Patterns of Behavioural Risk Factors and their Associations with Cardiovascular Outcomes in Ukraine

Oleh Kovalyshyn[1], Yaroslav Dorogyi[2], Vladyslav Ostash[1,3] and Nadiia Kasianchuk[1]
[1]*Department of Computer Science, Kyiv School of Economics, Shpaka Str. 3. 03113 Kyiv, Ukraine*
[2]*Department of Artificial Intelligence and Cybersecurity, Donetsk National Technical University, 82100 Drohobych, Ukraine*
[3]*Faculty of Biotechnology and Biotechnics, Igor Sikorsky Kyiv Polytechnic Institute, 03056 Kyiv, Ukraine*
oleh.kovalyshyn.25@kse.org.ua, yaroslav.dorohyi@donntu.edu.ua, vostash@kse.org.ua, nkasianchuk@kse.org.ua

Keywords: Ukraine, WHO STEPS, Cardiovascular Disease, Blood Pressure, Hypertension, Behavioural Risk Factors, Tobacco, Alcohol, Substance Use, Body Mass Index, Obesity, Logistic Regression.

Abstract: Evidence on behavioural and metabolic cardiovascular risk patterns in Ukraine remains limited beyond descriptive reporting from the WHO STEPS 2019 survey. We reconstructed the STEPS 2019 analytic framework and examined generational patterns of tobacco and alcohol use, body-mass index (BMI), and their associations with elevated blood pressure and self-reported cardiovascular history. We analysed de-identified microdata from the survey (adults aged 18-69; n = 4,409). The complex survey design was reconstructed in R and calibrated to Ukraine's 2019 age-sex population distribution. Prevalence estimates were produced for tobacco and alcohol indicators, including product-type patterns by age and sex. Elevated blood pressure (SBP ≥140 and/or DBP ≥90 mmHg) was compared across BMI categories. Predictors of self-reported cardiovascular history (stroke and/or heart disease event) were assessed using standard logistic regression, survey-weighted logistic regression, and Firth penalized logistic regression, with backward AIC-based selection; performance was evaluated using ROC AUC and F-score on complete cases (n = 2,279). Current tobacco use was 32.9% overall (49.9% men; 17.6% women), with younger adults more likely to use non-cigarette products. Current alcohol use was widespread (≈69%). Blood pressure differed significantly by BMI, while self-reported cardiovascular history did not. In multivariable modelling, the best-performing specification was Firth logistic regression (AUC = 0.721; F-score = 0.447), identifying age, BMI category, history of high blood sugar, and salt consumption frequency as significant predictors ($p < 0.05$); tobacco and alcohol use were not significant, consistent with potential reverse causation. Findings support prioritising sodium-reduction and weight-management strategies, while interpreting behavioural associations cautiously due to self-report and the cross-sectional design.

1 INTRODUCTION

According to the World Health Organization, non-communicable diseases (NCD) are responsible for 43 million deaths in 2021, of which 73% were in low- and middle-income countries. Cardiovascular diseases are the most common among NCD and cause 90% more deaths than cancers [1]. Moreover, recent findings suggest that younger adults are diagnosed with cardiovascular diseases more often than in previous years, potentially breaking the established pattern of them being associated with older people [2]. Finding socioeconomic and habitual patterns can increase understanding of the NCDs and help to form policies and treatments to reduce the risk for the population.

For that, it's important to understand the current risk factors, prevalence, generational trends and impact of local cultures on the level of NCDs in the population. Such analysis is lacking for Ukraine which is a nation with a transitioning economy and rapidly aging population.

STEPS is a survey, conducted by the World Health Organization that focuses on NCD risk factors among 18-69 years old adults. Conducted for the first, and, until today, only, time in Ukraine it provides insights into Ukrainian population's habits regarding alcohol and tobacco use, diet, sport activity, income and objective health measurements such as body-mass index (BMI), heart rate, blood pressure, cholesterol level.

The original report provides general analysis of the current state of Ukrainians' health and prevalence of different habits and risk factors. It doesn't try to look into dependencies between the data and connection between risk factors and measured parameters/habits. Analyzing the age and sex differences in habits and risk factors can lead to new insights regarding Ukrainian health and cultural norms.

