Innov Biosyst Bioeng, 2025, vol. 9, no. 4, 57-68
doi: 10.20535/ibb.2025.9.4.348257

Original article

DYNAMICS AND FORECAST OF SCARLET FEVER PREDICTION INCIDENCE IN UKRAINE

A.P. Podavalenko[1*], V.I. Zadorozhna[2], T.D. Nessonova[1], T.A. Sergeeva[2], N.V. Bilera[1]

[1]Kharkiv National Medical University, Kharkiv, Ukraine
[2]State Institution "L.V. Gromashevsky Institute of Epidemiology and Infectious Diseases of the National Academy of Medical Sciences of Ukraine", Kyiv, Ukraine

*Corresponding author: epid@ukr.net

Received 24 September 2025; Accepted 9 December 2025

Background. The epidemic situation with scarlet fever has become more complicated in Ukraine, which requires improving surveillance. Forecasting the intensity of the epidemic process plays an important role, which will allow for a prompt response to the situation, implementation of anti-epidemic measures.
Objective. Statistical forecasting of scarlet fever incidence rates in Ukraine and its regions based on the analysis of long-term time series.
Methods. The analysis of scarlet fever incidence for 2005–2024 was conducted in Ukraine and its regions: central-southern, eastern-northern and western. Generally accepted methods of applied statistics were used. To predict the incidence, ETS exponential smoothing models and Box-Jenkins ARIMA models were used.
Results. The intensity of the epidemic process of scarlet fever in Ukraine and its regions during 2005–2024 had common features, in particular, cyclicality with periods of 4–5 years, anomalous declines and an increase in morbidity. This may indicate the influence of similar internal and external factors on the process. The projected incidence of scarlet fever in 2025–2030 in Ukraine and the regions will not undergo significant changes, while in the central-southern region there is a possible tendency to stabilize; in the eastern-northern region to decrease; in the western region – to an increase in morbidity.
Conclusions. Based on the ETS and ARIMA models used to analyze the 20-year incidence of scarlet fever, a forecast of the intensity and trends of the epidemic process in Ukraine and the regions for 2025–2030 was made.

Key words: scarlet fever; epidemic process; cyclicity of incidence; incidence forecasting; statistical models.

Introduction

Scarlet fever is an infectious disease that is etiologically caused by *Streptococcus pyogenes* (*S. pyogenes*). Before the introduction of antibiotics, it was extremely common among the world's population and was accompanied by a high mortality rate [1, 2]. Better living conditions, the effectiveness of control and treatment of scarlet fever have significantly influenced the manifestations of its epidemic process. There was a significant decrease in the incidence rate, the structure of clinical forms was dominated by a mild course of infection, and mortality in scarlet fever was practically not recorded. But, despite the successes achieved, reports of complications of the epidemic situation with scarlet fever in many countries of the world on different continents [3, 4] have now begun to appear. Thus, outbreaks of scarlet fever were recorded in European countries [5, 6], Asian countries [7 – 9], Australia [10] and the USA [11], which were accompanied by a more severe course with a fatal outcome and seasonality atypical for this infection. This may indicate certain changes in the interaction of the population components of its parasitic system, which requires research. In modern conditions, a number of hypotheses are put forward regarding the cause of the increase in the incidence of scarlet fever and the evolution of the course of its epidemic process. It is assumed that this is due to the active circulation of highly virulent forms of the scarlet fever pathogen [12, 13], the increase in its resistance to antibiotics [8, 12, 14], the weakening of collective immunity in organized groups [6], the influence of social, natural and other environmental factors [15, 16].

Therefore, the evolution of the epidemic process of scarlet fever demonstrates the need to improve the epidemiological surveillance system in order to successfully identify and track the circulation of *S. pyogenes*, determine their properties, adequately treat patients and prevent negative consequences. An important role in this system is played by forecasting the intensity of the epidemic process, which will allow obtaining options for trends in its development in time and space, promptly responding to the situation and implementing appropriate preventive and anti epidemic measures aimed at preventing the spread of infection.

At the same time, a detailed mechanistic analysis of the impact of demographic, social and climatic factors remain a difficult task due to the limited amount of available parameters. In this study, we did not build mechanistic epidemiological models (such as SIR or their modifications). Our approach is based on methods of statistical analysis of time series, which allow us to identify inertial patterns of the epidemic process, reflecting the integrated impact of the listed factors in multi-year data.

The aim of the study is to statistically predict the incidence of scarlet fever in Ukraine and its regions based on the analysis of long-term time series.

Materials and Methods

The work used materials from official statistics of the Ministry of Health of Ukraine (form No. 2 "Reports on certain infections and parasitic diseases (annual)"). Based on these data, an analysis of the incidence of scarlet fever for 2005–2024 in Ukraine and its regions: central-southern, eastern-northern and western was conducted; epidemiological patterns in the long-term dynamics of the incidence per 100 thousand population were studied and a forecast for 2025–2030 was made. The following were included in the western region: Volyn, Zakarpattia, Ivano-Frankivsk, Lviv, Rivne, Ternopil, Khmelnytskyi and Chernivtsi areas; to the central-southern region: Vinnytsia, Zhytomyr, Kyiv, Mykolaiv, Odesa, Cherkasy, Kherson areas and the city of Kyiv; to the eastern-northern region: Dnipropetrovsk, Zaporizhia, Kirovohrad, Poltava, Sumy, Kharkiv and Chernihiv areas. Temporarily occupied territories (Donetsk, Luhansk areas and the Autonomous Republic of Crimea) were not taken into account, but the fact that pronounced demographic processes have been taking place since 2022, in particular international and internal population migration due to military operations in the territory of Ukraine, was taken into account.

The time series of scarlet fever incidence were analyzed using the following methods of applied statistics: descriptive statistics, non-parametric analysis of variance, methods for identifying trends, cycles and anomalous levels of time series. To predict the incidence, ETS exponential smoothing models (built using the Forecast Sheet tool in Microsoft Excel 2021) and Box-Jenkins ARIMA models (built using the Time Series/Forecasting tool in the Statistica 12 program) were used. The following principle is at the heart of these methods: most social phenomena are characterized by a certain inertia, that is, past values of indicators affect the following ones [17].

