International Journal of Mathematics and Statistics

ISSN (Online): 2208-2212
Volume 10 Issue 01 March 2024

DOI: https://doi.org/10.69980/f65ym554

STATISTICAL MODELING OF LIFE EXPECTANCY USING SOCIOECONOMIC, MORTALITY, AND HEALTH INDICATORS ACROSS COUNTRIES

Dr. Eleanor M. Whitfield[1*], Prof. Carlos R. Mendoza[2], Dr. Yuki A. Nakamori[3], Dr. Lukas F. Schneider[4]

[1]Department of Population Statistics and Global Health Analytics, Westbridge University, London, United Kingdom*
[2]Institute of Health Economics and Demographic Research, Universidad Central del Pacífico, Madrid, Spain
[3]School of Epidemiology and Quantitative Health Sciences, East Asia Institute of Public Health, Tokyo, Japan
[4]Department of Biostatistics and Population Studies, Rhine Valley University, Frankfurt, Germany

Abstract:

Life expectancy is the result of country differences in socioeconomic conditions, mortality and population health features. Publicly available country-level records were used for a quantitative observational analysis between 2000 and 2015. The association between life expectancy at birth and schooling, income composition of resources, gross domestic product, health expenditure, adult mortality, infant deaths, under-five deaths, HIV/AIDS, body mass index, alcohol consumption, and indicators of immunization were explored. For descriptive analysis, group comparisons, Pearson correlation analysis and multiple linear regression, 1,649 complete country-year observations were analyzed. The mean life expectancy for developed countries' observations was significantly higher than the mean life expectancy for developing country observations. Life expectancy had the largest positive socioeconomic association for schooling and income composition of resources; strong negative associations were found for adult mortality and HIV/AIDS. Longevity was positively correlated with immunization indicators and BMI, and negatively correlated with infant and under-five mortality. The regression model accounted for about 75% of the variation in life expectancy, and found that income composition, schooling, adult mortality, alcohol consumption, and infant deaths were significant independent predictors. Results suggest that these factors, the social development, education, mortality reduction, and health conditions are interrelated and influence longevity. Policies designed to enhance population survival should thus include an educational, an economic, a preventive health-care, and a mortality-reduction component, taking into account that the relationships observed are statistical associations but not necessarily causal effects.

Keywords: *life expectancy; socioeconomic determinants; mortality; health indicators; multiple linear regression*

1. Introduction:

One of the most common measures used to assess the health of a population, the conditions of survival, and social development generally are the people's life expectancy at birth. It serves as a succinct and comprehensive indicator of the total impact of death risks at various points in the lifespan and is useful for cross-national comparisons of health outcomes. Worldwide, there has been significant differences in longevity between regions and over time, suggesting that not all populations have benefited equally from improved survival [1]. The patterns of fertility and mortality growth and level and healthy life expectancy growth have also proven to reflect differences in national health conditions and growth pathways well [2].

There are socioeconomic, mortality and health factors that contribute to the difference in longevity. The number of years a person may live can be affected by their adult mortality, disease burden, living conditions, health behaviours and access to preventive services. The literature from other countries has demonstrated that inequitable access to key risk factors plays an important role in inequalities in life expectancy between different population groups and countries [3]. Additionally, socioeconomic development is also a central theme as enhancing income, education and living standards can increase access to healthcare, nutrition, sanitation, and other factors that contribute to longevity [4].

But differences in economic capacity only account for part of differences in longevity across countries. There are variations in life expectancy amongst countries with similar levels of income due to differences in social policy, healthcare organization, inequality, and resource distribution between countries [5]. The findings also from the ecological analyses have shown that there are multiple determinants operating simultaneously that should be analyzed in the context of life expectancy, but not a single economic or health determinant alone [6]. This reinforces the need for analytical tools that can analyze several explanatory indicators in one statistical framework.