2 METHODS

2.1 Data Source and Weighting

This study uses de-identified microdata from the WHO STEPS Ukraine 2019 survey, a nationwide multistage cluster household survey of adults aged 18-69 years conducted under the WHO STEPS protocol. The dataset includes questionnaire-based behavioral risk factors and sociodemographics (Step 1) and standardized physical measurements, including blood pressure and anthropometrics (Step 2). The analytic sample comprised 4,409 respondents drawn through the original sampling framework documented in the STEPS report [3].

We reconstructed the complex sample design in R using the *Survey* package. Individual-level analyses applied the provided sampling weights (*wstep1*). To align estimates with the national population structure, we additionally calibrated/post-stratified the weighted sample to Ukraine's 2019 age-sex distribution. Because the STEPS report does not specify the exact population source used to construct these demographic corrections, we used the 2019 population counts from the State Statistics Service of Ukraine as the reference distribution for calibration [4].

To validate the reconstruction, we compared reconstructed estimates against the STEPS report using pre-specified anchor indicators: active smoking prevalence, active alcohol use prevalence, and mean body mass index (BMI), evaluated overall and across age-sex cohorts. During validation, we assessed missing-data patterns in these anchor variables and identified substantial missingness in the alcohol-use indicator. To examine whether discrepancies with the published report could be driven by missing-data handling, we performed a sensitivity check that recoded missing alcohol-use responses as non-use (NA = false). Because this recoding imposes a strong and generally unjustified assumption, it was used only for sensitivity assessment; all subsequent substantive analyses treat missing alcohol-use responses as missing and use standard survey-based handling.

2.2 Substance Use Analysis: Tobacco and Alcohol

Tobacco and alcohol indicators were analysed by sex and WHO STEPS age groups (18-29, 30-44, 45-59, 60-69); tobacco analyses additionally considered place of residence (urban vs rural). All estimates accounted for the complex survey design and applied sampling weights (*wstep1*) using the R *survey* package, producing survey-weighted prevalence estimates with 95% confidence intervals.

Associations with sociodemographic factors were assessed using survey-weighted logistic regression (svyglm, binomial). For tobacco, current smoking status was modelled as a function of sex and age group, with urban/rural residence evaluated via stratification or as a covariate as appropriate. Among current tobacco users, product-type patterns were described using survey-weighted prevalence (95% CI) for combustible cigarettes, heated tobacco products, electronic cigarettes/vapes, hookah/waterpipe, and non-smoked tobacco, overall and by sex and age group.

For alcohol, we examined current use and lifetime abstinence (never consumed alcohol), each modelled separately by sex and age group using survey-weighted logistic regression. Alcohol analyses used outcome-specific complete-case handling (excluding respondents with missing values for the relevant alcohol outcome).

2.3 Body-Mass Index Analysis

Body-mass index (BMI; kg/m²) was summarised descriptively across respondent characteristics without survey weights to compare groups rather than produce population estimates. Mean BMI (95% CI) was reported by sex and WHO STEPS age groups (18-29, 30-44, 45-59, 60-69). Respondents were classified using WHO cut-offs: underweight (<18.5), normal (18.5-24.9), overweight (25.0-29.9), and obesity (≥30.0); these categories were used in all subsequent BMI-stratified analyses.

Physiological correlates of BMI were assessed by comparing heart rate and blood pressure across BMI categories. For each respondent, mean heart rate and mean systolic/diastolic blood pressure were calculated as the average of three repeated measurements; analyses used indicator-specific complete cases (excluding refusals/missing readings). Category-specific means (95% CI) were estimated overall and by sex.

We estimated the prevalence of elevated blood pressure (SBP ≥140 and/or DBP ≥90 mmHg) within each BMI category and tested differences in systolic and diastolic blood pressure across categories using the Kruskal-Wallis test; when significant, Dunn's post-hoc pairwise comparisons with Bonferroni adjustment were applied (two-sided α = 0.05). Cardiovascular morbidity was captured via self-reported history of heart attack, stroke, or heart-related chest pain; for each BMI category we reported prevalence and odds ratios relative to the normal-weight group.