Considering the relatively small number (20 years) and significant heterogeneity of data in the time series studied, forecasting models were built based on actual and theoretical (smoothed indicators) data. The time series smoothing procedure was used to eliminate anomalous values of the series for a clearer identification of its trend. Smoothing of actual data was carried out using a moving average with parameter 3 and using the Urbach formula to determine the first and last values of the series, that is, for each region a new time series (y) was constructed, in which: $y_1=(7*x_1+4*x_2-2*x_3)/9$, $y_{20}=(7*x_{20}+4*x_{19}-2*x_{18})/9$, $y_i=(x_{i-1}+x_i+x_{i+1})/3$, where i=2, 3,...19. This made it possible not to reduce after smoothing the number of values of the studied time series.

Thus, for each studied region and for Ukraine as a whole, 3 series of forecast values for the period 2025–2030 (6 years) were constructed and analyzed: 2 series according to the exponential smoothing model – according to actual and theoretical data and 1 series according to the ARIMA model – according to actual data.

The adequacy of the ETS model is guaranteed by the developers of Microsoft Excel, and the adequacy of ARIMA was determined by the following criteria: 1) the forecast values have a positive value, 2) all model coefficients are significant, 3) the autocorrelation of the residuals is insignificant, 4) the residuals of the model are distributed according to the normal law. To assess the accuracy of the forecast, MAPE (Mean Absolute Percent Error – the average percentage value of forecast errors) was used. The forecast accuracy was considered satisfactory when the MAPE values were less than or equal to 50%, and unsatisfactory when they were greater than 50%.

Forecasting trends in the epidemic process of scarlet fever in the regions was carried out according to the data of the summary table of forecast values for all models. The data elements of the summary table were designated $a_{i,j}$, where the rows of the matrix correspond to the years (i=1,...,6), the columns to the region and the forecasting method (j=1,...,12). Based on these data, a matrix of directions of change of predicted values of scarlet fever incidence indicators was constructed with elements $b_{i,j}$, the values of which were determined by the following rule: for the first row $b_{1,j}$="-", for the other rows

$$b_{i,j} = \begin{cases} empty\ if\ a_{i,j} = a_{i-1,j}\ (does\ not\ change\ over\ time) \\ -\ if\ a_{i,j} < a_{i-1,j}\ (decreases\ over\ time) \\ +\ if\ a_{i,j} > a_{i-1,j}\ (increases\ over\ time) \end{cases}$$

$$i = 2,\dots,6,\ j = 1,\dots,12.$$

Results

Based on the analysis of the long-term dynamics of scarlet fever incidence for 2005–2024, the average incidence rates (AI), coefficients of variation (CV) and average growth rates (AGR) were determined in the central-southern, eastern-northern, western regions and in Ukraine as a whole (Table 1).

Thus, in the central-southern region AI the incidence rate of scarlet fever in 2005–2024 was 25.42 per 100 thousand population, CV – 0.43% and AGR – (+1.03); in the western region – 26.75, 0.57% and (+1.09), respectively; in the eastern-northern region – 27.28, 0.46% and (+1.02), respectively; in Ukraine as a whole – 26.45, 0.44% and (+1.05), respectively.

Using the Kruskal-Wallis test, a comparative analysis of the average values of the incidence rates of scarlet fever by year and region was conducted. In 2023, a significant difference was found between the incidence rates in the western and eastern-northern regions (p=0.005), as well as an anomalous value of the incidence growth rate in all regions (Fig. 1).

For each region and Ukraine as a whole, Fourier spectral analysis established the cyclicity of the incidence, which for Ukraine, the central-southern, eastern-northern regions was 4 years, for the western – 5 years, and using the Irwin criterion, the years of anomalous rises and falls in the incidence of scarlet fever were identified. Thus, rises in the incidence were recorded in the central-southern region in 2011 and 2023, in the western – in 2023, in the eastern-northern region – in 2011 and 2024 and in Ukraine as a whole – in 2011, 2023 and 2024, and declines – in 2021–2022 in Ukraine and its regions.

Thus, in the long-term dynamics of scarlet fever incidence, we observed a sharp decrease in incidence in 2021–2022 and an increase in 2023–2024, which allowed us to assume the influence of unforeseen environmental factors on the course of the epidemic process in these years. Taking this fact into account, the time series of scarlet fever incidence indicators was divided into two periods: a period of relative stability (2005–2020) and a period of influence of unforeseen factors (2021–2024), and for each period the average incidence and average growth rates were determined. Using the Mann-Whitney test, a significant difference was found between the average incidence rates of scarlet fever in 2005–2020 and 2021–2024 only for the eastern-northern region (p=0.042). No significant difference was found between the rates of increase in the incidence of scarlet fever in these two periods (Table 2).

Table 1: Scarlet fever incidence per 100 thousand population in Ukraine and its regions (2005–2024)

Years	Regions			Ukraine
	Central-southern	Western	Eastern-northern	
2005	22.29	14.61	23.03	19.84
2006	25.21	21.27	24.34	23.57
2007	30.05	30.96	34.71	31.78
2008	36.51	32.65	46.93	38.34
2009	25.74	19.66	25.68	23.6
2010	26.66	21.36	33.88	27.01
2011	43.91	39.87	51.44	44.8
2012	31.61	24.51	42.87	32.57
2013	24.59	30.27	31.44	28.65
2014	31.07	26.80	27.83	28.6
2015	20.78	18.75	24.42	21.18
2016	15.49	18.44	19.87	17.85
2017	18.80	22.16	21.37	20.75
2018	26.10	31.92	26,41	28,22
2019	36.13	37.40	40.36	37.86
2020	13.77	14.37	17.89	15.23
2021	4.79	4.13	4.83	4.57
2022	2.30	3.75	2.74	2.94
2023	30.34	52.77	13.61	33.05
2024	42.39	69.34	31.93	48.58
AI	25.42	26.75	27.28	26.45
CV	0.43%	0.57%	0.46%	0.44%
AGR	1.03	1.09	1.02	1.05

