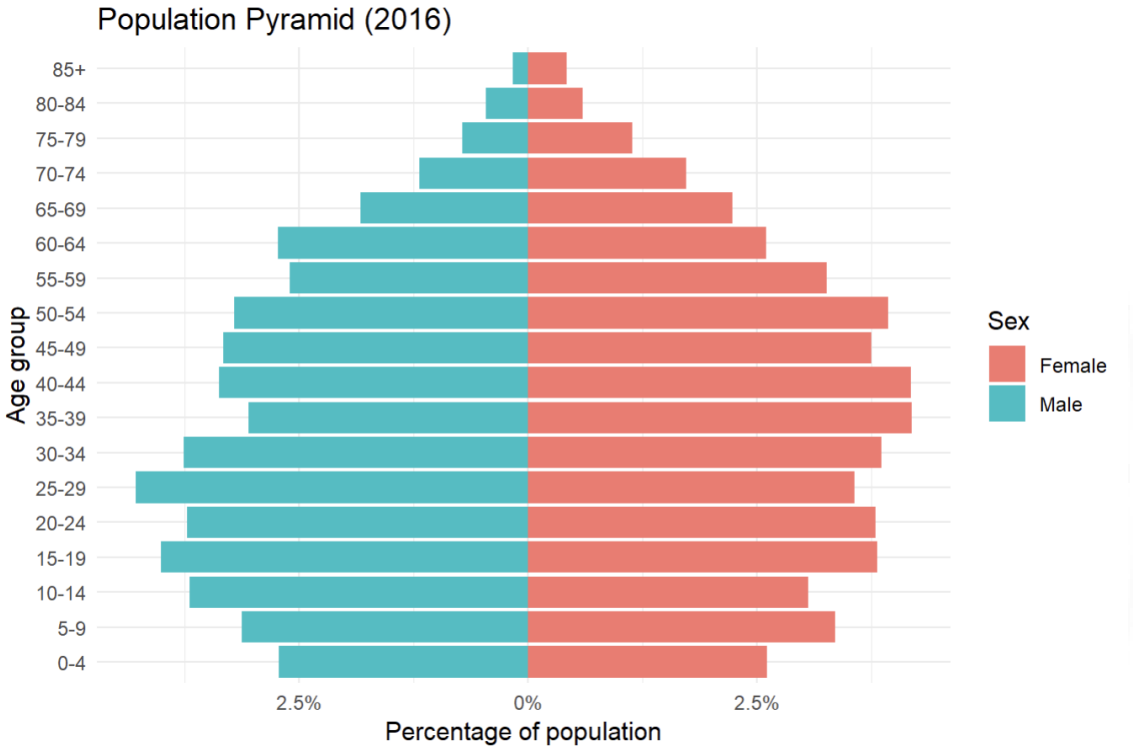


Figure 4.11: Population pyramid for the year 2016 generated using the dynamic synthetic population model

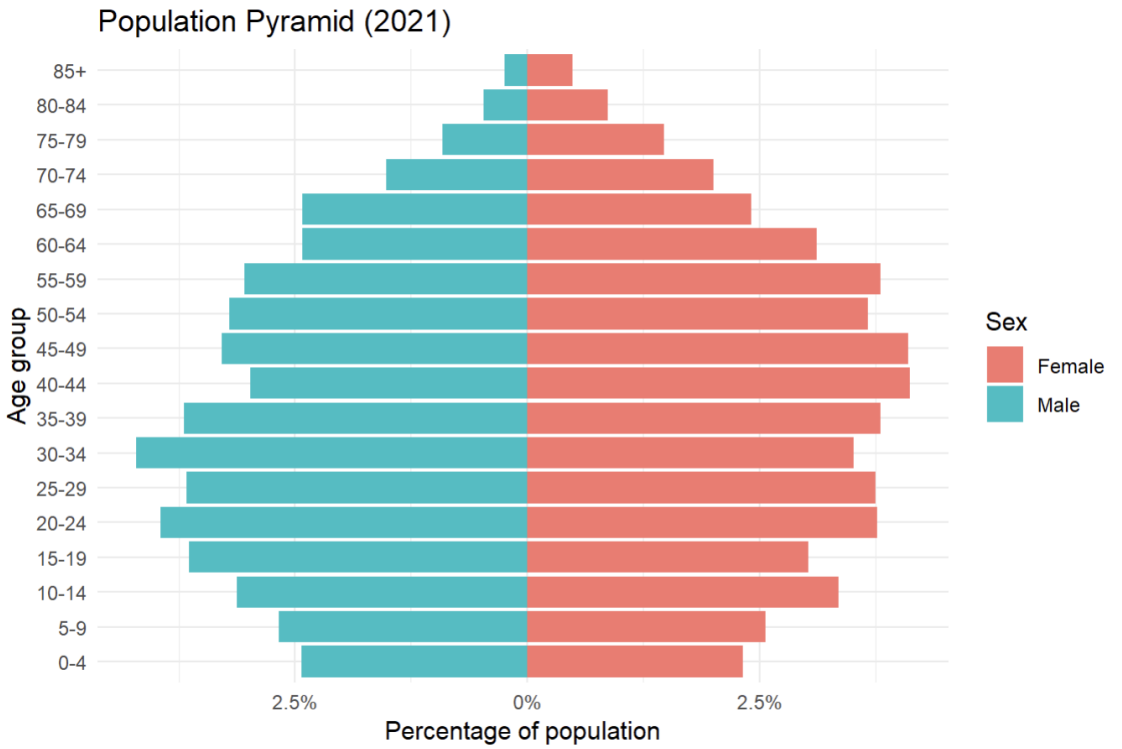


Figure 4.12: Population pyramid for the year 2021 generated using the dynamic synthetic population model

The 2026 pyramid assumed an increasingly column-like profile: the middle age groups constituted the widest segments, the proportions aged 15 years or younger reduced, and the upper age groups (65 years and above) expanded relative to earlier years, reflecting both declining fertility and improving survival at older ages. Sex distributions remained broadly symmetrical through middle age, while a females became more from about 70 years onward, consistent with higher female longevity.

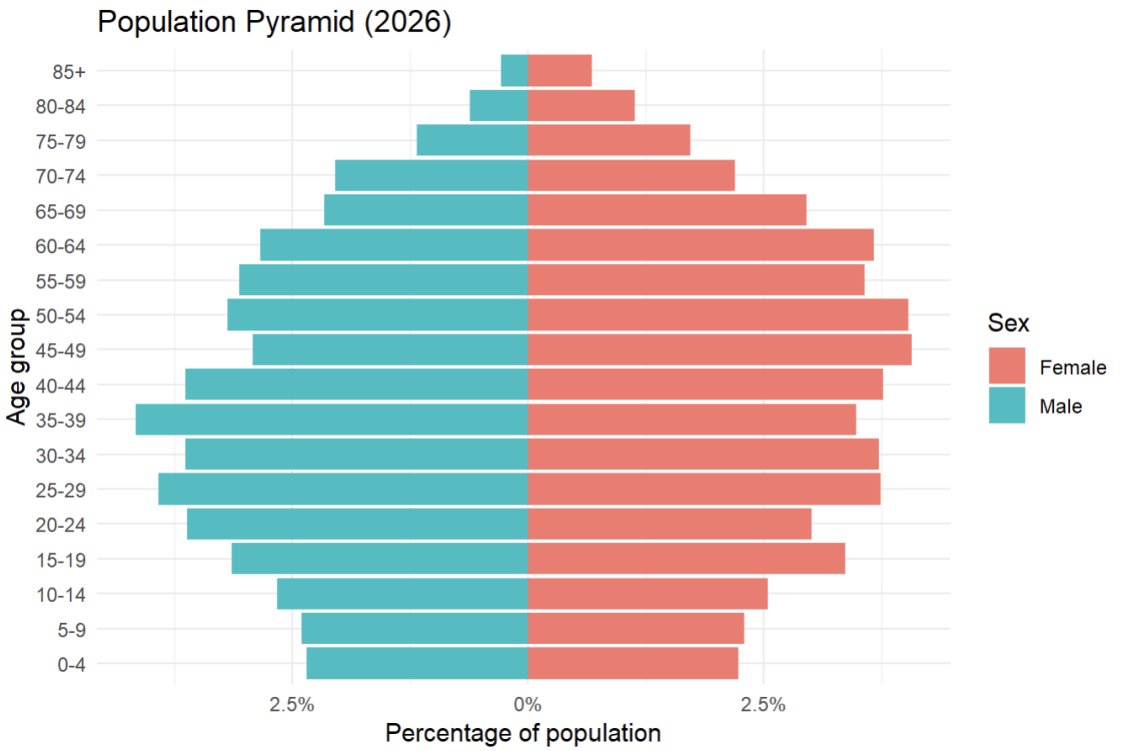


Figure 4.13: Population pyramid for the year 2026 generated using the dynamic synthetic population model

4.5 Model Calibration

Calibration was performed by varying the per-contact probability of infection, β, over the interval 0 – 1 and comparing the resulting weekly incident–case counts with Nikshay notifications. To shorten computation time, the agent-based model was executed on a random 20 percent sample of the synthetic population; all other parameters remained fixed at their baseline values.

Goodness of fit was assessed with the root-mean-square error (RMSE). As shown in Table 4.2, RMSE values ranged from 2.528 to 4.821 cases across the nine candidate settings. The lowest error (RMSE = 2.528) was obtained when $\beta = 0.2$, followed by $\beta = 0.3$ and $\beta = 0.4$, which produced marginally higher errors (RMSE $\approx$ 2.530–2.533).

Table 4.2: Root-mean-square error comparing simulated weekly TB notifications to Nikshay data for candidate values of the per-contact infection probability (β).

Model	RMSE
beta0.1	2.946899
beta0.2	2.528185
beta0.3	2.533980
beta0.4	2.530082
beta0.5	3.563337
beta0.6	3.041381
beta0.7	3.874682
beta0.8	4.821143
beta0.9	4.674961

Because $\beta = 0.2$ minimised the discrepancy between simulated and reported notifications, this value was adopted for all subsequent full-scale simulations.