2.4 Models for Heart Disease Prediction

Building on the descriptive results, we fit multivariable regression models to quantify associations between respondent characteristics and prior cardiovascular events. Finding significant associations gives better understanding of the social and behavioral context of the disease. We asked if any of the previous findings could translate into cardiovascular diseases in patients. The primary outcome was self-reported disease history, defined as reporting a previous stroke and/or heart disease event in the STEPS questionnaire. We compared three specifications: (i) standard logistic regression (GLM), (ii) survey-weighted logistic regression accounting for the complex design (R *survey* package), and (iii) Firth penalized logistic regression to reduce small-sample bias and separation (R *logistf* package).

Models were estimated on a complete-case subset excluding missing or ambiguous responses ("Refused", "Don't know", NA) for the outcome and all candidate predictors, yielding n = 2,279 (from 4,409), with 18.69% reporting the outcome. Predictors were selected based on prior analyses and included physical and lifestyle factors (e.g., BMI category, blood pressure indicators, smoking and alcohol measures, diet and physical activity variables) and core sociodemographics (age group, sex). Continuous variables were checked for plausibility and, where appropriate, transformed into clinically interpretable forms (e.g., BMI categories, hypertension indicator).

Variable selection was performed using backward stepwise elimination based on the Akaike Information Criterion (AIC), a model selection metric that balances goodness-of-fit against model complexity. Lower AIC values indicate a better trade-off between explanatory power and parsimony, reducing the risk of overfitting while retaining informative predictors.

Model discrimination was assessed via ROC analysis, reporting AUC and F-score, computed consistently across model families to compare the impact of survey weighting and penalized estimation on predictive performance.

3 RESULTS

3.1 Survey Reconstruction and Benchmarking Against the STEPS Report

To evaluate reconstruction accuracy and the effect of calibration to Ukraine's 2019 age-sex population structure, we benchmarked reconstructed estimates against the STEPS 2019 Ukraine report using three pre-specified anchor indicators: current (active) smoking prevalence, current (active) alcohol use, and mean BMI. Differences (reconstructed minus reported) are reported overall and by sex-age strata (Tables 1-3).

Table 1: Difference in reconstructed vs STEPS report estimates of current smoking prevalence and mean BMI by sex-age cohort (reconstructed − reported, percentage points), after calibration to the 2019 age-sex population distribution.

Cohort	Δ smoking, %	Δ BMI, kg/m^2
All ages and genders	-0.96	0.2
Male 18-29	0.87	0
Male 30-44	-0.58	0.2
Male 45-59	-0.3	0
Male 60-69	-0.07	0
Male 18-69	-0.4	0.1
Female 18-29	6.11	0
Female 30-44	0.75	0
Female 45-59	-0.03	0
Female 60-69	0.01	0
Female 18-69	0.94	0.2

Reconstructed smoking prevalence and mean BMI were broadly consistent with the report across most strata (overall smoking difference −0.96 percentage points; overall BMI difference +0.2 kg/m²), with the main deviation observed among females aged 18-29 (Table 1). In contrast, reconstructed estimates of current alcohol use exceeded report values by ~13-15 percentage points, while lifetime abstinence ("sober") closely matched (difference +0.09 percentage points) (Table 2). Inspection of the microdata indicated substantial missingness in the alcohol-use indicator (1,122 observations; 25% coded as NA), compared with no missingness for smoking and 5% missingness for BMI (233 observations), suggesting that differences in missing-data handling likely explain the alcohol discrepancy.

Table 2: Difference in reconstructed vs STEPS report estimates of current alcohol use prevalence (and lifetime abstinence, "sober") (reconstructed − reported, percentage points), after calibration to the 2019 age-sex population distribution.

Cohort	Difference, %
All ages and genders	13.45
Male 18-69	14.81
Female 18-69	13.1
Sober, all ages and genders	0.09

Table 3: Sensitivity analysis: difference in reconstructed vs STEPS report estimates of current alcohol use prevalence after recoding missing alcohol-use responses as non-use (NA → false); (reconstructed − reported, percentage points), after calibration to the 2019 age-sex population distribution.