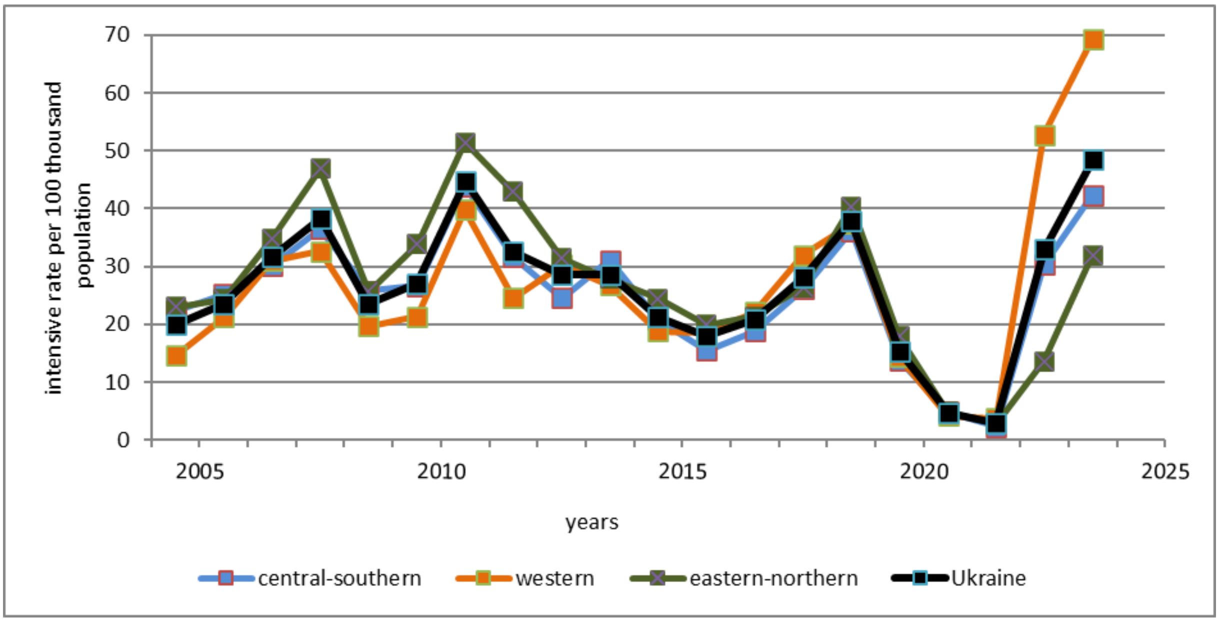

Figure 1: Long-term dynamics of scarlet fever incidence in Ukraine and its regions (2005–2024)

Table 2: Average incidence rates of scarlet fever (per 100 thousand population) and growth rates of its indicators (%) in Ukraine and its regions in different periods (2005–2020 and 2021–2024)

Years	Incidence/growth rate			
	Regions			Ukraine
	Central-southern	Western	Eastern-northern	
2005–2020	26.79 / 0.97	25.31 / 1.00	30.78 / 0.98	27.49 / 0.98
2021–2024	19.96 / 1.32	32.50 / 1.48	13.28 / 1.16	22.29 / 1.34

Therefore, in the period 2021–2024, there were changes in the structure of scarlet fever incidence rates in the regions of Ukraine, in particular, the share of incidence rates in the western region increased significantly (from 31% to 50%), and decreased in the eastern-northern region (from 37% to 20%) compared to the period 2005–2020 (Fig. 2).

For each region and Ukraine as a whole, the procedure for smoothing the actual incidence rates of scarlet fever was performed. The actual and smoothed data for the period 2005–2024 were examined for the presence of a trend using a test for comparing the average levels of the halves of the time series (Table 3).

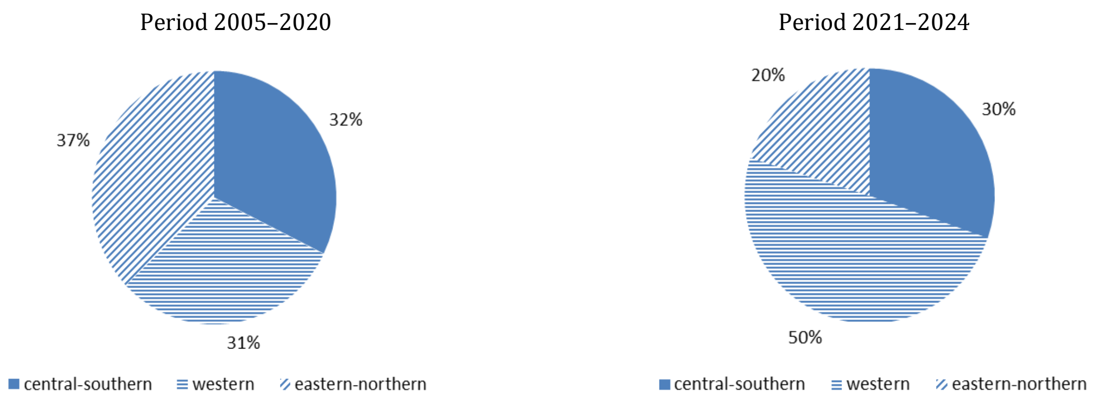

Figure 2: Comparison of the shares of average incidence rates of scarlet fever in the structure of incidence in regions for 2005–2020 and 2021–2024

Table 3: Average incidence rates of scarlet fever per 100 thousand population of the halves of the time series in Ukraine and its regions 2005-2014 and 2015–2024

Regions	Actual data				Smoothed data			
	2005–2014	2015–2024	P-value U-test	Tendency of incidence rates	2005–2014	2015–2024	P-value U-test	Tendency of incidence rates
Central-southern	29.762	21.087	0.088	stable	29.279	22.187	0.011	decrease
Eastern-northern	34.215	20.343	0.014	decrease	33.856	20.902	0.005	decrease
Western	26.194	27.303	0.623	stable	25.636	28.845	0.57	stable
Ukraine	29.87	23.023	0.185	stable	29.405	24.112	0.045	decrease

Given the lack of a clear trend for most regions and Ukraine as a whole, the augmented Dickey-Fuller test (DF-test) was used to test the stationarity of the time series of scarlet fever incidence rates for 2005-2024. The results of the analysis showed that the time series are non-stationary, but become stationary after taking the first difference of the data, as a result of which the number of observations decreases to 18.

Using actual and smoothed data, a correlation analysis was conducted between the time point (time series element number) and scarlet fever incidence rates. Significant Spearman correlation coefficients (r) were found only for the eastern-northern region, in particular for actual data r=(-0.44), for smoothed data r=(-0.5).

Autocorrelation and partial autocorrelation functions were constructed and analyzed for each of the time series. Based on the values of these functions, it can be stated that there are significant relationships between consecutive elements of the time series of scarlet fever incidence rates, therefore, exponential smoothing (ETS) models, as well as autoregression and integrated moving average (ARIMA) models were used for forecasting.

We built 2 forecast models using exponential smoothing algorithms based on actual (ETS1) and smoothed (ETS2) data for the period 2025–2030. The average forecast values of scarlet fever incidence indicators and their confidence limits in the regions and Ukraine as a whole were determined (Fig. 3).

