DOI: https://doi.org/10.31410/ERAZ.2023.291

Factors Associated with Self-Perceived Health Status in the Slovak Republic

Viera Labudová[1]

Received: August 16, 2023
Accepted: January 10, 2024
Published: March 16, 2024

Keywords:
Self-perceived health;
EU-SILC;
Logistic regression

Abstract: *Self-perceived health (SPH) expresses a subjective assessment by the respondent of his/her health. When assessing his/her health, the respondent states one of the following categories: 1: very good, 2: good, 3: fair, 4: bad, 5: very bad. This paper aimed to analyze the influence of individual characteristics of persons aged 16 and over in the Slovak Republic on the subjective perception of health and quantify the intensity of their impact. For this purpose, a logistic regression model was used to predict "good health" (categories 'very good' or 'good'). As explanatory variables for the logistic regression model have included these demographic or socio-economic indicators: marital status, age, the highest level of education attained based on the ISCED, main activity status, severe material deprivation, equalized household disposable income, degree of urbanization, and NUTS 3 region. The analysis was performed using the SAS Enterprise Guide on a sample of data from the EU SILC 2021 survey.*

1. INTRODUCTION

Health is an important attribute of the quality of life and general well-being of an individual. It has a functional value, and it is also important for the person's own identity (Blaxter, 2004). "Health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity" (World Health Organization, 1948).

The social conditions of people's lives strongly affect their chances for a healthy life. Poverty, social exclusion, discrimination, bad or unhealthy living conditions, and low employment status are some of the important determinants of most diseases, deaths, and health inequalities between and within countries (World Health Organization, 2018).

The existence of health inequalities leads to the belief that health is socially determined by factors that originate on different social levels (individual, structural). The efforts to conceptualize the effects of these determinants resulted in models of social health determinants. One of the most well-known is the Dahlgren-Whitehead model (Whitehead & Dahlgren, 2007). The concept of social determinants of health according to this model is shown in Figure 1. At its center are people, as part of a broader social system. They possess characteristics such as age, gender, and other personality traits, such as genetic predispositions, which can be considered constants at each moment. However, they are surrounded by variable factors. The first layer contains personal behavior factors, such as smoking, alcohol, and other addictive substances, as well as the physical activity of the person. Another layer of the concept of social health determinants consists of those characteristics that are related to the interaction of people with their peers or with the community in which they live. At a certain period of the person's life cycle, they start to have a major effect on the environment and working conditions (food supply, access to basic goods and services, access

[1] Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemská cesta 1, 852 35 Bratislava, Slovak Republic

to education, employment, living conditions, etc.). These are included in the third layer. The final layer consists of economic, cultural, and environmental factors. This model of conceptualization of the health determinants emphasizes interactions between the individual lifestyle, which is incorporated into social standards, and the living and working conditions, which are related to the broader socioeconomic and cultural environment (Whitehead & Dahlgren, 2007).