Cohort	Difference, %
All ages and genders	2
Male 18-69	-1.3
Female 18-69	-1.1

To assess this possibility, we conducted a diagnostic sensitivity analysis recoding missing alcohol-use responses as non-use (NA → false). Under this assumption, reconstructed alcohol-use estimates aligned much more closely with the report (overall difference +2 percentage points; −1.3 for men and −1.1 for women; Table 3). Because this recoding imposes a strong and generally unjustified missingness assumption, it was used for sensitivity assessment only and was not applied in subsequent substantive analyses.

3.2 Substance Use Analysis

Survey-weighted estimates indicate that current tobacco use in Ukraine in 2019 was 32.94% (design-based SE = 1.61%), with a large sex gap (49.91% in men vs 17.64% in women). Among women, tobacco use declined sharply with age, from 26.62% (18-29) and 25.05% (30-44) to 13.57% (45-59) and 3.61% (60-69). Among men, prevalence remained high across age groups, rising from 46.67% (18-29) to 50.52% (30-44) and peaking at 55.00% (45-59), before decreasing to 43.23% at ages 60-69 (Fig. 1).

We tested two hypotheses: (i) younger adults have lower overall tobacco use, and (ii) younger adults are more likely to use non-cigarette products [5], [6]. The first hypothesis was only weakly supported, as overall prevalence was not consistently lower in younger groups-particularly among men (Fig. 1). The second hypothesis was supported: although combustible cigarettes were the most common product in every age group, non-cigarette products (hookah/waterpipe, heated tobacco, and electronic cigarettes/vapes) were concentrated among younger adults and declined with age (Fig. 2). Because respondents could report multiple product types, product-specific prevalences are non-exclusive and should not be summed.

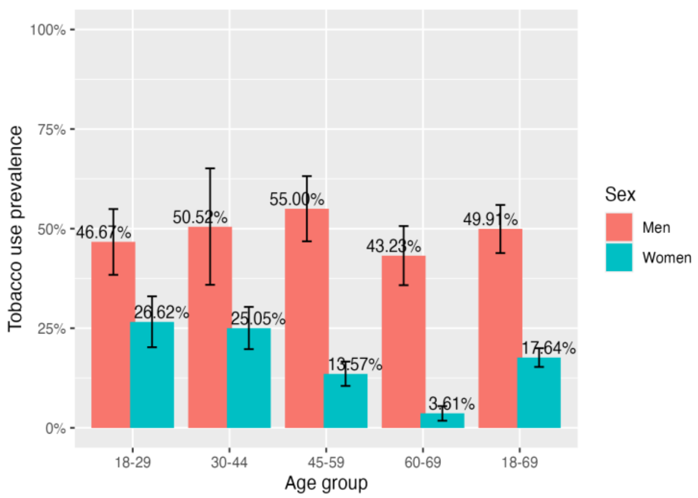

Figure 1: General tobacco use prevalence.

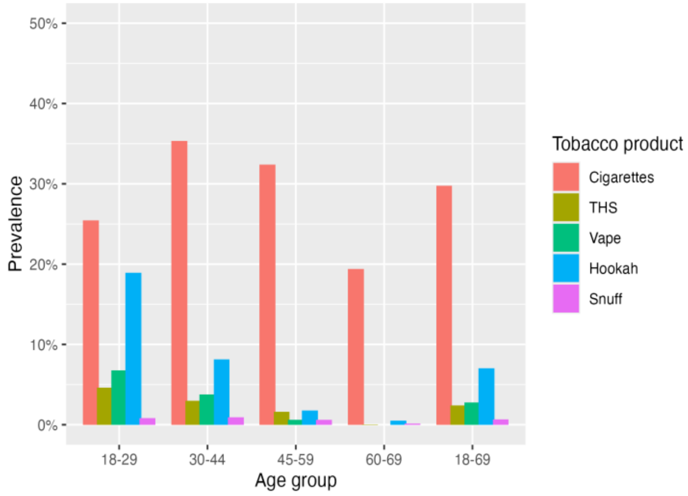

Figure 2: Survey-weighted prevalence of tobacco product types by age group (population prevalence). Product categories are non-exclusive; error bars indicate uncertainty (95% CI or SE).