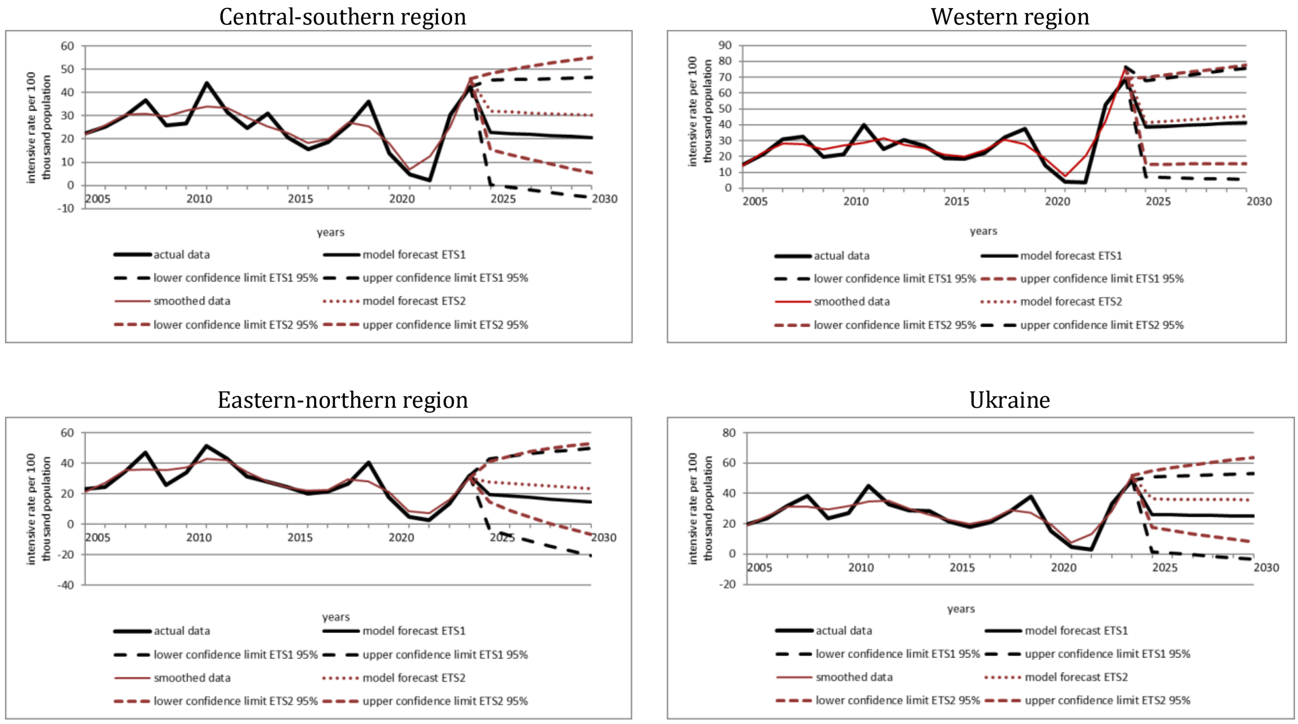

Figure 3: Forecasting the incidence of scarlet fever using ETS models in Ukraine and its regions for 2025-2030

Figure 3 shows that the forecast values based on smoothed data exceed the corresponding forecasts obtained based on actual data.

For each region and Ukraine as a whole, an ARIMA model was built based on actual data and a forecast was made for 2025–2030. The value of the parameter d in all models was equal to 1, since the first-order differentiation procedure transformed the actual data into a stationary time series. The actual and forecast values obtained using the ARIMA model, as well as their confidence intervals, are shown in (Fig. 4).

The main characteristics of the models and average of forecast values are given in Table 4.

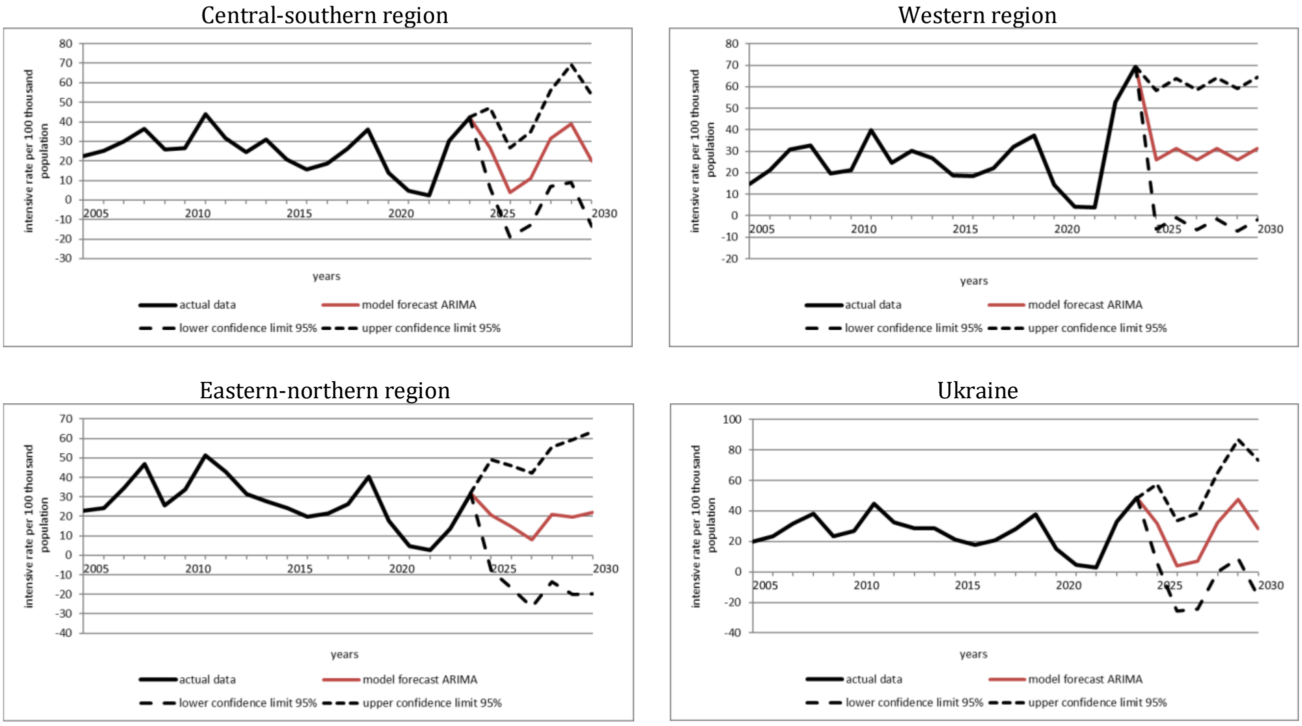

Figure 4: Forecasting the incidence of scarlet fever using ARIMA models in Ukraine and its regions for 2025–2030

