TB-DeepSeek-Agent: A Large Language Model for Tuberculosis Incidence Prediction with Web-Based Automatic Report

Mingming Chen, Yu Lu, and Tenglong Li(✉)

Xi'an Jiaotong-Liverpool University, Suzhou 215123, China
Tenglong.Li@xjtlu.edu.cn

Abstract. Tuberculosis (TB) remains a global health crisis, necessitating accurate prediction models to guide resource allocation and interventions. Traditional time-series and machine learning models often lack flexibility and precision in capturing complex TB trends. This study introduces TB-DeepSeek-Agent, a novel framework integrating a large language model (DeepSeek) with an Agent-based architecture to enhance TB incidence forecasting. The model leverages historical data, memory mechanisms, and reflective thinking to refine predictions dynamically. Evaluated on 135,802 TB cases from five Chinese cities (2011–2021), TB-DeepSeek-Agent achieved superior performance, with a 100% acceptance ratio (5% error margin), the lowest deviation ratio (0.0129), and RMSE (30.5876), outperforming traditional models (e.g., ARIMA, LSTM) and standalone LLM-based approaches. A web-based application was developed to automate simulations and generate actionable reports, streamlining TB research and policymaking. This work demonstrates the potential of Agent-enhanced AI models in public health analytics, offering a cost-effective solution for TB trend prediction and resource optimization.

Keywords: Tuberculosis Prediction · Large Language Model · Agent Reflective Mechanism · Public health analytics

1 Introduction

Tuberculosis (TB) persists as a global health priority, ranking as the second leading infectious cause of death worldwide [1–3]. Effective prevention and control strategies rely on timely and accurate TB data to inform policy and track epidemiological trends. Countries such as Nigeria have established surveillance systems to collect comprehensive data [4–7], yet the logistical complexity and high costs of large-scale screening impede data acquisition [8, 9]. Predictive modeling offers a viable alternative, leveraging historical data to enable cost-efficient forecasting and timely interventions such as active case-finding. This approach enhances resource optimization and targeted public health strategies, curbing transmission and mitigating outbreak risks [10]. Accurate prediction is critical: underestimation may lead to insufficient resource allocation, while overestimation risks economic strain. Thus, precise forecasting is essential for effective

H. Liu and D. Guo (Eds.): CIAC 2025, LNEE 1503, pp. 137–146, 2026.
https://doi.org/10.1007/978-981-95-4049-5_14

intervention design and understanding TB's full impact [11]. Traditional time-series models often lack flexibility to capture the complex, evolving dynamics of TB data [12–14]. Consequently, machine learning (ML) and artificial intelligence (AI) methods are increasingly explored to improve trend predictions [15]. Recent studies indicate that hybrid (e.g., CNN-LSTM) and ensemble models (e.g., XGBoost) show superior performance for TB incidence forecasting [16, 17].

Despite their potential, many ML and AI models still face challenges related to robustness, generalizability, and accuracy, often struggling to meet the diverse quality standards required by different health authorities [18, 19]. Additionally, the inconsistent temporal dynamics and limited feature availability in TB data complicate the training and performance of AI models. Leveraging the architecture of Agent, we can potentially predict TB cases with higher accuracy and thereby insight the TB situation correctly. This paper presents the development of an advance Agent-based forecasting model incorporating large language model DeepSeek for TB prediction (LLMTBP), coupled with an integrated application designed to generate accurate TB simulation data for research applications. The Agent-based mechanism can effectively integrate the predictive values generated by LLMTBP and historical TB records, leveraging memory to uncover deeper temporal patterns in TB trends. This paper presents the TB-DeepSeek-Agent model, combining Agent-based forecasting with DeepSeek for accurate TB prediction, and an integrated web-based application for generating simulation data for research. The principal contributions of this study are:

– We have developed a TB-DeepSeek-Agent computational model that predicts TB infection data more precisely.
– An interactive website has been designed to integrate our model, allowing researchers to perform TB data analysis accurately and conveniently.

