Mental Health Across Time and Stages of Economic Development in Europe

Kian Bagherlee, Lin Hui, Yun-Chung (Murphy) Liu, Afraa Noureen

2023-12-03

Abstract

Identifying factors associated with the prevalence of mental health conditions is crucial to address the associated harm and costs. The study aims to investigate contributing factors to mental-related disorders prevalence (MRDP) and suicide in Europe since the 1990s. The modulating effect of time and gross domestic product (GDP) per capita is also analyzed for its implication for policy making. The dataset analyzed is derived from the European Health for All database (HFA-DB)[1] from the World Health Organization in Europe. Ordinal regression was applied to model suicide rate group and linear regression was used to model MRDP. A priori variable selection and exploratory data analysis was conducted for variable selection. Youth unemployment rate, low GDP per capita, and accessibility to produce contribute to high suicide rate ($p=.01$, $.03$, $<.001$, respectively). Fat availability and GDP per capita are positively associated with MRDP ($p = <.001$) and years have a modulating effect on GDP per capita and unemployment rate on MRDP ($p<.001$). Our study identified contributing factors and revealed the modulating effect of years and GDP per capita, which can serve as reference for government policy making to promote mental health.

Introduction

Mental diseases are of great concern because of its prevalence and costs caused globally [2]. Around one fifth of the world population has been affected and one-thirds of the adults have experienced mental illnesses [3]. The mental diseases related expenses in 2010 were estimated to be at around 2.5 trillion US dollars [4]. In Europe, the cost was estimated to be 798 billion euros in the same year [5]. By 2030, the cost is expected to double [4]. Despite resource invested, mental diseases remain constantly a huge burden across countries, as expressed in disability-adjusted life years (DALYs) [6].

Mental health conditions can lead to suicide, which is associated to numerous social, economic, and nutritional factors. For example, when unemployment rate and inflation rate increase, suicide rate increases [7]. Increasing minimum wage, keeping it higher than the inflation rate lowers suicide rate [8]. Gross domestic product (GDP) per capita is reported to modulating effect on other factors (such as unemployment) on suicide rate and have temporal effect on reducing suicide rate [9]. Finally, dietary patterns, such as intake of vegetables, fruits and sweetened beverages are reported to be associated with depressive and anxiety symptoms [10].

Based on previous findings on factors affecting mental health, this study aims to investigate the contributions of different variables on mental health and the modulation effects of economic development stages (as measured by GDP per capita) and time. Data from European members of the World Health Organization were used for data analysis because of data availability and the diversity in multiple social economic factors across countries. Specifically, the aim is to answer the following questions:

1. What variables contribute to suicide rate for European countries in different stages of economic development?
2. What factors contribute to Mental-Related Disorder Prevalence (MRDP) in Europe? Do these factors influence the outcome differently across time?

Methods

Data and Preprocessing

The dataset, with 1860 observations and 88 columns, was obtained from the European Health for All database (HFA-DB)[1], containing demographic, disease prevalence, risk factors, and health care related investment data reported by governments of European states. The categorization of suicide rates into three groups (high, medium, and low) is determined by specific numerical thresholds. If the suicide rate is 15 or higher, it falls into the 'High' category. Conversely, if the rate is less than 5, it is categorized as 'Low.' For rates between 5 and 15 (inclusive), it is categorized as 'Medium.' Countries are classified as 'High GDP' if their Gross Domestic Product (GDP) per capita exceeds 14,944 US dollars , the lowest GDP per capita level of the International Monetary Fund (IMF)-defined advanced economies in Europe in 2023; otherwise, they are categorized as 'Low GDP.' Aggregated results data for country groups (e.g. overall Europe data) were removed from the dataset. Countries with missing data rate higher than 50 percent in any of the variables required were excluded. After excluding countries with high data missing rate, stochastic regression imputation was applied to impute missing data with variables included in the model.

Variable Selection

A priori variable selection was conducted based on previous research findings to model MRDP and suicide in Europe. Exploratory analysis was done for variable selection using scatter plot and boxplot. Interaction terms were included in the model based on observations of interaction effects during the exploratory analysis phase. Model assumptions and collinearity were assessed between independent variables. The final model includes variables yielding the best performance (by metrics detailed in the following section) and meeting models assumptions.

