A SPATIAL ANALYSIS OF LIFE EXPECTANCY: IS NEIGHBORHOOD MORE IMPORTANT THAN INCOME?

Makhliyo Aripova – Abbos Umarov – Ekaterina Kadochnikova – Nailya Bagautdinova[1]

ABSTRACT: *Life expectancy is a comprehensive indicator of a nation's well-being. The impact of population income, health situation, food consumption, and the environment on the former requires scientific confirmation to justify institutional action. The present study is based on the approaches of Preston (1975), Mackenbach (2012), and Beale et al. (2010), who argue that two measurements from geographically close locations are often more similar than measurements from more widely separated locations. The scientific novelty of the study is in measuring the relationship between life expectancy and median income in Russian regions, taking into account spatial relationships. Regression modeling was performed on the basis of 2022 regional historical data from the statistical abstract Regions of Russia. Socio-economic indicators. Spatial correlation was measured using Global Moran's I, and the choice of the SEM model was confirmed by Lagrange multiplier tests. In the context of income asymmetry and inequality, the results of the models obtained for 2022 demonstrate the spatial relationship of regions and reveal a positive association between life expectancy and the logarithm of median income for the entire country; however, no such relationship was found for specific regional clusters – a phenomenon known as spatial nonstationarity, which limits the reliability of conclusions based on crosssectional data. The study's findings confirm the need for a differentiated approach to managing public health in regions in different*

1 *Makhliyo Aripova* works at Bukhara State University, Uzbekistan; email address: m.s.aripova@buxdu.uz. *Abbos Umarov* is at Bukhara State University, Uzbekistan; email address: a.n.umarov@buxdu.uz. *Ekaterina Kadochnikova* is at Kazan Federal University, Russia; email address: kad-ekaterina@yandex.ru. *Nailya Bagautdinova* is at Kazan Federal University, Russia; email address: nailya.mail@mail.ru.

geographical locations, emphasizing the need for greater consideration of the spillover effects of latent factors from neighboring territories.
KEYWORDS: *life expectancy, average per capita income, Preston curve, income inequality, Moran's I, spatial econometric, SAR and SEM*

INTRODUCTION

The comprehensive nature of life expectancy makes it possible to assess the systemic quality of life in a country. It should be noted that this indicator is not without its paradoxes: regions with high incomes may underperform regions with lower incomes. The gap is sometimes explained by national traditions, the psychological climate, and the sense of social security. A number of studies have shown the dependence of life expectancy on various predictors: these include lifestyle, individuals' upbringing and education, the quality of nutrition and the healthcare system, the crime rate, and income level (Andreev–Shkolnikov 2018; Burkin et al. 2016; Kolosnitsyna et al. 2019). In the 1970s, S. Preston found that the relationship between income level and life expectancy weakens after reaching a certain income level, suggesting the impact of quality of healthcare, ecology, and lifestyle. Preston also noted a shift toward an increase over time: the higher the income distribution dispersion, the lower the life expectancy (Preston 1975). Therefore, high-income countries experience slower growth in life expectancy. Countries with mining-oriented economies tend to have lower health and education indicators than countries with better quality institutions (Edwards 2016). Important contextual predictors of the nature of the relationship between income and life expectancy are the uniformity of income distribution (Rodgers 2002; Pickett–Wilkinson 2015) and the source of income growth (Ryan 2016). Extensive research on within-country relationships between income and health using longitudinal and panel data suggests that within-country relationships are often weaker or differ from cross-sectional patterns: over time, changes in income may have no effect on health or even be negatively related. This is explained by income divergence and the convergence of health indicators due to technological diffusion over time (Spence–Lewis eds. 2009: 9). These comparisons have sparked interest in studying similar patterns in Russian regions with differentiated socioeconomic development. The purpose of the current article is to model the relationship between median per capita income and healthy life expectancy, taking into account spatial regional relationships. The research hypothesis is based on the assumption that spatial effects predominate over structural socioeconomic factors in influencing life expectancy in the regions of Russia.

LITERATURE REVIEW

Theoretical and empirical background

The analysis of the determinants of societal longevity has long attracted the interest of researchers. We examine three complementary approaches that explore why the income health relationship manifests differently in specific contexts.

In his well-known study, Preston demonstrated a nonlinear positive relationship using cross-sectional data: growth in GDP per capita rapidly increases life expectancy in poor countries, but these gains diminish in rich countries (Preston 1975). Preston's work demonstrated the limitations of explanations based solely on economic growth: the main contribution to the increase in life expectancy in the twentieth century was due to factors external to economic growth (medical knowledge and technology). For example, from 1870 to 2015, in a sample of 115 countries, the Preston curve does not flatten out but becomes more convex over time (as per capita income increases), which can be explained by advances in medical knowledge (De La Escosura 2023). Using a sample of data from 148 countries from 1970 to 2010, a weakening of the effect of GDP per capita on median life expectancy was found, while the impact of income inequality on health was found to be more significant in poor countries (Linden–Ray 2017). Moreover, a high level of income inequality has been associated with shorter life expectancy (Wilkinson–Pickett 2006; Curran–Mahutga 2018; Neumayer–Plümper 2016; Rambotti 2015). At the same time, data from a sample of 134 countries over the period 1960 to 2016 suggest that developing economies with lower life expectancy show a stronger upward trend in income (Wang et al. 2022).

In his seminal study, Mackenbach found that societies do not become healthier in a more equitable way as they develop economically, and concludes that inequality in health outcomes is not a direct and inevitable consequence of social inequality (Mackenbach 2012). To explain this paradox, the author draws on material and structural factors, personal and behavioral factors, and the theory of fundamental causes, arguing that people with higher socioeconomic status have access to resources that let them benefit from new medical knowledge and technologies. Research on the impact of extractive industries shows that income from natural resources does not always improve health as effectively as income from other economic sectors, highlighting the role of institutional quality and governance. For example, using data from Indonesia in 2009, it was shown that doubling the share of mining in the economy led to an increase in infant

mortality and a 5% decrease in life expectancy (Edwards 2016). Based on panel data from 88 countries from 2011 to 2020, classified into four income groups (poor, below average, above average, and rich), the differentiated influence of factors underlying life expectancy was identified: for poor countries, health care costs, internet access, and jobs were found to be significant; for middle-income countries, environmental protection and urbanization expenditure; while in high-income ones, harmful health behaviors were found to be significant (Vladimirskaya–Kolosnitsyna 2023).

Finally, Siordia, Saenz, and Tom showed that, following Tobler's law ("everything is related to everything else, but nearby things are more related than distant ones" (Tobler 1970)), the relationship between variables "depends on geographic location." For spatially structured data, the strength and even the sign of relationships can vary spatially due to the influence of local macro-contexts: for example, latent institutional quality or a historically established culture of health can simultaneously influence economic development and health, even creating a spurious relationship (Beale et al. 2010; Siordia et al. 2012).