2 Method

2.1 Data

Using the dataset sourced from Jiangsu Provincial Center for Disease Control and Prevention, we collected daily TB infection cases from Changzhou, Suzhou, Nanjing, Zhenjiang and Wuxi within Jiangsu Province, spanning from 2011 to 2021. After excluding the missing data, there are totally 135802 TB cases. For the purpose of this study, the data has been divided into two main sets: the training dataset and the testing dataset. The training dataset consists TB infection data from the years 2011 to 2017, which is used to train our predictive model. The testing dataset contains the data from 2018 to 2021, which serves to evaluate the performance of our model.

2.2 TB-DeepSeek-Agent Architecture

As illustrated in Fig. 1. The conceptual architecture of TB-DeepSeek-Agent, our TB-DeepSeek-Agent model consist of three core components: Large Language Model for TB Prediction (LLMTBP), the memory and the reflection flow (For the other three Agent-based models, each incorporates its own original approach to TB prediction while integrating memory and reflection flow within the same Agent framework as Fig. 1. The conceptual architecture of TB-DeepSeek-Agent). The TB data could be input into LLMTBP and undergoes preprocessing to ensure the dataset is cleaned where missing values are deleted. These preprocessing steps are essential for maintaining the integrity and reliability of the output generated by DeepSeek. It employs multi-scale spatiotemporal attention mechanisms to dynamically identify and prioritize critical TB features across diverse time horizons, thereby enhancing predictive decision-making [20, 21]. The following two systematically designed prompts were developed to generate summaries regarding TB-case distribution:

- Daily TB-case Summary: Records daily TB cases from 2018 to 2021, along with a daily summary.
- Monthly TB-case Summary: Captures monthly TB usage over the same period, summarized monthly.

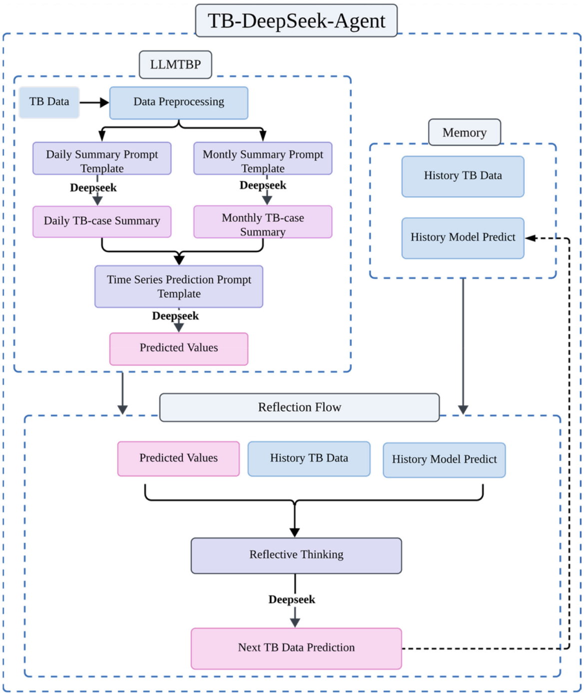

Fig. 1. The conceptual architecture of TB-DeepSeek-Agent

Equations 1 and 2 describe the methods for summarizing daily and monthly TB cases, respectively. (See parameter descriptions in Table 1)

$$\mathrm{RN}_{c,\mathrm{t_d}} = \mathrm{DSummarize}\left(\mathrm{P}_{c,\mathrm{t_d}}\right) \tag{1}$$

$$\mathrm{RN}_{c,\mathrm{t_m}} = \mathrm{MSummarize}\left(\mathrm{P}_{c,\mathrm{t_m}}\right) \tag{2}$$

These summaries are then combined for designed time intervals to provide further inputs based on the provided dataset to engineer the final time series prompts. After that, utilizing the time series prompt template, LLMTBP computes the predicted values defined by Eq. 3. (See parameter descriptions in Table 1)

$$\mathrm{PN}_{c,\mathrm{t_y}} = \Delta\mathrm{Predict}\big(\mathrm{RN}_{c,\mathrm{t_d}}, \mathrm{RN}_{c,\mathrm{t_m}}\big) \tag{3}$$