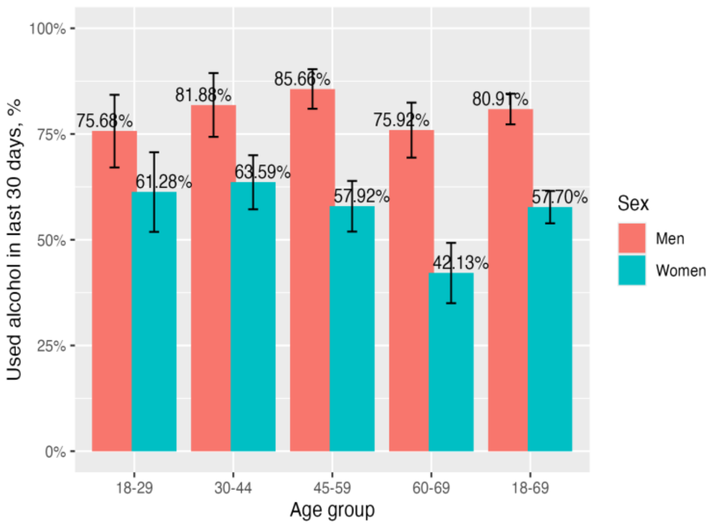

Figure 3: Active alcohol use prevalence.

Active use of alcohol was defined by the fact of consumption during the last 30 days at the time of survey. The metric is high among all age groups but is significantly higher in men vs women (80.91% vs 57.70%). The same pattern of the youngest and the oldest consuming less than middle-aged respondents is visible with active alcohol use (Fig. 3).

9.5% (SE = 0.85%) of Ukrainians have never tasted alcohol (self-reported), the pattern is inverse of that that we see for active use. Younger and older people are sober more than those whose youth coincided in time with the collapse of the Soviet Union and following economic shock. Also, women are sober more often than men (Fig. 4).

3.3 BMI Analysis

Consistent with prior evidence that BMI tends to increase with age, mean BMI in the STEPS 2019 sample rose across successive age groups for both sexes (Fig. 5) [7]. Among men, this upward trend plateaued with a small decline in the 60-69 group relative to ages 45-59, whereas among women the increase was more pronounced at older ages. The overall mean BMI in the sample was 25.0 kg/m² (design-based SE = 0.165), placing the average respondent at the threshold of the overweight category.

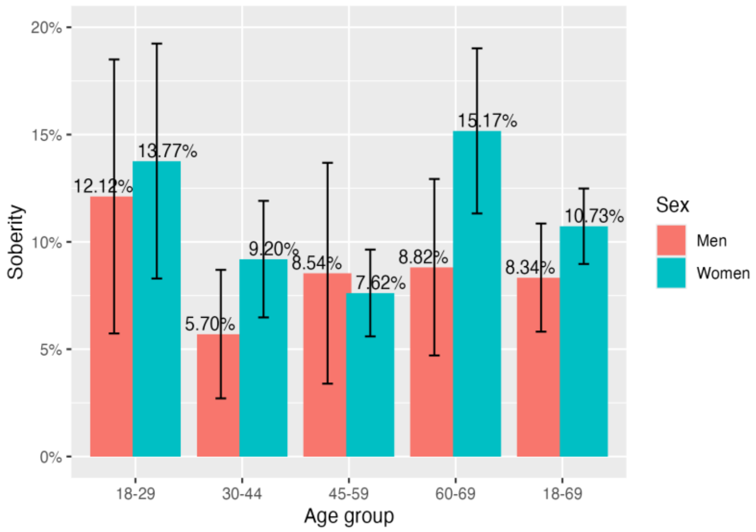

Figure 4: Sobriety prevalence.

