Tuberculosis Prevalence Trends from a Predictive Modelling Study — 10 High-Burden Countries, 1980–2035

Qiuping Chen[1,2,&]; Qiao Liu[3,&]; Kangguo Li[4,&]; Laurent Gavotte[5]; Roger Frutos[1,#]; Tianmu Chen[3,4,#]

Summary

What is already known about this topic?

Given the challenges presented by drug-resistant strains of tuberculosis (TB) and the rising mobility of the population, achieving the objective of eradicating TB appears uncertain.

What is added by this report?

The examination of TB incidence trends in 10 high-burden countries (HBCs) indicated a steady rise in cases, with India and China jointly accounting for nearly 70% of the burden. Projections for the future show diverse trajectories in these countries, with potential difficulties in reaching the TB elimination target, especially in Nigeria, Congo, and South Africa.

What are the implications for public health practice?

The number of TB cases is on the rise. It is crucial to learn from successful strategies to improve TB prevention and control worldwide through collaborative efforts.

The fight against tuberculosis (TB) is being severely hampered by the development of drug-resistant strains and increased human mobility, further compounded by the impact of the coronavirus disease 2019 (COVID-19) pandemic (*1*). The World Health Organization (WHO) Post-2015 End TB Strategy, which was endorsed by the World Health Assembly in 2014 and aligns with the Sustainable Development Goals, has the ambitious objective of eradicating the global TB epidemic (*2*). Nonetheless, there is growing concern regarding the adequacy of existing TB control strategies to meet these objectives. This study employed four predictive models to evaluate the likelihood of reaching the WHO's targets within a set timeframe, specifically examining the situation in the 10 high-burden countries (HBCs). Our findings indicate a worsening global TB burden, with numerous countries falling short of the aspirations to eliminate pulmonary TB. Therefore, it is imperative to draw on successful strategies and enact tailored interventions

within key nations in conjunction with bolstered international collaboration.

TB case data were extracted from the WHO Global Tuberculosis Programme's open data repository (https://www.who.int/teams/global-tuberculosis-programme/data), while demographic details, including population size, were gathered from the World Bank's open data platform. An incidence analysis of TB was carried out for 10 HBCs spanning from 1980 to 2021. It is predicted that China, Congo, India, Indonesia, Mozambique, Myanmar, Nigeria, the Philippines, South Africa, and Zambia will continue to rank among the top 30 HBCs until 2023 due to the persistent challenges of drug-resistant TB and TB/ human immunodeficiency virus (HIV) co-infection (*1*,*3*).

In order to predict TB incidence for the years 2023–2035, various predictive models were utilized, including the autoregressive integrated moving average (ARIMA) model (*4*), neural network model (*5*), bayesian structural model (*6*), and a hybrid model integrating ARIMA, exponential smoothing (ETS), and seasonal and trend decomposition using Loess (STL). Details on these models can be found in the Supplementary Material (available at https://weekly.chinacdc.cn/). Data analysis was conducted using R software (version 4.3.3, R Core Team, Vienna, Austria).