At this stage, the previously outlined illustration represents the complete structure of the LLMTBP model. In contrast, our proposed TB-DeepSeek-Agent model introduces significant advancements by enhancing the initial predictions generated by LLMTBP. Specifically, in the TB-DeepSeek-Agent model, the prediction process begins by calling LLMTBP to generate an initial set of predicted values for the TB cases. While LLMTBP serves as a strong foundation for the initial prediction, the TB-DeepSeek-Agent model incorporates additional mechanisms that refine and improve these predictions by leveraging past information and reflective thinking of DeepSeek. The first key addition is the memory component, which stores both historical TB data and the TB-DeepSeek-Agent's own previous predictions. This memory mechanism is essential, as it allows the model to not only consider the current epidemiological trends but also to recall and learn from past predictions and data. By integrating this historical information, the model is able to capture long-term trends and evolving patterns in TB incidence, which is crucial for improving the accuracy and reliability of predictions over time. Once the memory component has been called by Agent, this information is fed into the reflection flow. The reflection mechanism dynamically adjusts predictions by analyzing historical prediction errors from the LLMTBP model. Specifically, it calculates the 95% confidence interval (95%CI) of these errors over a sliding time window. Based on the upper and lower bounds of this interval, the system automatically generates a dynamic correction factor. This correction is applied to the raw LLMTBP output, enabling the TB-DeepSeek-Agent model to compensate for systematic biases and temporal uncertainties, ultimately producing optimized and statistically robust predictions. By considering the broader epidemiological context, the reflective flow enhances the model's ability to account for subtle shifts in the data that might otherwise be overlooked. It allows the system to dynamically adjust its predictions based on patterns observed over time, facilitating a more responsive and context-aware approach to TB prediction. The strength of the DeepSeek reflection lies in its capacity to continuously refine the model's understanding of complex, nonlinear trends, ensuring that the predictions are not static but instead reflect the most up-to-date and nuanced epidemiological insights. (See parameter descriptions in Table 1)

$$\mathrm{DPN}_{\mathrm{c},\mathrm{t_y}} = \Delta\mathrm{DPredict}\big(\mathrm{M}_{d,\mathrm{t_y}}, \mathrm{PN}_{c,\mathrm{t_y}}\big) \tag{4}$$

Additionally, the final predicted TB cases, as described above, are simultaneously provided as responses to the memory for its update.

Through the integration of LLMTBP, the memory and the reflective thinking, TB-DeepSeek-Agent provides more refined control over the data, ensuring that predicted TB cases are accurate. This refinement improves the accuracy of TB case predictions compared to the original LLMTBP model.

Table 1. Definitions of Parameters Used in Mathematical Formulas

Parameter	Description
RN_{c,t_d} :	Real number of TB cases of city c on day t_d.
RN_{c,t_m} :	Real number of TB cases of city c in month t_m.
P_{c,t_d} :	An individual being diagnosed as a TB case in city c on day t_d .
P_{c,t_m} :	An individual being diagnosed as a TB case in city c on day t_m .
DSummarize():	Function utilizes the LLMTBP model to summarize daily TB cases.
MSummarize():	Function utilizes the LLMTBP model to summarize monthly TB cases.
PN_{c,t_y} :	Predicted number of TB cases of city c for the year t_y using the LLMTBP model.
ΔPredict():	The prediction operation using the LLMTBP model.
M_{d,t_y} :	List of history TB Data d and history model predict of year t_y in the memory.
ΔDPredict():	The prediction operation using the DeepSeek model after applying reflective thinking.
DPN_{c,t_y} :	TB predicted number of city c for the year t_y

3 Result

To assess the robustness and reliability of the TB-DeepSeek-Agent predictive model, we conducted a comparative analysis of the forecasting accuracy of eight algorithms for the number of TB infections in five cities from 2018 to 2021. In evaluating the performance of these TB infection forecasting algorithms, we utilized four key metrics: acceptance ratio, deviation ratio, and Root Mean Square Error (RMSE).

The comprehensive performance comparison of the TB prediction models for the five cities presented in Table 2**.** Forecast results for 5 cities in Jiangsu from 2018 to 2021 highlights significant variations among the algorithms. Among traditional models, LR demonstrates relatively low prediction accuracy, with an acceptance ratio of 30% and considerable deviation from actual values evidenced by a high deviation ratio (0.0822) and RMSE (208.3093). In contrast, the ARIMA model shows a notable performance advantage, with an acceptance ratio of 85% and a reduced RMSE (86.81) compared to LR.