Features of empirical results for developed economies

Based on data from 17 regions of Spain in the 1980s and 1990s, it was found that there is no statistically significant relationship between life expectancy and average household income for men; for women, the relationship was significant in 1980, although this association had decreased or disappeared by 1990 after adjusting for poverty indicators. The results for women support the hypothesis that life expectancy in developed countries has become more detached from average income and more closely linked to income inequality. The absence of a relationship for men in 1990 may be due to the greater impact of premature mortality from disease in regions with the highest average total household income and/or the lowest income inequality (Regidor et al. 2003). Using data from 49 US states and the District of Columbia (2000–2010) as a sample, a stronger relationship between air pollution and life expectancy in states with higher income inequality was found (Hill et al. 2019), and an inverse relationship between income inequality and life expectancy for men and women (Hill–Jorgenson 2018). Based on structural modeling using samples from the USA, it was concluded that economic inequality further significantly reduces life expectancy due to absolute mobility, especially among men and low-income groups (Xiong–Wei 2025). Researchers identified spatially heterogeneous relationships between social and environmental

determinants of health across neighborhoods in Baltimore and Philadelphia (Mullen– Stewart 2025).

Data from 1968 to 2012, using the Chinese economy as an example, show that rising income inequality driven by regional differences and the urban-rural gap has reduced life expectancy. At the same time, gender differences have been identified – due to different gender roles in society, men's life expectancy loss is greater than women's due to unequal income distribution (Luo–Xie 2020). Also, based on data from the 2005 General Social Survey of China and social statistics at the regional level, a higher risk of health loss was found for regions with higher income inequality (Zhou–Qi 2012). However, based on data from 1991 to 2011 for China, there was no relationship between income inequality and health problems (Bakkeli 2016). However, spatial variation in the relationship between residential energy consumption and life expectancy was found in mainland China in 1990, 2000, and 2010 (Wang et al. 2019).

Features of empirical results for developing economies

In developing economies of mineral-exporting countries, people tend to be less healthy (Besley–Kudamatsu 2006) because rent-seekers can use political power to gain benefits from investments, exacerbating the effects of the "Dutch disease" (Auty 1997) and resisting industrialization and the even distribution of income. Panel data from 56 developing countries in North Africa, the Middle East, and Southeast Asia for the period 1990-2015 showed a positive relationship between economic growth and life expectancy (Gürler–Özsoy 2019). Panel data for low-income countries from 2004 to 2021 show that GDP per capita and government spending on healthcare have a positive impact on life expectancy (Gazilas 2024). One paper (Sofi–Yasmin 2025) showed that health expenses paid from personal sources are more negatively associated with life expectancy in low-income countries than in high-income countries, underscoring the need to reduce financial burdens and improve access to healthcare.

Empirical conclusions for Russian regions

Over the past 27 years, there has been a slight increase in life expectancy, despite difficult conditions leading to natural population decline, including "external reasons," a high demographic dependency burden, and a low birth rate (Fomicheva 2021). Based on panel data for 11 countries (including Russia) for the period from 2000 to 2013, a direct relationship between life expectancy

and income inequality was proven, as well as an inverse relationship with average per capita income and the share of government spending on healthcare (Nazarova–Borisenkova 2017). Based on regional panel data for the period from 1995 to 2017, the differentiation of regions in terms of growth rates in life expectancy has been demonstrated, along with its sensitivity in all regional clusters to the positive dynamics of the region's economic development and an increase in the per capita incomes of the population (Skipin et al. 2022). Based on the regional data of the Siberian Federal District for 2010-2023, it was concluded that the growth in per capita incomes of the population has a positive effect on life expectancy, and the presence of regional specificities in the manifestation of life expectancy factors differentiates the typology of regions (Pastukhova et al. 2025). For Russian regions in 2014, it was concluded that there are differences in life expectancy patterns for men and women, as well as for residents of the western and eastern regions of Russia: a significant positive relationship with income was found for men in the western part of Russia; in the eastern part, it was found for both men and women; taking into account spatial effects, a positive relationship with the income of the population was also confirmed, while the spatial dependence of life expectancy was not confirmed for the eastern regions (Zhukova et al. 2016).

DATA AND METHODS

This analysis used 2022 statistical data from the Statistical Abstract *Regions of Russia, Socio-Economic Indicators of 2025* (Rosstat 2025) for the 85 regions with available information (from a total of 89 regions, see Table 1).

Table 1. *Descriptive statistics of variables*

Variables	Mean	Median	Stand. dev.	Kurtosis	Asym.	Min.	Max.
Life expectancy, years	72	72	2	1	1.0	66	78
Median per capita monetary income per month, rubles	33,117	30,048	12.719	11	3.0	19,053	98,786
Number of doctors of all specialties per 10,000 population	49.070	47.650	10.118	2.833	1.255	30.900	89.100
Environmental protection expenditure per 10,000 population, in million rubles	61.779	40.711	64.421	8.214	2.729	1.009	337.136
Per capita consumption of meat and meat products per year, kg	75.976	75.500	12.991	0.757	0.528	46.000	115.000
Number of gyms and swimming pools per 10,000 people	7.577	6.546	5.767	23.628	4.732	2.579	40.009

Source: Compiled by the authors using data from the statistical abstract of Rosstat (2025)

As can be seen from Table 1, median per capita monthly income and the number of gyms and swimming pools per 10,000 people have the highest kurtosis values; they also have the highest absolute values of the asymmetry coefficient. Further, they have the largest positive (right-sided) skewness coefficients. This means that there is a small number of highly affluent regions that form a long right-hand tail in the income distribution. This confirms the spatial inequality in the distribution of median income for the entire country. Therefore, in the next stage of the study, we clustered regions by life expectancy and median income.

The k-median method was used to cluster regions (Hoon et al. 2019), in which the median is calculated to determine the centroid of each cluster. The objective function for *k*-medians is to find the *C(i)* distribution of observations *i* over clusters *h=1, ... k*, so that the sum of Manhattan distances between the members of each cluster and the median of the cluster is minimized (Anselin 2020):

$$\arg\min_{C(i)} \sum_{h=1}^{k} \sum_{i \in h} \left\| x_i - x_{h\,med} \right\|_{L1} \quad , \tag{1}$$

where the distance metric corresponds to the L1 norm (that is, the Manhattan distance).

To test spatial correlation, we used the Global Moran's I (Anselin 1988):

$$I(X) = \frac{N}{\sum_{i,j} w_{ij}} \cdot \frac{\sum_{i,j} w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_{i} (Xi - \bar{X})^2} \quad , \tag{2}$$

where N is the number of regions, is the average value of the indicator X, and w_{ij} are the elements of the weight matrix. In this study, we used two types of matrices: a queen-type contiguity matrix (a symmetric matrix based on common geographic boundaries) and a four-nearest-neighbors matrix (based on the four regions whose centroids are closest in Euclidean distance to the centroid of the given region).

Theoretically, Moran's I can take values in the range [–1, 1], where values from 0 to 1 indicate the presence of a positive spatial correlation, and values from –1 to 0 indicate a negative spatial correlation.

