Electronic Journal of General Medicine
2026, 23(1), em710
e-ISSN: 2516-3507
https://www.ejgm.co.uk/

MODESTUM

Original Article

OPEN ACCESS

Forecasting the future burden of major diseases in Kazakhstan using global burden of disease and time series models (ARIMA, Prophet, LSTM), 2019-2032

Tilektes Maulenkul [1], Temirgali Aimyshev [1], Ainash Oshibayeva [2], Adnan Yazici [3], Mohamad Aljofan [1], Aigerim Japparkulova [2], Abduzhappar Gaipov [1]*

[1] Department of Medicine, Nazarbayev University School of Medicine, Astana, KAZAKHSTAN
[2] Faculty of Medicine, Khoja Akhmet Yassawi Kazakh-Turkish International University, Turkistan, KAZAKHSTAN
[3] Department of Computer Science, School of Engineering and Digital Sciences, Nazarbayev University School of Medicine, Astana, KAZAKHSTAN
***Corresponding Author:** abduzhappar.gaipov@nu.edu.kz

Citation: Maulenkul T, Aimyshev T, Oshibayeva A, Yazici A, Aljofan M, Japparkulova A, Gaipov A. Forecasting the future burden of major diseases in Kazakhstan using global burden of disease and time series models (ARIMA, Prophet, LSTM), 2019-2032. Electron J Gen Med. 2026;23(1):em710. https://doi.org/10.29333/ejgm/17734

ARTICLE INFO

Received: 07 Nov. 2025

Accepted: 24 Dec. 2025

ABSTRACT

Background: Over the last years, it has become a widespread practice to use the global burden of disease (GBD) metrics for anticipating the disease patterns all over the world. Disability-adjusted life years (DALYs) is the main indicator to use for quantifying the losses in terms of health caused by disease. This study projects Kazakhstan's disease burden from 2019 to 2032 by applying and comparing four forecasting approaches: GBD projections, auto-regressive integrated moving average (ARIMA), Prophet, and long short-term memory (LSTM) neural networks.

Methods: DALY data for the top 10 disease categories in Kazakhstan were modeled using Python 3.11 with statistical and machine learning libraries. Each model was trained and validated for short- and medium-term forecasts, with performance compared across trajectory trends and disease ranking stability.

Results: GBD and LSTM models projected stable rankings among the top 10 DALY contributors through 2032, with only malignant neoplasms of the colon and rectum showing a decline, while ARIMA and Prophet exhibited greater temporal fluctuations, predicting a drop in lower respiratory infections. Across all models, noncommunicable diseases, particularly, cardiovascular and metabolic disorders, remain dominant drivers of Kazakhstan's future health burden.

Conclusions: Deep learning (LSTM) and GBD approaches yielded smoother, more robust long-term predictions, whereas ARIMA and Prophet captured short-term variability more sensitively, highlighting the benefit of integrating statistical and AI-based paradigms for comprehensive national health forecasting and policy design.

Keywords: forecasting, disease burden, DALY, global burden of diseases, time series, Kazakhstan

INTRODUCTION

Accurate forecasting of the national disease burden plays a critical role in health system planning, policy prioritization, and sustainable development. The quantification of disease burden at the population level is most commonly expressed through disability-adjusted life years (DALYs), which combine years of life lost due to premature mortality and years lived with disability to capture both fatal and nonfatal health outcomes in a single composite indicator [1, 2]. The global burden of disease (GBD) initiative, developed by the Institute for Health Metrics and Evaluation (IHME), has standardized the global use of DALYs in epidemiological research and policy modeling, providing harmonized datasets across countries and causes from 1990 onward [3, 4]. However, despite its global scope, the GBD model's proprietary nature and lack of methodological transparency have prompted growing concern about its predictive robustness and reproducibility, especially when applied to national-level health systems.

Kazakhstan, the largest nation in Central Asia, has undergone profound demographic, environmental, and socioeconomic transitions in its post-Soviet era, reshaping patterns of morbidity and mortality [5]. The burden of noncommunicable diseases (NCDs) such as ischemic heart disease, stroke, diabetes, and various cancers has risen sharply, surpassing communicable diseases such as chronic respiratory infections as the dominant contributors to DALYs [6]. This epidemiological transition mirrors global trends but poses unique challenges to Kazakhstan's healthcare infrastructure, labor productivity, and policy planning. Understanding the future trajectory of these disease burdens is therefore essential for designing effective prevention, intervention, and financing strategies aligned with the country's "digital Kazakhstan" and sustainable health initiatives.

While GBD projections are widely used to estimate disease burden globally, alternative time series forecasting methods can complement or challenge these results by providing transparent, data-driven predictions [7, 8]. In this study, three such models were employed alongside GBD's own forecasts:

(1) the auto-regressive integrated moving average (ARIMA) model, a classical statistical approach suited for short- to mid-term stationary series,

(2) Prophet, a decomposition-based model developed by meta that excels in identifying seasonal and trend components, and

(3) long short-term memory (LSTM) neural networks, a deep learning framework capable of learning nonlinear temporal dependencies and handling multivariate sequences [9-12].

This trio was chosen to represent 3 distinct "families" of time-series forecasting: statistical (ARIMA), decomposable/additive (Prophet), and deep learning (LSTM). While Bayesian models are excellent for quantifying uncertainty, they are often computationally expensive and require the specification of "priors". ARIMA/Prophet/LSTM trio is a standard "data-driven" benchmark comparison. Together, these methods provide a comprehensive comparison between traditional statistical and modern machine learning paradigms in forecasting the future health landscape.

This study aimed to utilize GBD, ARIMA, Prophet, and LSTM modeling to predict the trajectory of Kazakhstan's top-10 DALY contributors based on the historical GBD data. This is a first study from Kazakhstan to conduct a multi-model comparison based on the GBD data.

MATERIALS AND METHODS

This section outlines the methodological framework used to forecast Kazakhstan's disease burden up to 2032. The analysis integrates epidemiological data from the GBD database with statistical and machine learning-based time-series forecasting techniques. A multi-model approach was employed to ensure robustness and comparative validity, encompassing both classical statistical models and modern deep learning architectures. Specifically, the ARIMA, Prophet, and LSTM models were implemented to project future DALYs across the country's ten leading causes of disease burden. All models were developed, optimized, and evaluated within a unified Python-based computational environment to maintain reproducibility and comparability. The following subsections describe in detail the data sources, preprocessing steps, model configurations, and evaluation criteria adopted in this study.

Statistical Environment and Computational Framework

All analyses were performed using the Python 3.11 programming language within the Anaconda environment on a workstation running Windows 11 Pro (AMD Ryzen 9 7950X CPU, 64 GB RAM, NVIDIA RTX A5000 GPU). The study employed the following Python libraries:

- pandas (v2.2.2) for data wrangling and time-series manipulation,
- numpy (v1.26.4) and scipy (v1.13.1) for numerical and statistical operations,
- pmdarima (v2.0.4) for ARIMA model fitting and diagnostics,
- prophet (v1.1.5) for additive decomposition-based forecasting,
- torch (v2.3.0) for building and training the deep learning model (LSTM), and
- scikit-learn (v1.5.0) for normalization, model evaluation, and error metrics.

All scripts were version-controlled in GitHub and executed with fixed random seeds (42) to ensure full reproducibility.