Table 4: Main characteristics of ARIMA forecasting models of time series of scarlet fever incidence in Ukraine and its regions

Regions	Model type	Estimate	Average actual value	Average predictive value	MAPE (%)
Actual data (ARIMA1)					
Central-southern	ARIMA (2,1,0)	p1=0.49506 p2=-0.9178	25.43	22.00	61.18
Eastern-northern	ARIMA (2,1,0)	p1=0.45258 p2=-0.6999	27.28	17.65	65.4
Western	ARIMA (0,1,2)	q1=-0.1038 q2=0.87791	26.75	28.68	80.0
Ukraine	ARIMA (2,1,0)	p1=0.58056 p2=-0.9645	26.45	25.34	53.3
Smoothed data (ARIMA2)					
Central-southern	ARIMA (1,0,0)	const=26.313 p1=0.69217	25.73	34.73	30.3
Eastern-northern	ARIMA (1,0,0)	const=24.781 p1=0.79810	27.38	29.21	28.41
Western	ARIMA (0,0,1)	const=27.413 q1=-0.7074	27.24	32.19	28.9
Ukraine	ARIMA (1,0,0)	const=26.855 p1=0.66618	26.76	34.43	32.4

The adequacy conditions of the constructed models are met. Based on the residuals of the ARIMA models, the MAPE indicator was calculated. Its high value for models built on actual data can be explained by the insufficient number of observations in the time series, as well as the presence of anomalous values. Smoothed data exhibit stationarity problems. The operations required to make them stationary significantly reduce the length of the series. The best forecasts constructed from smoothed data outperform those obtained using the actual data, although they are characterized by an acceptable level of MAPE.

Based on the obtained results of forecasting the incidence of scarlet fever, a general forecast table was constructed (Table 5).

The data in Table 5 were used to construct a matrix of trends in the predicted incidence of scarlet fever according to all models in Ukraine and its regions (Table 6).

Table 5: Summary data on forecasts of the incidence of scarlet fever per 100 thousand population using ARIMA and ETS models in Ukraine and its regions

	Ukraine			Regions								
				Eastern-northern			Central-southern			Western		
Years	2005–2024	2005–2020	2021–2024	2005–2024	2005–2020	2021–2024	2005–2024	2005–2020	2021–2024	2005–2024	2005–2020	2021–2024
	26.45	27.49	22.29	27.28	30.78	13.28	25.43	26.79	19.96	26.75	25.31	32.50
	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA
	1	2	3	4	5	6	7	8	9	10	11	12
2025	26.17	36.31	32.08	19.54	27.74	20.68	22.79	31.88	26.74	38.41	41.43	26.03
2026	25.93	36.20	3.99	18.53	26.82	14.70	22.36	31.55	3.81	39.04	42.26	31.32
2027	25.70	36.10	7.13	17.52	25.90	7.93	21.92	31.23	10.94	39.67	43.08	26.03
2028	25.46	36.00	32.52	16.51	24.98	20.99	21.48	30.91	31.40	40.29	43.90	31.32
2029	25.23	35.89	47.75	15.50	24.06	19.70	21.04	30.58	39.10	40.92	44.73	26.03
2030	24.99	35.79	28.59	14.49	23.14	21.91	20.61	30.26	20.01	41.55	45.55	31.32
Average 2025–2030	25.58	36.05	25.34	17.02	25.44	17.65	21.70	31.07	22.00	39.98	43.49	28.68

Table 6: Matrix of trends in the predicted incidence of scarlet fever per 100 thousand population according to all models in the regions and in Ukraine as a whole

Years	Ukraine			Regions								
				Eastern-northern			Central-southern			Western		
	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA
	1	2	3	4	5	6	7	8	9	10	11	12
2025	-	-	-	-	-	-	-	-	-	-	-	-
2026	-	-	-	-	-	-	-	-	-	+	+	+
2027	-	+	+	-	-	-	-	-	+	+	+	-
2028	-	-	+	-	-	+	-	-	+	+	+	+
2029	-	-	+	-	-	-	-	-	+	+	+	-
2030	-	-	-	-	-	+	-	-	-	+	+	+

Given the results obtained, it is possible to assume an increase in the incidence of scarlet fever in the western region and, possibly, in 2028–2029 in the central-southern region and in Ukraine as a whole.

In order to confirm the forecast of the trend in the incidence of scarlet fever in the regions and in Ukraine as a whole, we used another statistical technique. Combining the predicted values with the actual data on the basis of which they were constructed, we obtained 12 new time series of 26 values (2005–2030). These series were examined for the presence of a trend using a test for comparing the average levels of the halves of the time series: the 1st half – 2005–2017, the 2nd half – 2018–2030 (Table 7).

Thus, if the forecasts are fulfilled, a trend towards an increase in morbidity rates is likely in the western region, and a trend towards a decrease in the northeastern region.

Discussion

The aim of the study was to predict the incidence of scarlet fever in Ukraine and its regions based on the analysis of the long-term dynamics of the epidemic process. The dependent variable was the incidence of scarlet fever per 100 thousand population for 2005–2024, which is a key indicator of the intensity of the epidemic process, and the independent variable was the year.

The intensity of the spread of infectious diseases is determined by a complex of demographic, socio-economic, environmental and medical-organizational factors. Building multivariate regression models requires significant resources and time, and many parameters that potentially affect the accuracy of the forecast are inaccessible or unclearly defined. To take into account regional characteristics, the areas of Ukraine were combined into three enlarged regions – central-southern, eastern-northern and western – taking into account the geographical location and approximately equal population size. Due to the lack of official data from the Ministry of Health of Ukraine for the Autonomous Republic of Crimea, Luhansk and Donetsk regions, these territories were excluded from the analysis. Thus, the second independent variable was the region of Ukraine with three levels.

Based on annual data, average regional morbidity rates were calculated and their graphical and statistical comparison was performed. To study intensive indicators, a time series analysis methodology was applied, which included data visualization; calculation of statistical characteristics (coefficient of variation, average absolute increase, average growth rate); determination of structural components (trend, cyclicality, anomalies); stationarity check; assessment of autocorrelation properties; construction of predictive models.