To test and represent the relationship between healthy life expectancy and predictors analytically, we used linear multiple regression without a spatial component, taking into account spatial relationships in the dependent variable (SAR) and in the model error (SEM). (Le Sage 2008):

$$Y_i = \alpha + \beta_j X_{i\,j} + \varepsilon_i \tag{3}$$

$$Y_i = \alpha + \rho W Y_i + \beta_j X_{i\,j} + \varepsilon_i \tag{4}$$

$$Y_i = \alpha + \beta_j X_{i\,j} + \lambda W u_i + \varepsilon_i \tag{5}$$

where Y_i is the healthy life expectancy in region i, α is the parameter to be evaluated; β_j is the regression coefficient; X_{ij} is the matrix of independent variables; W is the weight matrix; ρ is the spatial autoregression coefficient for the dependent variable, measuring the effect of the average expected life expectancy of neighboring regions; β_j are the coefficients for independent variables; λ is the spatial autoregressive error parameter which captures the influence of unobserved features (culture, traditions, infrastructure, recreational conditions, etc.) or omitted variables in neighboring regions on life expectancy in a given region; u is the spatially autoregressive error; and ε_i are random errors.

The dependent variable is healthy life expectancy. The variable of interest is median per capita monetary income per month. Aligned with previously published studies, we selected the following as control variables: per capita consumption of meat and meat products per year (You et al. 2022), the number of gyms and swimming pools per 10,000 people (Munaier–Tavares 2022), environmental protection costs (Khelfaoui et al. 2022; Lee et al. 2016), and the number of doctors of all specialties (Dutta et al. 2019; Lee et al. 2016; Majeed–Khan 2019). To eliminate excessive data heterogeneity, factors in logarithmic form were also used in the study.

Estimating spatial econometric SAR and SEM models requires symmetric weight matrices. Therefore, the models in this study use a queen-type associative link symmetric weighting matrix. The estimates were obtained using the GeoDa and R software environments.

To select the specification of the spatial econometric model (SAR or SEM), the study used the approach proposed by Anselin (Anselin et al. 1996). The need to extend the OLS model to account for spatial relationship effects is satisfied based on a collection of Lagrange multipliers: if one of the Lagrange multiplier tests (lag or error) is statistically significant (p-value < 0,1), then the corresponding model (SAR or SEM) is selected; if neither test is significant, we retain the results of the non-spatial ordinary least squares linear regression (OLS); if both Lagrange multiplier tests (lag and error) are statistically significant, then the corresponding robust criteria should be calculated; if only one robust criterion of the Lagrange multiplier (lag or error) is statistically significant, then we select the corresponding model; if both robust criteria are statistically significant, then we select the model with a higher robust LM statistics.

To assess the stability of the model, the study used the diagnosis of heteroscedasticity (the Breusch-Pagan test) and a map of the model's regression residuals – spatial clusters with very high or low residuals (anomalies) indicate regions that the model does not explain well.

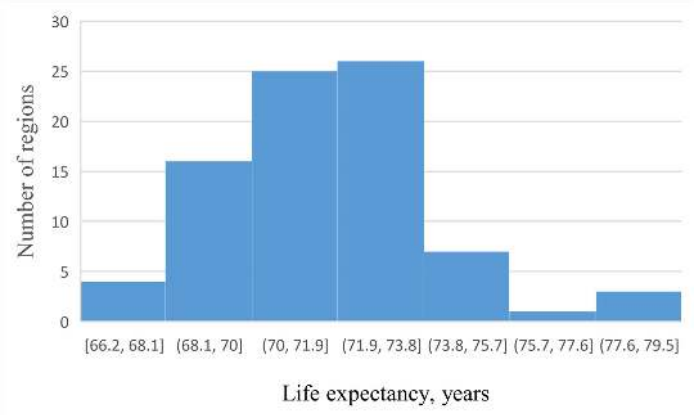
Taking into account the differentiation of Russian regions, the study assessed spatial regression models separately for each cluster of regions based on life expectancy.

RESULTS AND DISCUSSION

Regional differences in life expectancy and income

Life expectancy is characterized by a moderate right-sided asymmetry (Figure 1), that is, a small number of leading regions with high life expectancy form a long right-hand tail of the distribution. The largest number of regions fall within the range of 71.9–73.8 (26 federal subjects of the Russian Federation); a further four subjects lie within the range of 66.2–68.1; three subjects lie within the range of 77.6–79.5.

Figure 1. *Distribution of regions by life expectancy (histogram)*

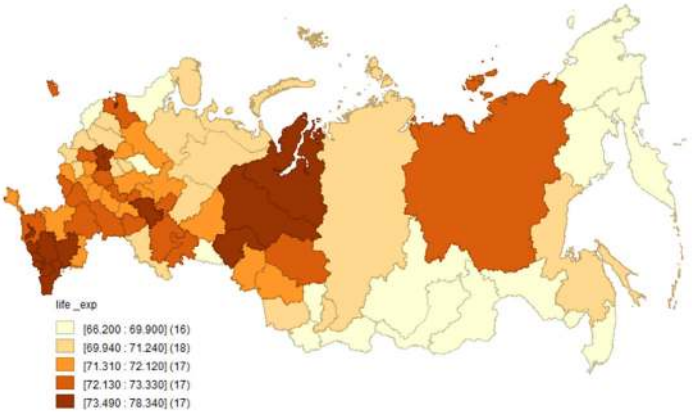


Source: Compiled by the authors using data from Rosstat (2025).

As shown in Figure 2, the lowest life expectancy values are observed in the Chukotka Autonomous Okrug, the Republic of Tyva, the Jewish Autonomous Oblast, Transbaikal Krai, Amur Oblast, and Magadan Oblast. The highest life expectancy is observed in the Kabardino-Balkarian Republic, St. Petersburg, Moscow, Republic of Dagestan, and Republic of Ingushetia.

The regions of the three clusters have the highest median life expectancy. These include Moscow and the Moscow Oblast, the Republic of Tatarstan, and the Tyumen Oblast (Figure 3). The high life expectancy in the regions of the Ural Federal District and the North Caucasus Federal District can be explained by national traditions and favorable climatic conditions.

Figure 2. *Choropleth map of life expectancy in the regions of the Russian Federation in 2022*



Source: compiled by the authors using data from Rosstat (2025).

Figure 3. *Choropleth map of regional clusters of life expectancy using the K-Medians method*



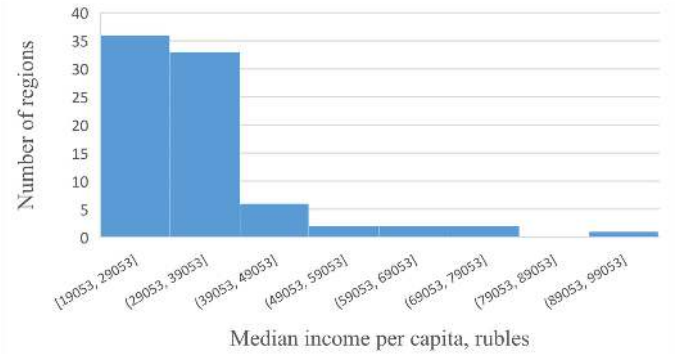
Source: Compiled by the authors using data from Rosstat (2025).

The high value of the inter-cluster sum of squares (71.8965) indicates variability between clusters. The largest value of the intra-cluster sum of squares is observed for the third cluster (7.2625), which indicates significant variability within this cluster. This may indicate the presence of outliers or a more complex

data structure. The ratio of total intra-cluster distance to total distance (0.3526, 35.26%) shows that intra-cluster distances make up a significant share of the total. This indicates a fairly clear segmentation of regions into groups.