Data Source and Extraction

The study utilized publicly available GBD data from the IHME VizHub portal [13]. DALY rates (per 100,000 population) were extracted for Kazakhstan over the period 1990-2019, including both sexes and all age groups. The extraction focused on the top ten causes of disease burden identified in 2019 based on total DALYs:

1. Ischemic heart disease
2. Stroke
3. Cirrhosis and other chronic liver diseases
4. Low back pain
5. Chronic obstructive pulmonary disease (COPD)
6. Diabetes mellitus
7. Infections of the lower respiratory tract
8. Headache disorders
9. Malignant neoplasms of the trachea, bronchi, and lungs
10. Malignant neoplasm of the colon and rectum

Data were downloaded in CSV format, cleaned, and aggregated by year. Missing values (if any) were linearly interpolated, and all rates were normalized using min-max scaling to ensure comparability across diseases. Min-max normalization was performed using parameters derived exclusively from the training subset during the validation stage. These parameters were then applied unchanged to the validation data and the full forecasting horizon to prevent information leakage from the hold-out period.

Reference Dataset and Baseline Comparison

For baseline comparison, the GBD official forecast available through the IHME VizHub dashboard was used as the reference projection for 2019-2032 [13]. This reference serves as the "status quo" model against which the performance and trend dynamics of all alternative forecasting methods were evaluated.

Time-Series Forecasting Models

To forecast DALY trajectories from 2019 to 2032, three time-series models were implemented in addition to the GBD reference: ARIMA, Prophet, and LSTM. Each model represents a distinct forecasting paradigm, classical statistical, additive decomposition, and neural sequence modeling, respectively.

ARIMA model

The ARIMA model was applied to capture linear temporal dependencies in DALY series [10].

- Optimal model orders *(p, d, q)* were selected automatically using the auto_arima() function in *pmdarima*, which minimizes the Akaike information criterion.

- Data were differenced to achieve stationarity, confirmed by the Augmented Dickey-Fuller (ADF) test ($p < 0.05$).
- Diagnostic checks on residuals ensured white-noise behavior (no autocorrelation in the Ljung-Box test).
- Forecasts were generated for the 2019-2032 horizon, and 95% confidence intervals were calculated.

The ARIMA model is particularly suited for short-term and moderate-term forecasts, where data exhibit stable autocorrelation but limited nonlinear complexity.

Prophet model

The Prophet forecasting model, developed by Meta (Facebook), was selected for its robustness to missing data, outliers, and seasonality [11]. Prophet decomposes the time-series into trend, seasonality, and holiday components using the Eq. (1):

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, \quad (1)$$

where $g(t)$ is the piecewise linear or logistic growth curve, $s(t)$ represents periodic seasonal effects, $h(t)$ denotes external regressors (none used here), and ϵ_t is the residual error term.

- Trend flexibility was set to "linear" to allow saturation effects in disease burden.
- Changepoint prior scale = 0.05 and seasonality prior scale = 10 were used to control overfitting.
- The yearly seasonality term was enabled.

The final Prophet implementation employed a linear growth specification. Exploratory analyses using logistic growth were conducted during preliminary testing but were discarded due to unstable saturation behavior and reduced long-term forecast plausibility. No external regressors were included.

LSTM model

The LSTM neural network, a subclass of recurrent neural networks, was implemented to capture nonlinear temporal dependencies and long-range patterns in DALY trends [12].

- Input sequences were windowed with a time lag of 5 years, predicting the next-year DALY rate.
- The model architecture consisted of one LSTM layer (hidden size = 50), followed by a fully connected layer (output size = 1).
- Activation function: *tanh*; optimizer: Adam (learning rate = 1×10^{-3}).
- Training used mean squared error as the loss function over 200 epochs, with early stopping based on validation loss (patience = 20).
- The dataset was divided 80:20 into training and validation subsets, with all inputs normalized to (0, 1).

Unlike ARIMA or Prophet, the LSTM model can capture nonlinear, multivariate interactions and adapt to complex, evolving health trends.

Model Evaluation and Comparison

All models (ARIMA, Prophet, LSTM, and GBD) were evaluated based on their forecast DALY trajectories for each disease through 2032.

- Forecast accuracy was assessed using root mean square error (RMSE) and mean absolute scaled error (MASE) for 2015-2019 (out-of-sample validation).
- A single hold-out period (2015-2019) was intentionally used to preserve an adequate training window and stabilize long-term trend extraction.
- Results were compared visually and statistically across methods.
- The relative ranking stability of the top-10 diseases was analyzed to determine whether each model preserved or altered the hierarchy of disease burden.

Forecast results were summarized in:

- Tables, listing 2032 DALY estimates and disease ranks for all models; and
- Figures showing historical data (1990-2019) and projected trajectories (2019-2032) for each disease and model comparison.

Ethical Considerations

The study used publicly available, aggregated data from IHME without any personally identifiable information. Therefore, ethical approval was not required under institutional research policies.

RESULTS

A comprehensive summary of the projected top-10 contributors to DALYs in Kazakhstan for the period 2019-2032 is presented in **Table 1**.

Table 1. Forecasting data on top-10 DALY contributors per 100,000 population in Kazakhstan: GBD and time series (ARIMA, Prophet, and LSTM)

Forecasting model		Rank in 2019	Rank in 2032	DALY in 2019	DALY in 2032 (95% CI, lower/upper)
Ischemic heart disease	GBD	1	1	3,443	3,231
	ARIMA	1	2	3,443	2,320 (from 9,710.6 to 14,657.7)
	Prophet	1	2	3,443	2,176 (from 572.1 to 4,601.7)
	LSTM	1	1	3,443	3,809.8 (from 3,809.8 to 3,809.8)
Stroke	GBD	2	2	2,791	2,563
	ARIMA	2	1	2,791	2,357 (from 593.6 to 4,197.0)
	Prophet	2	1	2,791	2,625 (from 1,837.5 to 3,570.6)
	LSTM	2	2	2,791	3,108.9 (from 3,108.9 to 3,108.9)
Cirrhosis and other chronic liver diseases	GBD	3	3	1,239	1,162
	ARIMA	3	4	1,239	999 (from -1,567.8 to 3,621.6)

Table 1 (Continued). Forecasting data on top-10 DALY contributors per 100,000 population in Kazakhstan: GBD and time series (ARIMA, Prophet, and LSTM)

Forecasting model		Rank in 2019	Rank in 2032	DALY in 2019	DALY in 2032 (95% CI, lower/upper)
	Prophet	3	3	1,239	1,971 (from 1,525.3 to 2,254.2)
	LSTM	3	3	1,239	1,305.5 (from 1,305.5 to 1,305.5)
Low back pain	GBD	4	4	1,070	1,134
	ARIMA	4	3	1,070	1,126 (from 1,076.1 to 1,158.7)
	Prophet	4	4	1,070	1,169 (from 1,111.7 to 1,191.7)
	LSTM	4	4	1,070	1,124 (from 1,124.0 to 1,124.0)
Chronic obstructive pulmonary disease	GBD	5	5	945	1,095
	ARIMA	5	6	945	841 (from 566.0 to 1,118.8)
	Prophet	5	5	945	843 (from 621.9 to 1,091.8)
	LSTM	5	6	945	886.6 (from 886.6 to 886.6)
Infections of the lower respiratory tract	GBD	6	7	837	668
	ARIMA	6	11	837	0
	Prophet	6	11	837	0
	LSTM	6	7	837	642,8 (from 642.8 to 642.8)
Diabetes	GBD	7	6	732	878
	ARIMA	7	5	732	908 (from 690.0 to 1,074.1)
	Prophet	7	6	732	839 (from 772.0 to 875.1)
	LSTM	7	5	732	984.1 (from 984.1 to 984.1)
Headache syndrome	GBD	8	8	590	585
	ARIMA	8	7	590	581 (from 563.9 to 592.5)
	Prophet	8	7	590	557 (from 535.4 to 587.7)
	LSTM	8	8	590	596 (from 595.8 to 595.8)
Malignant tumours of the trachea, bronchi, and lungs	GBD	9	9	439	444
	ARIMA	9	10	439	172 (from 80.6 to 342.0)
	Prophet	9	10	439	152 (from 158.9 to 227.9)
	LSTM	9	9	439	407.8 (from 407.8 to 407.8)
Malignant neoplasm of the colon and rectum	GBD	10	11	266	296
	ARIMA	10	9	266	259 (from 194.4 to 323.6)
	Prophet	10	8	266	188 (from 180.2 to 218.0)
	LSTM	10	11	266	256.3 (from 256.3 to 256.3)