4.6 Baseline (Current Program) Scenario

4.6.1 Pulmonary TB Incidence

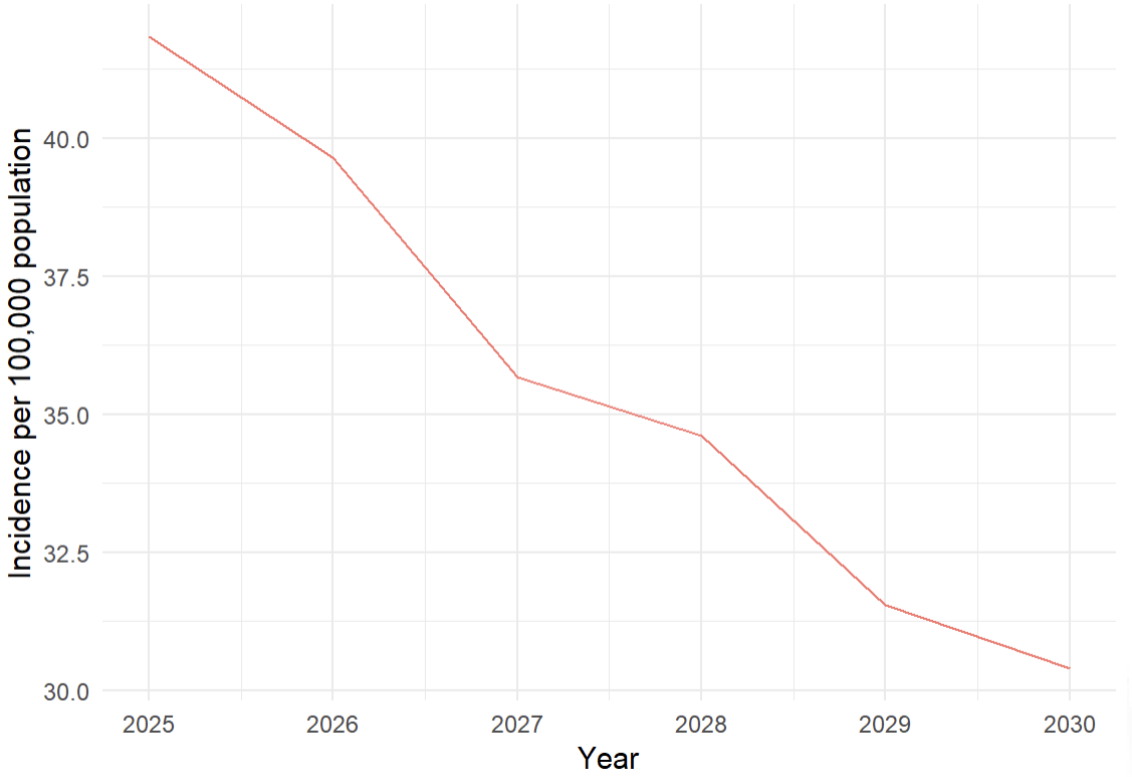


Figure 4.14: Projected annual pulmonary TB incidence per 100,000 population under the baseline scenario, 2025–2030.

Under the baseline scenario, the modelled annual pulmonary TB incidence declined from 41.84 cases per 100,000 population in 2025 to 30.38 cases per 100,000 population in 2030 (Figure 4.14 , Table 4.3). The largest year-on-year reduction was observed between 2026 and 2027, when incidence fell by 10.01 percent (from 39.64 to 35.67 per 100,000), corresponding to a 14.75 percent decrease relative to the 2025 baseline. A more modest decline of 2.97 percent occurred between 2027 and 2028, followed by an 8.86 percent decrease in 2029 and a further 3.68 percent reduction in 2030. Overall, incidence under the baseline assumptions was projected to decrease by 27.38 percent over the five-year period (Table 1).

Table 4.3: Annual pulmonary TB incidence per 100,000 population under the baseline scenario, with percentage decreases from the previous year and from the base year (2025).

Year	Incidence per 100,000	Decrease from previous year (%)	Decrease from base year (%)
2025	41.84	-	-
2026	39.64	5.27	5.27
2027	35.67	10.01	14.75
2028	34.61	2.97	17.28
2029	31.54	8.86	24.61
2030	30.38	3.68	27.38

4.6.2 Prevalence of Pulmonary TB

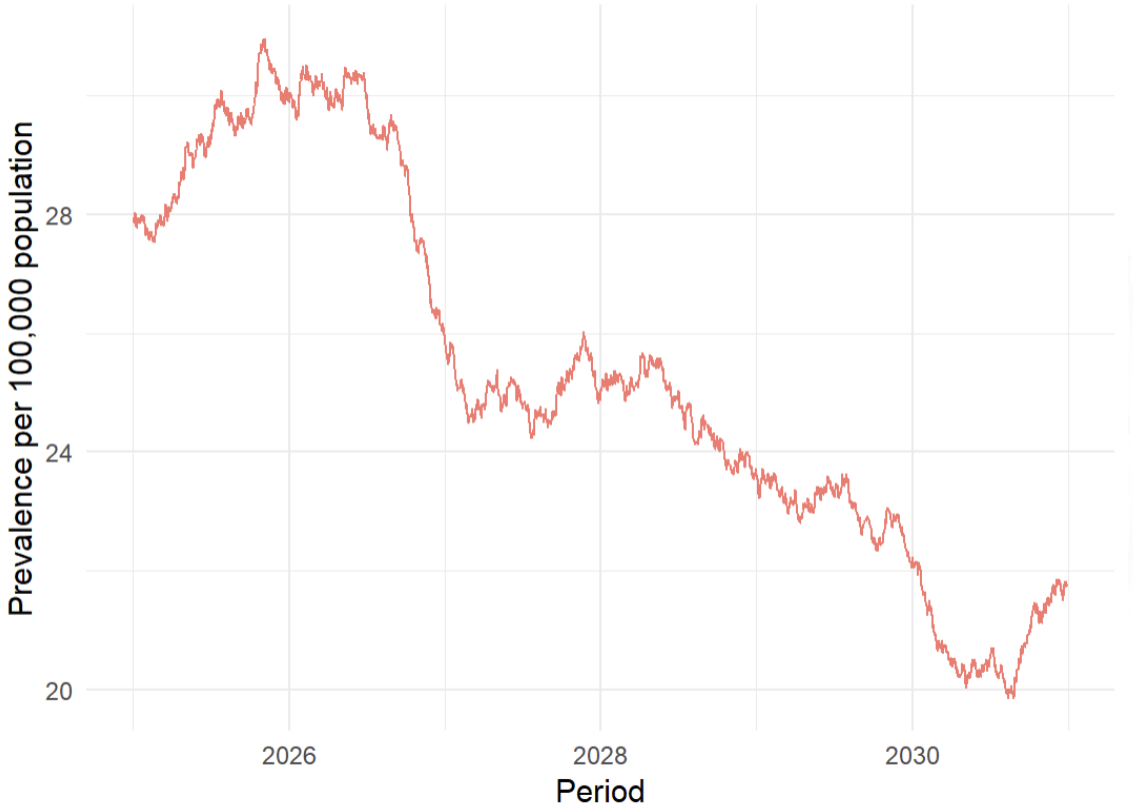


Figure 4.15: Projected temporal trend in the prevalence of active pulmonary TB per 100,000 population under the baseline diagnostic-delay (78 days) and household-contact screening (93%) scenario, 2025–2030.

Under the baseline scenario, pulmonary TB prevalence exhibited an initial increase before entering a sustained downward trajectory (Figure 4.15). Prevalence rose from ap-

proximately 27.9 cases per 100,000 population in early 2025 to a peak of 30.1 per 100,000 towards end of 2025. Thereafter, a marked decline happened, with prevalence falling to roughly 24.5 per 100,000 by mid 2027. Although short-term fluctuations were apparent, the overall trend remained downward, reaching a low of about 20.0 per 100,000 in early 2030 before a minor rebound to approximately 21.7 per 100,000 by late 2030.

4.7 Intervention Scenario Results

4.7.1 Delay in Diagnosing and Starting TB Treatment

4.7.1.1 Pulmonary TB Incidence

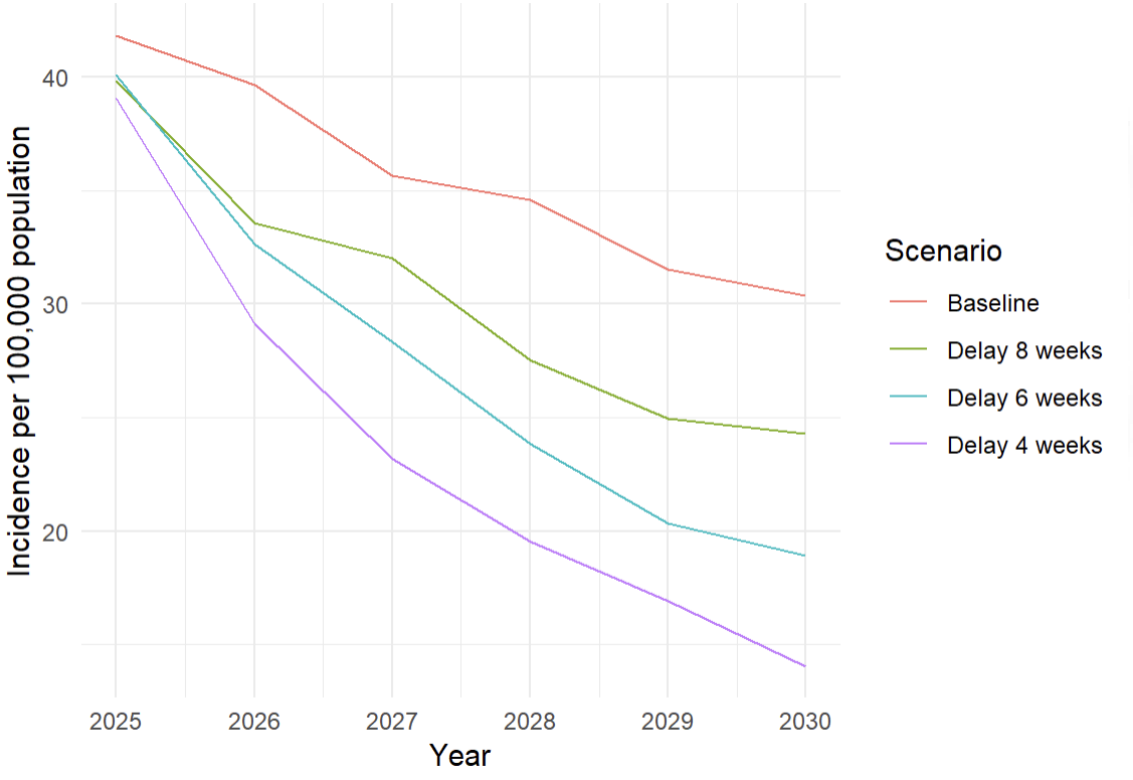


Figure 4.16: Projected annual pulmonary TB incidence per 100,000 population under scenarios of 4, 6 and 8-week diagnostic delays, 2025–2030.