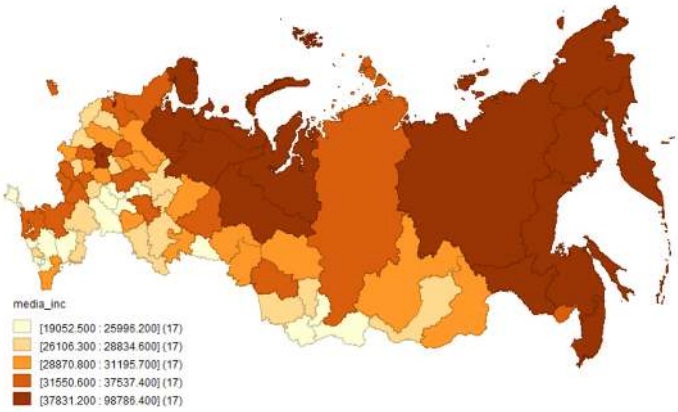
All independent variables have a right-sided asymmetry. This asymmetry is typical for median per capita incomes in most regions (Figure 4), which indicates income inequality. The highest median per capita incomes in 2022 were observed in the Chukotka Autonomous District, the Tyumen Oblast, the Magadan Oblast, and Moscow.

Figure 4. *Distribution of regions by median per capita income (histogram)*



Source: Compiled by the authors using the data from Rosstat (2025).

Figure 5. *Choropleth map of median per capita income in the regions of the Russian Federation in 2022*



Source: Compiled by the authors using the data from Rosstat (2025).

In terms of median per capita income, clear spatial differentiation emerges: the highest incomes are typical of the northern regions and the Far East, of Moscow, the Moscow Region, and St. Petersburg. Comparing the maps of life expectancy (Figure 2) and median per capita income (Figure 5) does not clearly show a negative or positive correlation.

Figure 6. *Choropleth map of regional clusters based on predictors of life expectancy identified using the K-medians method*



Source: Compiled by the authors based on the data Rosstat (2025).

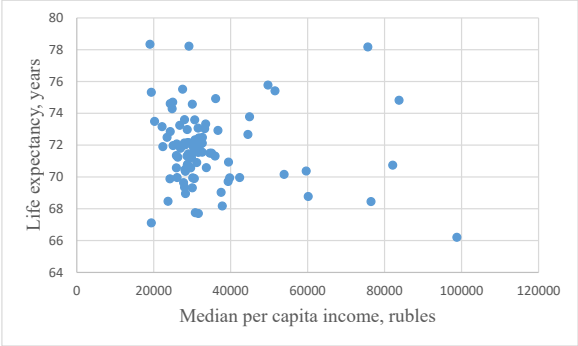
The highest median values of the predictors belong to the regions of Cluster 4. Figure 6 also does not show a close relationship between predictor clusters and life expectancy clusters (Figure 3).

Regression results

As can be seen in Figure 7, there is no relationship between life expectancy and median per capita income. The Pearson linear correlation coefficient was –0.104. A Student's t-test confirmed the absence of a statistical linear relationship between life expectancy and median per capita income ($t=0.957 < t_{0,1;83}=1.663$). A similar conclusion emerged from Spearman's rank correlation coefficient: with a coefficient of -0,131, the significance test also confirmed no relationship ($t= -1.205 < t_{0,1;83}=1.663$). The upper-left part of the graph shows regions with low incomes and fairly high life expectancy (the Republic of Dagestan, the Republic of Ingushetia, the Kabardino–Balkarian Republic, and the Karachay–

Cherkess Republic). Among them, the Republic of Tuva has the lowest life expectancy and one of the lowest median per capita incomes. There are five regions on the right side of the graph, two of which (Moscow and the Yamalo-Nenets Autonomous District) have high incomes and high life expectancy, while three regions (the Nenets Autonomous District, the Magadan Oblast, and the Chukotka Autonomous District) have high incomes and low life expectancy, which can be explained by the harsh climatic conditions.

Figure 7. *Scatter plot of median per capita income and life expectancy across Russian regions, 2022*



Source: Compiled by the authors based on the data from Rosstat (2025).

Table 2. *Global Moran's I for variables in 2022*

Variables	Index (associative connection)	Index (4 closest neighbors)
Life expectancy, years	0.518*** (0.001)	0.532*** (0.001)
Median per capita monetary income per month, rubles	0.499*** (0.001)	0.447*** (0.001)
Number of doctors of all specialties per 10,000 population	0.088* (0.088)	0.047 (0.194)
Environmental protection expenditure per 10,000 population, million rubles.	0.402*** (0.001)	0.421*** (0.001)
Per capita consumption of meat and meat products per year, kg	0.026 (0.289)	0.113** (0.046)
Number of gyms and swimming pools per 10,000 people	0.208*** (0.007)	0.185** (0.034)

Source: Compiled by the authors based on on the data from Rosstat (2025).
Note: The p-value for Moran's I is included in brackets.

Positive spatial correlations across all indicators (the "high-high" and "low-low" clusters) point to spatial inequality in development, the presence of prosperous donor regions, growth hotspots (the South, Central Russia, and oil and gas regions), and peripheral regions (some regions of Siberia, the Northwest, and the Far East). This is due to labor migration to wealthier regions, access to healthcare, and the cross-border effects of environmental costs. The "high-high" cluster is often associated with large agglomerations, where physician density, environmental costs, and life expectancy are higher. However, Figure 7 shows that high-income clusters do not necessarily have higher life expectancy. This demonstrates so-called "spatial non-stationarity," whereby, across the country as a whole, individuals in rich regions do not live longer than those in poor regions, while all indicators individually form stable positive clusters of "high-high" and "low-low." This means that regions are similar to their neighbors. However, this does not prove that income is linked to life expectancy within the region. The relationship between income and life expectancy varies radically across parts of Russia (long-lived regions with low incomes, peripheral agricultural regions, oil and gas regions of the North, the South, the Caucasus, and large metropolitan areas). This suggests heterogeneity in the mechanisms determining health across Russia. The income-health relationship does not operate universally, but only in certain spatial regimes that need to be identified separately. A positive spatial correlation reflects not interfactorial dependence but the inertia of territorial systems – historically established settlement patterns, climatic zones, budgetary centralization, and cultural traditions transmitted from region to neighbor. The spatial structure of indicators is more stable than their economic interrelationships. Thus, geographic proximity and general background conditions (climate, urbanization, culture) turn out to be stronger predictors of health than income level.