Table 2. Validation metrics: RMSE by nosology

Nosology	ARIMA	Prophet	LSTM	Best
Stroke	407.77	655.87	681.41	ARIMA
Low back pain	8.43	29.43	28.49	ARIMA
COPD	53.29	160.18	92.68	ARIMA
Colon and rectum cancer	9.20	9.26	9.42	ARIMA
Lower respiratory infections	100.64	316.22	156.79	ARIMA
Diabetes mellitus	10.87	6.34	42.28	Prophet
Cirrhosis	208.78	570.03	208.17	LSTM
Ischemic heart disease	134.65	921.74	860.30	ARIMA
Alzheimer's disease	9.36	12.11	9.33	LSTM
Headache disorders	2.51	1.92	13.88	Prophet
Lung cancer	48.81	42.39	14.69	LSTM

Table 3. Validation metrics: MASE by nosology

Nosology	ARIMA	Prophet	LSTM	Best
Stroke	2.524	4.713	4.437	ARIMA
Low back pain	1.701	6.862	6.647	ARIMA
COPD	0.852	3.084	1.646	ARIMA
Colon and rectum cancer	1.110	0.810	1.123	Prophet
Lower respiratory infections	1.003	3.556	1.532	ARIMA
Diabetes mellitus	0.527	0.384	2.649	Prophet
Cirrhosis	2.797	8.185	2.972	ARIMA
Ischemic heart disease	0.416	3.220	2.828	ARIMA
Alzheimer's disease	5.675	7.748	5.882	ARIMA
Headache disorders	0.862	0.649	4.890	Prophet
Lung cancer	1.597	1.418	0.510	LSTM

Apart from that, **Table 1** contains the 95% confidence intervals (CI) values for ARIMA, Prophet, and LSTM models. Notably, during the long-term time-series forecasting, the width of the CIs increased progressively over the projection horizon, reflecting the accumulation of uncertainty associated with extrapolation beyond observed data.

Table 4. Average metrics across all nosologies

Metric	ARIMA	Prophet	LSTM
RMSE	90.39	247.77	192.49
MAE	79.02	244.94	177.95
MASE	1.733	3.694	3.192

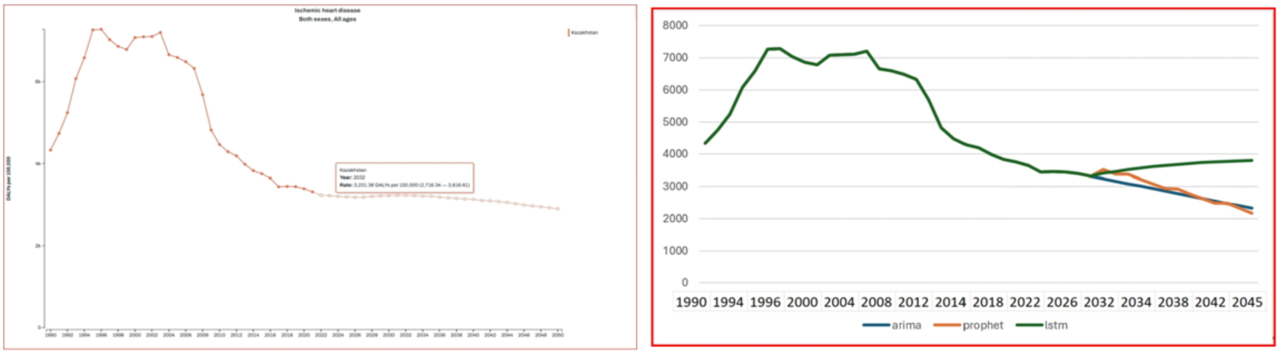
The comparative visualizations of model outputs are provided in part a of the figures for GBD forecasts and part b of the figures for the combined forecasts of ARIMA, Prophet, and LSTM models. These graphical results enable a visual interpretation of both historical and projected disease burden trends. To facilitate interpretation, the ten diseases were categorized into two major groups, NCDs and communicable diseases, and analyzed according to their observed forecasting patterns and model-specific behaviors.

Additionally, results of RMSE/MASE tests that enable comparison of the forecast accuracy, were presented in **Table 2**, **Table 3**, and **Table 4**. Formal statistical comparison using the Diebold-Mariano test was not applied to forecast horizon, as such tests require observed ground truth values and are therefore not meaningful for purely prospective projections.

NCDs

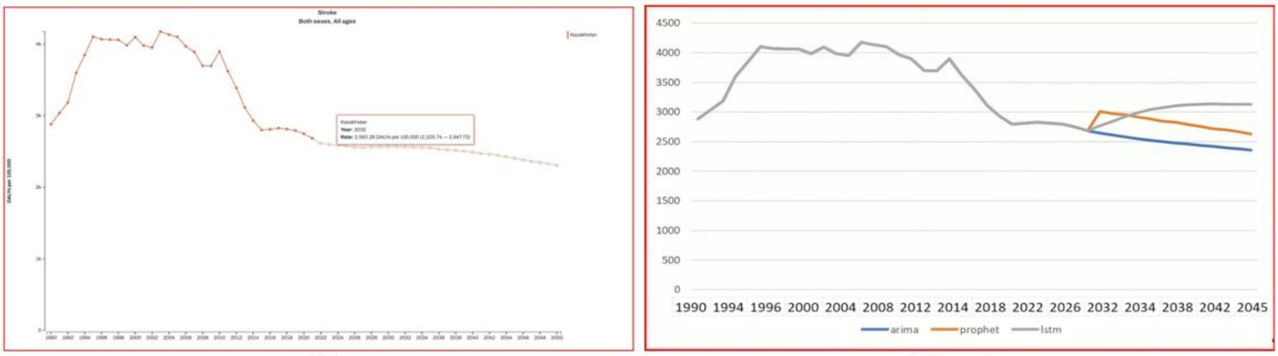
NCDs remain the dominant contributors to DALYs in Kazakhstan, collectively accounting for over 80% of the total disease burden. The analysis identified three distinct trend categories among NCDs:

(1) diseases with stable or declining DALY trajectories,



a. GBD b. Time series

Figure 1. Forecasting DALYs related to ischemic heart disease in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)



a. GBD b. Time series

Figure 2. Forecasting DALYs related to stroke in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)

(2) diseases with gradually increasing or plateauing trends, and

(3) those exhibiting divergence across models, reflecting uncertainty or transition in epidemiological dynamics.

Stable or declining DALY trends

The first subgroup comprises Ischemic heart disease, Cirrhosis and other chronic liver diseases, COPD, and Malignant neoplasms of the trachea, bronchi, and lungs. These conditions have long represented major chronic disease burdens but demonstrate relatively stable or decreasing DALY trajectories in recent years.