Model Fitting and Evaluation

Ordinal regression was used to model suicide rate groups. Akaike Information Criterion (AIC) values and Proportional Odds Assumption Test were used to assess the ordinal regression model. Variance Inflation Factor (VIF) was used to test collinearity, variables with VIF measure greater than 20 were evaluated to be excluded by the model. A confusion matrix was used to evaluate the performance of ordinal regression. Multiple linear regression (MLR) model was used for modeling MRDP. Residual vs fitted plot was used to test linearity and homoscedasticity. VIF was used to test collinearity. Cook's distance was used to assess the influence of outliers. R-squared was used to evaluate the performance of the linear regression model. All statistical analyses were done using R programming language (version 4.3.1).

Results

Overview of Included Data

Data reported from 1990 to 2019 were included in the analysis resulting in 1470 rows of data, each representing health related statistics of a given year of a given country in Europe. The overall data missing rate is 11.86% (average of all variables). Among the included data, 726 rows were considered data from advanced economies (GDP per capita $>$ 14,944), 744 were considered data from non-advanced economies; 443, 602, 163 rows were considered high, medium, and low suicide rate respectively. Finally, 490 rows of data were included in each of the 1990's, 2000's and 2010's.

Table 1: Characteristics of European States Included

	Suicide Rate	MRDP	GDP per Capita (thousand USD)
Median	12.28	22.34	16.03
1st Quartile	7.33	14.39	4.63
3rd Quartile	17.78	33.63	34.58
Mean	13.87	25.88	24.09

Research Question 1: Contributing Factors of Suicide Rate

The results of the final model for suicide rate is in table 2. *A priori* selected variables were included in the model fitting process. Performances of different models were compared to find best one. Then once multicollinearity was evaluated, the model had the predictors that comfortably answered the research question. From here, an interaction term between GDP per capita and another predictor was determined. After the final model was fit, assumptions were tested.

Table 2: Ordinal Regression Model Summary

	Variable	Estimate	Std Error	t-value	p-value
1	Youth Unemployment	0.03	0.01	2.46	0.01
2	GDP Group (Low)	1.04	0.39	2.67	0.03
3	Number of Hospitals	-2.33e-4	1.08e-4	-2.15	0.03
4	HDI	2.84	0.22	13.19	<0.001
5	Available Produce	-0.01	0.001	-9.78	<0.001
6	Youth Unemployment * GDP Group	-0.039	0.02	-2.17	0.03
7	Low Suicide | Medium Suicide	-2.86	0.23	-12.24	<0.001
7	Medium Suicide | High Suicide	0.43	0.19	2.29	<0.001

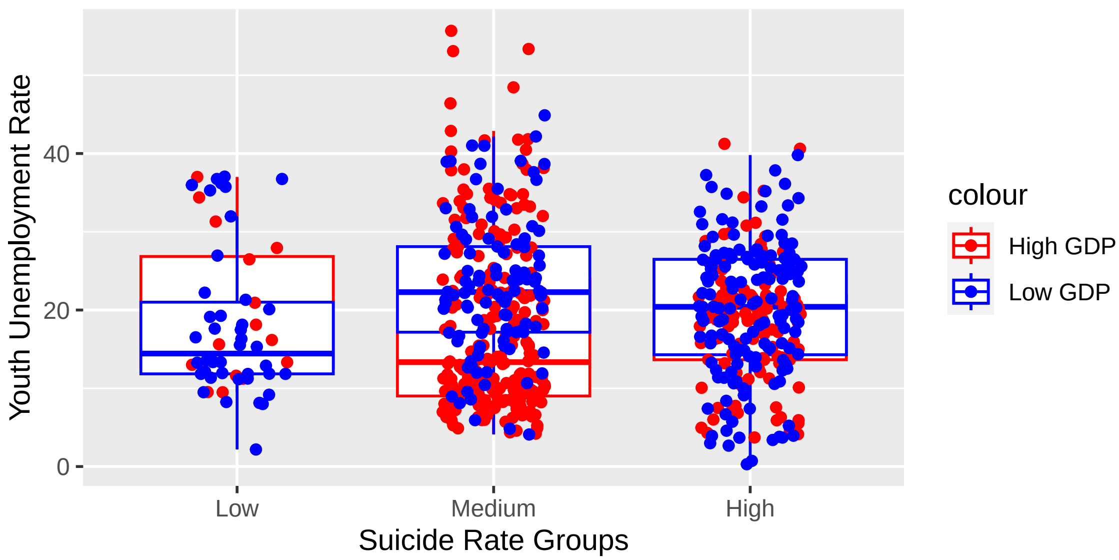
A priori selected variables were Youth Unemployment Level, GDP grouping of each county, Human Development Index (HDI), Number of Fruits and Vegetables per Person per Year, and Gross National Income (GNI). The imputed dataset was utilized, as an ordinal regression model cannot be fit with NA values present. GNI was forced to be removed as its values scaled would disrupt the code, preventing the model from ever being fit.