The regression analysis results (Table 3) revealed the logically expected positive relationship between the logarithm of median income and life expectancy. Contrary to theoretical expectations, the spatial SEM identified a negative relationship between life expectancy and the logarithms of environmental protection expenditures per 10,000 population and the logarithm of gyms and swimming pools per 10,000 people. This can be attributed to low population density in industrial and extractive regions characterized by harsh climatic conditions (Magadan Oblast; the Republics of Karelia, Sakha (Yakutia), and Tuva; the Yamalo-Nenets, Khanty-Mansi, and Chukotka Autonomous Okrugs; and Irkutsk, Krasnoyarsk, Kamchatka, and Sakhalin Oblasts). According to the Lagrange multiplier tests, the SEM model was preferable (the p-value for the Lagrange Multiplier (error) is $p<0.001$; for Robust LM (error), $p<0.001$). This model revealed a statistically significant positive effect of the unobserved

characteristics of neighboring regions on life expectancy in the given region. This indicates that macro-regional factors (climate, cultural norms, migration flows, the quality of institutions, and the degree of urbanization) have a greater impact on life expectancy in this region than the region's own spending on environmental protection. Management tools at the level of an individual region have limited effectiveness in the context of regional differentiation. Regions with good healthcare may have poor environmental conditions or high levels of stress (due to urbanization), creating a "compensation" effect. Previously, we found no relationship between median income and life expectancy in a cross-sectional analysis (Figure 1), but each indicator is spatially clustered (Figure 3 and 6). The regression results confirm this – spatial effects "crowd out" structural factors.

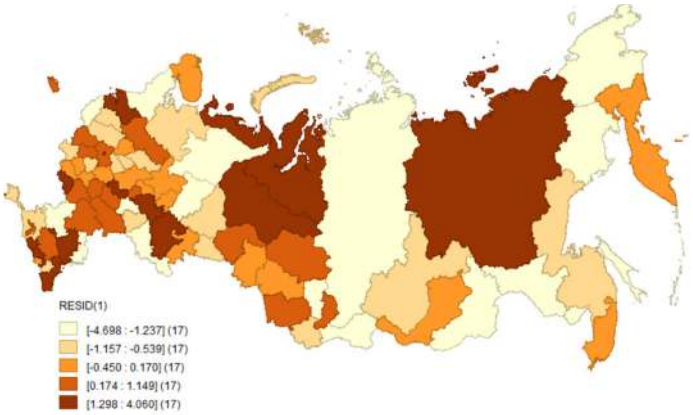
Table 3. *Results of the regression analysis of life expectancy by country*

Indipendent variables	Dependent variable: life expectancy		
	OLS	SEM (1) (associative connection)	SEM (2) (associative connection)
Median per capita monetary income per month (1) / Natural logarithm of the median monthly monetary income per capita (2)	–1.506e-06 (0.959)	1.620e-05 (0.471)	1.523* (0.068)
Number of doctors of all specialties per 10,000 population (1) / Natural logarithm of number of doctors of all specialties per 10,000 population (2)	0.026 (0.419)	0.011 (0.546)	1.076 (0.231)
Environmental protection expenditure per 10,000 population (1) / Natural logarithm of environmental protection expenditure per 10,000 population (2)	–1.924e-05 (0.995)	0.002 (0.503)	–0.548** (0.032)
Per capita consumption of meat and meat products per year (1) / Natural logarithm of per capita consumption of meat and meat products per year (2)	–0.012 (0.512)	–0.015 (0.223)	–0.302 (0.743)
Number of gyms and swimming pools per 10,000 people (1) / Natural logarithm of number of gyms and swimming pools per 10,000 people (2)	–0.393*** (0.000)	–0.291*** (0.000)	–2.790*** (0.000)
Constant	74.118*** (0.000)	73.702*** (0.000)	60.187*** (0.000)
Lambda (spatial autoregressive error parameter)	–	0.749*** (0.000)	0.687*** (0.000)
R^2	0.168	0.578	0.623
p-value (F)	0.011	-	-
p-value –Test Lagrange Multiplier (lag)	0.418	-	-
p-value – Robust LM (lag)	0.865	-	-
p-value –Test Lagrange Multiplier (error)	0.000	-	-
p-value – Robust LM (error)	0.000	-	-
Log likelihood	–185.187	–163.321	–157.105

Indipendent variables	Dependent variable: life expectancy		
	OLS	SEM (1) (associative connection)	SEM (2) (associative connection)
Akaike info criterion	382.374	338.641	326.21
p-value – Breusch-Pagan test	0.000	0.000	0.000
p-value – Likelihood Ratio Test (spatial error dependence)	0.000	0.000	0.000

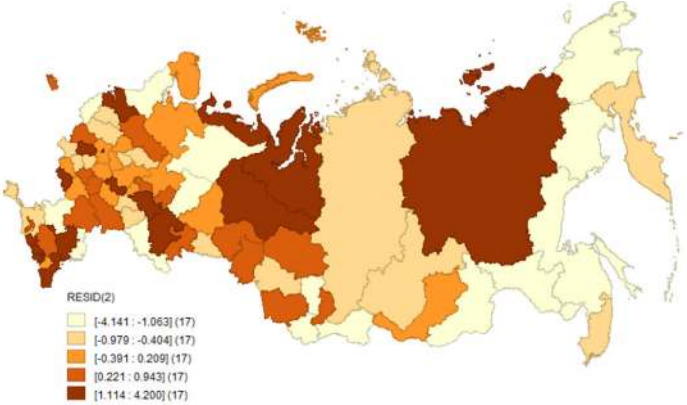
Source: Compiled by the authors based on the data from Rosstat (2025).
Note: p-values for Student's t-statistics are in brackets.

Figure 8. *Choropleth map of residuals from the Spatial Error Model (1)*



Source: Compiled by the authors based on the data Rosstat (2025).

Figure 9. *Choropleth map of residuals from the Spatial Error Model (2)*



Source: Compiled by the authors based on the data Rosstat (2025).

The Breusch-Pagan test for the SEM models (Table 3) indicates heteroskedasticity in the residuals, i.e., a problem with the model specification as a whole for the entire country. The residual maps for the SEM models (Figures 8 and 9) also reveal the presence of clusters of regions with anomalous residuals. These circumstances do not confirm the robustness of SEM models for the country as a whole.

Below are the spatial regression results for each life expectancy cluster (see Table 4.).

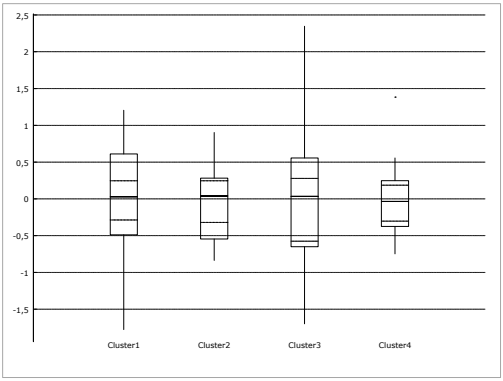
Table 4. *Results of regression analysis of life expectancy in regional clusters*