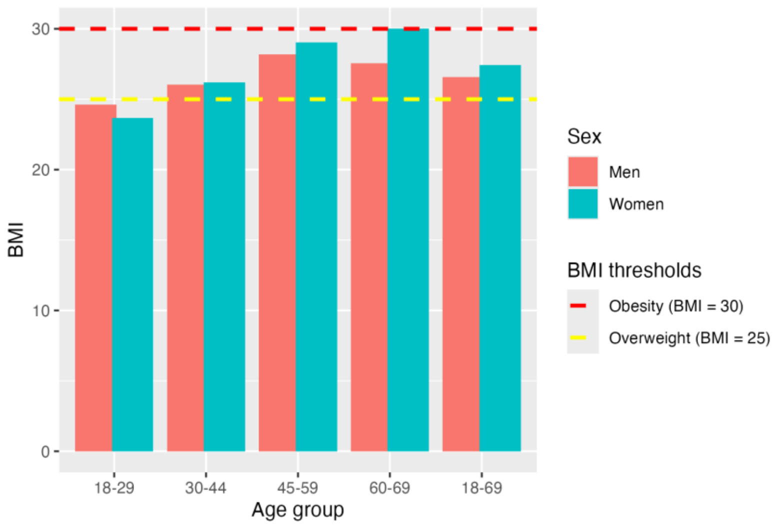

Figure 5: Mean BMI (kg/m²) by age group and sex.

Most age-sex strata exhibited mean BMI values above the overweight threshold (BMI $\geq$ 25), and women aged 60-69 had a mean BMI at approximately 30.0 kg/m² (SE = 0.33), corresponding to the obesity threshold. In survey-weighted GLM analysis of BMI, age group was strongly associated with BMI relative to the 18-29 reference group, whereas sex was not statistically significant at the 0.05 level (Table 4).

Given the well-established positive relationship between BMI and blood pressure [8], [9], we examined the prevalence of elevated blood pressure across BMI categories. The proportion of respondents meeting the elevated blood pressure criterion was substantially higher in the overweight and obese groups, and lowest in the underweight group (Fig. 6). This pattern was consistent with odds ratios comparing each BMI category to the normal-weight reference group: the odds of elevated blood pressure increased monotonically with the BMI category (Table 5).

Table 4: Survey-weighted GLM p-values for predictors of BMI.

Parameter	p-value
sex	0.098378
Age range 30-44	0.000182
Age range 45-59	< 2e-16
Age range 60-69	< 2e-16

We next assessed self-reported history of cardiovascular events/symptoms across BMI categories. The proportion of respondents reporting a cardiovascular history was higher in the overweight and obese groups, while normal and underweight groups appeared similar (Fig. 6). However, inference for the underweight category is constrained by limited sample size: after restricting to respondents with an unambiguous “yes/no” response, only 60 individuals were classified as underweight (including 8 reporting cardiovascular history), which reduces precision and statistical power. Consistent with this limitation, the Kruskal-Wallis test did not detect statistically significant differences across BMI categories for self-reported cardiovascular history ($\chi^2 = 3$, df = 3, p = 0.392). Odds ratios by BMI category (Table 6) should therefore be interpreted cautiously, particularly for the underweight group.

Table 5: Odds ratios for elevated blood pressure by BMI category (reference: normal weight).

Cohort	OR
underweight	0.410
normal	1
overweight	1.82
obesity	4.84

Table 6: Odds ratios for self-reported cardiovascular history by BMI category (reference: normal weight).

Cohort	OR
underweight	1.06
normal	1
overweight	1.54
obesity	2.70

3.4 Key Predictor Selection for Cardiovascular History Using Linear Models

The final multivariable model identified four significant independent predictors: age, BMI category, history of high blood sugar, and frequency of salt consumption (p<0.05). Other potential risk factors, including physical activity, household income, marital status, and education level, current tobacco and alcohol use, did not reach statistical significance in the adjusted model (p>0.05). This lack of association likely reflects reverse causality, where individuals diagnosed with cardiovascular conditions may have ceased substance use following their diagnosis.

Firth’s penalized logistic regression yielded the highest discriminative performance with an AUC of 0.721 and an F-Score of 0.447, slightly outperforming the standard survey-weighted GLM (AUC = 0.709) (Fig. 7).