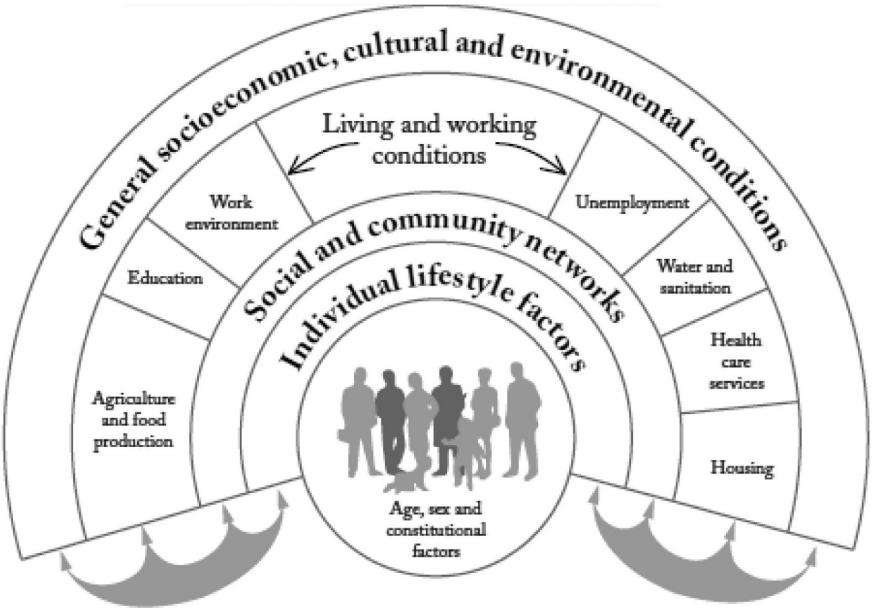

Figure 1. The Main Determinants of Health
Source: Whitehead & Dahlgren, 2007

Self-perceived health (SPH) (Self-rated health (SRH) or self-assessed health (SAH)) is a measure, in which people provide an assessment of their own general health. Self-perceived health can be defined as a general assessment of health done through subjective self-assessment. This indicator is considered an integrated indicator of health, which is related to the social, psychological, and biological aspects of the individual. The World Health Organization recommends it as a strong indicator of life expectancy and the health of the population (Silva et al., 2017). In the EU SILC and EHIS surveys, the Minimum European Health Module (MEHM) consists of 3 questions: subjective health assessment (SPH), the presence of a chronic disease or long-term health problem, and limitations in carrying out normal daily activities due to chronic diseases or long-term health problems.

The aim of this paper is to identify the factors that determine the subjective perception of health (Self-perceived health – SPH).

2. DATA AND METHODOLOGY

The analysis, the results of which are presented in this article, was based on data from the EU SILC survey (European Commission, 2021). We used the cross-sectional component of the EU SILC 2021 survey. Individual data were provided by the Statistical Office of the Slovak Republic. The dataset is composed of 11,755 records of respondents aged 16 and over.

A logistic regression model to find out the factors that influence self-perceived health is used. The modeled dependent variable was *self-perceived health*. The variable self-perceived health

was created from the variable general health. The variable general health has five categories: 1 = very good; 2 = good; 3 = fair; 4 = bad, and 5 = very bad. The values (categories) of this variable are the result of measurement on a 5-point scale. Respondents aged 16 and over answered the question: "*How do you assess your health condition overall*?" (European Commission, 2021). For the binary logistic regression, the categories 1 = very good, and 2 = good were merged, while a new category "good health" (value 1) was created. The categories 3 = fair, 4 = bad, and 5 = very bad were also merged, and a new category "poor health" (value 0) was created from them. This method of categorization states e. g. Manderbacka et al. (1999). A binary logistic regression model was used to predict "good health" (value 1).

The independent variables in the model were these demographic and socio-economic indicators:

- *age* (year of the survey minus year of birth minus 1). Age was grouped into four categories: 1 = 16–24 years; 2 = 25–44 years; 3 = 45–64 years; 4 = above 65 years,
- *degree of urbanization* with three categories: 1 = cities (densely populated area); 2 = towns and suburbs (intermediate area); 3 = rural areas (thinly populated area),
- *educational attainment level* with three categories: 1 = low (less than primary, primary and lower secondary education); 2 = medium (upper secondary and post-secondary non-tertiary education); 3 = high (tertiary education),
- *equalized household disposable income* with five categories: 1 = the first quintile; 2 = the second quintile; 3 = the third quintile; 4 = the fourth quintile; 5 = the fifth quintile,
- *main activity status* with four categories: 1 = employed; 2 = unemployed; 3 = student, pupil, fulfilling domestic tasks, compulsory military, or civilian service and other; 4 = retired or unable to work due to long-standing health problems),
- *marital status* with four categories: 1 = never married; 2 = married; 3 = separated, widowed, or divorced),
- *severe material deprivation* (at least 4 out of 9 deprivation items[2]) with two categories: 1 = household is deprived; 2 = household is not deprived),
- *NUTS 3* with eight categories[3]: 1 = Bratislava Region; 2 = Trnava Region; 3 = Trenčín Region; 4 = Nitra Region; 5 = Žilina Region; 6 = Banská Bystrica Region; 7 = Košice Region; 8 = Prešov Region (European Commission, 2021).