According to both the GBD and LSTM models, Ischemic heart disease remains the leading cause of DALYs in Kazakhstan from 2019 through 2032, while ARIMA and Prophet projections ranked it second by the end of the forecast horizon. As shown in part a in **Figure 1**, the GBD projection reveals a historical peak between 2002-2005, followed by a substantial decline, with DALY rates falling from 3443 in 2019 to a projected 3231 in 2032. The comparative models in part b in **Figure 1** exhibit similar general patterns, ARIMA and Prophet predict continued gradual decline, while LSTM indicates a temporary dip post-2019 followed by mild stabilization or rebound, suggesting potential saturation of prevention gains or aging population effects.

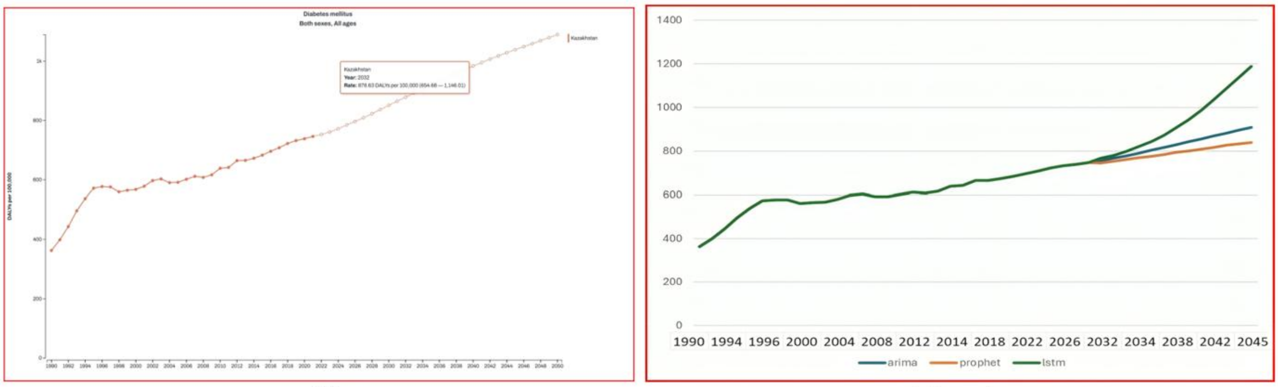
A comparable trajectory was observed for Cirrhosis and other chronic liver diseases, COPD, and lung malignancies, where all models, particularly Prophet and LSTM, yielded overlapping or near-consistent projections with only minor rank shifts. These results reflect the gradual impact of ongoing health reforms and improvements in early detection and risk-factor management.

Stable or rising trends

The second subgroup includes stroke, low back pain, and headache disorders, all of which showed either stability or moderate increase in DALYs across the projection period. Stroke was the second leading DALY contributor in 2019, and both GBD and LSTM models predicted their persistence at this rank through 2032. By contrast, ARIMA and Prophet models forecasted stroke to become the leading cause, despite a general decline in total burden. As illustrated in part a in **Figure 2**, the GBD data show a historical rise until the early 2000s, followed by a continuous decrease from 2791 in 2019 to 2563 in 2032. In part b in **Figure 2**, the ARIMA model projects a steeper downward slope, Prophet a gentler decline leveling off mid-forecast, while LSTM indicates an initial decrease followed by a mild resurgence toward 2032. The increasing trends in Low back pain and headache disorders reflect the growing recognition of musculoskeletal and neurological conditions in aging populations, with all models showing gradual or steady rises in their relative contribution.

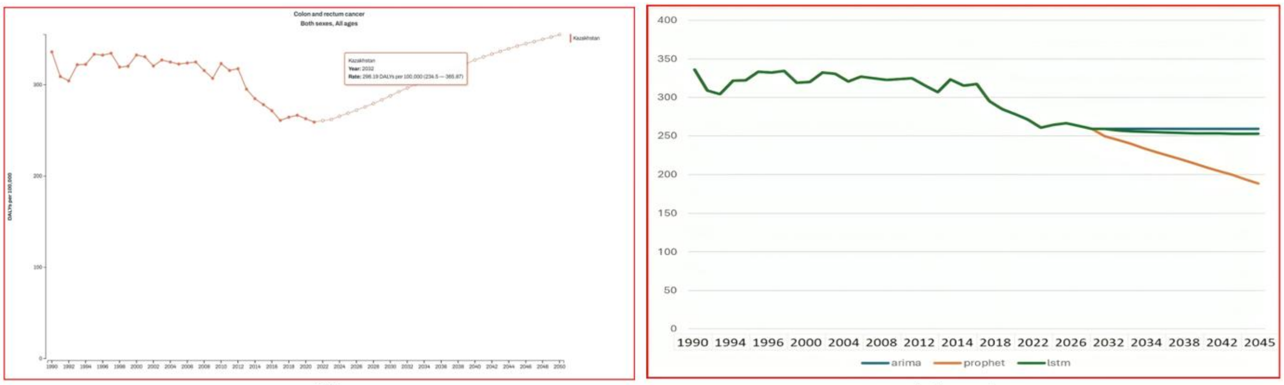
Steadily increasing trend–Diabetes

The most prominent upward trajectory among NCDs was observed for diabetes mellitus. As shown in part a in **Figure 3**, GBD model indicates a consistent increase across the historical period, projecting a rise from 732 DALYs in 2019 to 878 in 2032. The ARIMA, Prophet, and LSTM forecasts in part b in **Figure 3** corroborate this trend, with the LSTM model predicting the steepest growth, suggesting potential acceleration due to lifestyle factors, population aging, and rising obesity prevalence. This convergence of evidence underscores the growing public health importance of diabetes as a key target for preventive and therapeutic interventions in Kazakhstan.



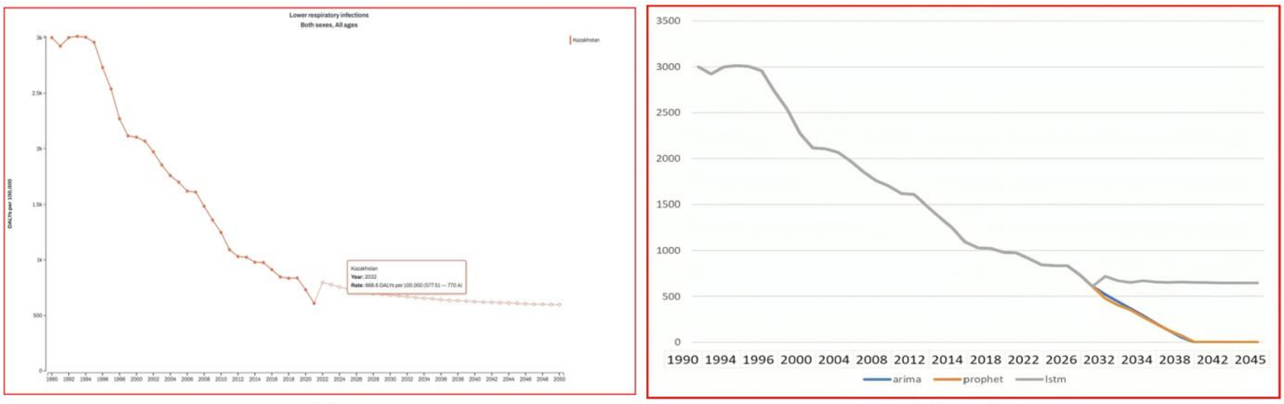
a. GBD b. Time series

Figure 3. Forecasting DALYs related to diabetes in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)