Under the different diagnostic-delay scenarios, annual pulmonary TB incidence declined in all cases, with the magnitude of reduction being inversely related to the length

of delay (Figure 4.16, Table 4.4). An 8-week delay to diagnosis lead to the decrease of incidence from 39.82 to 24.31 cases per 100,000 (38.96 % reduction). A 6-week delay yielded decline from 40.11 to 18.93 cases per 100,000 (52.81 % reduction). When the delay to diagnosis was shortened to 4 weeks, incidence fell from 39.09 cases per 100,000 in 2025 to 14.05 cases per 100,000 in 2030, corresponding to a 64.05 percent reduction from the base year.

Table 4.4: Annual pulmonary TB incidence per 100,000 population under differing diagnosis-delay scenarios, with percentage decreases from the previous year and from the base year (2025).

Scenario	Year	Incidence per 100,000	Decrease from previous year (%)	Decrease from base year (%)
Baseline	2025	41.84	-	-
Baseline	2026	39.64	5.27	5.27
Baseline	2027	35.67	10.01	14.75
Baseline	2028	34.61	2.97	17.28
Baseline	2029	31.54	8.86	24.61
Baseline	2030	30.38	3.68	27.38
Delay 8 weeks	2025	39.82	-	-
Delay 8 weeks	2026	33.58	15.66	15.66
Delay 8 weeks	2027	32.01	4.68	19.61
Delay 8 weeks	2028	27.51	14.07	30.92
Delay 8 weeks	2029	24.97	9.23	37.30
Delay 8 weeks	2030	24.31	2.65	38.96
Delay 6 weeks	2025	40.11	-	-
Delay 6 weeks	2026	32.62	18.68	18.68
Delay 6 weeks	2027	28.33	13.15	29.37
Delay 6 weeks	2028	23.86	15.79	40.52
Delay 6 weeks	2029	20.33	14.81	49.33
Delay 6 weeks	2030	18.93	6.87	52.81
Delay 4 weeks	2025	39.09	-	-
Delay 4 weeks	2026	29.14	25.44	25.44
Delay 4 weeks	2027	23.17	20.51	40.73
Delay 4 weeks	2028	19.54	15.65	50.01
Delay 4 weeks	2029	16.91	13.44	56.73
Delay 4 weeks	2030	14.05	16.92	64.05

4.7.1.2 Prevalence of Pulmonary TB

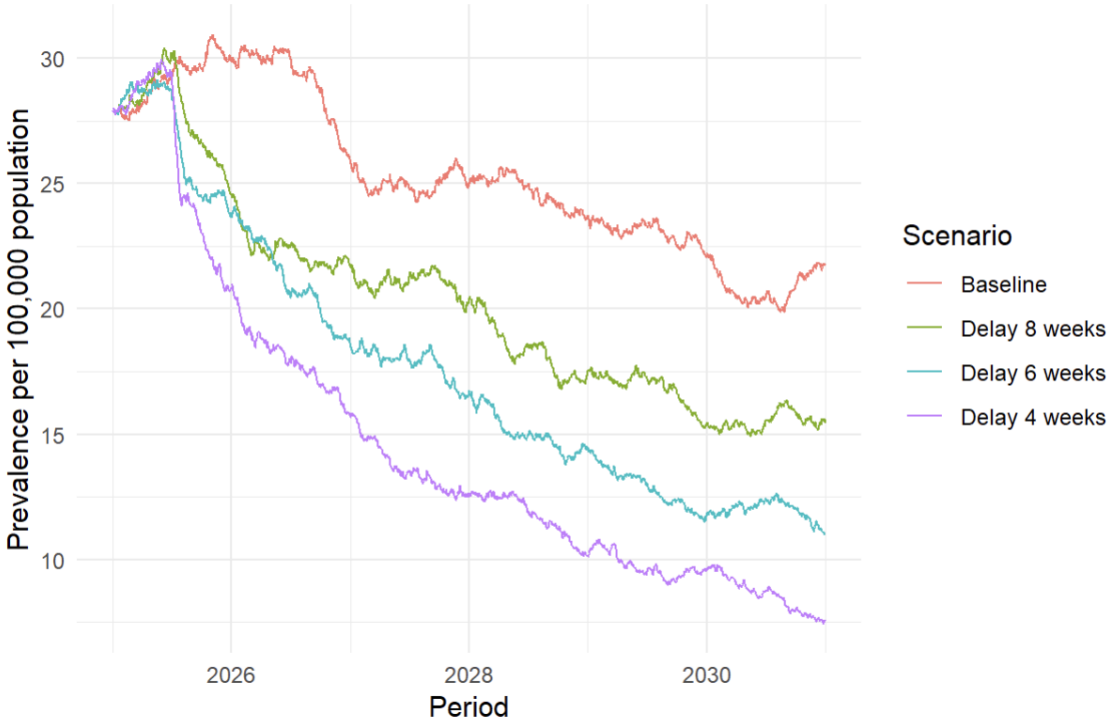


Figure 4.17: Projected prevalence of pulmonary TB per 100,000 population under baseline and reduced diagnostic-delay scenarios, 2025–2030.

Under all scenarios, pulmonary TB prevalence was projected to decline over the 2025–2030 period, with the rate and magnitude of decline being strongly dependent on the length of diagnostic delay (Figure 4.17). Under the baseline assumption (78 day delay), prevalence fell from approximately 27.9 cases per 100,000 population in early 2025 to 21.7 cases per 100,000 by late 2030, reflecting a gradual but consistent reduction. When the delay was shortened to four weeks, a markedly steeper decline was seen. Prevalence dropped to roughly 7.5 cases per 100,000 by 2030. Intermediate outcomes were observed for six- and eight-week delays, with prevalences of about 11.0 and 15.6 cases per 100,000, respectively towards late 2030.

4.7.1.3 Proportion of TBI

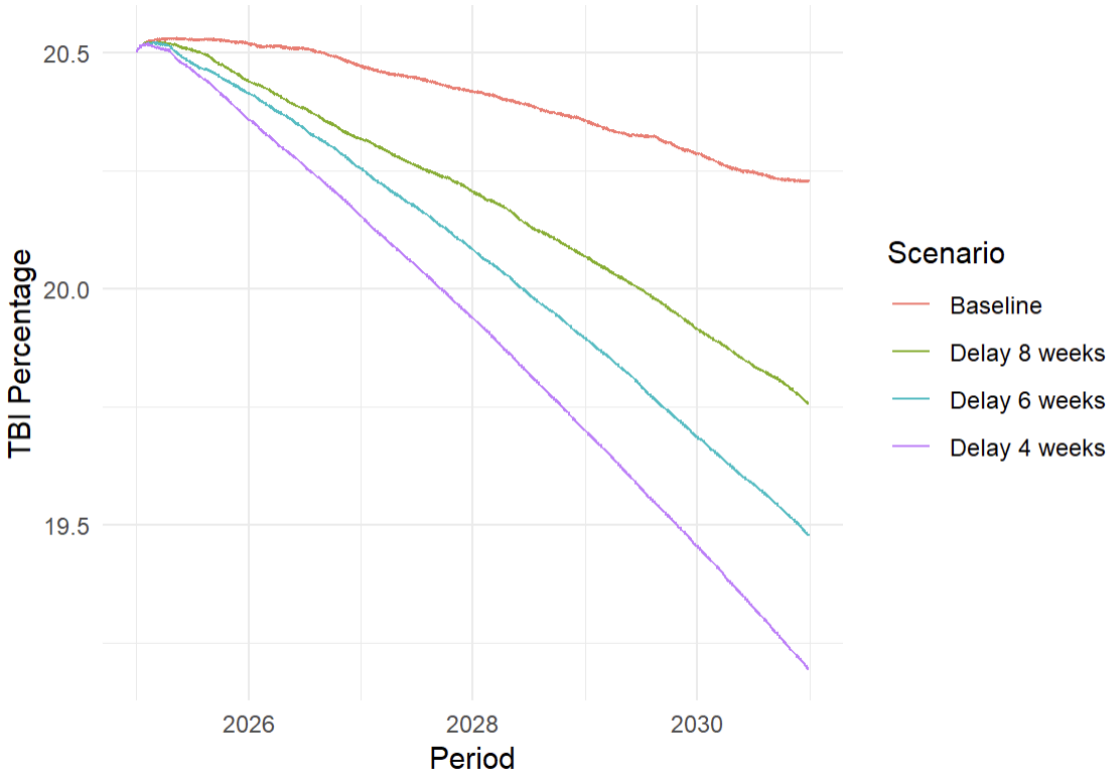


Figure 4.18: Projected proportion of the population with tuberculosis infection under baseline and reduced diagnostic-delay scenarios, 2025–2030.

Under all diagnostic-delay scenarios, the proportion of the population with tuberculosis infection was projected to decline over the 2025–2030 period, with the steepest reduction observed under the shortest delay (Figure 4.18). Under the baseline (78 day) delay, TBI prevalence was projected to decrease only marginally, from 20.5 percent in 2025 to 20.2 percent by 2030. When the delay to diagnosis was shortened to four weeks, the proportion infected fell more substantially, reaching 19.2 percent at the end of the simulation, a drop of 1.3 percentage points relative to the baseline year. Intermediate declines were seen under six- and eight-week delays, with final TBI prevalence of 19.5 percent and 19.8 percent, respectively.

4.7.2 Household Contacts Screened for TBI and Started on TPT

4.7.2.1 Pulmonary TB Incidence

The program claims a household contact screening of 93% currently, but there could be some discrepancies in this claim. Therefore some realistic scenario on household contact screening are attempted below.