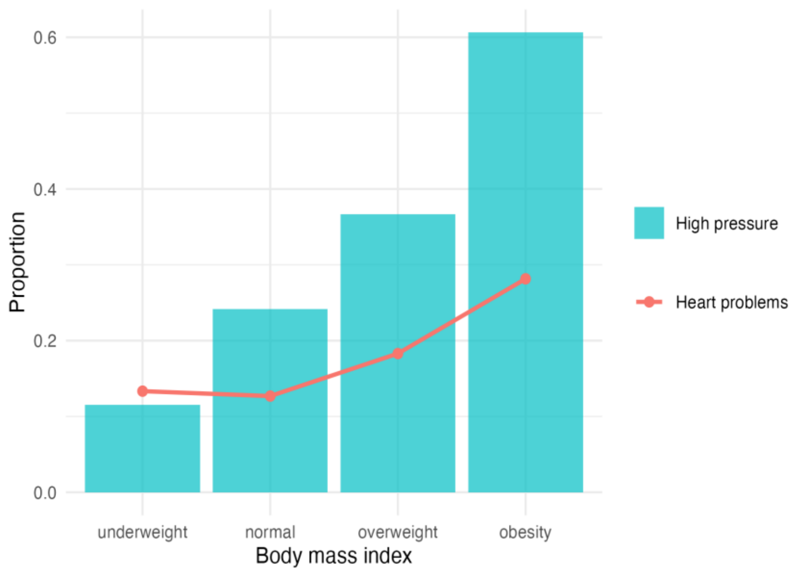

Figure 6: Prevalence of elevated blood pressure and self-reported heart problems by BMI category.

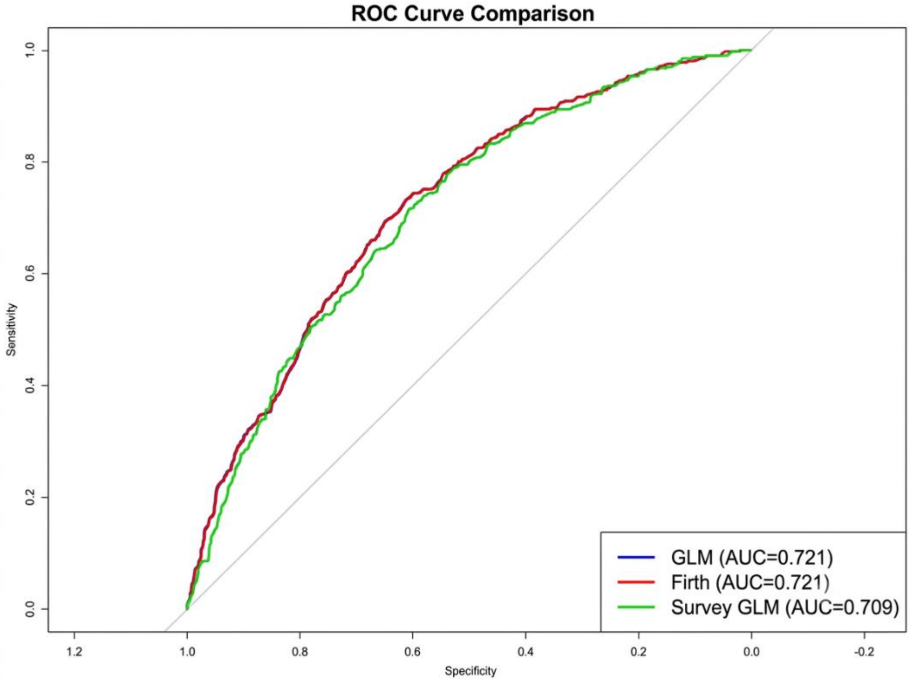

Figure 7: ROC curves comparison.

4 CONCLUSIONS

Increased BMI is an established factor of risk of cardiovascular diseases [10]. Habits such as alcohol and tobacco consumption are also related to increased mortality and lifetime disease risk [11], [12].

These findings have specific implications for public health policy in Ukraine. While alcohol and tobacco control remains essential, our multivariable model highlights that BMI and salt consumption frequency are the most robust predictors of existing cardiovascular conditions in this dataset. Unlike behavioral habits (smoking/drinking) which might be modified after a diagnosis (reverse causality), high BMI and dietary salt intake appear to be persistent risk drivers. Consequently, interventions targeting sodium reduction (e.g., labeling regulations, public awareness campaigns) and weight management strategies may yield the most immediate benefits for secondary prevention and reducing cardiovascular burden.

This study reveals that Ukrainians continue to be active smokers and drinkers despite educational efforts and generational differences. Trends in tobacco consumption seem to shift rather than decrease, so more studies might be needed regarding the influence of novel consumption methods on one’s health. The relation between high blood pressure (which is also one of risk factors of cardiovascular diseases) and body mass index is confirmed. Confirmed/established probable predictors of heart problems: age, BMI, high blood sugar, increased salt consumption.