From here, there were 2 other predictors of interest that needed to be evaluated: GINI value and Number of Hospitals. The approach was to fit 3 models: one with each of the predictors incorporated, and a third that housed both. The AIC value of each model was used at the main comparison metric, as this value is primarily used for model comparison. The lowest AIC value ended up being the model that utilized both predictors. Because AIC natively punishes more predictors being added, but the models' AIC is the lowest with both predictors, it is safe to assume that this is the best model to work with.

Then, multicollinearity became a prime target of interest. This however came back with grave results, as the VIF value for both HDI and GINI were over 20. Because of this clear violation, more in-depth research was conducted to determine the cause. It was revealed that the GINI index, paired with other variables, is used to calculate HDI, meaning they were directly proportional. To address this issue, a tempting solution was to make an interaction term with both of them, and then refit the model. A valiant effort, this exasperated the multicollinearity issue further. It became clear that both predictors could not be in the model. HDI was opted to be the survivor, because if HDI utilizes GINI and other variables, it is therefore a more holistic metric. Therefore, because of the extra information HDI can provide that GINI can't, it stayed and fixed the multicollinearity issue.

Finally, it was time to analyze a potential interaction term with the GDP Group variable. Analyzing the graphs created during prior EDA, as seen below, Youth Unemployment and GDP per capita seemed very tied together. After creating this interaction term and re-fitting the model, the AIC value decreased comfortably enough to leave the interaction term. Looking at the model summary above, the negative coefficient implies that as Youth Unemployment increases, a country will be less likely to be a high GDP country. Then, in conjunction, a country with higher GDP per capita has increased odds of being in a higher suicide rate group.

Figure 1. Youth Unemployment by Suicide Rate Groups for High and Low GDP.

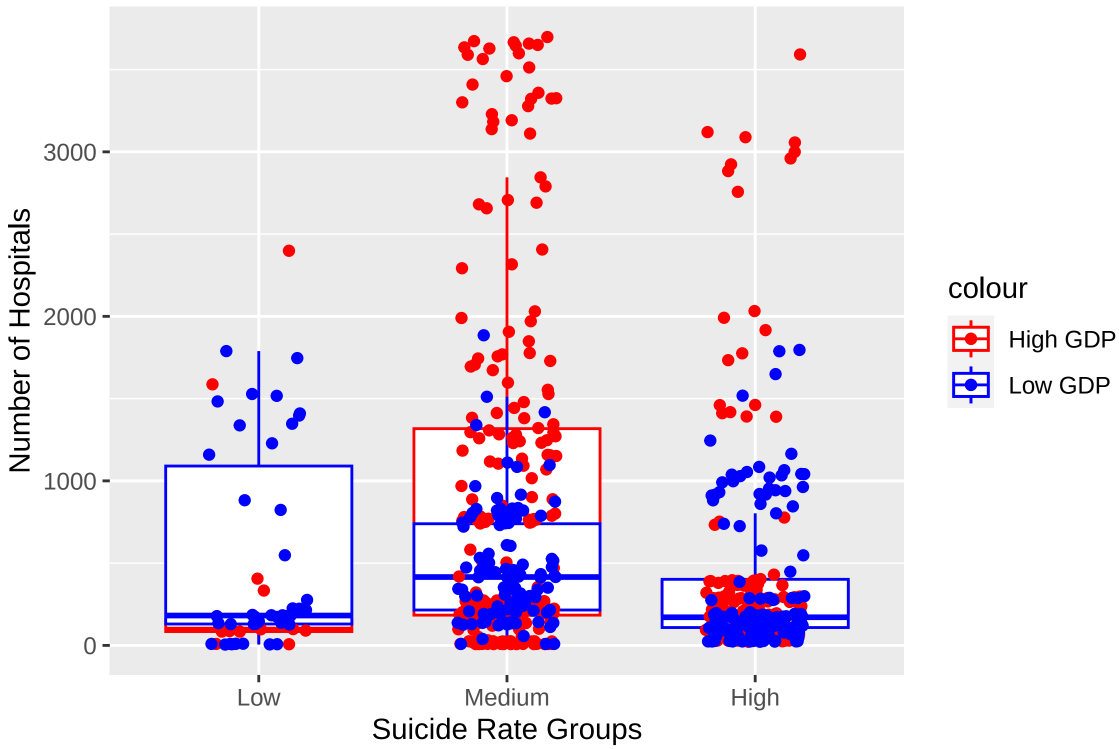


From here, all that was left was to check whether the created final model violated the Proportional odds assumption or not. What was being looked out for was if the predicted values from both the fit ordinal model, and a multinomial model, are similar. The team determined that if values started to differ by over 0.05, the assumption would be violated. However, the highest difference in predicted values was 0.03. With this, the model can safely be stated as not violating the proportional odds assumption.