a. GBD b. Time series

Figure 4. Forecasting DALYs related to malignant neoplasm of the colon and rectum in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)



a. GBD b. Time series

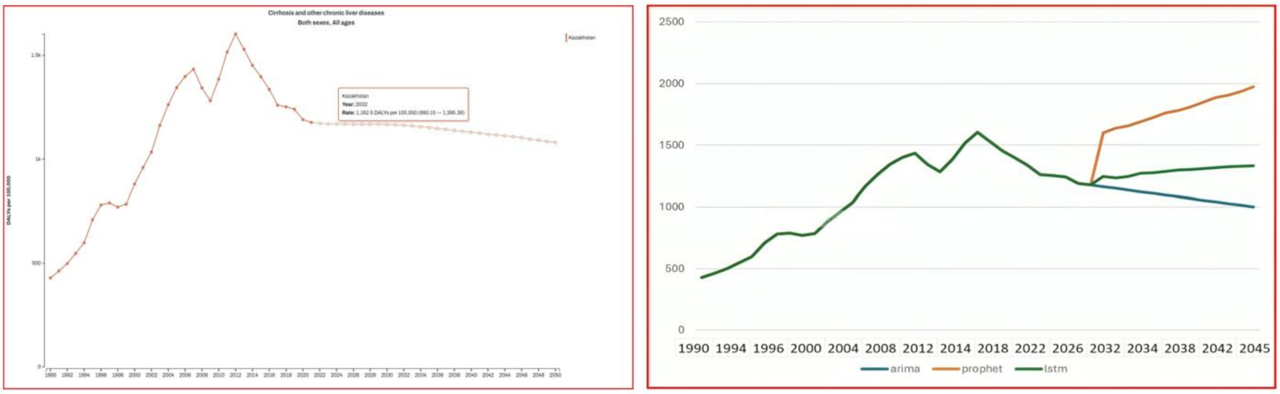
Figure 5. Forecasting DALYs related to infections of the lower respiratory tract in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)

Divergent projections–Colorectal cancer

An interesting divergence across models was found for malignant neoplasms of the colon and rectum. The GBD forecast predicted an increase in DALYs, from 266 in 2019 to 296 in 2032, as shown in part a in **Figure 4**. However, Prophet projected a moderate decline, while ARIMA and LSTM yielded intermediate trajectories remaining within the GBD-Prophet range (part b in **Figure 4**). This variation reflects uncertainty surrounding future screening rates, diagnostic coverage, and environmental risk factors. Overall, while the GBD model anticipates an upward trend, the data-driven models highlight potential stabilization or modest decline, suggesting a transitional phase influenced by improving preventive care.

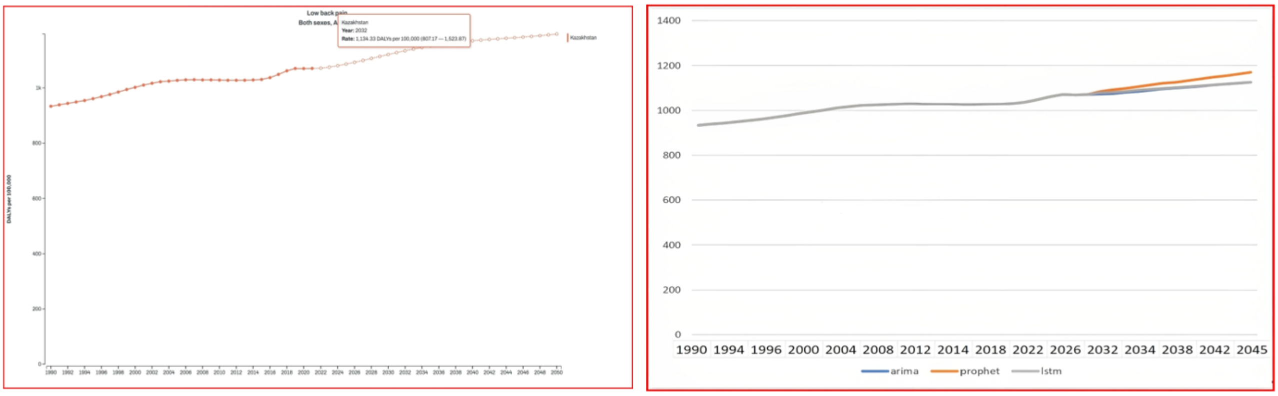
Communicable Diseases

The only communicable disease among the top-10 DALY contributors in Kazakhstan was Infections of lower respiratory tract, representing a key but diminishing component of the national disease burden. Part a in **Figure 5** presents the GBD forecast, showing a steep historical decline prior to 2010, followed by a slower yet persistent decrease, with DALYs falling from 837 in 2019 to 668 in 2032. This downward trend aligns with national improvements in vaccination coverage, healthcare access, and environmental health policies.



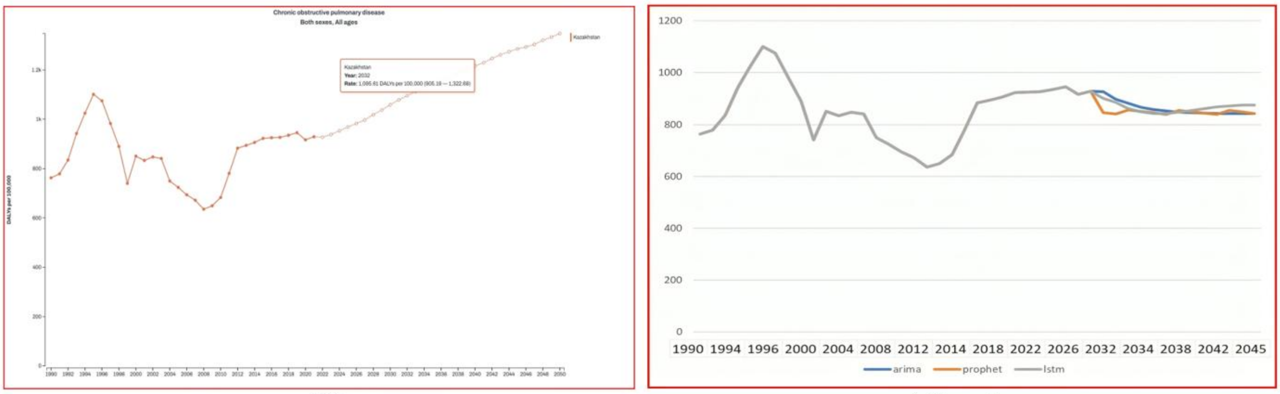
a. GBD b. Time series

Figure 6. Forecasting DALYs related to cirrhosis and other chronic liver diseases in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)



a. GBD b. Time series

Figure 7. Forecasting DALYs related to low back pain in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)



a. GBD b. Time series

Figure 8. Forecasting DALYs related to COPD in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)

In contrast, the ARIMA and Prophet models in part b in **Figure 5** forecast an even sharper reduction, approaching near-zero DALYs in the long term, which may represent statistical overfitting rather than epidemiological reality. The LSTM model, however, predicts a more gradual decline followed by stabilization, consistent with real-world expectations where respiratory infections persist at low but non-negligible levels.

Figure 6 shows forecasting DALYs related to cirrhosis and other chronic liver diseases in Kazakhstan.

Figure 7 depicts the forecasting DALYs related to low back pain in Kazakhstan.

