Application of Artificial Intelligence for Predicting the Risk of Infectious Diseases Based on Hygienic Environmental Factors

Akhmedova Saodat Tashboltaevna
Lecturer at the Department of Microbiology, Public Health, and Management of the Termez branch of the Tashkent State Medical University
E-mail: saodat140284@gmail.com

Tuychieva Mekhriniso Farkhodovna
Students of the Termez branch of the Tashkent State Medical University
E-mail: mehrinisotuychiyeva111@gmail.com

Ashurova Farzona Golibovna
Students of the Termez branch of the Tashkent State Medical University
E-mail: fashurova704@gmail.com

Abstract: Traditional epidemiological models are often unable to promptly and adequately account for the complex non-linear interactions between multiple hygienic factors and infectious morbidity. This work substantiates the feasibility and appropriateness of integrating artificial intelligence methods into the system of social-hygienic monitoring. Based on an analysis of contemporary literature, priority predictors of infectious risk have been identified, including indicators of drinking water quality, ambient air quality, meteorological conditions, and sanitary-living characteristics of territories. It is demonstrated that neural network algorithms can improve the accuracy of acute intestinal infection incidence forecasts by an average of 18–22 % compared to classical regression approaches. A conceptual architecture of a hybrid predictive system adaptable to the conditions of regional hygienic control is described. The conclusion is drawn regarding the necessity of phased implementation of machine learning technologies into the practice of the sanitary-epidemiological service of the Republic of Uzbekistan.

Keywords: Artificial Intelligence, Neural Networks, Hygienic Diagnostics, Infectious Diseases, Risk Factors, Epidemiological Forecasting.

1. Introduction

The current stage of development in hygienic science is characterized by the accumulation of vast datasets concerning environmental conditions and public health. The laboratory network of the sanitary-epidemiological service generates thousands of analytical results annually on ambient air, drinking water, soil, and food products. Concurrently, infectious and somatic morbidity rates are systematically recorded. In practice, however, these information flows exist in isolation from one another, and their analytical processing is largely confined to simple comparisons with established normative standards.

As early as the beginning of the 2000s, foreign researchers convincingly demonstrated the existence of statistically significant associations between environmental quality and the incidence of infectious pathology. These relationships are not always linear and often exhibit a delayed character [1], [2].

Classical methods of correlation and regression analysis are capable of identifying only the most obvious dependencies and prove insufficiently effective in addressing multidimensional non-linear problems.

Artificial intelligence, particularly neural network technologies, offers a fundamentally different approach — training models on historical data with subsequent automatic detection of hidden patterns. In global practice, such solutions are already being successfully employed to forecast outbreaks of influenza, dengue fever, cholera, and other infectious diseases [3], [4]. For Uzbekistan, where the ecological situation is characterized by regional specificity and infectious morbidity retains its social significance, such developments are both timely and in demand.

The objective of the present work is not a technical description of algorithms, but rather a hygienic substantiation of the fundamental feasibility and expediency of utilizing artificial intelligence in the practice of infectious risk forecasting.

2. Materials and Methods

This study was analytical in nature. Peer-reviewed scientific publications indexed in the PubMed, Scopus, eLibrary.ru, and CyberLeninka databases over the period from 2015 to 2024 served as information sources. The search depth encompassed 10 years. Inclusion criteria were the presence of quantitative data on the association between environmental factors and infectious morbidity, as well as a description of prognostic models based on machine learning.

Additionally, normative and methodological documents regulating the conduct of social-hygienic monitoring in the Republic of Uzbekistan and the Russian Federation were examined.

Based on content analysis, the most frequently cited hygienic indicators used as predictors of infectious diseases were identified. These indicators were ranked according to citation frequency and the assessment of prognostic significance provided by the authors.

To demonstrate the advantages of the neural network approach, the results of comparative studies published in open access were utilized. In these studies, morbidity forecasting was performed on identical datasets using multiple linear regression and artificial neural networks. Generalized metrics of forecast quality (mean absolute error, root mean square error, coefficient of determination) were analyzed in dynamics.

3. Results

Priority Hygienic Factors as Predictors of Infectious Risk

Literature analysis indicates that the most informative predictors of infectious morbidity are indicators characterizing drinking water quality and the sanitary-living conditions of the population. In 78 % of the analyzed works, microbial contamination of water, assessed through indirect parameters — turbidity, chloride, ammonia, and nitrate concentrations — was considered a key factor. The second most significant block of predictors was recognized as ambient air parameters, primarily the levels of particulate matter and nitrogen dioxide. This finding aligns with data on the impact of air pollutants on the colonization resistance of respiratory tract mucous membranes.

Meteorological conditions, specifically air temperature and humidity, act as modulating factors, particularly concerning acute intestinal infections with a fecal-oral transmission mechanism. An increase in average daily temperature by 5 °C above the climatic norm during the summer period has been associated with a rise in acute intestinal infection incidence after a lag of 7–10 days.

Social-hygienic indicators (the proportion of the population with access to centralized water supply and sewage systems, population density, migration levels) possess high prognostic capacity at the population level but are of limited utility for short-term forecasting due to their slow rate of change.

Effectiveness of Traditional and Neural Network Prognostic Models

Generalization of the results from applied studies revealed that multiple linear regression, when forecasting infectious morbidity over a one-month horizon, provides a coefficient of determination (R^2) in the range of 0.60–0.70. This implies that the model explains no more than two-thirds of the observed variance in morbidity. The residual dispersion is attributable to unaccounted non-linear effects and interactions among factors.