A significant limitation is the reliance on self-reported data for behavioral factors and medical history. This introduces the risk of social desirability bias, where respondents may underreport stigmatized behaviors such as alcohol consumption or tobacco use, and recall bias regarding medical diagnoses. Furthermore, the cross-sectional nature of the STEPS survey precludes causal inference, limiting our ability to distinguish between risk factors that caused the disease and lifestyle changes adopted post-diagnosis. We were also not able to assess the influence of socioeconomic factors (income, relocation due to war) because these data were too vague or missing. Another limitation is that the data were actual as for 2019 and no subsequent STEPS surveys were conducted to check the dynamics. Since the russian invasion and beginning of the full-scale war, many factors might have changed due to increased background and active stress level, uncertainty and massive population movements.

REFERENCES

[1] World Health Organization, "Noncommunicable diseases," World Health Organization, Dec. 23, 2024, [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.

[2] H. C. Gooding et al., "Challenges and Opportunities for the Prevention and Treatment of Cardiovascular Disease Among Young Adults: Report from a National Heart, Lung, and Blood Institute Working Group," Journal of the American Heart Association, vol. 9, no. 19, Oct. 2020, [Online]. Available: https://doi.org/10.1161/jaha.120.016115.

[3] "STEPS: prevalence of noncommunicable disease risk factors in Ukraine 2019," WHO Regional Office for Europe, Copenhagen, 2020, [Online]. Available: https://phc.org.ua/sites/default/files/users/user90/2019_STEPS_report_eng.pdf, [Accessed: Feb. 03, 2026].

[4] "Data Bank | State Statistics Service of Ukraine," Stat.gov.ua, 2026, [Online]. Available: https://stat.gov.ua/uk/explorer?md5=22fb4077f27cf51088adc40eb6e2caec, [Accessed: Feb. 11, 2026].

[5] K. E. Warner, "Kids No Longer Smoke Cigarettes. Why Aren't We Celebrating?" American Journal of Public Health, vol. 114, no. 11, pp. 1191-1194, Nov. 2024, [Online]. Available: https://doi.org/10.2105/AJPH.2024.307841.

[6] R. Dahal, A. Bhattarai, and K. Adhikari, "Age and sex-related patterns of electronic cigarette use in the general population: Supporting a de novo substance use pattern," Population Medicine, vol. 4, no. December, pp. 1-10, Dec. 2022, [Online]. Available: https://doi.org/10.18332/popmed/157270.

[7] T. Zhang et al., "Rate of change in body mass index at different ages during childhood and adult obesity risk," Pediatric Obesity, vol. 14, no. 7, p. e12513, Jan. 2019, [Online]. Available: https://doi.org/10.1111/ijpo.12513.

[8] L. Song et al., "Body Mass Index is Associated with blood pressure and vital capacity in medical students," Lipids in Health and Disease, vol. 22, no. 1, Oct. 2023, [Online]. Available: https://doi.org/10.1186/s12944-023-01920-1.

[9] L. Yuan et al., "Dose-response relationship between body mass index and hypertension: A cross-sectional study from Eastern China," Preventive Medicine Reports, vol. 46, p. 102852, Aug. 2024, [Online]. Available: https://doi.org/10.1016/j.pmedr.2024.102852.

[10] S. S. Khan et al., "Association of Body Mass Index with Lifetime Risk of Cardiovascular Disease and Compression of Morbidity," JAMA Cardiology, vol. 3, no. 4, pp. 280-287, Apr. 2018, [Online]. Available: https://doi.org/10.1001/jamacardio.2018.0022.

[11] M. Piano, "Alcohol's Effects on the Cardiovascular System," Alcohol Research: Current Reviews, vol. 38, no. 2, p. 219, 2017, [Online]. Available: https://doi.org/10.35946/arcr.v38.2.06.

[12] M. Rahman et al., "Cardiovascular Effects of Smoking and Smoking Cessation: A 2024 Update," Global Heart, vol. 20, no. 1, Jan. 2025, [Online]. Available: https://doi.org/10.5334/gh.1399.