Another key factor of longevity is the performance of health systems. International studies have associated higher quality health care utilization with longer life span, and suggested that broader social and economic factors play a role [7]. Educational attainment has a special significance because it can influence employment opportunities, health literacy, preventive behaviour, and access to health services, and unequal distribution of educational attainment can be a root cause of continued inequalities in life expectancy [8].

There is also more recent longitudinal evidence that suggests that economic conditions, healthcare resources, social characteristics, and population health indicators are all associated with life expectancy [9]. Other studies have revealed similar results: that demographical and socioeconomic factors can influence life expectancy as well as the various indexes of "good health" [10]. Other factors, such as education and income inequality, have also been important in understanding the variation between European populations in longevity [11]. Furthermore, investing in health, education and social spending over the long-term has been linked to better life-expectancy measures, demonstrating the linkages among these factors [12].

While much has been learned about individual determinants, comparative analyses, which incorporate socioeconomic conditions, mortality indicators, and health indicators in a unified statistical evaluation are still important. Thus, the cross country variation in life expectancy was analyzed and its association with the level of schooling, the composition of income, gross domestic product, mortality indicators, immunization coverage, HIV/AIDS, body mass index, alcohol consumption, and other indicators examined. Multiple linear regression was then undertaken to determine the factors that remained statistically independently associated with life expectancy.

Research Objectives

1.To describe and compare the distribution and variability of life expectancy, socioeconomic conditions, mortality measures, and health indicators across countries.

2.To examine the statistical relationships between life expectancy and selected socioeconomic, mortality, and health indicators.

3.To develop a multiple linear regression model for identifying the major statistical predictors of life expectancy across countries.

2. Materials and Methods

2.1 Research Design and Data Source

A quantitative observational design was used to analyze the observed differences between countries in their life expectancy and the relationship to variables of socio-economic, mortality and health indicators. The country-level indicators were taken from publicly available records from the WHO and other international sources for the period 2000 to 2015. One analytical observation was equal to one country-year. There was no primary survey, experiment or intervention undertaken. The design enabled cross-country statistical comparisons and facilitated the analysis of the selected indicators within a uniform analytical framework over countries and over the years.

2.2 Study Variables and Indicator Classification

The dependent variable was life expectancy at birth in years. Education (schooling), income composition of resources, gross domestic product, and percentage expenditure were included in the socioeconomic indicators. Adult mortality, infant deaths and under-fives deaths were included among the mortality measures. Health related indicators included HIV/AIDS, Hepatitis B immunization, Polio immunization, Diphtheria immunization, BMI, and Alcohol consumption. Variables were kept based on their conceptual relevance to longevity and the number of observations for which these variables were available for use in subsequent statistical procedures and analysis.

2.3 Data Preparation and Descriptive Statistical Analysis

Records were checked for missing, incomplete and implausible values before analysis was carried out. There was no imputation for missing observations. A descriptive summary was done only with the available values, while 1,649 complete country-year observations were used for the correlation and regression. Descriptive statistics were presented as the mean, the median, and the standard deviation (where applicable), minimum, maximum, and range. Categorical characteristics like development status were given frequencies and percentages. The main socioeconomic, mortality, and health indicators in the analysis were summarized in these procedures and described in terms of their distribution and variation.

2.4 Bivariate Statistical Analysis

Bivariate analysis was conducted to explore both the direction and intensity of the correlations between the selected continuous indicators and life expectancy. Socio-economic, mortality, and health variables were correlated using Pearson correlation coefficients. An independent-samples t-test was used to compare mean life expectancy of the developed vs. developing country groups. Both the magnitude and direction of each association was considered along with the statistical significance. Values of $p < 0.05$ were taken as statistically significant. These analyses were used to identify variables that had meaningful relationships with longevity without adjustment.

2.5 Multiple Linear Regression Analysis

Multiple linear regression was used to determine variables that were independently associated with life expectancy with simultaneous adjustment. The final specification contained variables from four domains (socioeconomic, mortality, behavioural) and did not include variables that are too similar. The regression coefficients, standard errors, t-statistics, probability values, R-squared and adjusted R-squared were reported along with the overall F-statistic. A p value of less than 0.05 was determined as statistical significance. Standard spreadsheet functions and the Microsoft Excel Data Analysis ToolPak were used for all calculations.