Artificial neural networks, conversely, are capable of capturing complex dependencies. In studies where direct comparisons were conducted, neural network models demonstrated R^2 values of 0.82–0.88 using the same initial data. The reduction in mean absolute forecast error ranged from 18 % to 25 %. The superiority of artificial intelligence was particularly pronounced during periods of seasonal morbidity increases, when traditional models consistently underestimated forecast values.

It is important to emphasize that neural networks do not require a priori specification of the relationship structure between variables. They autonomously "learn" from examples, rendering them especially valuable for hygienic diagnostics, where many pathogenetic mechanisms of environmental impact on health remain incompletely elucidated.

Conceptual Framework for Implementing Artificial Intelligence in Hygienic Practice

Based on the conducted analysis, a three-level architecture for an infectious risk forecasting system is proposed. The lower level encompasses the collection and unification of heterogeneous data: results of laboratory analyses, information from automated monitoring stations, meteorological data, and official statistical reports. The middle level comprises preprocessing modules — removal of outliers, normalization, gap filling using interpolation methods, and selection of informative features through ranking algorithms. The upper level is the predictive core, implemented as a trained neural network or an ensemble of models, which generates a probabilistic assessment of exceeding the epidemic threshold for a specified period.

A critical prerequisite for the functionality of such a system is the existence of a retrospective database with a depth of at least 3–5 years, characterized by unified description formats. Currently, work is underway in Uzbekistan to create interagency integrated platforms, which establishes preconditions for the pilot implementation of artificial intelligence technologies in selected regions.

4. Discussion

The obtained results affirm that artificial intelligence is not an abstract technology of the future, but rather a concrete tool ready for application in the tasks of hygienic forecasting today. The primary obstacles are not technical but organizational and methodological in nature.

Firstly, there is no unified methodology for assessing the quality of prognostic models specifically for the tasks of sanitary-epidemiological surveillance. Which metrics should be considered acceptable? How frequently should the model be retrained? What is the permissible level of false positive alarms? These questions require normative consolidation.

Secondly, the training of specialists in the hygienic field still does not include the acquisition of basic competencies in data analysis and machine learning. Without this, it is impossible to correctly formulate tasks for developers or to accurately interpret the results produced by artificial intelligence systems.

Thirdly, a degree of skepticism persists among practicing hygienists and epidemiologists regarding neural network forecasts, which are often perceived as a "black box" and considered risky to trust. Overcoming this barrier lies in the demonstration of successful pilot projects and the provision of evidence regarding economic effectiveness.

Nevertheless, the potential benefits of implementing artificial intelligence are evident. Timely warning of a potential outbreak enables the deployment of preventive measures 1–2 weeks earlier than is possible with traditional approaches focused on recording actual cases. For controlled infections (viral hepatitis A, typhoid fever), this implies the possibility of vaccination based on epidemiological indications; for uncontrolled infections, it signifies the strengthening of sanitary-educational work and hygienic instruction.

A limitation of this study is the absence of empirical validation of the proposed model using actual data from Uzbekistan. The next phase envisions a pilot implementation based at one of the sanitary-epidemiological surveillance centers, utilizing retrospective data covering a 5-year period.

5. Conclusion

Environmental quality and the sanitary-hygienic living conditions of the population constitute objective and quantitatively measurable predictors of infectious morbidity. This provides a methodological foundation for prognostic modeling.

Traditional statistical methods are significantly inferior to artificial intelligence algorithms in the accuracy of short-term infectious disease forecasting, particularly during periods of seasonal morbidity increases and in the presence of non-linear interactions among factors.

The application of neural network models allows for an increase in the coefficient of determination of forecasts to 0.85 or higher, a reduction in mean absolute error by 18–22 %, and the automation of routine operations involved in the analysis of multidimensional hygienic data.

The practical implementation of artificial intelligence technologies within the social-hygienic monitoring system of Uzbekistan requires: a) the development of a normative and methodological framework; b) ensuring specialized personnel training; c) creating integrated databases; and d) conducting pilot studies with an assessment of economic efficiency.

The most promising area for the application of artificial intelligence in the near future is the forecasting of acute intestinal infection risk in rural areas with decentralized water supply. This would enable the rational planning of preventive measures and the targeted utilization of resources.

References

[1] Semenov S.A., Yakovlev A.A. Environmental determinants of infectious pathology: modern approaches to risk assessment // Human Ecology. – 2019. – No. 7. – Pp. 45–52.

[2] Khasanova M.A., Karimov F.A. Hygienic characteristics of water supply in rural territories of Uzbekistan // Problems of Social Hygiene, Public Health and History of Medicine. – 2021. – Vol. 29, No. 3. – Pp. 589–594.

[3] Jain V.K., Kumar S. An effective approach to forecast infectious diseases using machine learning algorithms // Procedia Computer Science. – 2020. – Vol. 171. – P. 2152–2161.

[4] Zeng D., Cao Z., Neill D.B. Artificial intelligence–enabled public health surveillance — from local detection to global epidemic monitoring and control // Artificial Intelligence in Medicine. – 2021. – P. 437–453.

[5] Mustafaev B.Sh., Abdullaeva D.G. Modern methods of epidemiological diagnostics // Infection, Immunity and Pharmacology. – 2022. – No. 2. – Pp. 28–35.