Indipendent variables	Dependent variable: life expectancy			
	SEM (1 cluster) (associative connection)	SEM (2 clusters) (associative connection)	SEM (3 clusters) (associative connection)	SEM (4 clusters) (associative connection)
Natural logarithm of median monthly monetary income per capita	–1.199 (0.415)	–0.463 (0.592)	1.561 (0.170)	0.475 (0.689)
Natural logarithm of number of doctors of all specialties per 10,000 population	–0.385 (0.646)	0.938 (0.392)	1.868 (0.343)	–0.053 (0.962)
Natural logarithm of environmental protection expenditure per 10,000 population	–0.029 (0.877)	–0.005 (0.981)	–0.810** (0.033)	0.235 (0.579)
Natural logarithm of per capita consumption of meat and meat products per year	1.235 (0.263)	–0.965 (0.351)	–2.214** (0.048)	1.059 (0.244)
Natural logarithm of number of gyms and swimming pools per 10,000 people	–0.591 (0.496)	0.156 (0.750)	–0.292 (0.795)	–4.075 (0.100)
Constant	81.845*** (2.87e-06)	75.575*** (6.76e-010)	64.015*** (0.001)	66.145*** (0.000)
Lambda (spatial autoregressive error parameter)	0.547*** (0.000)	0.436*** (0.000)	0.428*** (0.000)	0.403*** (0.000)
R^2	0.125	0.164	0.564	0.631
Log likelihood	–31.376	–18.115	–24.429	–9.854
Akaike info criterion	74.752	48.230	60.858	31.708
p-value – Breusch-Pagan test	0.145	0.894	0.284	0.524
p-value – Jarque–Bera test	0.533	0.525	0.517	0.126
p-value – Likelihood Ratio test (spatial error dependence)	0.000	0.000	0.000	0.000

Source: Compiled by the authors based on the data from Rosstat (2025).
Note: p-values for Student's t-statistics are in brackets.

The Breusch-Pagan and Jarque–Bera tests indicate homoscedasticity in the regression residuals for the regional clusters, supporting the robustness of these models. The boxplots of the regression residuals for the regional clusters (Figure 10) show anomalous residuals for the third cluster of regions with the highest life expectancy. The model does not explain this indicator well for Moscow, the Republic of Dagestan, the Republic of Ingushetia, and the Tyumen Oblast.

Figure 10. *Boxplots of regression residuals for regional clusters*



The regression results for the clusters (Table 4) confirm that spatial effects "crowd out" structural factors. The negative correlation between environmental protection expenditure per 10,000 population and life expectancy in the third cluster can be explained by the low population density in the Khanty-Mansi and Yamalo-Nenets Autonomous Okrugs, where extractive industry enterprises are concentrated in the northern part of the country. A positive spatial autoregressive error indicates regional cooperation in terms of healthy life expectancy: an increase in unobserved predictors of life expectancy among neighbors leads to an increase in life expectancy in a given region. Regional dietary traditions, migration flows, and pollutant emissions are transmitted through proximity. The effectiveness of spending on healthcare and education varies due to the varying quality of regional institutions. Northern regions have higher mortality rates due to cold stress, short daylight hours, and low vitamin levels. Large cities have better healthcare, but worse environmental conditions. The balance of these effects is highly spatially clustered. The relationship between median monthly per capita income and life expectancy in regional clusters was not confirmed. This finding may be explained by previous research (Spence–Lewis eds. 2009: 9; Nazarova–Borisenkova 2017; Sofi–Yasmin 2025).

First, in the context of declining regional income inequality (Gluschenko 2022), there are higher income growth rates in less well-off regions compared to richer ones (Polbin–Ivakhnenko 2022), but given the persistence of a gap between poor and rich regions of Russia (Dmitriev 2024), spatial inertia is observed. Second, high incomes in eastern, mining-oriented regions with unfavorable climatic conditions, and lower incomes in southern agricultural regions supportive of a healthy life, lead to a "compensation effect" in the relationship between life expectancy and median income.

CONCLUSION

The study found a positive spatial correlation for each individual indicator, including life expectancy and its predictors: median monthly cash income per capita, the number of physicians of all specialties per 10,000 population, environmental protection expenditure per 10,000 population, annual meat and meat product consumption per capita, and the number of gyms and swimming pools per 10,000 population. Regression analysis (based on spatial sample data) identified a positive relationship between life expectancy and median income for the country as a whole, though it did not hold for individual regional clusters. A statistically significant positive influence of the unobserved characteristics of neighboring regions on life expectancy was identified. Population health is determined not by local budgetary policy, but by macroregional factors – climate, cultural norms, institutional quality, and the degree of urbanization – which vary smoothly across the territory.

The positive spatial correlation of life expectancy indicates that the current health status of the population in regions depends on a combination of underlying factors in neighboring regions. In the context of income convergence, the lack of correlation between life expectancy and income in certain regional clusters can be attributed to differences in socioeconomic and climatic conditions, traditions, migration, and mentality across regions. The health of a region depends on its neighbors, not just on its own policies, and traditional levers of control (spending on medicine, education) produce different effects depending on the quality of institutions. Nevertheless, the results allow us to formulate recommendations for institutions developing life expectancy improvement programs.

These include using macroregional healthcare clusters instead of regional boundaries for indicative management and developing federal target development programs, and approving separate sets of target indicators and tools for each cluster, taking into account climatic, cultural, and migration

characteristics. Also, taking into account the "health" or "ill-health" of neighboring regions when distributing federal healthcare subsidies to offset negative externalities (e.g., the hospitalization of patients from neighboring depressed regions or damage from transboundary pollution). Further, applying the principle of interregional financing to federal environmental programs: a region suffering from transboundary pollution could be entitled to compensation payments from the federal budget to cover additional medical expenses. We also advise including the indicator "results of regional healthcare expenditure" as the ratio of the increase in life expectancy to the volume of healthcare expenditures over a given period, adjusted for spatial effects, among the key performance indicators used by governors. Additionally, creating a federal rating of "quality of healthcare institutions" that includes indicators of procurement transparency, staffing levels, the accessibility of primary care, and public satisfaction. Due to regional differences in cumulative life expectancy factors, in addition to the federal strategy, regional programs targeting the priority needs of regions and settlements of different types are needed. In low-income regions, it is necessary to focus not only on the effective allocation of funds within the healthcare system and food security, but also on social policy, employment and the resettlement of the population, and information infrastructure. In medium- and high-income regions, healthy lifestyle programs can be implemented, and emissions of harmful atmospheric pollutants can be monitored and reduced.

For leading regions, shift the focus to the quality and accessibility of healthcare. Invest in urban environments that impact health: parks, bike paths, pedestrian areas, air purification (green spaces, electric transport). Develop programs to prevent stress-related diseases in metropolitan areas. For peripheral regions, focus on basic and primary healthcare; use telemedicine and mobile teams to compensate for staff shortages, attracting doctors from neighboring, more affluent regions on a part-time basis. Develop programs to retain young people to reduce the outflow of healthy people.

To model the relationship between income and life expectancy while accounting for spatial non-stationarity, future studies may use geographically weighted regression. Future research can develop approaches to decomposing the spatial lag into measurable components: climatic (temperature, solar radiation, daylight hours); cultural and behavioral (tobacco and beverage consumption, dietary patterns); institutional (government effectiveness, budgetary process quality); and migration (migration balance, share of newcomers, youth outflow). The modeling of healthy life expectancy can also be continued, taking into account the sectoral specifics of the regional economy and the level of development of information and communication technologies.

REFERENCES

Andreev, E. M. – V. Shkolnikov, V. (2018) The relationship between mortality and economic development levels in Russia and the regions. *Demograficheskoe obozrenie* [*Demografic Review*], Vol. 5, No.1, pp. 6–24, (in Russian) DOI: https://doi.org/10.17323/demreview.v5i1.7707.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models.* Dordrecht, NL, Kluwer.