Logistic regression extends the idea of linear regression to the situation where the dependent variable, Y, is binary (Agresti & Finlay, 2014). The predictors can be quantitative, qualitative, or both types (Agresti, 2019). Letting Y be the binary dependent variable, which can take only the values 0 or 1 and $X_1, X_2, \ldots, X_k$ is a set of k predictors (quantitative and dummy variables). Level 1 usually represents the occurrence of an event of interest. It is assumed that $P(Y = 1)$ is possibly dependent on $\mathbf{x} = (x_0, x_1, \ldots, x_k)$, a vector of k predictor's values. The goal is to model $p(\mathbf{x}) = P(Y = 1 \mid \mathbf{x})$. Modelling $p(\mathbf{x})$ is really modelling $E(Y) = E(Y \mid \mathbf{x}) = \pi$. The logistic regression model has a linear form for the logit of this probability (Agresti, 2019):

$$ln\ \frac{\pi}{1-\pi} = \beta_0 + \sum_{j=1}^{k} \beta_j x_j \tag{1}$$

An alternative formula for logistic regression refers directly to the probability:

$$\pi = \frac{exp\ \left(\beta_0 + \sum_{j=1}^{k} \beta_j x_j\right)}{1 + exp\ \left(\beta_0 + \sum_{j=1}^{k} \beta_j x_j\right)} \tag{2}$$

[2] A severe material deprivation is defined as the enforced lack of at least four out of nine material deprivation items in the economic strain and durables dimension (Statistical Office of the Slovak Republic, 2020).

[3] Variables refer to the region of the residence of the household at the date of interview.

When interpreting the estimated parameters of the model, logistic regression uses the odds ratio, abbr. *OR*:

$$OR_{x_1,x_0} = \frac{chance\ for\ the\ group\ x_{j1}}{chance\ for\ the\ group\ x_{j0}} \quad (3)$$

where x_{j1} represents one group of cases that differs from the other group x_{j0} only in the values of the variable X_j. (Agresti, 2019; Hosmer et al., 2013). The values of the other variables are the same for group x_{j1} as for group x_{j0}. The odds ratio is a measure that quantifies the effect of a presumed causal factor on the occurrence of one of two possible situations (two possible outcomes).

3. RESEARCH RESULTS

The results of the logistic regression model are presented in Tables 1–3. To assess the quality of the model, we used three statistics: AIC – Akaike's Information Criterion, SC – Schwarz Criterion and –2 log L – the logarithmic transformation of the likelihood function L (Table 1). Based on these measures, the estimated logistic regression model with predictors (column Intercept and Covariates) is better than a model that has only an intercept (column Intercept only). The values of these measures (AIC, SC, –2 log L) are lower for the estimated model with predictors (Allison, 2012).

Table 1. Fit Statistics of Model

Criterion	Intercept Only	Intercept and Covariates
AIC	16014.385	11125.679
SC	16021.757	11317.337
–2 Log L	16012.385	11073.679

Source: Own processing in SAS Enterprise Guide

Table 2 shows the results of testing the statistical significance of the model. The null hypothesis states that all parameters of the model are equal to zero (none of the input variables has a statistically significant effect on self-perceived health). The alternative hypothesis states that at least one parameter of the model is non-zero (at least one of the input variables has a statistically significant effect on self-perceived health). Three different tests were used: Likelihood Ratio test, Score test and Wald test. Each of the tests had a p-value less than 0.01. Therefore, we can reject the null hypothesis and accept the hypothesis that at least one input variable has a statistically significant effect on the modelled variable.