Nevertheless, the real breakthrough in performance comes with the introduction of the Agent-based mechanism, which substantially improves model accuracy. Specifically, combining the Agent with LR dramatically improves its acceptance ratio from 30% to 85%, with the RMSE dropping from 208.3093 to 82.9696. Though the performance difference between ARIMA and ARIMA-Agent is not significant, the higher acceptance ratio (from 85% to 90%) illustrates the Agent-based model's ability to adapt quickly to new patterns. A more noticeable improvement is seen with LSTM when adding the Agent mechanism: the acceptance ratio increases from 65% to 100%, the deviation ratio drops from 0.0407 to 0.0157, and RMSE reduces from 100.7956 to 50.2668. This demonstrates the effectiveness of integrating the Agent framework into deep learning models for accurate TB infection case prediction. Among all the models compared, TB-DeepSeek-Agent yielded the most accurate predictions, with the highest acceptance ratio (100%), the lowest deviation ratio (0.0129), and the lowest RMSE (30.5876).

Table 2. Forecast results for 5 cities in Jiangsu from 2018 to 2021

Algorithms	Acceptance Ratio	Deviation Ratio	RMSE
LR	30.00%(6/20)	0.0822	208.3093
LR-Agent	85.00%(17/20)	0.0277	82.9696
LSTM	65.00%(13/20)	0.0407	100.7956
LSTM-Agent	100.00%(20/20)	0.0157	50.2668
ARIMA	85.00%(17/20)	0.0271	86.8159
ARIMA-Agent	90.00%(18/20)	0.0205	75.3356
LLMTBP	35.00%(7/20)	0.0678	178.6214
TB-DeepSeek-Agent	100.00%(20/20)	0.0129	30.5876

Note: The four evaluation metrics used for model comparisons are explained as follows: Acceptance ratio is anchored in the requirement for predictions to maintain a 5% margin of error relative to true values, as this threshold is widely recognized as the standard for high-precision health surveillance analytics [22]*; Deviation Ratio refers to the average absolute deviation of predicted values from the true values, normalized by the true values, which provides a measure of relative predictive accuracy; RMSE measures the magnitude of the prediction errors by calculating the square root of the average squared differences between the predicted and observed values.*

We further examined the models' annual predictions for each city from 2018 to 2021 (Appendix Table 1). The TB-DeepSeek-Agent model consistently exhibited outstanding stability and accuracy, significantly surpassing the performance of all other models. For instance, it achieved perfect prediction accuracy in Zhenjiang (2020), with predicted and actual values both at 928 cases (ratio of predicted to true value: 1.0000). In Changzhou (2018), the model exhibited minimal deviation, predicting 1578 cases against the actual 1579 (ratio: 0.9994), demonstrating its ability to capture subtle epidemiological changes. In Nanjing (2021), the model predicted 2715 cases versus the actual 2712, showing only a 3-case difference (ratio: 1.0011), confirming its reliability in TB trend forecasting. The TB-DeepSeek-Agent model also demonstrated superior performance in complex epidemiological scenarios. In Changzhou (2019), it accurately predicted the upward trend, forecasting an increase from 1578 to 1610 cases (actual increase: 1579 to 1590, ratio: 1.0126), while other models failed to capture this trend. Similarly, in Nanjing's sustained growth period (2019–2020), TB-DeepSeek-Agent's predictions (2471 and 2522 cases) closely matched actual values (2485 and 2537), whereas other models miscalculated the trend. In Wuxi, where the downward trend persisted from 2018 to 2021, TB-DeepSeek-Agent consistently captured this reduction with minimal deviations (2255 to 2205 predicted versus 2205 to 2197 actual), outperforming ARIMA, which erroneously projected a 6.8% increase in 2021.

Finally, we developed an advanced web-based platform integrated with the TB-DeepSeek-Agent model for simulating TB data. This software is designed to transform TB research by providing a robust environment for comprehensive simulation experiments. In addition to automating TB simulations, the system generates detailed analytical reports, transforming complex datasets into actionable insights. By leveraging DeepSeek's advanced analytical capabilities, the platform produces well-structured reports that researchers can directly use, streamlining their workflow and enhancing the efficiency of their studies. This innovative tool offers researchers precision, depth, and convenience, making it an invaluable asset in the ongoing fight against TB (See Fig. 2).