Anselin, L. (2020) Cluster Analysis (3). Advanced Clustering Methods. Workbook, University of Chicago, US, Center for Spatial Data Science, https://geodacenter.github.io/workbook/7c_clusters_3/lab7c.html.

Anselin, L. – A. K. Bera – R. Florax – M. J. Yoon (1996) Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, Vol. 26, No.1., pp. 77–104, DOI: https://doi.org/10.1016/0166-0462(95)02111-6.

Auty, R. M. (1997) Natural resource endowment: The state and development strategy. *Journal of International Development*, Vol. 9, No. 4., pp. 651–663, DOI: https://doi.org/10.1002/(SICI)1099-1328(199706)9:4<651::AID-JID474>3.0.CO;2-4.

Bakkeli, N. Z. (2016) Income inequality and health in China: A panel data analysis. *Social Science & Medicine*, Vol. 157, pp. 39–47, DOI: https://doi.org/10.1016/j.socscimed.2016.03.041.

Beale, C. M. – J. J. Lennon – J. M. Yearsley – M. J. Brewer – D. A. Elston (2010) Regression analysis of spatial data. *Ecology Letters,* Vol. 13, No.2, pp. 246–264, DOI: https://doi.org/10.1111/j.1461-0248.2009.01422.x.

Besley, T. – M. Kudamatsu (2006) Health and democracy. *American Economic Review*, Vol. 96, No. 2., pp. 313–318, DOI: https://doi.org/10.1257/000282806777212053.

Burkin, M. M. – E.V. Molchanova – M.M. Kruchek (2016) Integral criterion of the influence of social, economic and environmental factors on the regional demographic processes. *Ekologiya cheloveka* [*Human Ecology*], Vol. 6, No. 6., pp. 39–46, DOI: https://doi.org/10.33396/1728-0869-2016-6-39-46.

Curran, M. – M. Mahutga (2018) Income inequality and population health: A global gradient? *Journal of Health and Social Behavior*, Vol. 59, No. 4., pp. 536–553, DOI: https://doi.org/10.1177/0022146518808028.

De La Escosura, L. P. (2023) Health, income, and the Preston curve: A long view. *Economics & Human Biology*, Vol. 48, No. 2, Art. No. 101212, DOI: https://doi.org/10.1016/j.ehb.2022.101212.

Dmitriev, S. G. (2024) Divergence and convergence of money income in Russian regions. *Creative Economy*, Vol.18, No. 2, pp. 487–510, DOI: https://doi.org/10.18334/ce.18.2.120553.

Dutta, U. P. – H. Gupta – P. P. Sengupta (2019) ICT and health outcome nexus in 30 selected Asian countries: Fresh evidence from panel data analysis. *Technology in Society*, Vol. 59, Art. No. 101184, DOI: https://doi.org/10.1016/j.techsoc.2019.101184.

Edwards, R. B. (2016) Mining away the Preston curve. *World Development*, Vol. 78, pp. 22–36, DOI: https://doi.org/10.1016/j.worlddev.2015.10.013.

Fomicheva, T. V. (2021) The dynamics of the life-span of citizens of the Russian Federation: the sociological aspect. *Problemi socialnoi gigieni, zdravookhranenia i istorii meditsini*, Vol. 29, No. 2., pp. 260–264, (in Russian) DOI: https://doi.org/10.32687/0869-866X-2021-29-2-260-264.

Gazilas, E. T. (2024) Factors influencing life expectancy in low-income countries: A panel data analysis. *Journal of Applied Economic Research*, Vol. 23, No. 3., pp. 580–601, DOI: https://doi.org/10.15826/vestnik.2024.23.3.023.

Gluschenko, K. P. (2023) Convergence patterns of regional income inequality levels in Russia. *Prostranstvennaya Ekonomika* [*Spatial Economics*], Vol. 19, No. 3., pp. 46–62, (in Russian) DOI: http://dx.doi.org/https://dx.doi.org/10.14530/se.2023.3.046-062.

Gürler M. – Ö. Özsoy (2019) Exploring the relationship between life expectancy at birth and economic growth in 56 developing countries. *Journal of Global Health Reports*, Vol. 3., e2019001, DOI: https://doi.org/10.29392/joghr.3.e2019001.

Hill, T. D. – A. K. Jorgenson (2018) Bring out your dead!: A study of income inequality and life expectancy in the United States, 2000–2010. *Health & Place*, Vol. 49, pp. 1–6, DOI: https://doi.org/10.1016/j.healthplace.2017.11.001.

Hill, T. D. – A.K. Jorgenson – P. Ore – K.S. Balistreri – B. Clark (2019) Air quality and life expectancy in the United States: An analysis of the moderating effect of income inequality. *SSM – Population Health*, Vol. 7., Art. no. 100346, DOI: https://doi.org/10.1016/j.ssmph.2018.100346.

Hoon, M. de – S. Imoto – S. Miyano (2019) *The C Clustering Library*. The University of Tokyo, Institute of Medical Science, Human Genome Center, Tokyo, Japan. http://bonsai.hgc.jp/~mdehoon/software/cluster/cluster.pdf

Khelfaoui, I. – Y. Xie – M. Hafeez – D. Ahmed – H. E. Degha et al. (2022) Information and communication technology and infant mortality in low-income countries: Empirical study using panel data models. *International Journal of Environmental Research and Public Health*, Vol. 19, No. 12., p. 7338. DOI: https://doi.org/10.3390/ijerph19127338

Kolosnitsyna, M. – T. Kossova – M. Sheluntsova (2019) Factors of the increase in life expectancy at birth: Cluster analysis across the countries in the world. *Demograficheskoe obozrenie* [*Demographic Review*], 2019, Vol. 6, No. 1., pp. 124–150, (in Russian).

Le Sage J. P. (2008) An Introduction to Spatial Econometrics. *Revue d'Économie Industrielle,* Vol. 123, No. 123., pp. 19–44, DOI: https://doi.org/10.4000/rei.3887.

Lee, M-H. – P-Y. Liu – M-C. Lio (2016) The impact of the diffusion of information and communication technology on health: A cross-country study. *Applied Research in Quality of Life*, Vol. 11, No. 2, pp. 471–491 DOI: https://doi.org/10.1007/s11482-014-9376-5.

Linden, M. – D. Ray (2017) Aggregation bias-correcting approach to the health–income relationship: Life expectancy and GDP per capita in 148 countries, 1970–2010. *Economic Modelling,* Vol. 61, pp. 126–136, DOI: https://doi.org/10.1016/j.econmod.2016.12.001.

Luo, W. – Y. Xie (2020) Economic growth, income inequality and life expectancy in China. *Social Science & Medicine*, Vol. 256, Art. No. 113046, DOI: https://doi.org/10.1016/j.socscimed.2020.113046.

Mackenbach, J. P. (2012) The persistence of health inequalities in modern welfare states: The explanation of a paradox. *Social Science & Medicine*, Vol. 75, No. 4., pp. 761–769, DOI: https://doi.org/10.1016/j.socscimed.2012.02.031

Majeed, M. T. – F.N. Khan (2019) Do information and communication technologies (ICTs) contribute to health outcomes? An empirical analysis. *Quality and Quantity*, Vol. 53, No.1, pp.183–206, DOI: https://doi.org/10.1007/s11135-018-0741-6.

