Results:

H1:
An ordinal logistic regression was conducted to predict smoking status (1 = never, 2 = former, 3 = current) from age and sex. Age was included as a continuous variable, and sex was included as a covariate.
Results indicated that age was significantly associated with smoking status, with older age being linked to a lower likelihood of current smoking. Specifically, the first age spline coefficient showed a negative association with current smoking (ns(age_con, df = 3)1: β = -0.192, p < .001). Another spline coefficient was not significant (ns(age_con, df = 3)2: β = 0.035, p = .183), while the third spline coefficient was strongly negative and highly significant (ns(age_con, df = 3)3: β = -1.481, p < .001), reflecting the nonlinear relationship between age and smoking status.
Sex was also highly significant: males had a substantially higher likelihood of being current smokers (sexMale: β = 3.520, p < .001).
The threshold estimates for the ordinal regression were 3.009 for “never smoked” vs. “former smoker” and 4.247 for “former smoker” vs. “current smoker.”
Visual inspection of model-based predicted probabilities suggested a non-linear age pattern, with the highest probability of current smoking observed around 40 years of age, followed by a gradual decline.

H2:
A multiple linear regression analysis was conducted to examine differences in body weight across smoking status while adjusting for age and sex. The overall model was statistically significant, F(4, 991,341) = 140,000, p < .001, explaining 36.1% of the variance in body weight (R^2 = .361). Age was negatively associated with body weight, such that weight decreased by approximately 0.13 kg per year of age, b = –0.13, SE < 0.01, p < .001. Sex was a strong predictor, with males weighing on average 13.75 kg more than females, b = 13.75, SE = 0.03, p < .001. Smoking status was significantly associated with body weight, showing both a linear trend, b = 0.37, SE = 0.02, p < .001, and a quadratic trend, b = –0.75, SE = 0.02, p < .001, indicating a non-linear relationship across smoking categories.
Body weight differed systematically by smoking status after adjusting for age and sex. The significant quadratic trend suggests that former smokers tend to have higher body weight compared to both never smokers and current smokers. In addition, older individuals weighed slightly less than younger individuals, whereas males were substantially heavier than females. Overall, these findings support the hypothesis that current smokers have lower mean body weight, while highlighting that weight gain is particularly pronounced among former smokers.
Effect sizes were examined using partial eta squared (η^2_p) to evaluate the practical relevance of predictors. Sex showed a large effect on body weight (η^2_p = .34), indicating that a substantial proportion of variance in weight was attributable to sex differences. Age showed a small-to-moderate effect (η^2_p = .06), whereas smoking status exhibited a very small but meaningless effect (η^2_p = .001). Although all predictors reached statistical significance due to the very large sample size, only sex demonstrated a practically meaningful association with body weight. The effects of age and smoking status were statistically reliable but negligible in magnitude.

H3:
Results
A multivariate analysis of covariance (MANCOVA) was conducted to examine the association between smoking status and lipid profile (HDL and LDL cholesterol), adjusting for age and sex. Using Pillai’s trace as the multivariate test statistic, a significant multivariate effect of smoking status on the combined lipid outcomes was observed, V = .030, F(4, 1,982,682) = 7,562, p < .001. Age, V = .017, F(2, 991,340) = 8,430, p < .001, and sex, V = .042, F(2, 991,340) = 21,831, p < .001, also showed significant multivariate effects.
Follow-up linear regression analyses were conducted separately for HDL and LDL cholesterol to examine the direction and magnitude of these effects while additionally adjusting for drinking status (DRK_YN).
Smoking status was significantly associated with HDL cholesterol, F(2, 991,293) = [value], p < .001. Both a significant linear trend, b = –1.61, SE = 0.03, p < .001, and a significant quadratic trend, b = –0.83, SE = 0.03, p < .001, were observed, indicating a non-linear association across smoking categories. Increasing age was associated with lower HDL cholesterol, b = –0.12, SE < 0.01, p < .001. Males had substantially lower HDL levels than females, b = –10.06, SE = 0.04, p <

.001. Participants who reported drinking alcohol showed higher HDL cholesterol levels, b = 4.75, SE = 0.03, p < .001. The model explained 12.9% of the variance in HDL cholesterol (R^2 = .129). Smoking status was also significantly associated with LDL cholesterol, F(2, 991,293) = [value], p = .007 (Holm-adjusted). A small but significant linear trend was observed, b = −0.23, SE = 0.08, p = .003, whereas the quadratic trend was not significant, b = 0.07, SE = 0.08, p = .36. Increasing age was associated with slightly higher LDL cholesterol, b = 0.05, SE < 0.01, p < .001. Male sex was associated with marginally higher LDL levels, b = 1.05, SE = 0.09, p < .001, whereas alcohol consumption was associated with lower LDL cholesterol, b = −3.03, SE = 0.08, p < .001. Overall, the model explained a small proportion of variance in LDL cholesterol (R^2 = .003).

To account for multiple testing across the two lipid outcomes, p-values for the effect of smoking status were adjusted using the Holm method. After adjustment, the association between smoking status and HDL cholesterol remained highly significant (p < .001), and the association with LDL cholesterol also remained statistically significant (p = .007).

Smoking status showed a significant association with lipid profile at both the multivariate and univariate levels. The strongest and most consistent effects were observed for HDL cholesterol, where smoking status was associated with a pronounced non-linear pattern, suggesting that HDL levels differ substantially across smoking categories, with particularly unfavorable profiles among current smokers. In contrast, the association between smoking status and LDL cholesterol was statistically significant but small in magnitude, indicating limited clinical relevance despite the large sample size. Age and sex were important covariates, with older age associated with lower HDL and slightly higher LDL cholesterol, and males exhibiting lower HDL and marginally higher LDL levels compared to females. Overall, these findings suggest that smoking status is more strongly related to HDL cholesterol than to LDL cholesterol, highlighting differential effects of smoking on components of the lipid profile. Visualization of the effects showed a big difference in how smoking status affects HDL and LDL cholesterol in men and women differently. Female smokers and former smokers exhibited a much lower HDL cholesterol than non smokers whereas this effect could not be observed in males.