The analysis showed that the time series are heterogeneous, which is confirmed by the high coefficient of variation. The average growth rate of regional indicators in 2005-2020 approached 1, which made it difficult to identify clear trends. The detected anomalous rises and falls made it possible to distinguish two time periods: a period of relative stability (2005–2020) and a period of influence of unforeseen external factors (2021–2024).

Table 7: Average incidence rates of scarlet fever per 100 thousand population of the halves of the time series in Ukraine and its regions (2005–2017 and 2018–2030)

Years	Ukraine			Regions								
				Eastern-northern			Central-southern			Western		
	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA	ETS1	ETS2	ARIMA
	1	2	3	4	5	6	7	8	9	10	11	12
2005–2017	27.58	27.60	27.58	31.37	31.31	31.37	27.13	27.21	27.13	24.72	24.74	24.72
2018–2030	24.92	30.21	24.81	18.45	22.55	18.74	22.00	26.72	22.14	34.89	37.24	29.67
P-value U-test												
	0.758	0.199	0.918	0.002	0.035	0.004	0.182	0.959	0.505	0.045	0.024	0.282
Tendency of incidence rates												
	stable	Stable	stable	decr.	decr.	decr.	stable	stable	stable	incr.	incr	stable

The epidemic process of scarlet fever in Ukraine was characterized by a cyclicity with periods of 4–5 years, which corresponds to natural fluctuations in population susceptibility and seasonality of infections with an aerosol-aspiration transmission mechanism [18]. In 2021–2022, abnormally low levels of morbidity were recorded, which coincided with the period of strict quarantine restrictions during the COVID-19 pandemic. Similar dynamics were characteristic of other infections with a similar mechanism of transmission of the infectious agent [19 - 21]. In 2023-2024, there was a sharp increase in the incidence, which may be associated with the lifting of quarantine, the accumulation of susceptible individuals, and increased internal migration processes in war conditions [7].

In 2023, the incidence in the western region was significantly higher than in the eastern-northern region, which is likely due to an increase in population density due to the displacement of the population from regions of active hostilities. A comparison of the average values between the periods 2005–2020 and 2021–2024 revealed a statistically significant difference only in the eastern-northern region (p=0.042), which may reflect depopulation and migration changes.

According to the results of the test of comparison of the average levels of the halves of the series in the eastern-northern region, the actual data showed a pronounced trend towards a decrease in morbidity. In contrast, the smoothed time series demonstrated stable trends towards a decrease in indicators in all regions, except for the western region (Table 3).

Different approaches are used to forecast medical data: the basic epidemiological model SIR (Susceptible-Infected-Recovered) [22], exponential smoothing models [23, 24], polyharmonic model [25], ARIMA [24], and LSTM neural networks [26]. It is believed that the best short-term forecasts are provided by exponential smoothing (ETS) and ARIMA models. Unlike the compartmental SIR model, which requires a preliminary determination of the coefficients of β-transmission of infection and γ-recovery of a person after transferring the disease, using the ETS and ARIMA models, it is possible to build forecasts of individual indicators only on the basis of observations of this indicator accumulated over time. After comparing forecasts built by different methods, more reliable final conclusions are possible. Therefore, these approaches were chosen for modeling. The abbreviation ARIMA (p, d, q) includes the following elements: p – the number of lags in the autoregression part; d – the order of differentiation, which ensures integration; q – the number of lags in the moving average part. ARIMA models built on raw data (ARIMA1) demonstrated high MAPE values (> 50%), which indicates their limited accuracy for quantitative interpretation of short-term forecasts. This may be due to the small number of observations in the time series (n=20), the presence of sharp anomalies in 2021–2024, and significant data heterogeneity. At the same time, models built on smoothed data (ARIMA2) showed significantly lower MAPEs (< 30–32 %).

A comparison of exponential smoothing models showed differences between forecasts based on actual data (ETS1) and forecasts obtained from smoothed series (ETS2). In all regions, the forecast for ETS2 was higher than for ETS1. The average forecast values were in the central-southern region – 21.70 versus 31.07; in the western region – 39.98 versus 43.49; in the eastern-northern region – 17.02 versus 25.44; in Ukraine as a whole – 25.58 versus 36.05 per 100 thousand population. We can assume the reliability of the specified forecast, the adequacy of which is guaranteed by the Microsoft Excel program itself, but provided that there is no influence of random factors and modern demographic processes are taken into account.

The ARIMA1 and ARIMA2 forecast models constructed for each region and Ukraine as a whole, as well as the forecast values for 2025–2030, also differed. Thus, the forecast indicators of scarlet fever incidence based on actual data (ARIMA1 model) and smoothed data (ARIMA2 model) for the western and eastern-northern regions almost coincided, and for the central-southern region and Ukraine as a whole, the forecast based on smoothed data exceeds the forecast based on actual data. At the same time, the calculated forecast accuracy using MAPE turned out to be satisfactory only for the ARIMA2 model. Thus, according to the ARIMA2 model, the highest average indicators are predicted in general in Ukraine and the central-southern region (34.43 and 34.73, respectively), and the lowest – in the eastern-northern region (27.46 per 100 thousand population).

Since the main goal of forecasting is to identify the general trend of changes in indicators, a combined series consisting of actual data for 2005–2024 and forecasts for 2025–2030 was tested. The results showed that in 2025–2030, a further increase in the incidence of scarlet fever is expected in the western region and a decrease in the eastern-northern region.

Forecasting trends in the epidemic process of scarlet fever in Ukraine and its regions using ETS and ARIMA models indicates a stabilization of the trend in Ukraine and the central-southern region, while a slight increase in the incidence is expected in the western region and, possibly, a decrease in the eastern-northern region. This may also be associated with demographic processes (natural and migratory movement), the influence of social factors and changes in the specific immunological stratum of the population.

In conclusion, we would like to emphasize that the relevance of predicting the intensity of the epidemic process of infectious diseases and developing its sensitive methods is growing from year to year, especially taking into account the probability of new epidemics and pandemics associated with both emergent and re-emergent infections [27 – 31]. This is extremely necessary to ensure readiness for such biological challenges and the appropriate operational and prospective response [32]. At the same time, it is still impossible to predict the impact on the epidemic process of infectious diseases of various factors of both natural and social nature, in particular those that may arise in unforeseen situations.

Conclusions

According to the analysis of the long-term dynamics of scarlet fever incidence (2005–2024), a similar intensity of the course of its epidemic process was determined both in Ukraine and its regions (central-southern, eastern-northern and western) with a cyclicity of 4–5 years. This may indicate a direct or mediated effect on the epidemic process of same environmental factors. During the period of increased influence of certain factors, for example, during the COVID-19 pandemic (2021–2022) and the war in Ukraine (2022–2024), as well as an increase in the susceptible population due to a previous sharp decrease in the circulation of *S. pyogenes*, there were anomalous decreases and increases in incidence rates.