Figure 1 depicts the trend of TB incidence in the 10 HBCs from 1980 to 2022. These 10 nations collectively account for 117,314,803 TB cases, representing 60% of the global total. The number of TB cases in the 10 HBCs has been increasing annually, with the exception of a decline in 2020 and 2021 due to the COVID-19 pandemic. Among these countries, nearly 70% of TB cases are concentrated in India and China, highlighting the urgent need to address the TB burden in these two nations. India has consistently shown the highest TB incidence over the past 43 years, with a rising trend since the publication of the TB report in 1997. In contrast, while China has a considerable TB incidence rate, it has been steadily
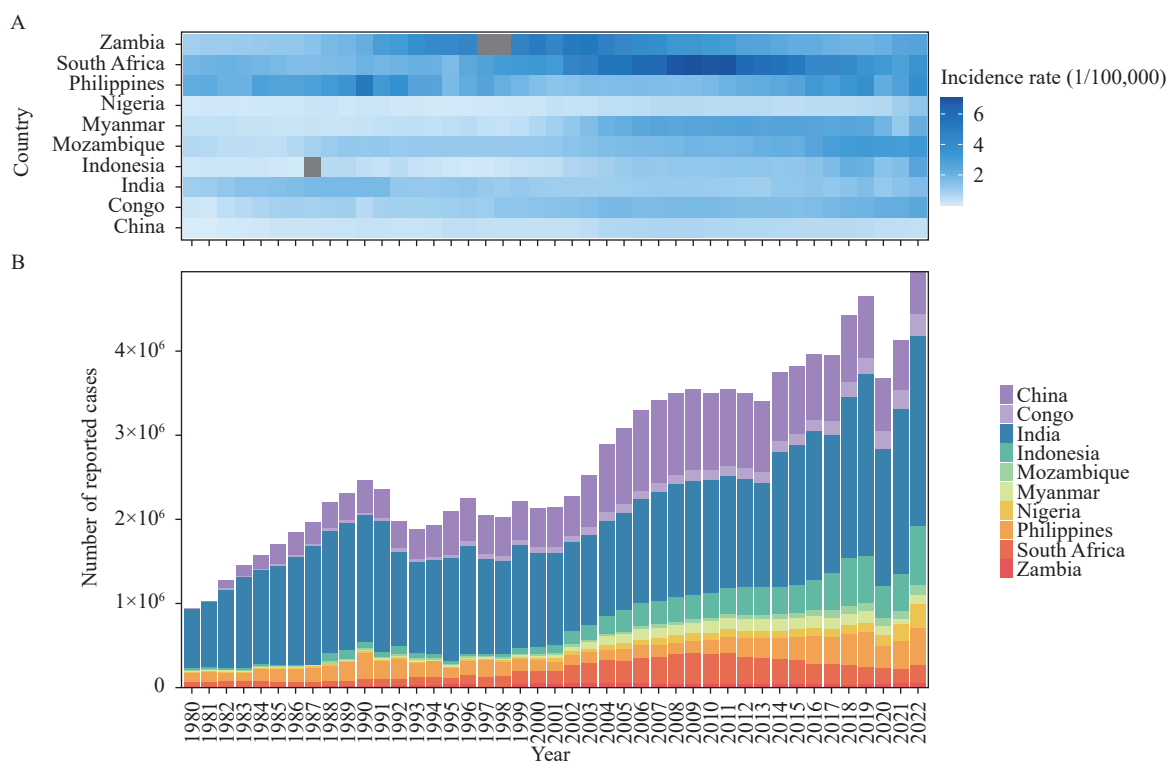
FIGURE 1. The spatiotemporal distribution characteristics of global tuberculosis.

decreasing since 2009, as evidenced in the graph depicting incidence patterns (Figure 1B).

To conduct a thorough analysis of TB incidence trends, we utilized four different models to predict TB cases from 2023 to 2035 (Figure 2). The results revealed a decreasing trend in future TB cases in China, as projected by three of the four models, with the exception of the neural network model. Conversely, Congo is expected to see a rise in TB cases over the next 13 years, regardless of the forecasting model employed. Similar patterns were observed for India, the Philippines, and Zambia, where the bayesian model anticipated a significant increase, while the other models predicted relatively stable case numbers. In the case of Mozambique and Nigeria, the ARIMA, bayesian, and hybrid models foresee a surge in TB cases, while the neural network model predicts minimal changes. South Africa's future TB incidence shows divergent forecasts, with the ARIMA model suggesting stability and the remaining models indicating an increase. Myanmar's projections also vary, with the bayesian model forecasting a decline while the other models predict minimal changes. Indonesia poses significant differences among the models, with the ARIMA and bayesian models forecasting a notable increase, while the hybrid and Neural Network models suggest periodic fluctuations

in TB cases. When considering forecasts based on pre-COVID-19 data using the four models (Supplementary Figure S1, available at https://weekly.chinacdc.cn/), it was noted that apart from Congo, Mozambique, and Nigeria, the forecast outcomes for other countries were influenced to varying degrees by the COVID-19 pandemic. For instance, the impact of COVID-19 has altered the upward trend in India and lessened the decline in China, and all models for Indonesia now predict an increase in TB cases.

Figure 3 and Supplementary Figure S2 (available at https://weekly.chinacdc.cn/) present a detailed overview of the projected TB case trends in the 10 HBCs aimed at reaching the goal of TB elimination. Besides China, which is forecasted to meet the TB elimination target according to the bayesian model, the other countries are unlikely to reach the desired goal across all forecasting models due to the impact of COVID-19. Excluding the influence of COVID-19, only the forecast results for South Africa based on the bayesian model indicate potential success in achieving this goal.