With confidence in our model, we now utilized a confusion matrix to check the accuracy of the model itself. The value that we retrieved is an accuracy score of 0.5996, which while there is room for improvement, the model itself is well made.

Our model identifies factors associated with suicide rate in Europe. First for suicide rate, the odds of going from one suicide group to a higher one increases per unit increase of youth unemployment, HDI, and available produce. However, the odds of going from one suicide group to a higher one decreases per unit increase of the number of hospitals. This helps explain the figure below. These conclusions are supported by the incredibly low p-values seen throughout the model. As well, GDP was found to be correlated with the suicide rate, through its interaction with youth unemployment.

Figure 2. Number of Hospitals by Suicide Rate Groups for GDP Group.



Research Question 2: Contributing Factors of MRDP

Unemployment rate (%), inflation rate (%), GDP per capita (thousand USD), HDI, fat availability (g/person/day) and year group were selected a priori for exploratory data analysis (figure 3). HDI was excluded because of the high missing rate (25.6% percent). Fat availability and GDP per capita are positively associated with MRDP. Inflation rate and unemployment rate are negatively associated with MRDP. Years have a modulating effect on GDP per capita and unemployment rate on MRDP. The interaction between year group (90's, 00's, 10's) and unemployment rate is included because one can clearly observe their differential influence on MRDP (figure 4). More specifically, unemployment rate is *positively* associated with MRDP in the 1990's and *negatively* associated with MRDP in the 2000's and 2010's.