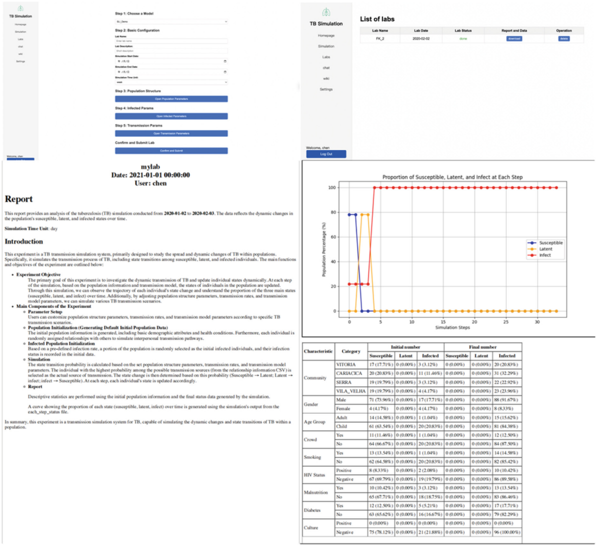

Fig. 2. The Website and Report of Our Application

4 Conclusion

This study demonstrates the efficacy of TB-DeepSeek-Agent, a novel framework integrating Agent architecture, reflective workflow with memory, and a large language model. Evaluated on TB infection records from five Jiangsu Province cities, the model significantly enhances prediction accuracy. Comparative analysis against seven baseline models revealed substantial error reduction, particularly outperforming the Large Language Model for TB Prediction (LLMTBP). The TB-DeepSeek-Agent architecture shows strong potential for improving TB forecasting, offering critical advantages for optimizing policy development and resource allocation in data collection and control programs. However, this study acknowledges limitations including potential data biases from missing value handling in historical records and unmeasured external confounding factors (e.g., healthcare disruptions during the COVID-19 pandemic). Future work should explore its generalizability across diverse regions and extend its application to broader prediction tasks.

Appendix

Appendix Table 1. Forecasting results for 5 cities in Jiangsu from 2018 to 2021

		2018	2019	2020	2021
Changzhou	True	1579	1590	1350	1241
	LR	1724(1.0918)	1762(1.1082)	1539(1.1400)	1365(1.0999)
	LR-Agent	1589(1.0063)	1652(1.0390)	1429(1.0585)	1240(0.9992)
	LSTM	1706(1.0804)	1731(1.0887)	1395(1.0333)	1319(1.0629)
	LSTM-Agent	1597(1.0114)	1586(0.9975)	1402(1.0385)	1222(0.9847)
	ARIMA	1489(0.9430)	1612(1.0138)	1366(1.0119)	1216(0.9799)
	ARIMA-Agent	1608(1.0184)	1598(1.0050)	1363(1.0096)	1262(1.0169)
	LLMTBP	1664(1.0538)	1566(0.9849)	1474(1.0919)	1336(1.0766)
	TB-DeepSeek-Agent	1578(0.9994)	1610(1.0126)	1405(1.0407)	1210(0.9750)
Nanjing	True	2480	2485	2537	2712
	LR	2758(1.1121)	2715(1.0926)	2471(0.9740)	2601(0.9591)
	LR-Agent	2584(1.0419)	2523(1.0153)	2453(0.9669)	2875(1.0601)
	LSTM	2519(1.0157)	2544(1.0237)	2443(0.9629)	2767(1.0203)
	LSTM-Agent	2536(1.0226)	2466(0.9924)	2498(0.9846)	2761(1.0181)
	ARIMA	2466(0.9944)	2378(0.9569)	2454(0.9673)	2874(1.0597)
	ARIMA-Agent	2520(1.0161)	2475(0.9960)	2604(1.0264)	2903(1.0704)
	LLMTBP	2290(0.9234)	2243(0.9026)	2245(0.8849)	2339(0.8625)
	TB-DeepSeek-Agent	2517(1.0149)	2471(0.9944)	2522(0.9941)	2715(1.0011)
Suzhou	True	4095	3597	3120	3292
	LR	4283(1.0459)	4028(1.1198)	3479(1.1151)	3185(0.9675)
	LR-Agent	4220(1.0305)	3528(0.9808)	3203(1.0266)	3456(1.0498)
	LSTM	4337(1.0591)	3734(1.0381)	3210(1.0288)	3238(0.9836)
	LSTM-Agent	4183(1.0215)	3498(0.9725)	3218(1.0314)	3287(0.9985)
	ARIMA	3967(0.9687)	3388(0.9419)	3175(1.0176)	3254(0.9885)
	ARIMA-Agent	4297(1.0493)	3632(1.0097)	3117(0.9990)	3288(0.9988)
	LLMTBP	3891(0.9502)	3857(1.0723)	3448(1.1051)	3144(0.9550)
	TB-DeepSeek-Agent	4141(1.0112)	3586(0.9969)	3125(1.0016)	3237(0.9833)
Wuxi	True	2411	2396	2205	2197
	LR	2719(1.1277)	2566(1.0710)	2447(1.1098)	2450(1.1152)
	LR-Agent	2472(1.0253)	2378(0.9925)	2340(1.0612)	2252(1.0250)