DISCUSSION

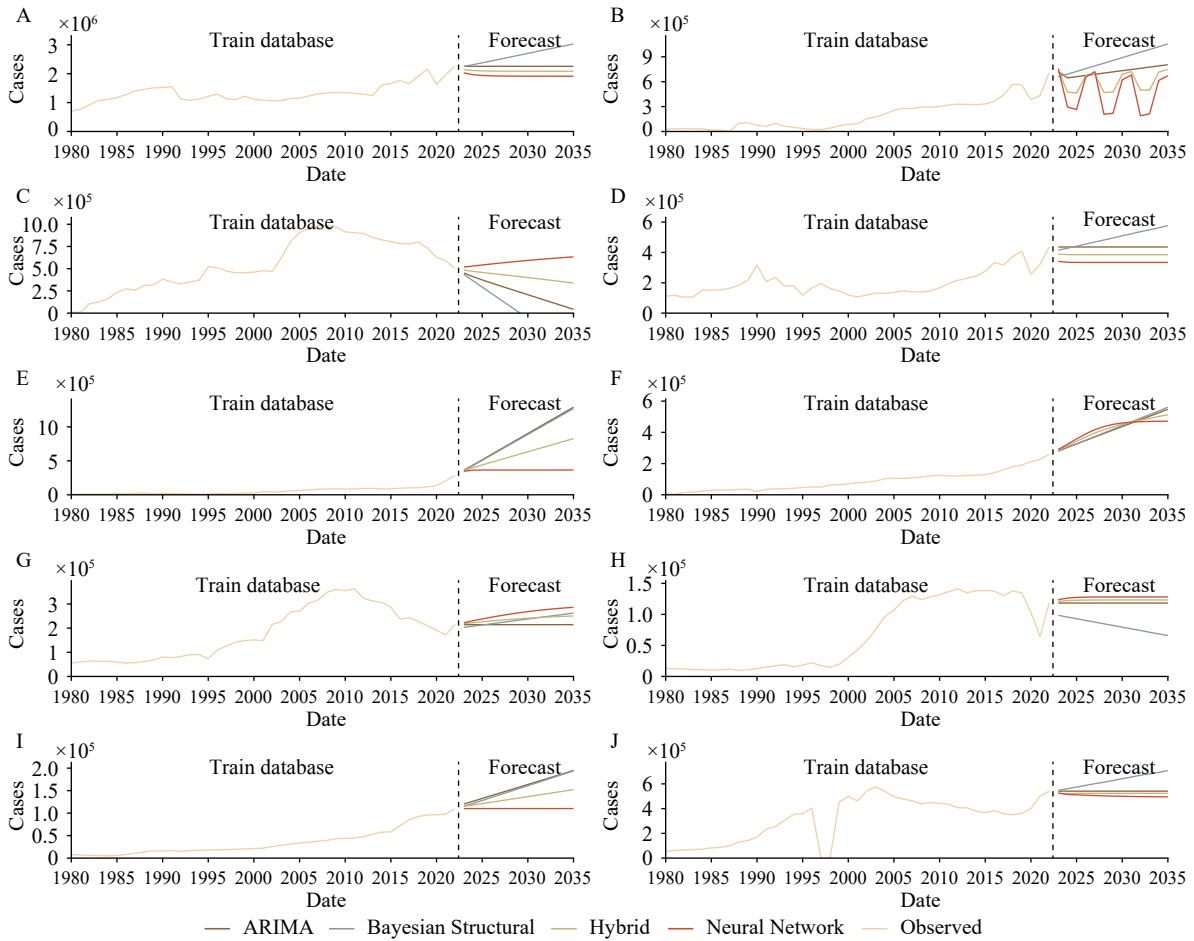This study suggests that a majority of countries are