3. Results

3.1 Characteristics of the Student Sample

Table 1 presents the descriptive profile of the study indicators. Across 2,928 non-missing life-expectancy observations, mean life expectancy was 69.22 years (SD = 9.52), with a median of 72.10 and a range of 36.30–89.00 years. Mean adult mortality was 164.80 per 1,000, while infant deaths averaged 30.30 per 1,000. The Mean GDP (US$7,483.16), schooling (11.99 years) and income composition (0.628) were reported. The coverage for Hepatitis B, polio and diphtheria was 80.94%, 82.55% and 82.32% respectively. The distribution of life expectancy is summarized in figure 1.

Table 1. Descriptive Statistics of Selected Socioeconomic, Mortality, and Health Indicators

Indicator	Valid n	Mean	SD
Life expectancy (years)	2,928	69.22	9.52
Adult mortality	2,928	164.80	124.29
Infant deaths	2,938	30.30	117.93
GDP (US$)	2,490	7,483.16	14,270.17
Schooling (years)	2,775	11.99	3.36
Income composition of resources	2,771	0.628	0.211
Hepatitis B coverage (%)	2,385	80.94	25.07
Polio coverage (%)	2,919	82.55	23.43
Diphtheria coverage (%)	2,919	82.32	23.72

Figure 1. Distribution of life expectancy across country-year observations

3.2 Cross-Country and Development-Status Differences

Longevity was significantly different between the development status groups. The mean life expectancy was 79.2 years (SD = 3.93) for developed country observations and 67.1 years (SD = 9.01) for developing country observations, which is about a 12.1 year difference ($p < 0.001$). The developed group thus had not only a longer average life span, but also significantly less dispersion, as shown in Table 2 and Figure 2. For the complete-case analytical sample 242 country-year observations were found for developed countries and 1,407 country-year observations were found for developing countries, with both groups well represented in subsequent analysis.

Table 2. Life Expectancy According to Development Status

Development status	Mean life expectancy (years)	SD	Mean difference	p-value
Developed	79.2	3.93		
Developing	67.1	9.01	12.1	<0.001