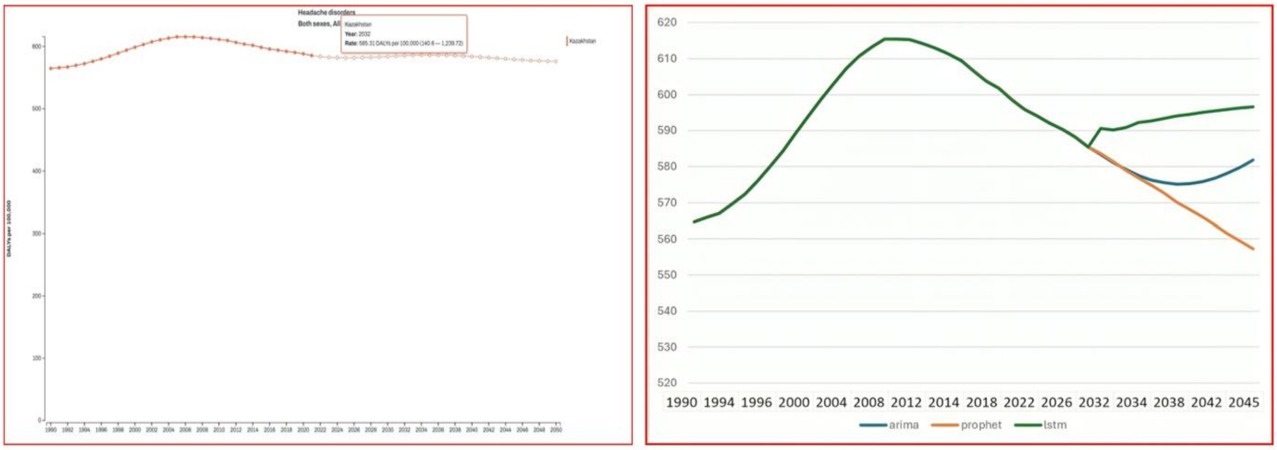
Figure 8 shows the forecasting DALYs related to COPD in Kazakhstan.

Figure 9 depicts the forecasting DALYs related to headache syndrome in Kazakhstan.

Figure 10 depicts the forecasting DALYs related to malignant tumors of the trachea, bronchi, and lungs in Kazakhstan.

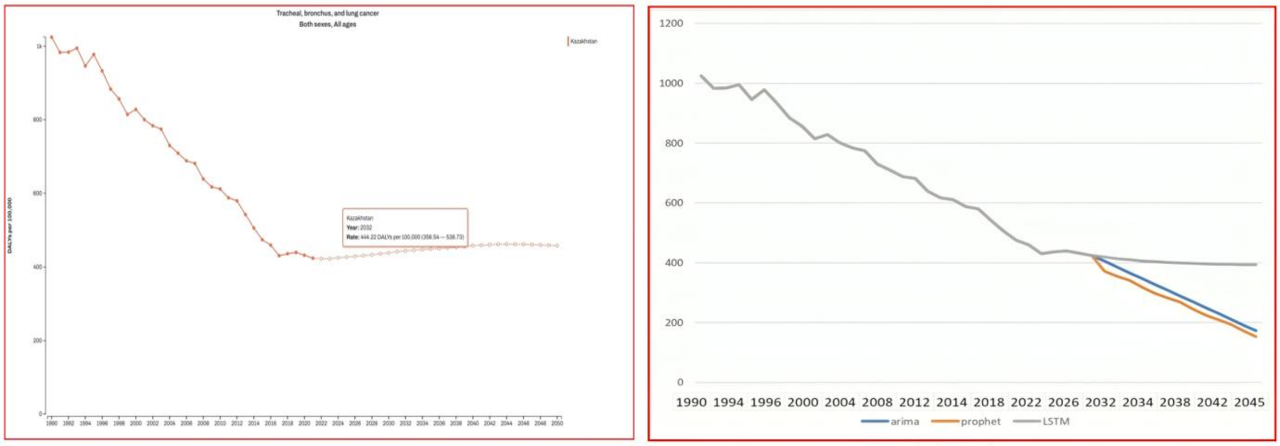
Model Behavior and Disease Ranking

Across all forecasting methods, the average deviation in projected DALY values for the top-10 diseases by 2032 remained within 5-8% between the LSTM and GBD models, indicating strong alignment in long-term trend estimation.



a. GBD b. Time series

Figure 9. Forecasting DALYs related to headache syndrome in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)



a. GBD b. Time series

Figure 10. Forecasting DALYs related to malignant tumors of the trachea, bronchi, and lungs in Kazakhstan (per 100,000 population) (Source: IHME. Vizhub)

The ARIMA and Prophet models, while more variable, deviated by up to 12% on average, primarily in diseases with higher inter-annual volatility such as lower respiratory infections and liver disorders. Regarding the near-zero DALY projections of ARIMA and Prophet toward the end of the forecast horizon, these outputs represent statistical degeneration at the lower boundary of the modeled distribution rather than epidemiologically plausible disease elimination and should be acknowledged accordingly. Despite these quantitative differences, the relative disease ranking across all models was largely preserved, nine of ten top contributors maintained their positions through 2032 in both LSTM and GBD forecasts.

Overall, the comparative modeling demonstrated consistent agreement across methods for high-burden chronic diseases such as ischemic heart disease and stroke, while greater variability appears in diseases with lower DALY magnitudes or complex etiological pathways such as colon cancer and respiratory infections. LSTM and GBD models provided smoother, long-term stable forecasts, whereas ARIMA and Prophet captured short-term fluctuations more sensitively.

Additional outcomes are shown in **Appendix A**.

DISCUSSION

This study represents the first comprehensive forecasting analysis of the top ten disease burden contributors (DALYs) in Kazakhstan, projecting trends up to 2032 using the GBD dataset in conjunction with three complementary time-series forecasting techniques: ARIMA, Prophet, and LSTM. The comparative approach allowed for the identification of consistent patterns across models and for the exploration of divergences attributable to methodological assumptions or data limitations.

The main findings indicate a clear epidemiological transition in Kazakhstan, with NCDs, notably ischemic heart disease, stroke, diabetes, and chronic respiratory conditions, projected to increase in relative burden, while communicable diseases, particularly lower respiratory tract infections, are expected to continue declining. This transition mirrors global trends described by the World Health Organization, which notes that NCDs now account for nearly three-quarters of all global deaths, primarily due to cardiovascular disease, cancer, and diabetes [14, 15]. The results align with other international studies that have identified the growing dominance of NCDs as a defining feature of global health in the 21st century [16, 17].

The projections for diabetes mellitus are particularly concerning. All four models forecast a continued and sharp rise in its DALY contribution through 2032, with the LSTM model predicting the steepest increase–from 732 DALYs per 100,000 in 2019 to 1187 DALYs in 2032–placing diabetes among the top five national health threats. This pattern is consistent with findings from comparable studies in low- and middle-income countries, such as Indonesia, where diabetes prevalence and mortality were predicted to nearly double by 2045 [18]. Such projections emphasize the escalating demand on healthcare systems, workforce capacity, and pharmaceutical supply chains, underscoring the need for proactive diabetes prevention, screening, and management programs.