FIGURE 2. Time trends forecasted by four models in ten HBCs. (A) India; (B) Indonesia; (C) China; (D) Philippines; (E) Nigeria; (F) Congo; (G) South Africa; (H) Myanmar; (I) Mozambique; (J) Zambia.
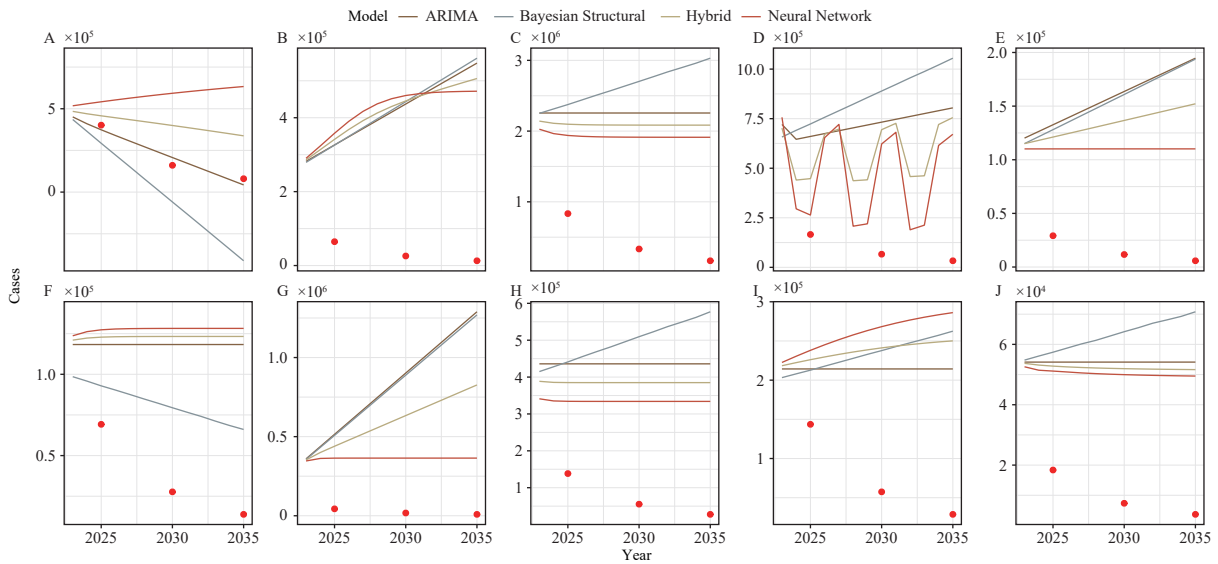Abbreviation: HBC=high-burden countries.

FIGURE 3. Accessibility projections for ten HBCs to achieve the goal of ending tuberculosis based on four models. (A) China; (B) Congo; (C) India; (D) Indonesia; (E) Mozambique; (F) Myanmar; (G) Nigeria; (H) Philippines; (I) South Africa; (J) Zambia.

projected to experience a sustained rise in TB cases from 2023 to 2035. The endeavor to eliminate pulmonary TB poses a significant challenge, underscoring the imperative for enhanced implementation of pertinent interventions.

Within the 10 HBCs, the projections for TB are discouraging. These countries, which are on all three of the major burden lists, contribute to over 60% of the global TB cases reported. India, in particular, has the highest TB incidence rate and is experiencing an annual increase. This underlines the urgent need for India to enhance its TB management strategies. The 2023 Global TB Report also highlights that China continues to be a country heavily afflicted by TB, with an estimated 748,000 cases in 2022, trailing only behind India, with 2.82 million cases, and Indonesia, with 1.06 million cases. Although models project a reduction in TB occurrences in China, a considerable divergence persists from the target to eliminate TB. Furthermore, projections indicate an increase in TB cases for the remaining eight countries, posing a serious challenge for TB prevention and control. Model predictions reflect varying degrees of influence by the COVID-19 pandemic, which may be attributed to alterations in population behaviors, implementation of preventive strategies, and data collection and quality issues during the pandemic. These factors introduce potential uncertainties and biases into the predictive models. Consequently, it is critical to meticulously consider the impact of the COVID-19 pandemic when estimating future TB rates to ensure that the projections are accurate and dependable (*1*,*3*).

The burden of TB in the 10 HBCs is significant, further complicated by issues like drug-resistant strains and the co-infection of TB with HIV. These nations are vital in the global fight against TB, with their struggles and achievements holding crucial insights for worldwide TB control endeavors. Between 2016 and 2020, Russia made notable strides in TB prevention and control, meeting the WHO's targets for 2020. This progress was linked to the implementation of new diagnostic tools, successful treatment protocols, expanded population coverage for TB screening, and the establishment of healthcare programs supporting TB patients, particularly those with HIV co-infection (*7*). With the challenges faced by these HBCs in achieving the goals of the End Tuberculosis Strategy, there is an immediate need to learn from their accomplishments and fortify global TB prevention and control measures. This entails embracing technological innovations, enforcing effective regulations, refining strategic frameworks, and enhancing service-oriented schemes.