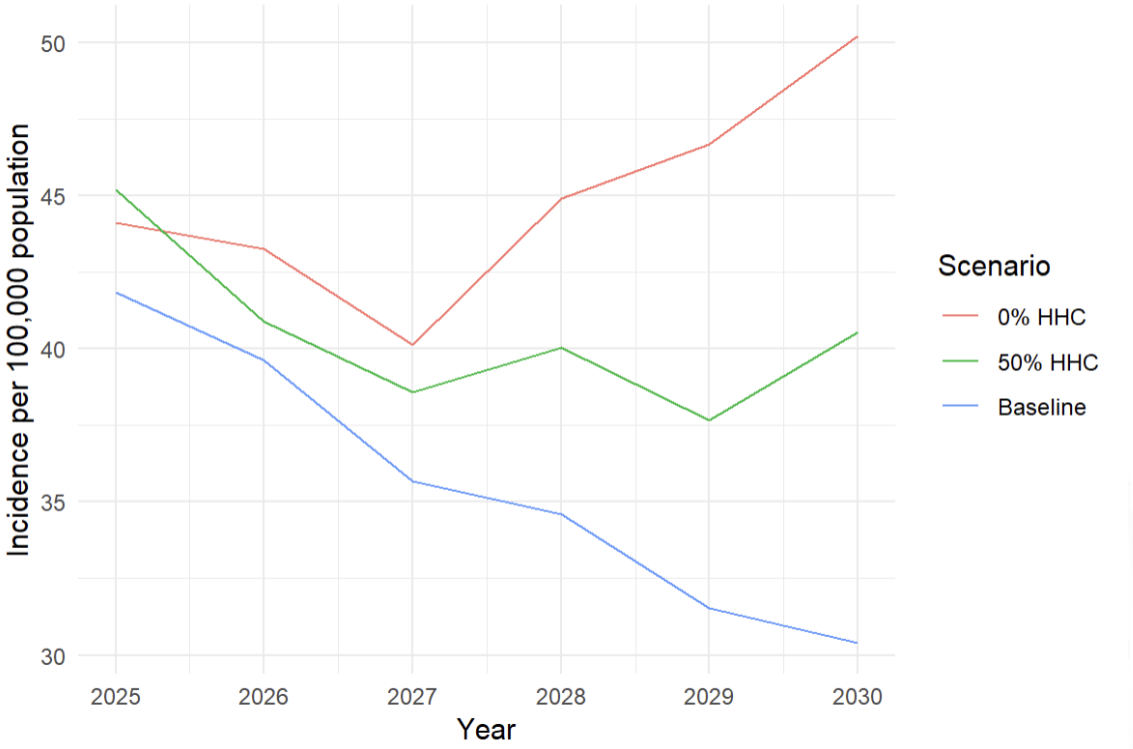


Figure 4.19: Projected annual pulmonary TB incidence per 100,000 population under scenarios of 0 percent and 50 percent household-contact screening, 2025–2030 (baseline scenario for comparison).

As shown in Figure 4.19 and Table 4.5, HHC screening reduced to 0 percent, incidence initially declined from 44.12 to 40.13 cases per 100,000 between 2025 and 2027 (-1.95% in 2026; -7.23% in 2027), but then increased by 11.91 percent in 2028, 3.90 percent in 2029, and 7.61 percent in 2030, yielding a net increase of 13.81 percent relative to the 2025 baseline. In contrast, a 50 percent HHC-screening coverage scenario produced an early decline to 38.57 per 100,000 by 2027 (–9.52% in 2026; –5.64% in 2027), followed

by a 3.77 percent rise in 2028, a further 5.92 percent drop in 2029, and a 7.63 percent increase in 2030. Overall, the 50 percent coverage scenario achieved a 10.30 percent reduction from the base year.

Table 4.5: Annual pulmonary TB incidence per 100,000 population under 0 percent and 50 percent household-contact screening scenarios, with percentage changes from the previous year and from the base year (2025).

Scenario	Year	Incidence per 100,000	Decrease from previous year (%)	Decrease from base year (%)
Baseline	2025	41.84	-	-
Baseline	2026	39.64	5.27	5.27
Baseline	2027	35.67	10.01	14.75
Baseline	2028	34.61	2.97	17.28
Baseline	2029	31.54	8.86	24.61
Baseline	2030	30.38	3.68	27.38
50% HHC	2025	45.18	-	-
50% HHC	2026	40.87	9.52	9.52
50% HHC	2027	38.57	5.64	14.63
50% HHC	2028	40.02	-3.77	11.41
50% HHC	2029	37.65	5.92	16.65
50% HHC	2030	40.52	-7.63	10.30
0% HHC	2025	44.12	-	-
0% HHC	2026	43.26	1.95	1.95
0% HHC	2027	40.13	7.23	9.04
0% HHC	2028	44.91	-11.91	-1.79
0% HHC	2029	46.67	-3.90	-5.77
0% HHC	2030	50.22	-7.61	-13.81

4.7.2.2 Prevalence of Pulmonary TB

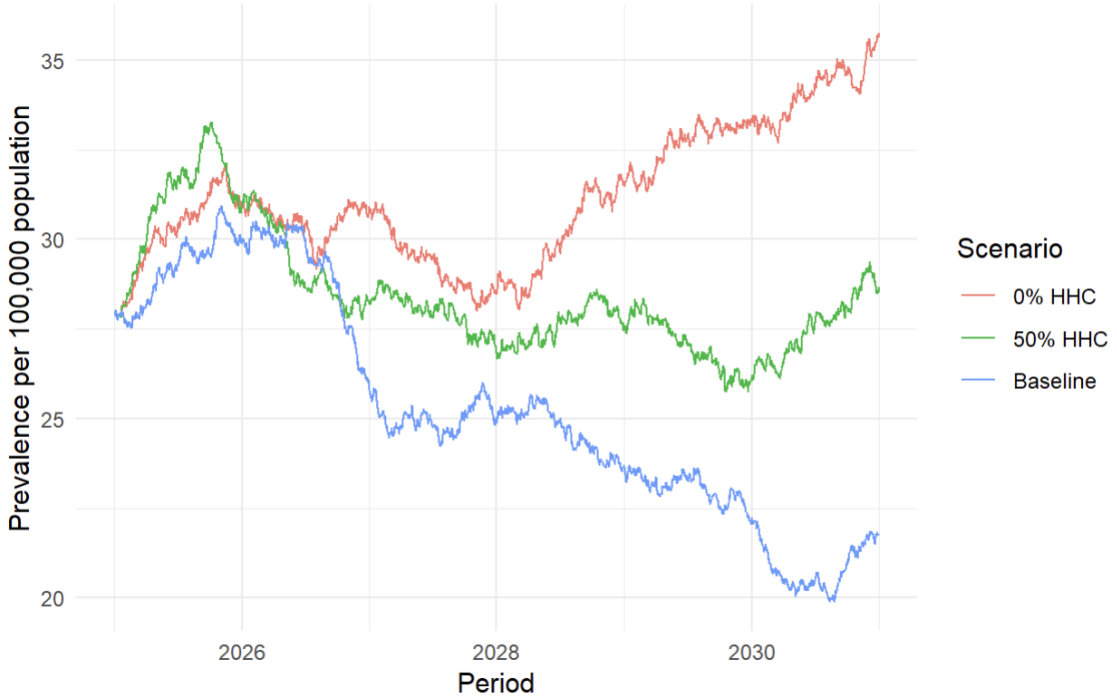


Figure 4.20: Projected prevalence of pulmonary TB per 100,000 population under scenarios of 0 percent, 50 percent, and 93 percent household-contact screening, 2025–2030.

Under scenarios of reduced household-contact screening, the projected prevalence of active pulmonary TB diverged markedly from the steady decline seen under baseline conditions (93% HHC screening; Figure 4.20). Initially, all scenarios began at approximately 27.9 cases per 100,000 in early 2025. Under the baseline assumption, prevalence peaked near 30.9 per 100,000 in mid-2026 before declining steadily to reach 21.7 per 100,000 by late 2030. When HHC screening was halved (50% coverage), prevalence rose to about 32.8 per 100,000 in mid-2026, then fell to a trough of roughly 26.1 per 100,000 in 2029, and subsequently rebounded to 28.7 per 100,000 by the end of the simulation. In the absence of any HHC screening, prevalence climbed to around 31.0 per 100,000 in 2026, dipped modestly to about 28.2 per 100,000 by 2027, and then increased continuously, culminating at 35.8 per 100,000 in 2030.

4.7.2.3 Proportion of TBI

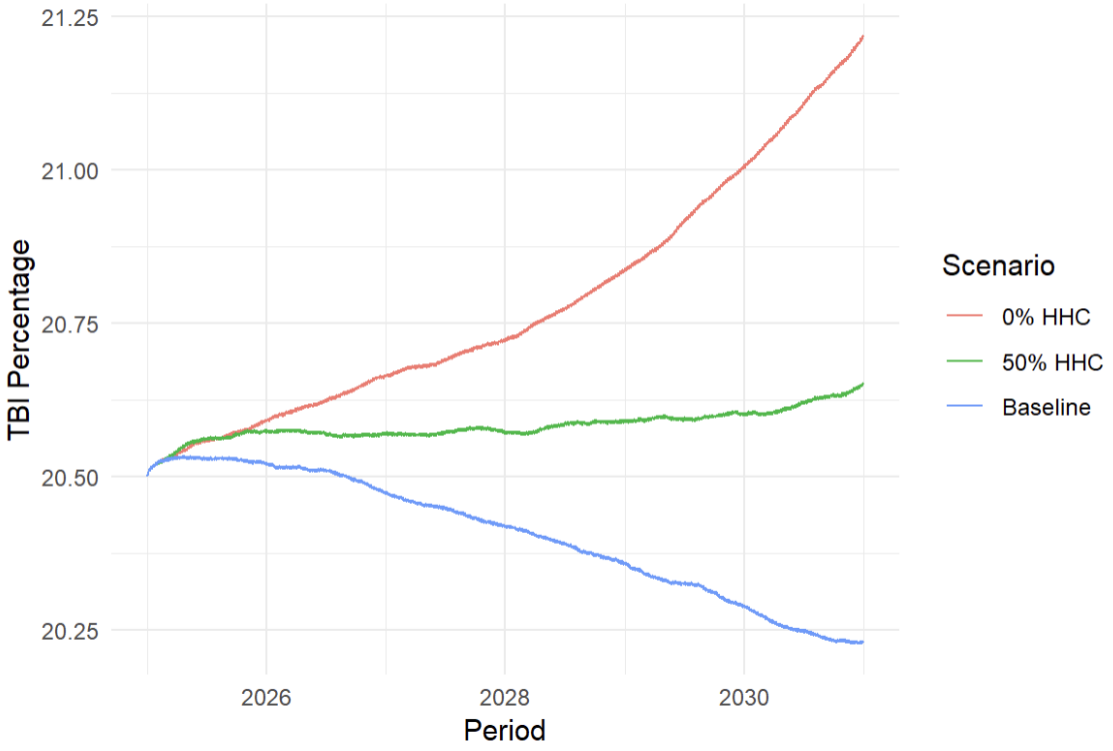


Figure 4.21: Projected proportion of the population with tuberculosis infection under scenarios of 0 percent, 50 percent, and 93 percent (baseline) household-contact screening, 2025–2030.