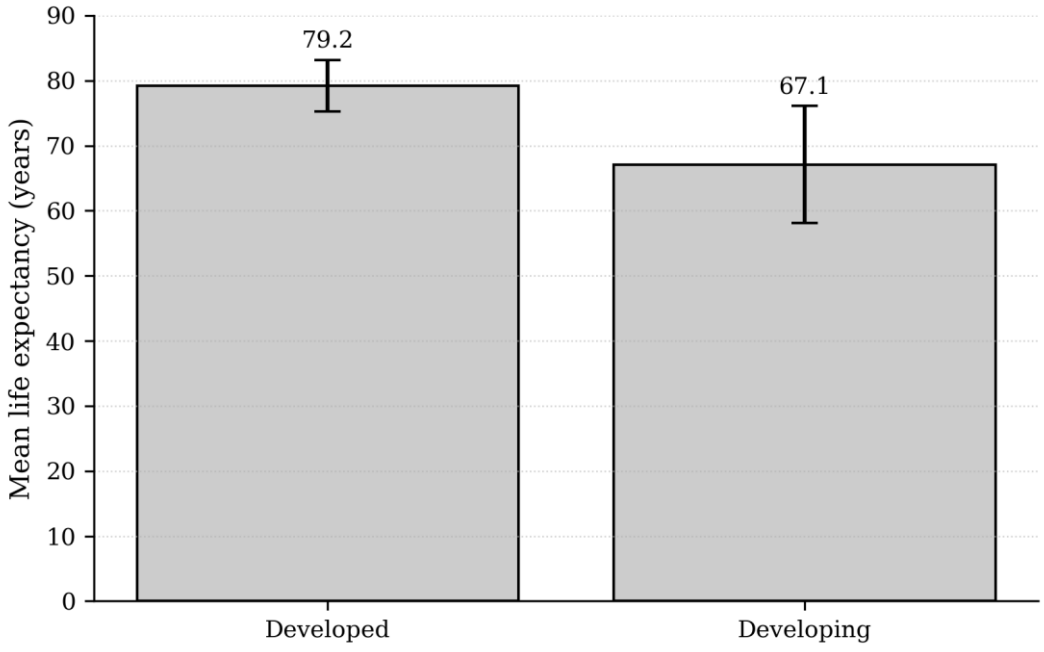

Figure 2. Mean life expectancy by development status (error bars represent ±1 SD)

3.3 Correlation of Life Expectancy with Socioeconomic Indicators

In the 1,649 full values, the life expectancy was favorably associated with the socioeconomic indicators. The strongest correlation ($r = 0.728$, $p < 0.001$) was with schooling followed by income composition of resources ($r = 0.721$, $p < 0.001$). Life expectancy was moderately correlated with GDP ($r = 0.441$, $p < 0.001$) and also percentage expenditure revealed positive correlation ($r = 0.410$, $p < 0.001$). The bivariate associations with longevity were significantly stronger for education and income composition than for the monetary indicators (Table 3). The pattern was consistent and pronounced.

Table 3. Pearson Correlations Between Life Expectancy and Socioeconomic Indicators

Socioeconomic indicator	Pearson r	p-value	Direction
Schooling	0.728	<0.001	Positive
Income composition of resources	0.721	<0.001	Positive
GDP	0.441	<0.001	Positive
Percentage expenditure	0.410	<0.001	Positive

3.4 Correlation of Life Expectancy with Mortality and Health Indicators

Mortality and health indicators showed contrasting associations. Adult mortality had the strongest negative correlation with life expectancy ($r = -0.703$, $p < 0.001$), followed by HIV/AIDS ($r = -0.592$, $p < 0.001$). There was a positive correlation between BMI ($r = 0.542$) and alcohol consumption ($r = 0.404$). Diphtheria ($r = 0.341$), polio ($r = 0.327$), and hepatitis B coverage ($r = 0.200$) were positive, whereas under-five deaths ($r = -0.192$) and infant deaths ($r = -0.169$) were negative. The results for all the associations were significant at $p<0.001$ as detailed in Figure 3.

