



# Forecasting tuberculosis through mechanistic learning of transmission dynamics: Insights from a case study in India

Adrita Ghosh<sup>a</sup>, Parthasakha Das<sup>b,\*,\*</sup>, Susanta Kumar Das<sup>a</sup>, Pritha Das<sup>a</sup>, Ranjit Kumar Upadhyay<sup>c</sup>

<sup>a</sup> Department of Mathematics, Indian Institute of Engineering Science and Technology, Shibpur, West Bengal 711103, India

<sup>b</sup> Department of Mathematics, Rajiv Gandhi National Institute of Youth Development, Sriperumbudur, Tamil Nadu 602105, India

<sup>c</sup> Department of Mathematics and Computing, Indian Institute of Technology (Indian School of Mines), Dhanbad, Jharkhand 826007, India

## ARTICLE INFO

Dataset link: [NI-kshay](#)

### Keywords:

Tuberculosis modeling  
Parametric calibration  
Sensitivity analysis  
Bifurcation analysis  
Epidemiology-guided deep learning

## ABSTRACT

Tuberculosis continues to pose a significant global health issue, with India facing the most considerable burden. We develop and calibrate an SEIR-type model to understand disease dynamics and perform sensitivity analysis to pinpoint crucial parameters. The basic reproduction number acts as the epidemic threshold, while backward bifurcation highlights the dangers of reinfection and superinfection. To enhance forecasting, we merge mechanistic modeling with deep learning techniques, utilizing feedforward, recurrent, and memory-based networks. Evaluated on TB case data from India, these hybrid models outperform the SEIR baseline, with the gated recurrent unit best capturing residual trends and the feedforward network demonstrating robust generalization. This integrated framework enhances predictive accuracy and interpretability, providing a valuable tool for TB forecasting and aiding in targeted interventions.

## 1. Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, is among the oldest infectious diseases, with evidence tracing back over 17,000 years [1]. Despite major advances in diagnostics and treatment, TB remains one of the world's top infectious killers, second only to HIV/AIDS among single-agent infections. Transmission occurs primarily through airborne droplets when infected individuals cough or sneeze, with symptoms including persistent cough, chest pain, hemoptysis, fatigue, weight loss, and loss of appetite [2]. TB disproportionately affects low- and middle-income countries, with South and Southeast Asia — especially India, Bangladesh, and Nepal — bearing a heavy burden [3]. In 2020, India alone accounted for 26% of global TB cases, with an incidence of 192 per 100,000 population [4]. Despite being curable, the effective management of tuberculosis is hindered by factors such as late diagnosis, issues with adherence to treatment, drug resistance, and the reactivation of dormant infections [5]. In response, the WHO declared TB a global emergency in 1993 and launched the End TB strategy, targeting a 50% incidence reduction by 2025 [6]. Yet, global cases surged to 10.6 million in 2022, partly due to setbacks from COVID-19, which worsened comorbidities and disrupted treatment [4]. With 5% – 10% of the global population at risk of progressing from latent to active TB, reinfections and recurrences remain major concerns [5]. India's TB-Free

India Campaign (targeted for 2035) reflects multisectoral efforts, but persistent gaps in surveillance and forecasting underscore the need for mechanistic modeling and predictive analytics [7]. Such approaches, widely used in other infectious diseases like dengue, cholera, and influenza, can capture the complex biological, social, and environmental drivers of TB dynamics [8].

Classical and modern approaches have examined TB transmission from diverse perspectives. Early mechanistic studies explored vaccine effects on dormancy [9], intrinsic transmission dynamics [10], and multi-drug resistance with heterogeneous fitness [11]. Subsequent work addressed latent infection risks in close contacts [12], occupational hazards for healthcare workers in LMICs [13], and strategies toward zero TB deaths [14]. Other contributions investigated post-exposure interventions [15], heterogeneity in adherence and transmission [16], bifurcation and relapse dynamics [17,18], reproduction number estimation [19], and diagnostic challenges in resource-limited settings [20]. More recent studies have linked public health interventions to host susceptibility [21], assessed post-exposure vaccines for resistant TB [22], examined TB–HIV co-epidemics [23], evaluated private provider engagement in India [24], and analyzed the role of behavioral change, media, reinfection, and delayed treatment in shaping endemic persistence [25–30]. Collectively, these studies underscore the value of

\* Corresponding author.

E-mail address: [parthasakha87das@gmail.com](mailto:parthasakha87das@gmail.com) (P. Das).