Under reduced household-contact screening, the TB infection diverged markedly from the steady decline observed under the baseline assumption (Figure 4.21). In the baseline scenario, TBI prevalence fell slightly from approximately 20.5 percent in early 2025 to 20.2 percent by late 2030. When HHC screening was halved to 50 percent, infection prevalence initially remained near 20.6 percent and then climbed upward, reaching about 20.7 percent at the end of the simulation period. With no household-contact screening, the proportion infected rose continuously from roughly 20.6 percent in 2025 to 21.2 percent in 2030.

Chapter 5

Discussion

The present study set out to examine how diagnostic delays and preventive therapy coverage among contacts influence pulmonary TB transmission in a synthetic population of Thiruvananthapuram, Kerala. In this chapter, we interpret our findings in light of the study objectives and current literature, and discuss their implications for TB control policy. We first summarize key aspects of the input data (Nikshay notifications and other secondary sources) and the construction of the synthetic population. We then discuss how diagnostic delay and tuberculosis preventive therapy (TPT) screening were modeled and how these factors impacted transmission dynamics in our simulations. Next, we compare our findings with similar agent-based modeling studies from India and internationally (2018–2025) and highlight the strengths and limitations of our approach. Finally, we offer conclusions and recommendations for policy and future research.

5.1 Nikshay and Secondary Data Analysis

We based our model inputs on an analysis of TB notification data from the Nikshay portal and related secondary data sources. In 2021-24, Thiruvananthapuram district's annual pulmonary TB notification rate approximately 40 per 100,000 population. This relatively

low incidence reflects Kerala's better-than-national average TB control performance. The Nikshay data revealed that most pulmonary TB cases occurred in older adults and a higher burden among men (consistent with national patterns where 55% of TB cases are in men (Global Tuberculosis Report, 2024)). These patient characteristics have implications for transmission. An aging population may mean a larger pool of TB infected individuals at risk of reactivation of TB, while widespread diabetes in Kerala (Anjana et al., 2023) could increase susceptibility to active disease.

We defined diagnostic delay as the time from symptom onset to confirmed diagnosis and treatment initiation. The literature suggested an average total delay of 78 days (Sebastian et al., 2021) for TB patients in Kerala. This figure includes both patient delays (hesitation or slow health-seeking after symptom onset) and health system delays (time to diagnose after presentation). A mean delay on the order of months is consistent with other studies in India showing substantial diagnostic lags. For example, a systematic review reported mean delays of about 51 days in India, though with wide variation (Yasobant et al., 2023). Our estimated mean delay is longer, suggesting a skewed distribution where some patients experience very long delays. Such prolonged delays are problematic because an untreated infectious TB case can continue to transmit the organism during that period.

Another key input from program data was the coverage of TPT among household contacts of TB cases. Kerala has been a strong performer in contact tracing – in 2022, India identified about 3.36 million household contacts of TB cases, which was ~93% of the estimated contacts, and screened 94% of those identified (Shah et al., 2025). Our baseline scenario assumed 93% of notified cases household contacts are screened for latent TB infection, in line with national reports. This high coverage is better than the global average (globally only 71% of estimated contacts were identified and 84% of those screened in

2022 (Shah et al., 2025). It indicates that the majority of exposed family members in our setting receive evaluation and, if eligible, preventive therapy. Such data from the India TB Report were directly integrated into the model's assumptions.

5.2 Synthetic Population Generation

To realistically simulate TB transmission in the community, a synthetic population for Thiruvananthapuram district was generated using an existing Iterative Proportional Updating (IPU)-based framework and computer program. The resulting synthetic population comprised approximately 3.3 million agents, closely matching the actual district population. Validation checks confirmed that the synthetic population closely reflected the real census profile, with differences of less than one percent across all five-year age groups and the largest deviation observed in the 20–24-year group. This high degree of demographic fidelity supports the suitability of the synthetic population for modeling TB transmission dynamics in Thiruvananthapuram.

A unique aspect of our model was making this synthetic population dynamic over the 2011–2030 period. Instead of keeping the population static, we incorporated births, deaths, and aging using a demographic projection approach. We applied the Lee–Carter model (Lee and Carter, 1992) to official vital statistics (Sample Registration System data for 2011–2020) to project age-specific fertility and mortality rates from 2021 through 2030. The Lee–Carter method, commonly used in demography, enabled us to capture trends such as declining fertility and improving life expectancy. For instance, our projections showed a shifting fertility pattern (with peak child-bearing age moving from 20–24 to 25–29 years and an overall decline in birth rates). Each year of the simulation, new agents were "born" into the population according to the projected age specific fertility rates, and some agents

died based on age and sex specific mortality probabilities. We did not include migration, meaning the population was closed aside from births and deaths. This is a simplification discussed later as a limitation. Even so, the demographic changes modeled were quite meaningful. Kerala's population growth is slow and skewed toward older ages, which has implications for TB (fewer young people means potentially fewer new infections, but more elderly means a larger pool at risk of reactivation TB).

The synthetic population approach, with its detail on individual households and daily mobility, allowed us to simulate TB spread in a very granular way. Unlike compartmental models that treat the population as homogeneous fractions, our agent-based model (ABM) lets each person interact in specific settings – mainly households, workplaces, and schools. We assigned daily schedules to agents so that they mix with family members at home and with peers in workplaces or educational institutions during the day. This structured contact network captures the reality that TB transmission risk is highest among close contacts. Such realism is a major strength of ABMs. They explicitly represent heterogeneity in contact patterns and population structure, which is critical for diseases like TB that have clustered transmission. Indeed, a recent review noted that many TB ABMs incorporate household structures to better reflect transmission hotspots (Bui et al., 2024). To our knowledge, this study is the first to develop a full-scale synthetic population ABM for TB in India, addressing a significant gap in modeling methods identified in the literature. By leveraging open data and modern computational tools, we demonstrated that it is feasible to simulate an entire district's population. We implemented the model in Julia, a high-performance programming language, which enabled us to run simulations with millions of agents.

5.3 Modeling Delay and TPT Screening

Using the above framework, we incorporated the two main factors of interest – diagnostic delay and TPT coverage – into the model dynamics and ran multiple scenario simulations. The baseline scenario represented the status quo: a mean diagnostic delay of 78 days and household contact screening coverage of 93% (with TPT given to TB infected). For diagnostic delay, we ran scenarios where the delay was reduced to 8 weeks, 6 weeks, or 4 weeks (while keeping contact screening at baseline levels). These mimic interventions that could speed up diagnosis – for example, intensified active case finding, better community awareness leading to quicker care-seeking, or health system improvements (like rapid molecular testing and quicker referral) that cut down the time to start treatment. For TPT, we ran scenarios where the proportion of household contacts screened and started on preventive therapy was lowered to 50%, or to 0% (no preventive therapy), while keeping the diagnostic delay at 78 days. These represent potential shortfalls in the contact investigation program – for instance, if resources for contact tracing were halved or if the program was absent – allowing us to gauge how much of the current progress depends on high TPT coverage. In all scenarios, we tracked outcomes from 2025 through 2030, including the annual TB incidence (new active TB cases per year), the prevalence of active TB, and the prevalence of latent TB infection (TBI) in the population. The model's disease progression parameters remained the same across scenarios.

Shortening the diagnostic delay had a substantial impact on TB transmission in the model. Under the baseline scenario, TB burden declined slowly. However, when the average delay to diagnosis was reduced, both pulmonary TB prevalence and incidence declined much more rapidly. Faster diagnosis limited the infectious period of each case, reducing onward transmission and reducing the pool of TB infections over time. Even

moderate reductions in delay led to meaningful improvements, with shorter delays consistently producing greater epidemiological benefits.

Other modeling work supports the critical role of early diagnosis. A South African TB model recently underscored that diagnosing people more quickly and starting treatment is crucial for epidemic control (Low, 2024). Our findings provide quantitative backing to this finding in the Kerala context.

Lowering the coverage of TPT through contact screening had a significant negative impact in the model. When household contacts were not screened, initial gains from other control measures were quickly reversed, and pulmonary TB incidence began to rise, indicating a loss of control over transmission. This was driven by the unchecked accumulation of latent infections, many of which progressed to active disease. In the scenario with only 50% contact screening, the trend was more mixed. Initial declines in incidence were followed by a rebound, suggesting that partial coverage was insufficient to maintain progress. In contrast, high coverage (as in the baseline scenario) sustained a steady decline in both incidence and latent infection.

Our model shows that if this prevention step is removed or weakened, the long-term epidemic trajectory worsens considerably. The undetected cases that TPT would have prevented gradually emerge over time, causing a resurgence in TB incidence. This finding aligns with prior modeling and empirical studies that highlight the importance of TPT in high-risk groups (Bui et al., 2024). It is encouraging that India's program has already scaled up contact TPT far above the global average. Globally, only 21% of eligible household contacts received preventive treatment in 2023 (Global Tuberculosis Report, 2024), whereas India achieved far higher coverage. The model indicates that this investment is yielding epidemiological dividends in Kerala. Conversely, it warns that any reduction of

the contact tracing effort could quickly translate into lost progress.