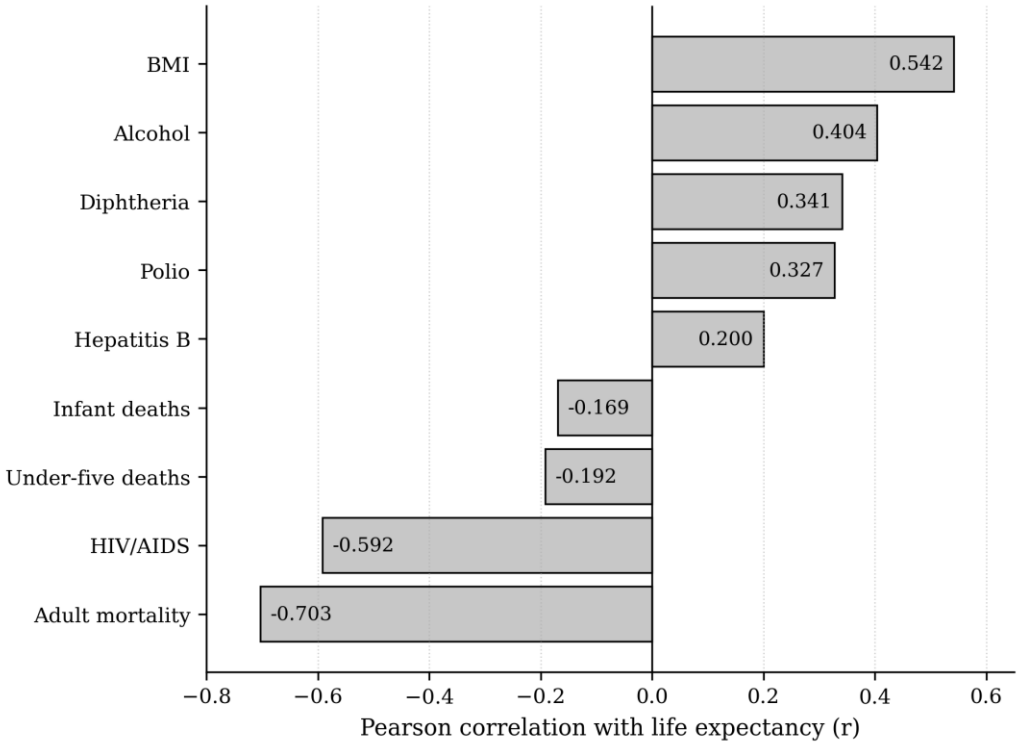

Figure 3. Correlations of major mortality and health indicators with life expectancy

3.5 Multivariate Regression Results

The multiple linear regression model was significant, $F(5, 1643) = 978.1$, $p < 0.001$, explaining 74.85% of variation in life expectancy ($R^2 = 0.7485$; adjusted $R^2 = 0.7478$). Income composition ($B = 13.233$), schooling ($B = 1.085$), adult mortality ($B = -0.0310$) and alcohol consumption ($B = -0.102$) were positive predictors, while infant deaths ($B = -0.00322$) was a negative predictor (Table 4). All predictors were found to be significant. The fitted predictors with the largest absolute t-statistic ($t = -31.33$) were adult mortality. The direction and contribution of the fitted predictors is summarised in Figure 4.

Table 4. Multiple Linear Regression Model for Life Expectancy

Predictor	B	Standard error	t-value	p-value
Constant	53.5805	0.7010	76.435	<0.001
Income composition of resources	13.2329	0.9967	13.277	<0.001
Schooling	1.0850	0.0691	15.694	<0.001
Adult mortality	−0.0310	0.0010	−31.329	<0.001
Alcohol consumption	−0.1022	0.0352	−2.903	0.0037
Infant deaths	−0.00322	0.00092	−3.484	0.0005

Model statistics: $R^2 = 0.7485$; Adjusted $R^2 = 0.7478$; $F(5, 1643) = 978.1$; $p < 0.001$.