Figure 3. Effects of variables associated with MRDP

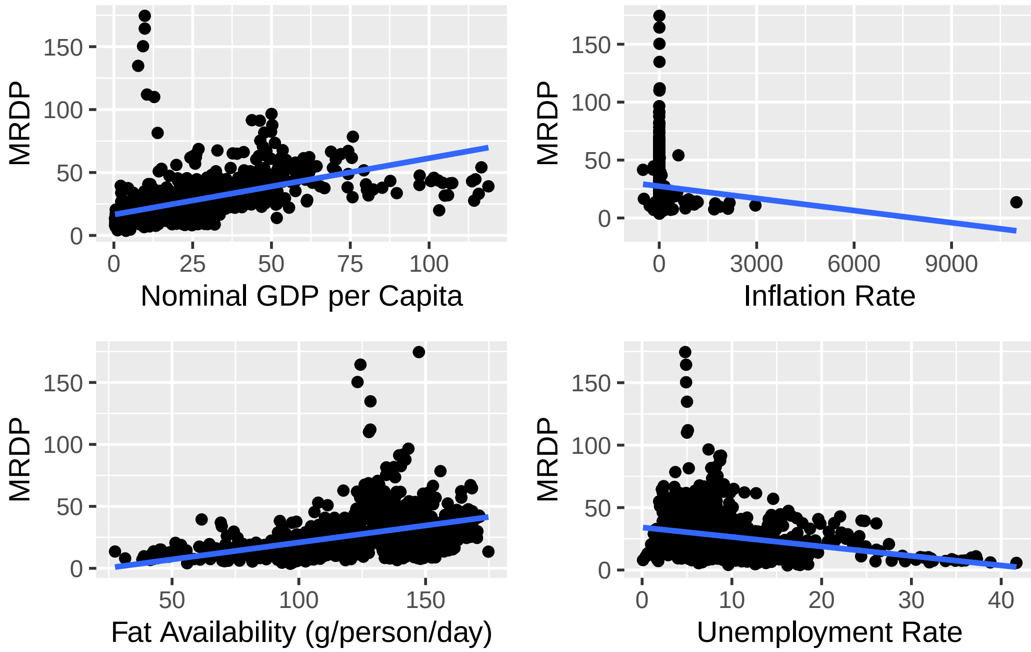
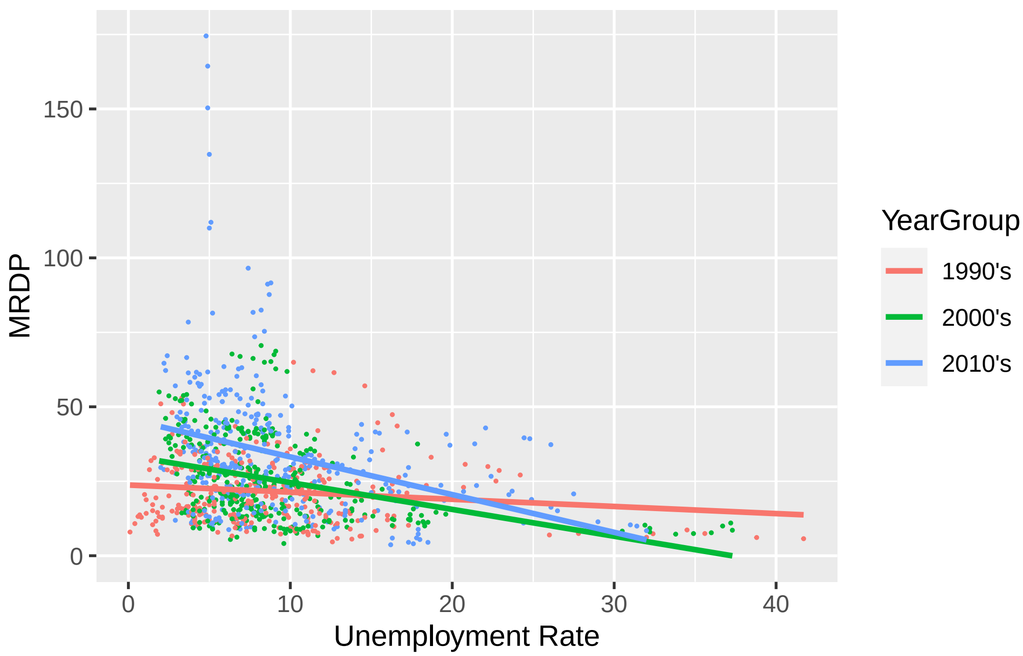
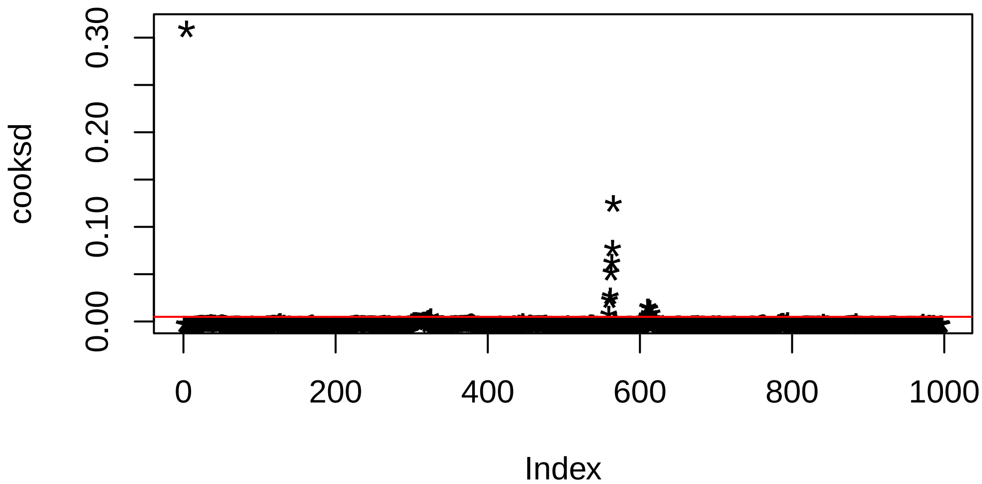


Figure 4. Interaction between Unemployment Rate and Year Group



Cook's distance was evaluated to detect influential observations. The analysis revealed a singular data point that markedly diverged from the others. Upon re-fitting the model without this outlier, no significant changes were observed in the results. Consequently, the decision was made to retain this observation, as its presence neither adversely affects nor improves the model's performance.

Figure 5. Influential points measured by Cook's distance.



Unemployment rate, inflation rate, GDP per capita, fat availability (g), year group and interaction between year group and unemployment rate are included in the final model. Year Group 2010's, GDP per capita and fat availability increases the predicted MRDP ($p < .001$). This finding may suggest that economic growth could be associated with factors that negatively impact mental health, such as stress related to a fast-paced lifestyle or social isolation. The interaction between year group 2010's and unemployment rate is statistically significant ($p < .001$). The model achieved an adjusted R-squared of 0.340.