For stroke, another leading cause of DALYs, all models forecasted persistence among the top-ranked conditions, though they diverged in directionality. The GBD and LSTM models projected stable or slightly declining trends, while ARIMA and Prophet indicated that stroke could surpass ischemic heart disease as the leading DALY contributor by 2032. Notably, recent research by Akhmedullin et al. reported discrepancies between GBD-estimated stroke mortality and data from Kazakhstan's unified electronic healthcare system (UNEHS) [19]. The UNEHS, a nationwide digital health registry, has proven a valuable source for population-level analyses [20-23]. In that study, ARIMA, Bayesian structural time series, and extreme gradient boosting models were used to project stroke mortality, all showing gradual decline–consistent with our findings. However, significant differences in absolute values across models and data sources suggest that GBD-based projections may not fully capture local epidemiological realities, warranting greater integration of national registry data into future forecasting studies.

The only communicable disease among Kazakhstan's top-10 DALY contributors–infections of the lower respiratory tract–showed a consistent downward trend across all forecasting methods. While GBD and LSTM predicted a gradual decline toward 2032, both ARIMA and Prophet generated extreme downward slopes, with Prophet occasionally producing negative DALY estimates that were programmatically adjusted to zero. This instability has been documented in prior studies using ARIMA-type models, where sensitivity to noise and small sample variation can produce biologically implausible forecasts [24]. Prophet's difficulty handling long-term, nonlinear patterns with outliers may also explain its erratic performance in this case [9]. These discrepancies illustrate the importance of model selection based on data structure and highlight the value of hybrid approaches that balance interpretability and robustness.

Overall, this study reaffirms that forecasting population health trends remains a complex and uncertain process. The GBD model remains a global standard due to its comprehensiveness, yet its limited disclosure of modeling assumptions and methodological part hinder independent validation [7]. In contrast, statistical (ARIMA and Prophet) and machine learning (LSTM) models, though data-dependent, provide transparent and adaptable frameworks for localized forecasting. The observed convergence across models, particularly between LSTM and GBD–supports the credibility of the general trend: a sustained increase in chronic, NCDs, and a continued decline in infectious disease burden. It is worth noting that GBD uses a highly smoothed, covariate-driven global model, ARIMA extrapolates a linear auto-regressive structure in differentiated time series, Prophet decomposes the signal into trend and seasonality with marked breakpoints, and LSTM learns nonlinear temporal dependencies from normalized sequences. These structural differences make ARIMA/Prophet more sensitive to short-term fluctuations and noise, while GBD and LSTM produce smoother long-term trends, particularly for volatile causes such as liver disease and lower respiratory tract infections.These findings have important policy implications, emphasizing the need to shift healthcare resources toward prevention and management of NCDs, integrate national health data systems with predictive analytics, and support capacity-building for data-driven decision-making.

Limitations

Despite its contributions, this study has several important limitations.

First, all forecasting models are inherently constrained by the quality and completeness of input data [25]. Historical DALY data may include underreporting, classification inconsistencies, and lags in health system reporting–particularly in developing contexts. Consequently, forecasts must be interpreted as indicative trends rather than precise quantitative predictions.

Second, forecasting models cannot account for unforeseen shocks such as the COVID-19 pandemic, geopolitical disruptions, or future epidemics, which can significantly alter mortality and morbidity dynamics. Such unpredictable events introduce structural breaks that traditional time-series models struggle to accommodate.

Third, data noise and measurement bias, common in large-scale health datasets, can distort model performance. As demonstrated by the ARIMA and Prophet anomalies, models trained on imperfect data can propagate these imperfections, resulting in over- or under-estimation of future burdens [26]. In the case with Infections of the lower respiratory tract, both ARIMA and Prophet decided to go negative without any interferences. Then, the negative values were manually corrected to zero.

Fourth, given the relatively short duration of the annual DALY series and the large capacity of the LSTM, the risk of overfitting persists, even when separating training and validation data, normalizing the data, stopping training prematurely, and selecting the model based on RMSE/MASE. Therefore, the steepest trajectories (e.g., for diabetes) should be considered trend extrapolations rather than accurate point predictions. Future work should incorporate more robust regularization and cross-validation on larger datasets to further reduce overfitting.

Finally, the models used in this study are agnostic to causality. They extrapolate patterns based solely on observed temporal trends, without incorporating underlying determinants such as behavioral, environmental, or socioeconomic factors. Thus, while they provide valuable insights into directional trends, they should be complemented by causal and mechanistic models for policy translation.

CONCLUSION

While highlighting the shifting landscape of major diseases in Kazakhstan over the past three decades, this study emphasizes forecasting the leading contributors to DALYs using multiple predictive modeling approaches. The results of forecasting suggest that the burden of NCDs will rise in

Kazakhstan over the coming years, thus, replicating the trends in all over the world. Among them is a diabetes. It is projected to surge significantly and deserves the attention of health policy makers. Key healthcare stakeholders should prioritize disease prevention and treatment strategies, with a special focus on such NCDs as diabetes. In this sense, forecasting the trajectory of diseases can be widely implemented in the current health policy processes as a vital element of efficient planning and resource allocation. Accurately forecasting the trajectory of key health threats is essential for anticipating future healthcare challenges, resource needs, and policy responses. Equally important is a clear understanding of the underlying principles, complexity, and methodological differences among forecasting models. Therefore, continued research is needed to systematically evaluate and compare these forecasting approaches in healthcare, ensuring more transparent, interpretable, and context-sensitive disease prediction systems.

Author contributions: TM, TA, & AG: validation, investigation, and writing–original draft; **TA & AJ:** resources; **TA & AG:** methodology; **TA:** data curation and visualization; **AO & MA:** writing–review & editing; **AY & AG:** project administration and funding acquisition; **AY:** formal analysis; & **AG:** conceptualization and supervision. All authors have agreed with the results and conclusions.

Funding: This study was conducted under the grant support from Ministry of Science and Higher Education of the Republic of Kazakhstan and under the consortium with Khoja Akhmet Yassawi International Kazakh-Turkish University, Kazakhstan (Grant title: Development of innovative technologies and creation of modern infrastructure for the sustainable development of the South Kazakhstan region. Grant number: BR24992814). Additionally, this study was partially supported by the Ministry of Science and Higher Education of the Republic of Kazakhstan through the grant titled “Smart-care: Innovative multi-sensor technology for elderly and disabled health management”, Grant number: AP23487613, awarded to Adnan Yazici.

Acknowledgments: The authors would like to thank Polpharma Santo for the support throughout this study.

Ethical statement: The authors stated that the study does not require any ethical approval since the study used publicly available, aggregated data from IHME without any personally identifiable information.

AI statement: The authors stated that no AI tools were used for data analysis, statistical processing, figure generation, or interpretation of results. The authors retain full responsibility for the integrity, originality, and scientific accuracy of the work.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