Mullen H. – K. Stewart (2025) Assessing the legacy of redlining on spatial inequities in social and environmental determinants of health. *Applied Geography*, Vol. 179, Art. no. 103637, DOI: https://doi.org/10.1016/j.apgeog.2025.103637

Munaier, C. G. e S. – L. W. Tavares (2022) Gym access: Proposal of a synthetic variable in the fight against non-communicable chronic diseases. *RISUS – Journal on Innovation and Sustainability,* Vol. 13, No. 4., pp. 68–82, DOI: http://dx.doi.org/10.23925/2179-3565.2022v13i4p68-82.

Nazarova V.V. – K.A. Borisenkova (2017) Assessment of the efficiency of the health care system in Russia. *Narodonaselenie* [*Population*], Vol. 4, pp. 119–134, DOI: 10.26653/1561-7785-2017-4-9.

Neumayer, E. – T. Plümper (2016) Inequalities of income and inequalities of longevity: A cross-country study. *American Journal of Public Health*, Vol. 106, No.1, pp. 160–165, DOI: https://doi.org/10.2105/AJPH.2015.302849.

Pastukhova, E. Y. – I.V. Korchagina – E. A. Morozova (2025) Dynamics and life expectancy socio-economic factors: Regional aspect. *Management Issues*, Vol.19, No.1, pp. 42–55. DOI: https://doi.org/10.22394/2304-3369-2025-1-42-55.

Pickett, K.E. – R.G. Wilkinson (2015) Income inequality and health: A causal review. *Social Science & Medicine*, Vol. 128, pp. 316–326, DOI: https://doi.org/10.1016/j.socscimed.2014.12.031.

Polbin A.V. – T.Y. Ivakhnenko (2022) Convergence of income inequality in Russia's regions. *Prostranstvennaya Ekonomika* [*Spatial Economics*], Vol. 18, No. 4., pp. 68–92, DOI: https://dx.doi.org/10.14530/se.2022.4.068-092

Preston, S. H. (1975) The changing relation between mortality and level of economic development. *Population Studies,* Vol. 29, No. 2., pp. 231–248, DOI: https://doi.org/10.1080/00324728.1975.10410201.

Rambotti, S. (2015) Recalibrating the spirit level: An analysis of the interaction of income inequality and poverty and its effect on health. *Social Science & Medicine*, Vol. 139, pp. 123–131, DOI: https://doi.org/10.1016/j.socscimed.2015.02.026.

Regidor, E. – M. E. Calle – P. Navarro – V. Domínguez (2003) Trends in the association between average income, poverty and income inequality and life expectancy in Spain. *Social Science & Medicine*, Vol. 56, No. 5., pp. 961-971, DOI: https://doi.org/10.1016/S0277-9536(02)00107-7.

Rodgers, G. B. (2002) Income and inequality as determinants of mortality: An international cross-sectional analysis. *International Journal of Epidemiology*, Vol. 31, No. 3., pp. 533–538, DOI: https://doi.org/10.1093/ije/31.3.533.

Rosstat (2025) *Regions of Russia. Socio-Economic Indicators of 2025.* Moscow, Russia, Rosstat, Federal State Statistics Service, https://rosstat.gov.ru/folder/210/document/13204.

Ryan, B. E. (2016) Mining away the Preston curve. *World Development*, Vol. 78, pp. 22–36, DOI: https://doi.org/10.1016/j.worlddev.2015.10.013.

Siordia C., – J. Saenz – S. E. Tom (2012) An introduction to macro- level spatial nonstationarity: A Geographically weighted regression analysis of diabetes and poverty. *Human Geography*, Vol. 6, No. 2., pp. 5–13.

Skipin, D. L. – Y. A. Yukhtanova – O. A. Kryzhanovskii – E.G. Tokmakova (2022) Life expectancy in Russia's regions. *Economic and Social Changes: Facts, Trends, Forecast*, Vol.15, No. 2., pp.156–171, DOI: https://doi.org/10.15838/esc.2022.2.80.10.

Sofi, S.A. – E. Yasmin (2025) Determinants of life expectancy: Evidence from World Bank income groups using a panel dummy interaction approach. *International Journal of Health Care Quality Assurance*, Vol. 38, No. 3., pp. 181–197, DOI: https://doi.org/10.1108/IJHCQA-03-2025-0026.

Spence, M. – M. Lewis (eds.) (2009) Health and Growth. Washington D.C., Commission on Growth and Development and The World Bank Group, https://openknowledge.worldbank.org/bitstreams/3ac886d3-8072-5a12-abdc-809d6c06f629/download.

Tobler, W. R. (1970) A computer movie simulating urban growth in the Detroit region. *Economic Geography*, Vol. 46, pp. 234–240, DOI: https://doi.org/10.2307/143141.

Vladimirskaya, A. A. – M.G. Kolosnitsyna (2023) Factors in life expectancy: A cross-country analysis. *Voprosy Statistiki*, Vol. 30, No. 1., pp. 70–89, (in Russian) DOI: https://doi.org/10.34023/2313-6383-2023-30-1-70-89.

Wang S. – Y. Liu – C. Zhao – H. Pu (2019) Residential energy consumption and its linkages with life expectancy in mainland China: A geographically weighted regression approach and energy-ladder-based perspective. *Energy*, Vol. 177, pp. 347–357, DOI: https://doi.org/10.1016/j.energy.2019.04.099.

Wang, S. – Z. Ren – X. Liu – Q. Yin (2022) Spatiotemporal trends in life expectancy and impacts of economic growth and air pollution in 134 countries: A Bayesian modeling study. *Social Science & Medicine*, Vol. 293, Art. No. 114660, DOI: https://doi.org/10.1016/j.socscimed.2021.114660.

Wilkinson, R. G. – K. E. Pickett (2006) Income inequality and population health: A review and explanation of the evidence. *Social Science & Medicine*, Vol. 62, No.7., pp. 1768–1784, DOI: https://doi.org/10.1016/j.socscimed.2005.08.036.

Xiong, N. – Y.D. Wei (2025) Economic inequality, intergenerational mobility, and life expectancy. *Social Science & Medicine*, Vol. 366, p. 117682, DOI: https://doi.org/10.1016/j.socscimed.2025.117682.

You W. – R. Henneberg – A. Saniotis – Y. Ge – M. Henneberg (2022) Total meat intake is associated with life expectancy: A cross-sectional data analysis of 175 contemporary populations. *International Journal of General Medicine*, Vol. 15, pp. 1833–1851, DOI: https://doi.org/10.2147/ijgm.s333004.

Zhou, B. – Y. Qi (2012) Income inequality and individual health: An empirical analysis of the 2005 Chinese General Social Survey. *Chinese Journal of Sociology/Shehui*, Vol. 32, No. 5, pp. 130–150.

Zhukova A.K. – A. M. Silaev – M.V. Silaeva (2016) Analysis of life expectancy taking into account spatial dependence by regions of Russia. *Prostranstvennaya Ekonomika* [*Spatial Economics*], No. 4, pp. 112–128, DOI: http://dx.doi.org/10.14530/se.2016.4.112-128.