Table 3: MRDP MLR Summary Table

Characteristic	Beta	SE	95% CI	p-value
Unemployment Rate (%)	-0.039	0.129	-0.292, 0.215	0.766
Fat Availability (g/person/day)	0.112	0.021	0.070, 0.153	< 0.001
Inflation Rate (%)	0.000	0.001	-0.002, 0.002	0.902
GDP per capita (thousand USD)	0.280	0.030	0.222, 0.338	< 0.001
factor(YearGroup)				
1990's	—	—	—	
2000's	3.361	2.04	-0.640, 7.363	0.100
2010's	13.02	2.15	8.799, 17.24	< 0.001
Unemployment Rate (%) * factor(YearGroup)				
Unemployment Rate (%) * 2000's	-0.269	0.188	-0.639, 0.100	0.153
Unemployment Rate (%) * 2010's	-0.711	0.190	-1.085, -0.338	< 0.001

Model assumption is assessed by residual vs. fitted plot. We observed patterns suggesting slight non-linearity and variance inconsistencies, but these were not substantial enough to preclude the use of linear regression. None of the VIF scores of the final model exceeded 2, indicating that multicollinearity was not a concern.

Conclusion

This study identified factors influencing suicide rates and MRDP since the 1990s in Europe. Using data from the European Health for All database, key insights are revealed, emphasizing the need for tailored interventions based on unique socio-economic landscapes. For suicide rates, youth unemployment, HDI, produce accessibility, and the number of hospitals play significant roles, with an interaction between GDP per capita and youth unemployment rate. Regarding MRDP, positive associations with fat availability and GDP per capita are highlighted, alongside the modulating effects of years and unemployment rates on MRDP.

Despite limitations of missing data and the existence of strong outliers for various potential predictors, our study underscores the nuanced relationships between these factors and mental health outcomes. This study contributes to a comprehensive understanding of mental health factors in Europe, providing insights for evidence-based policymaking.

Moving forward, it is essential to address these limitations, especially by enhancing data collection for countries, such as Albania, Azerbaijan, Belarus, Switzerland, etc., experiencing data quality challenges. Future research should focus on refining analyses and exploring mental health challenges across diverse European contexts. By addressing these issues, this study lays the foundation for effective interventions to tackle mental health issues in the region.

References

1. World Health Organization (WHO). n.d. "European Health for All Database (HFA-DB)." https://gateway.euro.who.int/en/datasets/european-health-for-all-database/.
2. Kählke, F., Buntrock, C., Smit, F. *et al.* (2022). Systematic review of economic evaluations for internet- and mobile-based interventions for mental health problems. *npj Digit. Med.* 5, 175.
3. Steel, Z. et al (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *Int. J. Epidemiol.* 43, 476–493.
4. Bloom, D. E. et al. (2011). The global economic burden of non-communicable diseases. http://www3.weforum.org/docs/WEF_Harvard_HE_GlobalEconomicBurdenNonCommunicableDiseases_2011.pdf.
5. Gustavsson, A. et al. (2011). Cost of disorders of the brain in Europe 2010. Eur. *Neuropsychopharmacol.* 21, 718–779.
6. GBD 2015 DALYs and HALE Collaborators. Global, Regional, and National Disability-adjusted Life Years (Dalys) for 315 diseases and injuries and healthy life expectancy (Hale), 1990-2015: a systematic. *Lancet* 388, 1603–1658 (2016).
7. Stack, S. (2021). Contributing factors to suicide: Political, social, cultural and economic. *Preventive Medicine*, 152, 106498.
8. Gertner, A. K., Rotter, J.S., Shafer P.R. (2019). Association between state minimum wages and suicide rates in the U.S. *Am. J. Prev. Med.*, 56 (5), 648-654.
9. Meda, N., Miola, A., Slongo, I., Zordan, M. A., & Sambataro, F. (2022). The impact of macroeconomic factors on suicide in 175 countries over 27 years. *Suicide and Life-Threatening Behavior*, 52(1), 49-58.
10. Chen, H., Cao, Z., Hou, Y., Yang, H., Wang, X., & Xu, C. (2023). The associations of dietary patterns with depressive and anxiety symptoms: a prospective study. *BMC medicine*, 21(1), 307.