Taken together, these findings highlight that timely diagnosis and high TPT coverage act synergistically to reduce TB transmission. In the model, the two interventions target different stages of the disease process. Early treatment interrupts transmission, while preventive therapy halts progression from latent to active TB. When both are implemented effectively, the decline in TB is steepest, suggesting the potential for local elimination over time. Neglecting either intervention weakens the overall impact. Reducing diagnostic delay without TPT allows latent infections to accumulate, while strong TPT efforts may be offset by continued transmission if diagnosis remains delayed. Kerala's current program performs well on contact screening but still has room to improve diagnostic timeliness, an area where strategic investment could yield significant additional gains.

5.4 Comparisons Similar Studies

Notably, there is a relative paucity of published TB ABMs focusing on Indian settings in the literature. Most modeling studies for India's TB epidemic have used compartmental models or static epidemiological models to evaluate interventions, rather than full individual-based simulations. One earlier individual-based simulation by Suen et al. (2015) evaluated the cost-effectiveness of improving TB diagnosis and treatment in India (Suen et al., 2015). Although that study is slightly older and used a different modeling approach, it similarly found that enhancing diagnostic access could be highly impactful and cost-effective in reducing TB burden in India. Our results are consistent with Suen et al.'s implication that investing in diagnostic infrastructure yields significant epidemiological benefits. However, unlike our ABM, their model did not explicitly represent social contacts or households – thus it could not examine the spatial or network aspects

of transmission.

Another study with relevance to India is the multi-country modeling by Ragonnet et al. (2019), which included India as one of the high-burden countries analyzed (Ragonnet et al., 2019). Ragonnet and colleagues used an individual-based framework to profile TB transmission in the five highest TB burden countries. They highlighted the importance of demographic and epidemiologic heterogeneity between countries – for instance, their model suggested that in India a large fraction of the disease burden is due to recent transmission (as opposed to reactivation of old infections). This supports the notion that interventions curbing ongoing transmission (like rapid diagnosis and treatment of infectious cases) can have a big impact in India, which aligns with our findings on delay reduction. Additionally, Ragonnet et al. underscored differences in age distribution of TB. India's younger population means transmission-driven epidemics, whereas countries like China with older populations face more reactivation TB. Kerala, while in India, has a demography closer to China's (older average age), which might mean a relatively larger role of reactivation. Our model explicitly incorporated Kerala's demography, and indeed we see that without interventions the decline in TB would be slow, partly because older latent infections fuel ongoing cases. Thus, our work extends Ragonnet's insights by looking within India at a demographically distinct state. It suggests that even within a country, sub-national variations matter for TB dynamics – agent-based models are well suited to capturing these nuances.

Numerous agent-based or individual-based TB models have been applied to other countries and specific aspects of TB control in recent years. Our findings align with many of these studies on the general principle that reducing infectious periods and treating latently infected individuals are effective strategies, though the magnitude of impact can

vary by setting. Abdollahi et al. (2022) created an ABM for the Canadian Inuit population in Nunavut (Abdollahi et al., 2022). Their model evaluated various TB elimination strategies, including intensive preventive therapy and improvements in diagnosis and treatment. They found that very high coverage of LTBI treatment along with sustained case detection efforts would be needed to substantially reduce TB in the Inuit community. This aligns strongly with our findings for Kerala: high coverage of contacts and reduced delays together yielded the best outcomes. Abdollahi et al. also emphasized the role of risk factors like housing and social determinants in the Inuit TB burden. Another sophisticated modeling framework by Renardy and Kirschner (2020) looked at network-based TB transmission using synthetic social networks (Renardy and Kirschner, 2020). They demonstrated the feasibility of constructing synthetic contact networks for TB and how network structure influences outbreak dynamics. Our work confirm the value of using synthetic population or contact data for TB modeling, as also recommended in that study. By capturing actual population structure, we can more reliably predict which interventions reach the people who need them. For example, network models show that households form clusters of infection. Our model took advantage of that by applying TPT to whole clusters (households) as the program does, which is more effective than random distribution of preventive therapy.

Finally, it is worth noting a general point that emerges from many ABM studies including ours. The complexity of ABMs must be balanced with transparency and calibration. The 2024 systematic review by Bui et al. observed that only 8 of 26 ABM studies provided public code and that many models lacked clarity on some assumptions (Bui et al., 2024). We have aimed to document our model clearly and plan to share key components openly, contributing to the growing practice of open-source modeling in TB research. The

challenges of calibrating ABMs are non-trivial, as each added layer of realism introduces new parameters that may be uncertain. We calibrated our model to baseline notification data and demographic trends, similar to how other studies have matched prevalence or incidence targets.

Comparing our Kerala-focused ABM with similar studies indicates a high degree of qualitative agreement. Rapid diagnosis and thorough preventive therapy are cornerstones of TB control in any context, and ABMs worldwide have reinforced this. Our contribution lies in quantifying these effects for a specific Indian setting and demonstrating an approach to incorporate routinely collected programmatic data and demography into an ABM.

5.5 Strengths and Limitations

The use of an agent-based modeling framework allowed us to capture individual-level heterogeneity and realistic contact patterns that are impossible to represent in aggregate compartmental models. By explicitly modeling households, workplaces, and schools, we were able to reflect the natural clustering of TB. The synthetic population approach means the model can be extended to examine interventions targeted by demographic group or location (for instance, focusing on older adults or certain panchayats) with relative ease, something homogeneous models can't do. Additionally, by relying on routinely collected programmatic data sources such as Nikshay notifications, census records, and Sample Registration System (SRS) estimates for model initialization and parameterization, the approach offers a practical framework that can be readily adapted to other settings.

Another strength is the agent-based design's flexibility to incorporate various interventions and delays as we demonstrated. We could model a specific intervention (like contact tracing) quite naturally, turning it on or off per scenario. The ability to capture

non-linear effects and emergent dynamics (like how reducing TB leads to fewer future latent infections, reinforcing decline) is an inherent strength of ABMs. We also highlight the computational implementation in Julia programming language as a strength. Julia's speed and capacity for handling large arrays enabled us to simulate nearly 3.5 million agents over 6 years with half-daily steps.

Despite these strengths, there are several limitations to acknowledge. One major limitation is that we did not model migration or mobility across long distances. The people did not move in or out of the district. In reality, Kerala has significant migration – both out-migration (to the Gulf countries) and in-migration (workers from other Indian states). Migration can influence TB dynamics by introducing new cases or infections from outside and by changing population density. Similarly, we did not explicitly model day-to-day mobility beyond home/work/school. Everyone was assumed to stay within their set locations. We couldn't simulate occasional travel or large gatherings, which could cause spurts of transmission. In future, incorporating more complex mobility patterns or at least stochastic mixing events could improve realism.

Another limitation is the reliance on notification data to calibrate infection prevalence and incidence, which might underestimate true burden if there is under-detection. Kerala's case detection is relatively strong, but if there are missing cases, our baseline might be low. This means our projections could be slightly skewed. The real incidence might be higher. Moreover, our model assumes that notification equals incidence. If some fraction of cases never get diagnosed at all, they would drop out of the model accounting.

The simplifications in model parameters also constitute a limitation. We modeled TB natural history in a simplified SEIR-like manner without differentiating by smear status, drug resistance, or re-infection effects. In reality, not all TB cases are equally infectious.

Smear positive pulmonary TB is most infectious, whereas smear negative cases contribute less to transmission. If a large share of cases in Kerala are smear negative, our model might overestimate transmission a bit. However, since we calibrated roughly to actual incidence trends, this may be minor. We also didn't explicitly model drug-resistant TB as a separate category. We assumed all infections are equally treatable and that treatment failure loops back to a susceptible-like state or re-treatable state.

From an interventions standpoint, a limitation is that we only modeled two interventions (diagnostic delays and household contact TPT) in isolation. In practice, many interventions occur concurrently. For example, improved diagnostics might come with better treatment supervision or with enhanced screening. We did not model other strategies like active case finding in the community. But the model is designed so that we can easily implement other interventions if needed.

5.6 Conclusion

This study demonstrates how an agent-based model built on a synthetic population can yield actionable insights into TB transmission dynamics and control strategies in Kerala. By simulating a realistic population of Thiruvananthapuram district, we were able to evaluate the impact of diagnostic delays and household TPT interventions on TB incidence. The results showed that reducing delays in detecting and treating infectious TB cases has a substantial effect on curbing transmission.

In the context of Kerala's TB elimination efforts, our findings suggest that the state's strategy of intensive case finding and contact management is on target, but further strengthening, especially in reducing diagnostic delays, could accelerate the decline. Kerala has already achieved commendable success in tracing contacts. Our model indicates this must

continue without any fail. Meanwhile, more can be done to shorten patient and system delays.

More broadly, this study represents a methodological advance by integrating rich empirical data with agent-based modeling to address a public health challenge. It paves the way for similar models in other districts or states of India, which can incorporate region-specific data to guide tailored interventions. It also underscores the importance of considering demographic changes in planning TB control.

5.7 Recommendations

The model findings highlight the critical need to address delays in tuberculosis diagnosis and treatment initiation. Reducing these delays would significantly lower transmission within the community. This could involve expanding community-based screening for people with cough, integrating private providers so that no patient experiences delays due to multiple consultations, and ensuring rapid molecular diagnostic tests are accessible at the point of care. If such measures can reduce the typical workup time to a few weeks, the payoff would be significantly fewer new infections occurring.

Household-level screening remains a vital strategy for TB control. Routine screening of household contacts, especially for patients with pulmonary TB, should be sustained and expanded. The agent-based model developed in this study can be further used to simulate the impact of targeted household screening interventions on interrupting transmission chains, particularly within high-risk households.

Enhancing model inputs by integrating broader datasets would improve the realism and usefulness of future simulations. Strengthening data capture within the Nikshay system would further enhance future modeling and surveillance efforts. While Nikshay data

provides a strong foundation, incorporating additional variables such as housing conditions, nutritional status, socioeconomic status, duration of patient delays and access barriers would allow for more nuanced simulations and better planning of targeted interventions.

References

Abdollahi E, Keynan Y, Foucault P, et al. (2022) Evaluation of TB elimination strategies in Canadian Inuit populations: Nunavut as a case study. *Infectious Disease Modelling* 7(4): 698–708.