This study is subject to some limitations. The analysis utilizes notification data from the WHO online database, which could fluctuate due to changes in national surveillance systems, potentially impacting underreporting and missed diagnoses. Consequently, these data may not entirely reflect the actual trends in TB incidence, particularly over prolonged periods. Additionally, our predictive modeling is assumption-based and susceptible to influences such as policy changes and healthcare infrastructure.

In conclusion, the burden of TB is increasing in the 10 HBCs, particularly in India, Indonesia, and China. Variations exist among prediction models, with the predictions being impacted to varying degrees by the COVID-19 pandemic. Attaining the target of eradicating TB is proving to be difficult for many countries. These results highlight the necessity of tailored interventions and continuous initiatives to manage TB, especially in countries with substantial upward trends like Indonesia, Nigeria, and Congo. Additionally, they stress the importance of enhanced international collaboration.

Conflicts of interest: The funder had no role in the study design, data collection and analysis, decision to publish, or manuscript preparation. No other conflicts of interest.

Corresponding authors: Tianmu Chen, chentianmu@xmu.edu.cn; Roger Frutos, frutossmt@gmail.com.

1 CIRAD, Intertryp, Montpellier, France; 2 Université de Montpellier, Montpellier, France; 3 State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, National Innovation Platform for Industry-Education Integration in Vaccine Research, Xiamen University, Xiamen City, Fujian Province, China; 4 State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory, School of Public Health, Xiamen University, Xiamen City, Fujian Province, China; 5 Espace-Dev, Université de Montpellier, Montpellier, France.
& Joint first authors.

REFERENCES

1. World Health Organization. Global tuberculosis report 2022. 2022. https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022. [2023-12-21].
2. World Health Organization. The end TB strategy. 2015. https://www.who.int/teams/global-tuberculosis-programme/the-end-tb-strategy.

[2023-12-21].

3. World Health Organization. Global tuberculosis report 2023. 2023. https://www.who.int/publications/i/item/9789240083851.[2023-12-21].

4. Yan CQ, Wang RB, Liu HC, Jiang Y, Li MC, Yin SP, et al. Application of ARIMA model in predicting the incidence of tuberculosis in China from 2018 to 2019. Chin J Epidemiol 2019;40(6):633 – 7. https://doi.org/10.3760/cma.j.issn.0254-6450.2019.06.006.

5. Li ZM, Li YN. A comparative study on the prediction of the BP artificial neural network model and the ARIMA model in the incidence of AIDS. BMC Med Inform Decis Mak 2020;20(1):143. https://doi.org/10.1186/s12911-020-01157-3.

6. Paramasivan K, Sudarsanam N. Impact of COVID-19 pandemic on road safety in Tamil Nadu, India. Int J Inj Contr Saf Promot 2022;29(2):265 – 77. https://doi.org/10.1080/17457300.2021.2007134.