1. Murray CJ, Lopez AD, Jamison DT. The global burden of disease in 1990: Summary results, sensitivity analysis and future directions. Bull World Health Organ. 1994;72(3):495-509.
2. Salomon JA. Disability-adjusted life years. In: Encyclopedia of health economics. Elsevier; 2014: 200-3. https://doi.org/10.1016/B978-0-12-375678-7.00511-3
3. IHME. Global burden of disease (GBD). Institute for Health Metrics and Evaluation; 2025. Available at: https://www.healthdata.org/research-analysis/gbd (Accessed: 6 November 2025).
4. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: A systematic analysis for the global burden of disease study 2019. Lancet. 2020;396(10258):1204-22. https://doi.org/10.1016/S0140-6736(20)30925-9 PMid:33069326
5. Maulen A, Kadyrzhanov R, Mussatayev S. Post-Soviet ethnodemographic dynamics as a factor of national building in Kazakhstan. Nationalism Ethn Polit. 2025;31(1): 138-55. https://doi.org/10.1080/13537113.2024.2422196
6. IHME. Kazakhstan. Institute for Health Metrics and Evaluation; 2025. Available at: https://www.healthdata.org/research-analysis/health-by-location/profiles/kazakhstan (Accessed: 6 November 2025).
7. Mahajan M. The IHME in the shifting landscape of global health metrics. Global Policy. 2019;10(S1):110-20. https://doi.org/10.1111/1758-5899.12605
8. Mathers CD. History of global burden of disease assessment at the World Health Organization. Arch Public Health. 2020;78:77. https://doi.org/10.1186/s13690-020-00458-3 PMid:32850124 PMCid:PMC7443850
9. Yadav S. A comparative study of ARIMA, Prophet and LSTM for time series prediction. J Artif Intell Mach Learn Data Sci. 2023;1(1):1813-6. https://doi.org/10.51219/JAIMLD/sandeep-yadav/402
10. Noble J. What are ARIMA models? IBM; 2025. Available at: https://www.ibm.com/think/topics/arima-model (Accessed: 6 November 2025).
11. Facebook. Prophet. Facebook; 2025. Available at: https://facebook.github.io/prophet/ (Accessed: 6 November 2025).
12. DiPietro R, Hager GD. Deep learning: RNNs and LSTM. In: Handbook of medical image computing and computer assisted intervention. Elsevier; 2020: 503-19. https://doi.org/10.1016/B978-0-12-816176-0.00026-0
13. IHME. Vizhub. Institute for Health Metrics and Evaluation; 2025. Available at: https://vizhub.healthdata.org/gbd-foresight/ (Accessed: 6 November 2025).
14. WHO. Noncommunicable diseases. World Health Organization; 2025. Available at: https://www.who.int/health-topics/noncommunicable-diseases#tab=tab_1 (Accessed: 6 November 2025).
15. Pan American Health Organization. Noncommunicable Diseases. Pan American Health Organization; 2025. Available at: https://www.paho.org/en/topics/noncommunicable-diseases (Accessed: 6 November 2025).
16. Alimbayev A, Zhakhina G, Gusmanov A, et al. Predicting 1-year mortality of patients with diabetes mellitus in Kazakhstan based on administrative health data using machine learning. Sci Rep. 2023;13(1):8412. https://doi.org/10.1038/s41598-023-35551-4 PMid:37225754 PMCid: PMC10206549
17. Arupzhanov I, Alimbayev A, Seyil T, et al. Methodological note on predicting one-year mortality for chronic diseases using administrative data. Epidemiol Health Data Insights. 2025;1(4):ehdi015. https://doi.org/10.63946/ehdi/17159
18. Wahidin M, Achadi A, Besral B, et al. Projection of diabetes morbidity and mortality till 2045 in Indonesia based on risk factors and NCD prevention and control programs. Sci Rep. 2024;14(1):5424. https://doi.org/10.1038/s41598-024-54563-2 PMid:38443384 PMCid:PMC10914682

19. Akhmedullin R, Aimyshev T, Zhakhina G, et al. Stroke mortality in Kazakhstan. JACC Asia. 2025;5(12):1618-27. https://doi.org/10.1016/j.jacasi.2025.07.023 PMid: 40996387
20. Spataev Y. Digital Journey: Kazakhstan's healthcare. Government of Kazakhstan; 2020. Available at: https://www.gov.kz/memleket/entities/dsm/press/article/details/4848?lang=en (Accessed: 6 November 2025).
21. Gusmanov A, Zhakhina G, Yerdessov S, et al. Review of the research databases on population-based registries of unified electronic healthcare system of Kazakhstan (UNEHS): Possibilities and limitations for epidemiological research and real-world evidence. Int J Med Inform. 2023;170:104950. https://doi.org/10.1016/j.ijmedinf.2022.104950 PMid:36508752
22. Aimyshev T, Zhakhina G, Yerdessov S, et al. Mortality trends in Kazakhstan: Insights from a million of deaths from 2014 to 2022. BMC Public Health. 2025;25(1):2312. https://doi.org/10.1186/s12889-025-23346-3 PMid: 40611028 PMCid:PMC12225197
23. Saduyeva K, Zhakhina G, Karsybay A, et al. Predictors of all-cause mortality among hospitalized HIV patients in Kazakhstan: A retrospective study. J Clin Med Kaz. 2024;21(5):27-34. https://doi.org/10.23950/jcmk/15488
24. Li Y, Wu K, Liu J. Self-paced ARIMA for robust time series prediction. Knowl Based Syst. 2023;269:110489. https://doi.org/10.1016/j.knosys.2023.110489
25. Sandhya Arora MK. Forecasting the future: A comprehensive review of time series prediction techniques. J Electr Syst. 2024;20(2s):575-86. https://doi.org/10.52783/jes.1478
26. Fatima SSW, Rahimi A. A review of time-series forecasting algorithms for industrial manufacturing systems. Machines. 2024;12(6):380. https://doi.org/10.3390/machines12060380

APPENDIX A

Table A1. Augmented Dickey-Fuller test results

Nosology	ADF statistics	p	Stationary?	ADF (1st differene)	p	Stationary?	d
Stroke	-0.518	0.8885	No	-3.166	0.0221	Yes	1
Low back pain	-1.294	0.6320	No	-3.349	0.0128	Yes	1
COPD	-4.633	0.0001	Yes	-3.499	0.0080	Yes	0
Colon and rectum cancer	-0.452	0.9010	No	-6.077	0.0000	Yes	1
Lower respiratory infections	-0.568	0.8782	No	-2.991	0.0357	Yes	1
Diabetes mellitus	-3.681	0.0044	Yes	-2.637	0.0856	No	0
Cirrhosis	-2.122	0.2359	No	-3.379	0.0117	Yes	1
Ischemic heart disease	-1.030	0.7421	No	-2.464	0.1244	No	2
Alzheimer's disease	-2.238	0.1929	No	-3.390	0.0113	Yes	1
Headache disorders	-3.960	0.0016	Yes	-0.222	0.9359	No	0
Lung cancer	-0.821	0.8128	No	-6.894	0.0000	Yes	1

Note. d = 0 **stationary at level:** 3/11 (COPD, diabetes, and headache); d = 1 **require first differencing:** 7/11 series; & d = 2 **require second differencing:** 1/11 (IHD)

Table A2. Kendall's tau rank correlations

Comparison	Kendall's τ	p	Significance
ARIMA vs. Prophet	0.8909 strong	< 0.0001	***
ARIMA vs. LSTM	0.6727 moderate	0.0031	**
Prophet vs. LSTM	0.6364 moderate	0.0057	**

Note. ***p < 0.001 & **p < 0.01

Table A3. Spearman rank correlations

Comparison	Spearman's ρ	p	Significance
ARIMA vs. Prophet	0.9727 strong	< 0.0001	***
ARIMA vs. LSTM	0.8455 strong	0.0010	**
Prophet vs. LSTM	0.8364 strong	0.0013	**

Note. ***p < 0.001 & **p < 0.01

Friedman Test

Chi-square (χ^2) is 0.2162 and p is 0.8975.

Conclusion

No statistically significant difference in ankings among models (p = 0.90).

Interpretation

The high rank correlations (Kendall's τ: 0.64-0.89, Spearman's ρ: 0.84-0.97) indicate **moderate to strong agreement** in how the three models rank the different nosologies.

The non-significant Friedman test (p = 0.90) suggests that while models may differ in absolute predictions, they tend to **rank diseases similarly** in terms of relative burden.