Table 2. Results of global null hypothesis testing

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4938.7060	25	< .0001
Score	4362.3225	25	< .0001
Wald	3166.0453	25	< .0001

Source: Own processing in SAS Enterprise Guide

In Tab. 3 are point estimates of model parameters (Parameter Estimate), standard errors of estimates (Standard Error) point estimates of odds ratios (Odds Ratio Estimate) and results of testing their statistical significance (Wald Chi-Square; p-value). Each categorical variable that had m categories was replaced by m-1 dummy variables. The estimated model has 25 parameters, of which 14 are statistically significant. Odds ratio estimates were used to interpret the values of the estimated parameters.

Table 3. Analysis of Maximum Likelihood Estimates and Odds Ratio Estimates

Effect	Parameter Estimate	Odds Ratio Estimate	Standard Error	Wald Chi-Square	p-value
Age					
age (16-24)	2.2818	9.794	0.1906	143.2561	<.0001
age (25-44)	1.0746	2.929	0.1029	108.9710	<.0001
age (45-64)	-0.1153	0.891	0.0796	2.0963	0.1477
age (65+)	reference category				
Main activity status					
employed	1.8285	6.224	0.0762	575.5442	<.0001
unemployed	1.4062	4.080	0.1274	121.7832	<.0001
student, pupil, fulfilling domestic tasks, and other	2.3450	10.433	0.1672	196.7735	<.0001
retired or unable to work	reference category				
Equalized disposable income					
2nd quintile	-0.0351	0.965	0.0789	0.1979	0.6564
3rd quintile	0.1136	1.120	0.0793	2.0522	0.1520
4th quintile	0.2471	1.280	0.0806	9.3905	0.0022
5th quintile	0.5279	1.695	0.0844	39.1166	<.0001
1st quintile	reference category				
Severe deprivation					
not deprived	0.5602	1.751	0.1095	26.1929	<.0001
deprived	reference category				
Marital status					
never married	-0.0381	0.963	0.1001	0.1446	0.7038
married	0.1230	1.131	0.0834	2.1738	0.1404
widowed	-0.4847	0.616	0.1110	19.0743	<.0001
divorced	reference category				
The highest level of education attained based on the ISCED					
medium	0.4136	1.512	0.0825	25.1423	<.0001
high	0.8573	2.357	0.0997	73.9177	<.0001
low	reference category				
Degree of urbanization					
cities	0.1502	1.162	0.0711	4.4671	0.0346
towns and suburbs	-0.0479	0.953	0.0550	0.7576	0.3841
rural area	reference category				
Region					
Bratislava Region	0.5056	1.658	0.1033	23.9554	<.0001
Trnava Region	0.1508	1.163	0.0981	2.3620	0.1243
Trenčín Region	0.1958	1.216	0.0955	4.2055	0.0403
Nitra Region	0.3146	1.370	0.0914	11.8547	0.0006
Žilina Region	-0.0691	0.933	0.0905	0.5837	0.4449
Banská Bystrica Region	0.0574	1.059	0.0895	0.4116	0.5211
Košice Region	0.1598	1.173	0.0934	2.9255	0.0872
Prešov Region	reference category				

Source: Own processing in SAS Enterprise Guide

The difference in the perception of good health was manifested mainly when comparing different age categories of respondents and when comparing respondents who differ in main activity status and education level. The age categories (16–24) (odds = 9.794) and (25–44) (odds = 2.929) had a significant influence on good self-perceived health. Pensioners and those unable to work perceive their health the least positively. The chance that an employed respondent evaluates his overall health condition as good (very good or good) is 6.224 times higher, and the chance for the category of student, pupil, fulfilling domestic tasks, and other is even 10.433 times higher