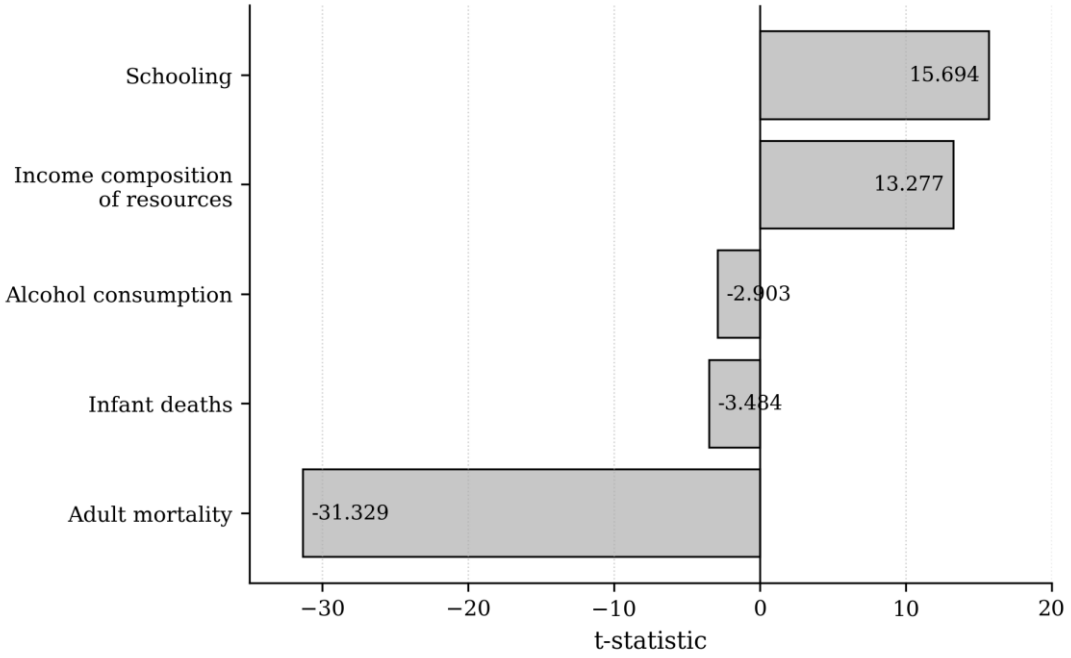

Figure 4. t-statistics for predictors in the multiple linear regression model

4. Discussion

The results revealed significant inter-country differences in life expectancy, and demonstrated that life expectancy was consistently related with socioeconomic position, mortality burden and certain health indicators. Observations from developed countries had a mean life expectancy of nearly 12 years more than those from developing countries, reflecting an enduring disparity in survival outcomes between developed and developing countries. These differences are similar to the results of studies indicating a significant contribution of socioeconomic position, such as level of education, to differences in life expectancy [13]. The same educational disparities in life expectancy have been described at the population level, which further emphasizes the long-term importance of education as a health advantage [14].

Schooling and income composition of resources had the highest positive score with life expectancy among the socioeconomic indicators. Both retained in the multivariable model where both were positive significant predictors. The education may affect the lifespan by affecting the health literacy, employment prospects, preventive behaviour and social resources. Previous longitudinal studies have also found education and economic factors to be key determinants of life expectancy and healthy life expectancy [15]. There was also a positive bivariate association, but less strong than that of schooling and income composition, between national economic capacity and improved survival, as measured by GDP. Similar data have correlated increased economic growth with increased life expectancy, and have noted the effects of broader environmental and social factors [16].

The percentage expenditure was found to be positively correlated with life expectancy, suggesting the role of healthcare spending on population health. An increase in spending on the public sector can increase the availability of health services, preventive care, treatment facilities, and access to necessary interventions, but the effectiveness of increases depends on the efficiency of each allocation of resources [17]. The relationship between income and health expenditure, too, can be helpful in understanding the relationship between country-level results in life expectancy and the level of economic resources available to the country [18].

Theoretically similar patterns were observed for health-related indicators. A positive correlation was observed between life expectancy and immunization coverage for diphtheria, polio and hepatitis B. High sustained vaccination coverage both decreases preventable childhood morbidity and mortality and shows the progress made on basic health services delivery, and has been shown in the past to be correlated with socioeconomic and health-system conditions internationally [19]. Adult mortality had the highest correlation with longevity and was the most statistically significant negative predictor in the regression model, in contrast. The types of deaths that significantly drive inequalities in life expectancy have been shown elsewhere in relation to socio-economic status [20].

The increase of life expectancy with income composition and socioeconomic development is also compatible with findings among Eastern European populations, where there was a strong covariance of socioeconomic conditions and survival outcomes [21]. Similarly, recent studies in South Asia have demonstrated that spending on health care, income and education are all associated with differences in life expectancy [22]. This difference in development status may thus be due to a combination of disadvantages beyond healthcare. A range of impacts of socioeconomic disadvantage, including reduced life expectancy and increased premature deaths, have been linked to multidimensional deprivation [23].

BMI was positively associated with life expectancy within the range observed, while HIV/AIDS was negatively associated with it. After the adjustments, the negative predictor remained the infant death rate. These findings indicate that mortality and disease burden continued to influence longevity even when socioeconomic variables were considered simultaneously. There have also been recent indications that socioeconomic inequalities and cause-specific mortality patterns are intertwined in determining life-expectancy inequalities [24]. The significant disparities in longevity across regions in low and middle-income countries also emphasize the role of structural and socio-economic factors in understanding inequalities in longevity [25].