7. Starshinova A, Dovgalyk I, Beltukov M, Zinchenko Y, Glushkova A, Starshinova AY, et al. Tuberculosis in the Russian federation: dynamics of the epidemic indicators before and after COVID-19 pandemic. Life 2022;12(10):1468. https://doi.org/10.3390/life12101468.
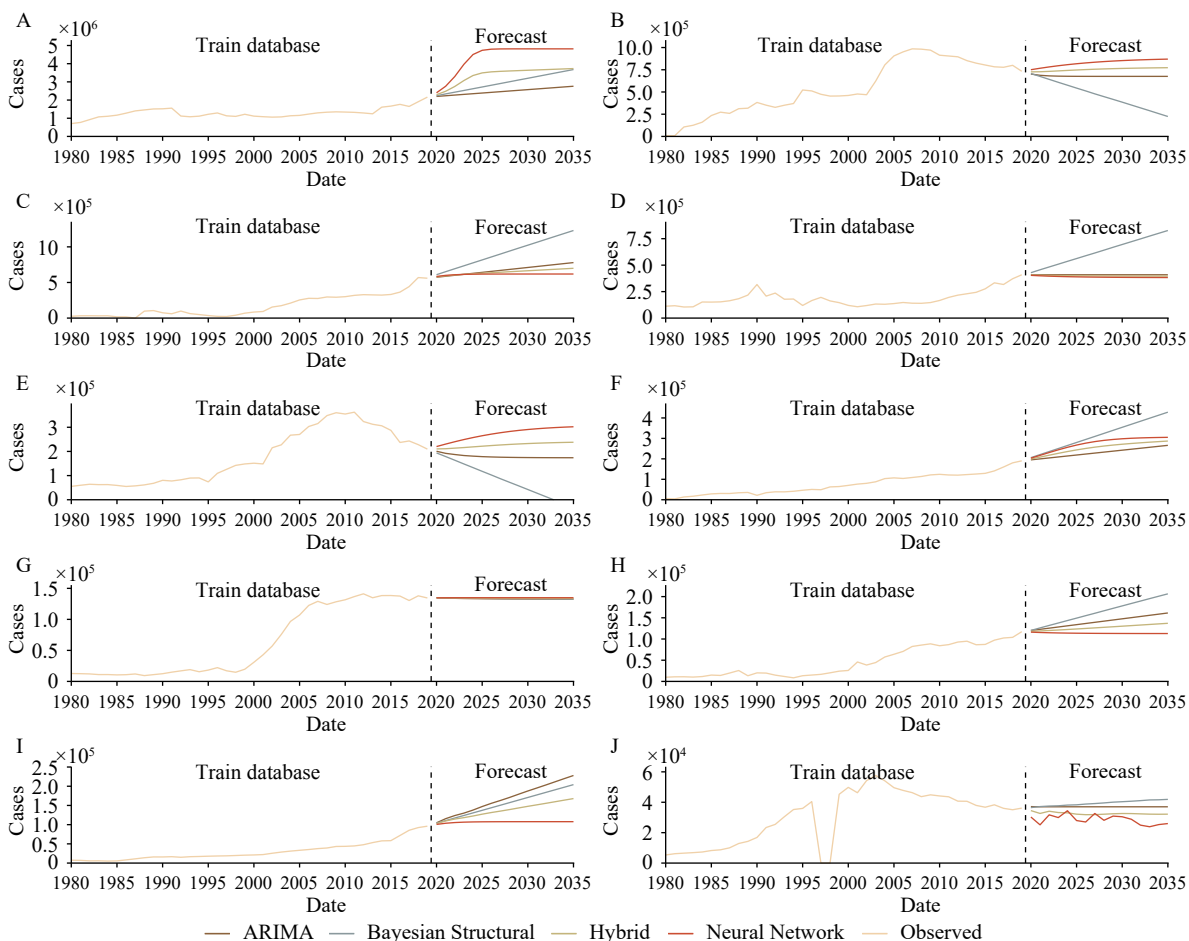
SUPPLEMENTARY MATERIAL

Methods

The Autoregressive Integrated Moving Average (ARIMA) model was specified as ARIMA (p, d, q), where p represents autoregressive terms, d is the number of differences for stationarity, and q shows the moving average terms. Model selection was based on the Akaike Information Criterion. The bayesian structural model was implemented using Markov chain Monte Carlo sampling to estimate the posterior distribution of the time series parameters. The neural network model comprised an ensemble of 20 networks, each with randomly initialized weights, trained for one-step forecasting and recursively applied for multi-step predictions.

When attempting to forecast future trends in tuberculosis case reports using time series models, the selection of appropriate models is crucial. The rationale behind choosing the "bayesian structural," "ARIMA," and the hybrid model combining elements of "neural network" with seasonal autoregressive integrated moving average (SARIMA), exponential smoothing (ETS), and seasonal and trend decomposition using Loess (STL) components is based on their unique strengths and capabilities in capturing different aspects of time series data.

The bayesian structural model was chosen for its ability to incorporate prior knowledge and uncertainty into the modeling process. By leveraging bayesian inference, this model can provide more robust estimates and predictions, especially when dealing with complex data patterns and structural changes over time.