than the chance for the reference category (retired or unable to work). There is a statistically significant difference in health assessment expressed through the odds ratio between respondents with high and respondents with low education (odds ratio (high vs. low) = 2.357). Based on the results of the analysis, we can conclude that a higher equivalent disposable income is strongly associated with good SPH status (odds ratio (fifth vs. first quintile) =1.695, odds ratio (fourth quintile vs. first quintile) = 1.280). The respondent's marital status also affects the self-assessment of health. The chance of evaluating your health as good is 1.131 times higher for married people than for divorced people (ceteris paribus). Another statistically significant factor affecting self-assessment of health is the indicator of severe material deprivation. If a person faces an enforced lack of at least three out of nine material deprivation items in the economic strain and durables dimension, his chance of evaluating his health status as good is only 57 percent of the chance of a materially non-deprived person.

4. CONCLUSION

The objective of the paper was to identify the factors that have a statistically significant effect on the subjective perception of self-perceived health in the Slovak population aged 16 and over in 2021. A logistic regression model was used to identify and quantify the strength of their effect. The dependent variable was the binary variable self-rated health, which was created by merging the categories of the variable general health. Considering the article (Jindrová & Labudová, 2020), we used the following explanatory indicators in the model: age, degree of urbanization, educational attainment level, equalized household disposable income, main activity status, marital status, severe material deprivation, and NUTS 3 region. The estimated parameters and odds ratios correspond to the model predicting the value of good health.

Acknowledgment

The paper was supported by a grant agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic VEGA no. 1/0431/22 Implementation innovative approaches of modeling and managing risks in internal models of insurance companies in accordance with the Solvency II.

References

Agresti, A. (2019). *An Introduction to Categorical Data Analysis*. New York: Wiley.

Agresti, A., & Finlay, B. (2014). *Statistical Methods for the Social Sciences*. Pearson Education Limited.

Allison, P. D. (2012). *Logistic Regression Using SAS. Theory and application*. USA: SAS Institute, Inc.

Blaxter, M. (Ed.). (2004). *Health*. Cambridge: Polity Press.

European Commission. (2021). *Methodological guidelines and descriptions of EU-SILC, target variables*. https://ec.europa.eu/eurostat/documents/203647/203704/DOC65.pdf/434b2180-33b3-0d8c-ed1e-2da912d6a685?t=1655461990699

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression. In *Wiley series in probability and statistics*. https://doi.org/10.1002/9781118548387

Jindrová, P., & Labudová, V. (2020). The impact of socio-economic and demographic determinants on self-perceived health. *E+M. Ekonomie a Management*, *23*(4), 68–88. https://doi.org/10.15240/tul/001/2020-4-005

Manderbacka, K., Lundberg, O., & Martkainen, P. (1999). Do risk factors and health behaviours contribute to self-ratings of health? *Social Science & Medicine*, *48*(12), 1713–1720. https://doi.org/10.1016/S0277-9536(99)00068-4

Silva, J., Truzzi, A., Schaustz, F., De Barros, R. D., Santos, M., & Laks, J. (2017). Impact of insomnia on self-perceived health in the elderly. *Arquivos de Neuro-Psiquiatria*, *75*(5), 277-281. https://doi.org/10.1590/0004-282x20170031

Statistical Office of the Slovak Republic. (2020). *EU SILC 2019 Poverty Indicators and Social Exclusion (material deprivation)*. Bratislava: ŠÚ SR. www.statistics.sk

Whitehead, M., & Dahlgren, G. (2007). *European strategies for tackling social inequities in health: Levelling up Part 2*. Copenhagen: WHO Regional Office for Europe. http://www.euro.who.int/__data/assets/pdf_file/0018/103824/E89384.pdf

World Health Organization. (1948). *Summary Reports on Proceedings Minutes and Final Acts of the International Health Conference held in New York from 19 June to 22 July 1946*. https://apps.who.int/iris/handle/10665/85573

World Health Organization. (2018). *WHO Housing and Health Guidelines*. Geneva: World Health Organization. https://iris.who.int/bitstream/handle/10665/276001/9789241550376-eng.pdf?sequence=1