(*continued*)

Appendix Table 1. (*continued*)

		2018	2019	2020	2021
	LSTM	2446(1.0145)	2368(0.9883)	2383(1.0807)	2300(1.0469)
	LSTM-Agent	2392(0.9921)	2385(0.9954)	2309(1.0472)	2210(1.0059)
	ARIMA	2414(1.0012)	2330(0.9725)	2229(1.0109)	2356(1.0724)
	ARIMA-Agent	2453(1.0174)	2351(0.9812)	2350(1.0658)	2212(1.0068)
	LLMTBP	2322(0.9631)	2333(0.9737)	2279(1.0336)	2230(1.0150)
	TB-DeepSeek-Agent	2377(0.9859)	2366(0.9875)	2255(1.0227)	2205(1.0036)
Zhenjiang	True	1234	1179	928	949
	LR	1207(0.9781)	1250(1.0602)	973(1.0485)	1010(1.0643)
	LR-Agent	1279(1.0365)	1181(1.0017)	936(1.0086)	956(1.0074)
	LSTM	1250(1.0130)	1258(1.0670)	980(1.0560)	968(1.0200)
	LSTM-Agent	1266(1.0259)	1185(1.0051)	929(1.0011)	946(0.9968)
	ARIMA	1214(0.9838)	1150(0.9754)	932(1.0043)	927(0.9768)
	ARIMA-Agent	1257(1.0186)	1216(1.0314)	931(1.0032)	967(1.0190)
	LLMTBP	1122(0.9092)	1089(0.9237)	981(1.0571)	870(0.9168)
	TB-DeepSeek-Agent	1227(0.9943)	1226(1.0399)	928(1.0000)	930(0.9800)

References

1. World Health Organization: Global Tuberculosis Report 2024. World Health Organization, Geneva (2024)
2. Dheda, K., Perumal, T., Moultrie, H., Perumal, R., Esmail, A., Scott, A.J., et al.: The intersecting pandemics of tuberculosis and COVID-19: population-level and patient-level impact, clinical presentation, and corrective interventions. Lancet Respir. Med. **10**(6), 603–622 (2022)
3. World Health Organization: Tuberculosis response recovering from pandemic but accelerated efforts needed to meet new targets. https://www.who.int/news/item/07-11-2023-tuberculosis-response-recovering-from-pandemic-but%2D%2Daccelerated-efforts-needed-to-meet-new-targets, last accessed 2025/02/07
4. Kusimo, O.C., Ugwu, C.I., Aduh, U., Okoro, C.A.: Implementing TB surveillance in Nigeria: best practices, challenges and lessons learnt. J. Tuberc. Res. **8**(4), 199–208 (2020)
5. Puyat, J.H., Brode, S.K., Shulha, H., Romanowski, K., Menzies, D., Benedetti, A., et al.: Predicting risk of tuberculosis (TB) disease in people who migrate to a low-TB incidence country: development and validation of a multivariable, dynamic risk-prediction model using health administrative data. Clin. Infect. Dis. **ciae561** (2024)
6. Engel, N., Ochodo, E.A., Karanja, P.W., Schmidt, B.M., Janssen, R., Steingart, K.R., et al.: Rapid molecular tests for tuberculosis and tuberculosis drug resistance: a qualitative evidence synthesis of recipient and provider views. Cochrane Database Syst. Rev. **4** (2022)
7. Gabrielian, A., Engle, E., Harris, M., Wollenberg, K., Juarez-Espinosa, O., Glogowski, A., et al.: TB DEPOT (Data Exploration Portal): A multi-domain tuberculosis data analysis resource. PLoS One. **14**(5), e0217410 (2019)