Amadi M, Shcherbacheva A and Haario H (2021) Agent-based modelling of complex factors impacting malaria prevalence. *Malaria Journal* 20(1): 185.

Anjana RM, Unnikrishnan R, Deepa M, et al. (2023) Metabolic non-communicable disease health report of India: The ICMR-INDIAB national cross-sectional study (ICMR-INDIAB-17). *The Lancet Diabetes & Endocrinology* 11(7): 474–489.

Annual Report, Periodic Labour Force Survey (2024) Ministry of Statistics; Programme Implementation, Government of India. Available at: https://www.mospi.gov.in/sites/default/files/publication_reports/AnnualReport_PLFS2023-24L2.pdf.

Bandini S, Manzoni S and Vizzari G (2009) Agent Based Modeling and Simulation. In: Meyers RA (ed.) *Encyclopedia of Complexity and Systems Science*. New York, NY: Springer, pp. 184–197. Available at: https://doi.org/10.1007/978-0-387-30440-3_12 (accessed 10 November 2024).

Bhargava A, Bhargava M, Meher A, et al. (2023) Nutritional supplementation to prevent tuberculosis incidence in household contacts of patients with pulmonary tuberculosis in India (RATIONS): A field-based, open-label, cluster-randomised, controlled trial. *The Lancet* 402(10402): 627–640.

Borgdorff MW, Floyd K and Broekmans JF (2002) Interventions to reduce tuberculosis mortality and transmission in low- and middle-income countries. *Bulletin of the World Health Organization* 80: 217–227.

Brauer F (2008) Compartmental Models in Epidemiology. *Mathematical Epidemiology* 1945: 19.

Bui VL, Hughes A, Ragonnet R, et al. (2024) Agent-based modelling of Mycobacterium tuberculosis transmission: A systematic review. Available at: https://www.researchsquare.com/article/rs-3952348/v1 (accessed 16 May 2024).

CDC (2025) History of World TB Day. *World TB Day*. Available at: https://www.cdc.

gov/world-tb-day/history/index.html (accessed 17 April 2025).

Census tables | Government of India (2011). Available at: https://censusindia.gov.in/census.website/data/census-tables (accessed 20 October 2024).

Chapuis K, Taillandier P and Drogoul A (2022) Generation of Synthetic Populations in Social Simulations: A Review of Methods and Practices. *Journal of Artificial Societies and Social Simulation* 25(2): 6.

Cuddington K, Beisner B, San P, et al. (2005) Ecological Paradigms Lost Routes of Theory Change. Epub ahead of print January 2005.

Desai S and Vanneman R (2018) India Human Development Survey-II (IHDS-II), 2011-12. Inter-university Consortium for Political; Social Research. Available at: https://www.icpsr.umich.edu/web/DSDR/studies/36151/versions/V6 (accessed 20 October 2024).

Dharmadhikari AS, Mphahlele M, Venter K, et al. (2014) Rapid impact of effective treatment on transmission of multidrug-resistant tuberculosis. *The International Journal of Tuberculosis and Lung Disease: The Official Journal of the International Union Against Tuberculosis and Lung Disease* 18(9): 1019–1025.

Dong Z, Wang Q-Q, Yu S-C, et al. (2022) Age–period–cohort analysis of pulmonary tuberculosis reported incidence, China, 2006–2020. *Infectious Diseases of Poverty* 11(1): 85.

Ferguson NM, Cummings DAT, Fraser C, et al. (2006) Strategies for mitigating an influenza pandemic. *Nature* 442(7101): 448–452.

Fox GJ, Barry SE, Britton WJ, et al. (2013) Contact investigation for tuberculosis: A systematic review and meta-analysis. *The European Respiratory Journal* 41(1): 140–156.

Geard N, McCaw JM, Dorin A, et al. (2012) Synthetic Population Dynamics: A Model of Household Demography. *Journal of Artificial Societies and Social Simulation* 16(1): 8.

Global Tuberculosis Report (2024). Available at: https://www.who.int/publications/i/item/9789240101531 (accessed 6 November 2024).

Government of India | Office of the Registrar General & Census Commissioner, India (2025). Available at: https://censusindia.gov.in/census.website/data/SRSSTAT (accessed 29 April 2025).

Grassly NC and Fraser C (2008) Mathematical models of infectious disease transmission. *Nature Reviews Microbiology* 6(6): 477–487.

Gridded Population of the World (GPW), v4 | SEDAC (2016). Available at: https://sedac.

ciesin.columbia.edu/data/collection/gpw-v4 (accessed 20 October 2024).

Halloran ME, Ferguson NM, Eubank S, et al. (2008) Modeling targeted layered containment of an influenza pandemic in the United States. *Proceedings of the National Academy of Sciences* 105(12): 4639–4644.

Halloran ME, Auranen K, Baird S, et al. (2017) Simulations for designing and interpreting intervention trials in infectious diseases. *BMC Medicine* 15(1): 223.

Hamer WH (1928) *Epidemiology, Old and New*. London : K. Paul, Trench, Trubner & Co., 1928. Available at: https://wellcomecollection.org/works/tnsj9kpb (accessed 10 November 2024).

Hammond RA (2015) Considerations and Best Practices in Agent-Based Modeling to Inform Policy. In: *Assessing the Use of Agent-Based Models for Tobacco Regulation*. National Academies Press (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK305917/ (accessed 10 November 2024).

Hershkovitz I, Donoghue HD, Minnikin DE, et al. (2015) Tuberculosis origin: The Neolithic scenario. *Tuberculosis* 95. Supplement issue: Tuberculosis in Evolution: S122–S126.

Houben RMGJ and Dodd PJ (2016) The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLOS Medicine* 13(10): e1002152.

Huynh N, Namazi-Rad M-R, Perez P, et al. (2013) Generating a synthetic population in support of agent-based modeling of transportation in Sydney. *Faculty of Engineering and Information Sciences - Papers: Part A*: 1357–1363.

India TB Report (2023).

Iranzo V and Pérez-González S (2021) Epidemiological models and COVID-19: A comparative view. *History and Philosophy of the Life Sciences* 43(3): 104.

Jonas DE, Riley SR, Lee LC, et al. (2023) Screening for Latent Tuberculosis Infection in Adults: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA* 329(17): 1495–1509.

Keeling MJ, Keeling MJ and Rohani P (2008) *Modeling Infectious Diseases in Humans and Animals*. Princeton, NJ: Princeton Univ. Press.

Kendall DG (1956) Deterministic and Stochastic Epidemics in Closed Populations. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Health*. University of California Press, pp. 149–166. Available at: https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Third-Berkeley-Symposium-on-Mathematical-Statistics-and/chapter/Deterministic-and-Stochastic-Epidemics-in-Closed-Populations/bsmsp/1200502553 (accessed 10

November 2024).

Kermack WO, McKendrick AG and Walker GT (1927) A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115(772): 700–721.

Kerr CC, Stuart RM, Mistry D, et al. (2021) Covasim: An agent-based model of COVID-19 dynamics and interventions. *PLOS Computational Biology* Marz M (ed.) 17(7): e1009149.

Lee RD and Carter LR (1992) Modeling and Forecasting U. S. Mortality. *Journal of the American Statistical Association* 87(419): 659–671.

Liyew AM, Gilmour B, Clements AC, et al. (2023) Comparative effectiveness of interventions for preventing tuberculosis: Systematic review and network meta-analysis of interventional studies. *eClinicalMedicine* 64: 102209.

Low M (2024) A complex, high-performance agent-based model used to explore tuberculosis and COVID-19 case-finding interventions. Epub ahead of print 2024.

Matteelli A, Lovatti S, Sforza A, et al. (2023) Programmatic management of tuberculosis preventive therapy: Past, present, future. *International Journal of Infectious Diseases* 130. Commemorating World Tuberculosis Day, March 24th, 2023: 'Yes! We Can End TB': S43–S46.

Miller PB, Zalwango S, Galiwango R, et al. (2021) Association between tuberculosis in men and social network structure in Kampala, Uganda. *BMC Infectious Diseases* 21(1): 1023.

Mohan A, Indu PS, Nair S, et al. (2022) Prevalence of Tuberculosis Infection among Adults of Thiruvananthapuram District of Kerala as Measured by Interferon Gamma Release Assay – A Cross-Sectional Study. *Indian Journal of Community Medicine: Official Publication of Indian Association of Preventive & Social Medicine* 47(4): 501–505.

National TB Prevalence Survey in India 2019-2021 (2023) Central Tuberculosis Division. Available at: https://tbcindia.mohfw.gov.in/2023/06/06/national-tb-prevalence-survey-in-india-2019-2021/ (accessed 6 November 2024).

Neekhra B, Kapoor K and Gupta D (2023) SYNTHPOP++: A HYBRID FRAMEWORK FOR GENER- ATING A COUNTRY-SCALE SYNTHETIC POPULATION. Epub ahead of print 2023.

Ni-kshay Reports (2025). Available at: https://reports.nikshay.in/ (accessed 29 April 2025).

NTEP | Knowledge Base (2025). Available at: https://ntep.in/ (accessed 29 April 2025).

Panda S (2023) Tuberculosis: National survey in India & elimination challenges reflected through global learning. *The Indian Journal of Medical Research* 157(2-3): 111–113.

Population Projections for India and States, 2011-2036 (2019).

Prathiksha G, Daniel BD and Natrajan M (2019) Active case-finding for tuberculosis in India. *The National Medical Journal of India* 32: 90.

Ragonnet R, Trauer JM, Geard N, et al. (2019) Profiling Mycobacterium tuberculosis transmission and the resulting disease burden in the five highest tuberculosis burden countries. *BMC Medicine* 17(1): 208.

Renardy M and Kirschner DE (2020) A Framework for Network-Based Epidemiological Modeling of Tuberculosis Dynamics Using Synthetic Datasets. *Bulletin of Mathematical Biology* 82(6): 78.

Richard DM and Lipsitch M (2024) What's next: Using infectious disease mathematical modelling to address health disparities. *International Journal of Epidemiology* 53(1): dyad180.

Ross R and Hudson HP (1916) An application of the theory of probabilities to the study of a priori pathometry.—Part III. 93(650): 225–240.