SUPPLEMENTARY FIGURE S1. Results of Four Model Fits and Predictions 10 HBCs during the pre-COVID-19 Era. (A) India; (B) China; (C) Indonesia; (D) Philippines; (E) South Africa; (F) Congo; (G) Myanmar; (H) Nigeria; (I) Mozambique; (J) Zambia.
Abbreviation: HBC=high-burden countries; ARIMA=autoregressive integrated moving average; COVID-19=coronavirus disease 2019.
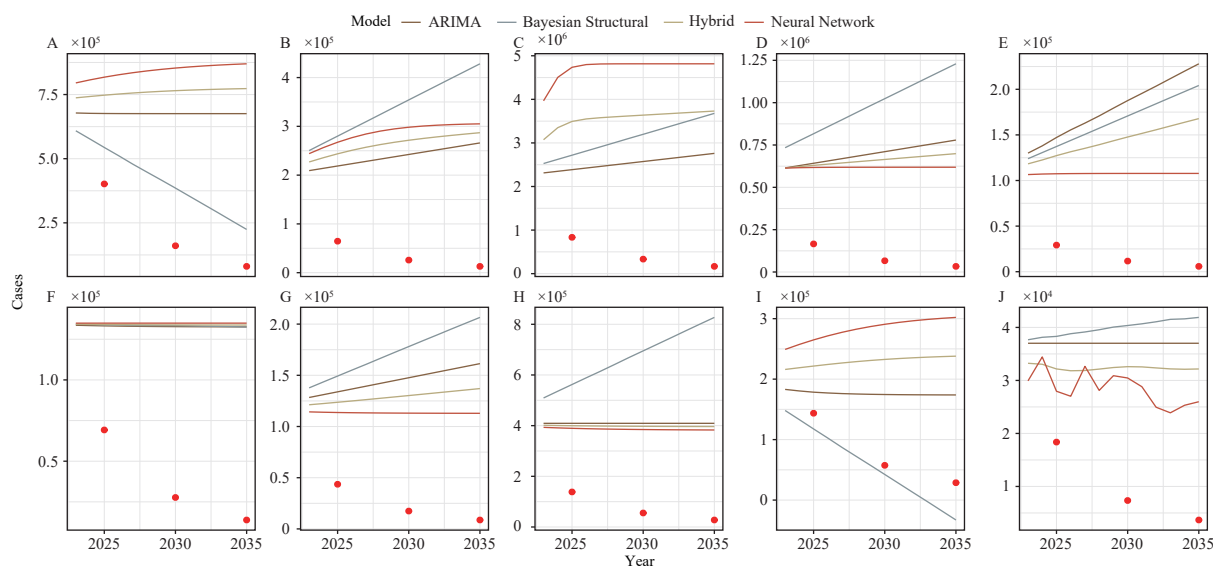
SUPPLEMENTARY FIGURE S2. Feasibility of Ending Tuberculosis during pre-COVID-19 predictions for 10 high-burden countries. (A) China; (B) Congo; (C) India; (D) Indonesia; (E) Mozambique; (F) Myanmar; (G) Nigeria; (H) Philippines; (I) South Africa; (J) Zambia.
Abbreviation: HBC=high-burden countries; ARIMA=autoregressive integrated moving average; COVID-19=coronavirus disease 2019.

ARIMA model is a classic choice for time series forecasting due to its effectiveness in capturing linear dependencies and stationary patterns in the data. It is particularly useful when the data exhibit trends and seasonal variations that can be modeled through autoregressive and moving average components.

The neural network model was chosen for its ability to capture nonlinear relationships and complex dependencies in the data, offering flexibility to model patterns that may not be captured effectively by traditional time series models like ARIMA.

The hybrid model integrates SARIMA, ETS, and STL components with neural network components to leverage the strengths of each approach. SARIMA and ETS can handle seasonality and trend components well, while STL can capture irregular patterns. The neural network component enhances the model's capability to capture nonlinear relationships and complex dependencies in the data.

In summary, the selection of the bayesian structural model, ARIMA model, neural network model, and the hybrid model that combines SARIMA, ETS, and STL with neural network components is based on the diverse strengths of these models in capturing different aspects of the tuberculosis case reports data, including linear dependencies, nonlinear relationships, seasonal patterns, trends, irregularities, and structural changes. This comprehensive approach aims to improve the accuracy and robustness of the forecasting process for tuberculosis case reports.