8. Cole, B.: Essential components of a public health tuberculosis prevention, control, and elimination program: recommendations of the Advisory Council for the Elimination of Tuberculosis and the National Tuberculosis Controllers Association. MMWR Recomm. Rep. **69** (2020)
9. Sarkar, M.: Tuberculosis infection prevention and control. Indian J. Tuberc. (2024)
10. Gharamaleki, O.G., Colijn, C., Sekirov, I., Johnston, J.C., Sobkowiak, B.: Early prediction of Mycobacterium tuberculosis transmission clusters using supervised learning models. Sci. Rep. **14**(1), 27652 (2024)
11. Siawaya, J.F.D., Bapela, N.B., Ronacher, K., Veenstra, H., Kidd, M., Gie, R., et al.: Immune parameters as markers of tuberculosis extent of disease and early prediction of anti-tuberculosis chemotherapy response. J. Infect. **56**(5), 340–347 (2008)
12. Maipan-Uku, J.Y., Cavus, N.: Forecasting tuberculosis incidence: a review of time series and machine learning models for prediction and eradication strategies. Int. J. Environ. Health Res., 1–16 (2024)
13. Aryee, G., Kwarteng, E., Essuman, R., Agyei, A.N., Kudzawu, S., Djagbletey, R., et al.: Estimating the incidence of tuberculosis cases reported at a tertiary hospital in Ghana: a time series model approach. BMC Public Health. **18**, 1–8 (2018)
14. Liao, Z., Zhang, X., Zhang, Y., Peng, D.: Seasonality and trend forecasting of tuberculosis incidence in Chongqing, China. Interdiscip. Sci. Comput. Life Sci. **11**, 77–85 (2019)
15. Frauenfeld, L., Nann, D., Sulyok, Z., Feng, Y.S., Sulyok, M.: Forecasting tuberculosis using diabetes-related Google Trends data. Pathog. Global Health. **114**(5), 236–241 (2020)
16. Abdualgalil, B., Abraham, S., Ismael, W.M., George, D.: Modeling and forecasting tuberculosis cases using machine learning and deep learning approaches: a comparative study. In: International Conference on Data Management, Analytics & Innovation, pp. 157–171. Springer, Singapore (2022)
17. Hamna Mariyam, K.B., Jose, S.A., Jirawattanapanit, A., Mathew, K.: A comprehensive study on tuberculosis prediction models: integrating machine learning into epidemiological analysis. J. Theor. Biol. **597**, 111988 (2025)
18. Al Sailawi, A.S.A., Kangavari, M.R.: Utilizing AI for extracting insights on post WHO's COVID-19 vaccination declaration from X (Twitter) social network. AIMS Public Health. **11**(2), 349–378 (2024)
19. Awasthi, R., Nagori, A., Nasri, B.: Global Generalisability of AI-Driven COVID-19 Vaccination Policies: A Cross-Sectional Observational Study. medRxiv 2023.01 (2023)
20. Guo, D., et al.: Deepseek-r1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. arXiv preprint arXiv:2501.12948 (2025)
21. Liu, A., et al.: Deepseek-v2: A Strong, Economical, and Efficient Mixture-of-Experts Language Model. arXiv preprint arXiv:2405.04434 (2024).
22. Wang, Y., et al.: Temporal trends analysis of tuberculosis morbidity in mainland China from 1997 to 2025 using a new SARIMA-NARNNX hybrid model. BMJ Open. **9**(7), e024409 (2019)