Using two models (ETS and ARIMA), which are based on different approaches to the analysis of time series of scarlet fever incidence over a 20-year period, a forecast was made for 2025–2030. In the absence of the influence of random and unexpected factors, it is assumed that the incidence of scarlet fever in the regions and in Ukraine as a whole will not change significantly.

At the same time, a trend is predicted towards stabilization of the epidemic process of scarlet fever in Ukraine as a whole and in the central-southern region, towards a decrease in its intensity in the eastern-northern region and towards growth in the western region.

Although this study used statistical time series forecasting models (ETS, ARIMA), a promising direction is the integration of these approaches with alternative methods of epidemiological modeling that can improve understanding of the epidemic process and increase forecasting accuracy.

Prospects for further scientific research

Scientific research will be aimed at identifying the leading environmental factors in the development and spread of scarlet fever in the regions of Ukraine.

Conflict of interest

The authors declare the absence of a conflict of interest.

Information on funding

The epidemiological study is a fragment of the initiative scientific research work of the Ministry of Health of Ukraine "The role of environmental factors in the spread of scarlet fever and other group A streptococcal infections" (state registration number 0125U001089).

References

[1] InformedHealth.org. Overview: Scarlet fever [Internet]. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2006 [cited 2025 Sep 19]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK279620/

[2] Gale AH. A century of changes in the mortality and incidence of the principal infections of childhood. Arch Dis Child. 1945 March;20(101):2-21. DOI: 10.1136/adc.20.101.2

[3] Brockmann SO, Eichner L, Eichner M. Constantly high incidence of scarlet fever in Germany. Lancet Infect Dis. 2018 May;18(5):499-500. DOI: 10.1016/S1473-3099(18)30210-x

[4] WHO. Disease Outbreak News. Increased incidence of scarlet fever and invasive Group A Streptococcus infection - multi-country [Internet]. WHO; 2022 [cited 2025 Sep 19]. Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON429

[5] Lamagni T, Guy R, Chand M, Henderson KL, Chalker V, Lewis J, et al. Resurgence of scarlet fever in England, 2014-16: a population-based surveillance study. Lancet Infect Dis. 2018 Feb;18(2):180-87. DOI: 10.1016/S1473-3099(17)30693-X

[6] Erkmen A, Tüzün N, Erkmen O. Scarlet fever epidemics, deaths, effects on society and precautions taken in the Ottoman Empire and the Early Turkish Republic. Discov Public Health [Internet]. 2024;21(1). DOI: 10.1186/s12982-024-00144-7

[7] Wu R, Xiong Y, Wang J, Li B, Yang L, Zhao H, et al. Epidemiological changes of scarlet fever before, during and after the COVID-19 pandemic in Chongqing, China: a 19-year surveillance and prediction study. BMC Public Health [Internet]. 2024;24(1):2674. DOI: 10.1186/s12889-024-20116-5

[8] Sobh EK, Kumarasamy TA, Sobh ZK. Retrospective analysis of an outbreak of scarlet fever in United Arab Emirates. East Mediterr Health J. 2024;30(5):363-68. DOI: 10.26719/2024.30.5.363

[9] Luk E, Lo J, Li A, Lau M, Cheung T, Wong A, et al. Scarlet fever epidemic, Hong Kong, 2011. Emerg Infect Dis. 2012;18(10):1658-61. DOI: 10.3201/eid1810.111900

[10] Walker MJ, Brouwer S. Scarlet fever makes a comeback. Lancet Infect Dis. 2018 Feb;18(2):128-29. DOI: 10.1016/S1473-3099(17)30694-1

[11] Dunne EM, Hutton S, Peterson E, Blackstock AJ, Hahn CG, Turner K, et al. Increasing Incidence of Invasive Group A Streptococcus Disease, Idaho, USA, 2008-2019. Emerg Infect Dis. 2022 Sep;28(9):1785-95. DOI: 10.3201/eid2809.212129

[12] Brouwer S, Rivera-Hernandez T, Curren BF, Harbison-Price N, De Oliveira DMP, Jespersen MG, et al. Pathogenesis, epidemiology and control of Group A Streptococcus infection. Nat Rev Microbiol. 2023 Jul;21(7):431-47. DOI: 10.1038/s41579-023-00865-7

[13] Lynskey NN, Jauneikaite E, Li HK, Zhi X, Turner CE, Mosavie M, et al. Emergence of dominant toxigenic M1T1 Streptococcus pyogenes clone during increased scarlet fever activity in England: a population-based molecular epidemiological study. Lancet Infect Dis. 2019 Nov;19(11):1209-18. DOI: 10.1016/S1473-3099(19)30446-3

[14] Sun L, Xiao Y, Huang W, Lai J, Lyu J, Ye B, et al. Prevalence and identification of antibiotic-resistant scarlet fever group A Streptococcus strains in some paediatric cases at Shenzhen, China. Journal of Global Antimicrobial Resistance. 2022. Sep;30:199-04. DOI: 10.1016/j.jgar.2022.05.012

[15] Cordery R, Purba AK, Begum L, Mills E, Mosavie M, Vieira A, et al. Frequency of transmission, asymptomatic shedding, and airborne spread of Streptococcus pyogenes in schoolchildren exposed to scarlet fever: a prospective, longitudinal, multicohort, molecular epidemiological, contact-tracing study in England, UK. Lancet Microbe. 2022 May;3(5):e366-e375. DOI: 10.1016/s2666-5247(21)00332-3

[16] Duan Y, Yang LJ, Zhang YJ, Huang XL, Pan GX, Wang J. Effects of meteorological factors on incidence of scarlet fever during different periods in different districts of China. Sci Total Environ. 2017 Mar1;581-582:19-24. DOI: 10.1016/j.scitotenv.2017.01.010

[17] Dzhyhyrei IM, Skladannyi DM. Special sections of mathematics. Statistical data analysis in the STATISTICA environment [Internet]. Kyiv: Igor Sikorsky Kyiv Polytechnic Institute; 2019 [cited 2025 Sep 10]. 74 p. Available from: https://ela.kpi.ua/bitstream/123456789/28228/1/2019srm2.pdf