The overall explanatory power of the regression model was relatively high, using about 75% of the variation in life expectancy. The associations should not be considered as causal relationships since the analysis was based on observational country-year data. Other institutional, environmental, behavioural and healthcare factors may have contributed to the variability not explained by the measured factors.

5. Conclusion

Life expectancy was significantly different by country and was correlated with socioeconomic development, mortality burden and selected health measures. The observations from developed countries revealed significantly greater longevity than those from developing countries, and the best positive socioeconomic correlates were schooling and income composition of resources. HIV/AIDS, adult mortality, and immunization coverage and BMI were among the largest negative health-related associations, while BMI and immunization coverage were positively correlated with longevity. A large amount of variation in the observed data was accounted for by the multiple linear regression model and income composition, schooling, adult mortality, alcohol consumption and infant deaths were identified as significant separate predictors. The results suggest that increase in longevity is associated not only with capacity for economic development, but also with education, lowering of death rates, better preventive health care and wider social development. The findings thus favor an integrated policy approach focusing on investments in education, access to health care, immunization, disease control, and reducing mortality. The relationships found in this analysis should be viewed as statistical associations, due to the observational nature of the country-year information used. The model could be further developed in future to include environmental, institutional and behavioural indicators to account for further cross-country variation in life expectancy.

References

1.GBD 2017 Mortality Collaborators. Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392(10159):1684–1735. doi:10.1016/S0140-6736(18)31891-9.

2.GBD 2019 Demographics Collaborators. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1160–1203. doi:10.1016/S0140-6736(20)30977-6.

3.Mackenbach JP, Rubio Valverde J, Bopp M, Brønnum-Hansen H, Deboosere P, Kalediene R, et al. Determinants of inequalities in life expectancy: an international comparative study of eight risk factors. Lancet Public Health. 2019;4(10):e529–e537. doi:10.1016/S2468-2667(19)30147-1.

4.Miladinov G. Socioeconomic development and life expectancy relationship: evidence from the EU accession candidate countries. Genus. 2020;76:2. doi:10.1186/s41118-019-0071-0.

5.Freeman T, Gesesew HA, Bambra C, Giugliani ERJ, Popay J, Sanders D, et al. Why do some countries do better or worse in life expectancy relative to income? An analysis of Brazil, Ethiopia, and the United States of America. Int J Equity Health. 2020;19(1):202. doi:10.1186/s12939-020-01315-z.

6.Paramita SA, Yamazaki C, Koyama H. Determinants of life expectancy and clustering of provinces to improve life expectancy: an ecological study in Indonesia. BMC Public Health. 2020;20:351. doi:10.1186/s12889-020-8408-3.