Sakula (1982) Robert Koch: Centenary of the discovery of the tubercle bacillus, 1882. *Thorax* 37(4).

Sebastian NM, Haveri SP and Nath AS (2021) Delay in diagnosis of tuberculosis and related factors from a district in Kerala, India. *Indian Journal of Tuberculosis* 68(1): 59–64.

Shah H, Patel J, Rai S, et al. (2025) Advancing tuberculosis elimination in India: A qualitative review of current strategies and areas for improvement in tuberculosis preventive treatment. *IJID Regions* 14: 100556.

Shah NS, Yuen CM, Heo M, et al. (2014) Yield of contact investigations in households of patients with drug-resistant tuberculosis: Systematic review and meta-analysis. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 58(3): 381–391.

Siettos C, Anastassopoulou C, Russo L, et al. (2015) Modeling the 2014 Ebola Virus Epidemic – Agent-Based Simulations, Temporal Analysis and Future Predictions for Liberia and Sierra Leone. *PLoS Currents* 7: ecurrents.outbreaks.

Sterling TR, Villarino ME, Borisov AS, et al. (2011) Three Months of Rifapentine and Isoniazid for Latent Tuberculosis Infection. *New England Journal of Medicine* 365(23): 2155–2166.

Suen S, Bendavid E and Goldhaber-Fiebert JD (2015) Cost-Effectiveness of Improve-

ments in Diagnosis and Treatment Accessibility for Tuberculosis Control in India. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 19(9): 1115–xv.

Tolles J and Luong T (2020) Modeling Epidemics With Compartmental Models. *JAMA* 323(24): 2515–2516.

Trauer JM, Dodd PJ, Gomes MGM, et al. (2019) The Importance of Heterogeneity to the Epidemiology of Tuberculosis. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 69(1): 159–166.

Turner RD and Bothamley GH (2015) Cough and the transmission of tuberculosis. *The Journal of Infectious Diseases* 211(9): 1367–1372.

White PJ and Enright MC (2012) Mathematical models in infectious disease epidemiology. *Infectious Diseases*: 70.

WHO Consolidated Guidelines on Tuberculosis: Tuberculosis Preventive Treatment (2020) World Health Organization.

Willem L, Verelst F, Bilcke J, et al. (2017) Lessons from a decade of individual-based models for infectious disease transmission: A systematic review (2006-2015). *BMC Infectious Diseases* 17(1): 612.

Yasobant S, Shah H, Bhavsar P, et al. (2023) Why and where?—Delay in Tuberculosis care cascade: A cross-sectional assessment in two Indian states, Jharkhand, Gujarat. *Frontiers in Public Health* 11.

Zelner JL, Murray MB, Becerra MC, et al. (2014) Age-specific risks of tuberculosis infection from household and community exposures and opportunities for interventions in a high-burden setting. *American Journal of Epidemiology* 180(8): 853–861.

Zhu K, Yin L, Liu K, et al. (2024) Generating synthetic population for simulating the spatiotemporal dynamics of epidemics. *PLOS Computational Biology* Althouse B (ed.) 20(2): e1011810.

Annexure

Annexure I: GitHub Repository

The complete code base used for simulation, data processing and visualization in this thesis are available at the following GitHub repository: github.com/razinoushad/mph_thesis This repository includes scripts for model setup, data processing, scenario simulation, and output visualization.

Annexure II: IEC Permission Letter

श्री चित्रा तिरुनाल आयुर्विज्ञान और प्रौद्योगिकी संस्थान, त्रिवेन्द्रम
तिरुवनन्तपुरम - ६९५०११, केरल, इंडिया
SREE CHITRA TIRUNAL INSTITUTE FOR MEDICAL SCIENCES AND TECHNOLOGY, TRIVANDRUM
Thiruvananthapuram - 695 011, Kerala, India
(An Institute of National Importance under Govt. of India)
Grams : Chitramet, Phone : +91-471-2443152, Fax : +91-471-2550728 / 2446433, E-mail : sct@sctimst.ac.in, Website : www.sctimst.ac.in

Institutional Ethics Committee

CDSCO Registration No: ECR/189/Inst/KL/2013/RR-24
DHR Registration No:EC/NEW/INST/2022/2775

SCT/IEC/2328/DECEMBER/2024 11.04.2025

Dr. Razi N
MPH Student, AMCHSS
SCTIMST, Thiruvananathapuram

Dear Dr. Razi,

The Institutional Ethics Committee held on 30th December, 2024, reviewed and discussed your application to conduct the study titled "**ASSESSING THE ROLE OF DELAYS IN PULMONARY TB DETECTION AND CONTROL STRATEGIES ON TRANSMISSION DYNAMICS: A SYNTHETIC POPULATION APPROACH IN THIRUVANANTHAPURAM DISTRICT, KERALA" (IEC /2328)** .

Principal Investigator	Dr. Razi N, MPH Student, AMCHSS, SCTIMST
Co-Principal Investigator(s)	Dr. Biju Soman, Professor, AMCHSS, SCTIMST
Duration of the study	6 months

The following members of the Ethics Committee were present at the meeting held on 30th December, 2024

SL. No.	Member Name	Highest Degree	Gender	Scientific /Non Scientific	Affiliation with Institution(s)
1.	Dr. Manoj Murhekar	MD (Preventive & Social Medicine)	Male	Social Scientist	No
2.	Dr. Annapurna.Y	MD	Female	Basic Medical Scientist	No
3.	Adv. Priya Kaimal	LLM, MBL	Female	Legal Expert	No
4.	Dr. Indu PV	MBBS, DPM, MPhil, PhD	Female	Clinician	No
5.	Dr. Rejnish Kumar	MBBS,MD ,DNB	Male	Clinician	No
6.	Smt. Rekha Radhakrishnan	M.A (Eng.); B.Ed	Female	Lay person	No
7.	Dr. Biju Soman	MBBS,MD, DPH, MSc, DLSHTM	Male	Basic Medical Scientist	Yes
8.	Dr. Syam K	MBBS, MD, DM	Male	Clinician	Yes
9.	Dr. Srinivas G	PhD	Male	Basic Medical Scientist (Member Secretary)	Yes

1

Annexure II: IEC Permission Letter

SCT/IEC/2328/DECEMBER-2024

The following documents were reviewed:

Original submission

1. *IEC Application Form*
2. *Covering letter addressed to the Chairperson, IEC, SCTIMST forwarded by HOD*
3. *Funding details and Budget*
4. *Data Extraction Template*
5. *Letter to State TB officer for permission*
6. *Checklist Form*
7. *Proposal*
8. *Declaration Form*
9. *CV of PI and Co-PI*

Revised submission

1. *Covering letter addressed to the Member Secretary, IEC, SCTIMST dated 13.01.2025*
2. *IEC Recommendation letter dated 09.01.2025*
3. *Proposal*
4. *CV of PI*

Revised submission on 26.03.2025

1. *Covering letter addressed to the Member Secretary, IEC, SCTIMST dated 25.03.2025 regarding the change of the title and objective of the study*
2. *Proposal*

IEC Decision
The IEC approved the conduct of the study in the present form.

Remarks:
The Institutional Ethics Committee expects to be informed about the progress of the study, any SAE occurring in the course of the study, any changes in the protocol and patient information/informed consent and asks to be provided a copy of the final report.

There was no member of the study team / Guide who participated in voting / decision making process. The ethics committee is organized and operated according to the requirements of Good Clinical Practice and the requirements of the Indian Council of Medical Research (ICMR).

Sincerely,

Dr. G. Srinivas
Member Secretary, IEC

MEMBER SECRETARY
INSTITUTIONAL ETHICS COMMITTEE (IEC)
SCTIMST, THIRUVANANTHAPURAM

Annexure III: State TB Office Permission Letter

GOVT OF KERALA

Directorate of Health Services, Kerala

State TB Cell
Red Cross Road, General Hospital Jn
Thiruvananthapuram PIN.695 035
Kerala State

No. 089/STC/DHS/2024 (4)

Thiruvananthapuram
27.01.2025

From,

Additional Director of Health Services (TB)

To

The District TB Officers
District TB Centre
Thiruvananthapuram

Dear Doctor,

Sub: **Sanction accorded to avail NTEP data from District TB Centre -reg**
Ref: 1, Razi.N, MPH Scholar, AMCHSS, Thiruvananthapuram
2. Approval from minutes of State OR Committee meeting on 10.01.2025

With reference to the above, permission to Razi. N, MPH Scholar, AMCHSS, Thiruvananthapuram for collecting the details of the patients who are registered under NTEP from Thiruvananthapuram for the use of his study title "***Assessing the Role of Delays in TB Detection and Control Strategies on Transmission Dynamics: A Synthetic Population Approach in Thiruvananthapuram, Kerala***". Please ensure that the above thesis study should be conducted as per the following conditions.

- NTEP guidelines should be adhered to throughout the period of study
- The said study does not affect the routine functions of the institutions or NTEP activities.
- The study does not disturb (in any way) to the patients or general public.
- No additional investigations would be done for the patient, and if needed for the study purpose, should be done free of cost to the patient with the consent of the State TB authorities.
- Privacy and confidentiality of the patient should be strictly maintained.
- Patients would not be forced to travel to the study centre or would be compensated for such travel.
- The content of study should be presented to the undersigned before submitting the same to any other institutions/journals
- The study should be for a short-period and **no financial assistance will be provided by this institution until the funds are available and sanctioned by the State NHM.**
- An undertaking may be submitted

Dr. Rajaram Kizhakkekkandiyil
Addl. Director of Health Services-TB/ State TB Officer

STATE TB CELL

Copy to : Razi. N

Annexure IV: Plagiarism Report

Assessing the Role of Delays in Pulmonary TB Detection and Control Strategies on Transmission Dynamics: A Synthetic Population Approach in Thiruvananthapuram District, Kerala

ORIGINALITY REPORT

1 %
SIMILARITY INDEX

PRIMARY SOURCES

1	dspace.sctimst.ac.in Internet	36 words — < 1%
2	www.coursehero.com Internet	30 words — < 1%
3	www.health.govt.nz Internet	27 words — < 1%
4	wish.org.qa Internet	24 words — < 1%
5	ntep.in Internet	22 words — < 1%
6	www.ncbi.nlm.nih.gov Internet	21 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 15 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 20 WORDS