[18] Ryabokon EV, Onishchenko TE, Furyk EA, Zadiraka DA. Infection diseases with air-drop mechanism of transmission (Module 1) : manual for practical training and independent work for students of the 5th year of medical faculty Specialty 222 «General Medicine» Master's level of training. Zaporizhzhia: ZSMU, 2020. 138 p. Available from: https://dspace.zsmu.edu.ua/bitstream/123456789/13138/1/45.%20%20INFECTION%20DISEASES%20WITH%20AIR-DROP.pdf

[19] Ather B, Mirza TM, Edemekong PF. Airborne Precautions [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 [updated 2023 Mar 13, cited 2025 Sep 22]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK531468/

[20] Cai J, Zhou X, Zhang C, Jiang Y, Lv P, Zhou Y, et al. Ongoing epidemic of scarlet fever in Shanghai and the emergence of M1UK lineage group A Streptococcus: a 14-year surveillance study across the COVID-19 pandemic period. Lancet Reg Health West Pac. 2025 May;58:101576. DOI: 10.1016/j.lanwpc.2025.101576

[21] Facciola A, Lagana A, Genovese G, Romeo B, Sidoti S, D'Andrea G, et al. Impact of the COVID-19 pandemic on the infectious disease epidemiology. J Prev Med Hyg. 2023 Nov;64(3):274-82. DOI: 10.15167/2421-4248/jpmh2023.64.3.2904

[22] Solovyov S, Dziublyk I, Mintser O. Predictive model of the epidemic process of coronavirus infection COVID-19 in Ukraine. Medical Informatics and Engineering. 2020;70–78. DOI: 10.11603/mie.1996-1960.2020.2.11176

[23] Panda M. Application of ARIMA and Holt–Winters forecasting model to predict the spreading of COVID-19 for India and its states. DOI: 10.1101/2020.07.14.20153908

[24] Shokeralla AAA, Sameeh FRI, Musa AGM, Zahrani S. Prediction of the daily number of confirmed COVID-19 cases in Sudan using ARIMA and Holt–Winters exponential smoothing. International Journal of Development Research. 2020;10(8):39408–39413. DOI: 10.37118/ijdr.19811.08.2020.

[25] Podavalenko AP, Nessonova TD, Zadorozhna VI, Hrytsenko LM. Epidemiological analysis of pertussis morbidity in Ukraine. Wiadomosci Lekarskie. 2021;74(7):1628-33. DOI: 10.36740/WLek202107114

[26] Kumar M, Gupta S, Kumar K, Sachdeva M. Spreading of COVID-19 in India, Italy, Japan, Spain, UK and the US: a prediction using ARIMA and LSTM models. Digital Government: Research and Practice. 2020;1(4). DOI: 10.1145/3411760

[27] Zhang Y, Yu Y, Fong PSW, Shen J. Addressing unforeseen public health risks via the use of sustainable system and process management. Front Public Health. 2023 Nov;11:1249277. DOI: 10.3389/fpubh.2023.1249277

[28] De Gaetano S, Ponzo E, Midiri A, Mancuso G, Filippone D, Infortuna G, et al. Global Trends and Action Items for the Prevention and Control of Emerging and Re-Emerging Infectious Diseases. Hygiene. 2025;5(2):18. DOI: 10.3390/hygiene5020018

[29] Keshavamurthy R, Pazdernik KT, Ham C, Dixon S, Erwin S, Charles LE. Meeting Global Health Needs via Infectious Disease Forecasting: Development of a Reliable Data-Driven Framework. JMIR Public Health Surveill. 2025;11:59971. DOI: 10.2196/59971

[30] Qian Y, Zhang K, Marty E, Basu A, O'Dea EB, Wang X, et al. Physics-informed deep learning for infectious disease forecasting. ArXiv. 2025 Apr;2501:09298. DOI: 10.48550/arXiv.2501.09298

[31] Fujita S, Akutsu T. Enhancing epidemic forecasting with a physics-informed spatial identity neural network. PLoS One. 2025 Sept;20(9):0331611. DOI: 10.1371/journal.pone.0331611

[32] Zhan Z, Dong W, Lu Y, Yang P, Wang Q, Jia P. Real-Time Forecasting of Hand-Foot-and-Mouth Disease Outbreaks using the Integrating Compartment Model and Assimilation Filtering. Sci Rep. 2019 Feb 25;9(1):2661. DOI: 10.1038/s41598-019-38930-y

А.П. Подаваленко[1], В.І. Задорожна[2], Т.Д. Нессонова[1], Т.А. Сергеєва[2], Н.В. Білера[1]

[1]Харківський національний медичний університет, Харків, Україна

[2]Державна установа «Інститут епідеміології та інфекційних хвороб ім. Л.В. Громашевського НАМН України», Київ, Україна

ДИНАМІКА ТА ПРОГНОЗ ЗАХВОРЮВАНОСТІ НА СКАРЛАТИНУ В УКРАЇНІ

Передумови. В Україні ускладнилася епідемічна ситуації зі скарлатини, що потребує удосконалення епідеміологічного нагляду. Важливу роль має прогнозування інтенсивності епідемічного процесу, що дозволить оперативно реагувати на ситуацію, впроваджувати протиепідемічні заходи.
Мета. Статистичне прогнозування показників захворюваності на скарлатину в Україні та її регіонах на основі аналізу довгострокових динамічних рядів.
Методи. Аналіз захворюваності на скарлатину за 2005–2024 рр. було проведено по Україні та її регіонах: центрально-південному, східно-північному та західному. Застосовували загальноприйняті методи прикладної статистики. Для прогнозування захворюваності використовували моделі експоненційного згладжування ETS та моделі Бокса-Дженкінса ARIMA.
Результати. Інтенсивність епідемічного процесу скарлатини в Україні та її регіонах протягом 2005–2024 років мала спільні риси, зокрема, циклічність з періодами 4-5 років, аномальні спади та зростання захворюваності. Це може свідчити про вплив подібних внутрішніх та зовнішніх факторів на процес. Прогнозована захворюваність на скарлатину у 2025–2030 роках в Україні та регіонах не зазнає суттєвих змін, водночас у центрально-південному регіоні можлива тенденція до стабілізації; у східно-північному регіоні до зниження; у західному – до зростання захворюваності.
Висновки. На основі моделей EST та ARIMA, використаних для аналізу 20-річної захворюваності на скарлатину, здійснено прогноз інтенсивності й тенденцій епідемічного процесу в Україні та регіонах на 2025–2030 рр.
Ключові слова: скарлатина; епідемічний процес; циклічність захворюваності; прогнозування захворюваності; статистичні моделі.