7.Zarulli V, Sopina E, Toffolutti V, Lenart A. Health care system efficiency and life expectancy: a 140-country study. PLoS One. 2021;16(7):e0253450. doi:10.1371/journal.pone.0253450.
8.Danler C, Pfaff K. The impact of an unequal distribution of education on inequalities in life expectancy. SSM Popul Health. 2021;16:100954. doi:10.1016/j.ssmph.2021.100954.
9.Roffia P, Bucciol A, Hashlamoun S. Determinants of life expectancy at birth: a longitudinal study on OECD countries. Int J Health Econ Manag. 2023;23:189–212. doi:10.1007/s10754-022-09338-5.
10.Varbanova V, Hens N, Beutels P. Determinants of life-expectancy and disability-adjusted life years (DALYs) in European and Organisation for Economic Co-operation and Development (OECD) countries: a longitudinal analysis (1990–2019). SSM Popul Health. 2023;24:101484. doi:10.1016/j.ssmph.2023.101484.
11.Sart G, Bayar Y, Danilina M. Impact of education and income inequalities on life expectancy: insights from the new EU members. Front Public Health. 2024;12:1397585. doi:10.3389/fpubh.2024.1397585.
12.Aytemiz L, Sart G, Bayar Y, Danilina M, Sezgin FH. The long-term effect of social, educational, and health expenditures on indicators of life expectancy: an empirical analysis for the OECD countries. Front Public Health. 2024;12:1497794. doi:10.3389/fpubh.2024.1497794.
13.Ingleby FC, Woods LM, Atherton IM, Baker M, Elliss Brookes L, Belot A. Describing socio-economic variation in life expectancy according to an individual's education, occupation and wage in England and Wales: an analysis of the ONS Longitudinal Study. SSM Popul Health. 2021;14:100815. doi:10.1016/j.ssmph.2021.100815.
14.Welsh J, Bishop K, Booth H, Butler D, Gourley M, Law HD, et al. Inequalities in life expectancy in Australia according to education level: a whole-of-population record linkage study. Int J Equity Health. 2021;20:178. doi:10.1186/s12939-021-01513-3.
15.Aanegola R, Nakamura Sakai S, Kumar N. Longitudinal analysis of the determinants of life expectancy and healthy life expectancy: a causal approach. Healthc Anal. 2022;2:100028. doi:10.1016/j.health.2022.100028.
16.Rahman MM, Rana R, Khanam R. Determinants of life expectancy in most polluted countries: exploring the effect of environmental degradation. PLoS One. 2022;17(1):e0262802. doi:10.1371/journal.pone.0262802.
17.Anwar A, Hyder S, Mohamed Nor N, Younis M. Government health expenditures and health outcome nexus: a study on OECD countries. Front Public Health. 2023;11:1123759. doi:10.3389/fpubh.2023.1123759.
18.Ahmad N, Raid M, Alzyadat J, Alhawal H. Impact of urbanization and income inequality on life expectancy of male and female in South Asian countries: a moderating role of health expenditures. Humanit Soc Sci Commun. 2023;10:552. doi:10.1057/s41599-023-02005-1.
19.Lanza-León P, Cantarero-Prieto D, Pascual-Sáez M. Exploring trends and determinants of basic childhood vaccination coverage: empirical evidence over 41 years. PLoS One. 2024;19(3):e0300404. doi:10.1371/journal.pone.0300404.
20.Tetzlaff F, Sauerberg M, Grigoriev P, Tetzlaff J, Mühlichen M, Baumert J, et al. Age-specific and cause-specific mortality contributions to the socioeconomic gap in life expectancy in Germany, 2003–21: an ecological study. Lancet Public Health. 2024;9(5):e295–e305. doi:10.1016/S2468-2667(24)00049-5.
21.Țarcă V, Țarcă E, Moscalu M. Social and economic determinants of life expectancy at birth in Eastern Europe. Healthcare (Basel). 2024;12(11):1148. doi:10.3390/healthcare12111148.
22.Dhungana BR, Singh JK, Dhungana S. Life expectancy and health care spending in South Asia: an econometric analysis. PLoS One. 2024;19(12):e0310153. doi:10.1371/journal.pone.0310153.
23.Das J, Mohanty SK. Estimates of life expectancy and premature mortality among multidimensional poor and non-poor in India. BMC Public Health. 2024;24:3546. doi:10.1186/s12889-024-20966-z.
24.Timonin S, Adair T, Welsh J, Canudas-Romo V. Socioeconomic inequalities in life expectancy in Australia, 2013–22: an ecological study of trends and contributions of causes of death. Lancet Public Health. 2025;10(7):e599–e608. doi:10.1016/S2468-2667(25)00138-0.
25.Kyriacou AP, Miranda-Lescano R, Muinelo-Gallo L, Roca-Sagales O. Subnational life expectancy disparities in low- and middle-income countries: measurement and determinants. Health Econ Rev. 2025;15:96. doi:10.1186/s13561-025-00676-x.