Full Article:

Theoretical Framework:

Cigarette smoking remains a leading cause of preventable morbidity and mortality worldwide, with well-documented impacts on cardiovascular and metabolic health (World Health Organization [WHO], 2023). Despite global declines in smoking prevalence, significant demographic and behavioral differences persist, emphasizing the need to understand how smoking status interacts with age, sex, and physiological parameters such as body weight and lipid metabolism (Albanes et al., 1987; Wakabayashi, 2008).

Smoking prevalence tends to decline with age, as older adults are more likely to have quit or avoided smoking altogether (Albanes et al., 1987; Bobo et al., 2018). Age-related reductions in nicotine dependence and health-motivated cessation may explain the observed inverse relationship between age and current smoking status, while sex differences remain salient — men generally exhibit higher smoking rates than women across age groups (Plurphanswat & Rodu, 2014; Mammas et al., 2003).

Metabolic consequences of smoking have also been widely documented. Nicotine exposure influences energy balance by suppressing appetite and increasing metabolic rate, resulting in lower mean body weight among current smokers compared to nonsmokers (Akbartabartoori et al., 2005; Sneve & Jorde, 2008). Conversely, smoking cessation is frequently followed by modest weight gain, which may deter attempts to quit (Albanes et al., 1987).

Furthermore, smoking is associated with adverse alterations in lipid metabolism, increasing cardiovascular risk. Current smokers exhibit higher low-density lipoprotein cholesterol (LDL-C) and lower high-density lipoprotein cholesterol (HDL-C) compared to nonsmokers, even when controlling for age, sex, and alcohol use (Li et al., 2018; Zaid et al., 2018; Wakabayashi, 2008). Nicotine-induced oxidative stress and lipid peroxidation may underlie these dyslipidemic effects (Phillips et al., 1981).

Taken together, existing evidence indicates that smoking behavior exerts measurable effects on physiological outcomes across multiple domains. The present study aims to examine the relationship between smoking status and (a) age, (b) body weight, and (c) lipid profiles, controlling for relevant demographic and lifestyle covariates. Specifically, it hypothesizes that:

(H1) older age is associated with a lower likelihood of current smoking;

(H2) current smokers exhibit lower mean body weight than nonsmokers; and

(H3) smokers have higher LDL-C and lower HDL-C compared to nonsmokers, controlling for age, sex, and alcohol use.

Methods:

This study uses a cross-sectional, correlational design based on secondary data from the Korean National Health Insurance Service (NHIS), an open-access administrative dataset covering adults aged 20–85 years from the Korean population. No primary data collection was conducted. The analytic sample comprises 991,299 individuals with complete information on smoking status, age, anthropometrics, lipid parameters, and relevant covariates.

The pre-analysis focused on data preparation, descriptive inspection, and systematic testing of model assumptions prior to hypothesis testing. Smoking status was recoded into three ordered categories (never smoked, former smoker, current smoker). Body mass index (BMI) was computed from measured height and weight. Distributions of all variables were examined visually (histograms, Q–Q plots), and no missing values were detected in the analytic dataset. Extremely implausible lipid values (LDL > 1000 mg/dL; HDL > 500 mg/dL) were excluded for analyses involving lipid metabolism, as these values are physiologically unlikely and would unduly influence regression estimates.

For H1 (age and smoking status), ordinal regression models were evaluated, including tests of linearity using spline terms and likelihood-ratio comparisons. A spline specification for age provided superior fit, while the proportional-odds assumption was met. For H2 (smoking status and body weight), linear regression assumptions were assessed via residual diagnostics, Q–Q plots, variance inflation factors, and homoscedasticity checks using hexbin residual plots; all assumptions were adequately satisfied. For H3 (smoking status and lipid profiles), linearity, residual normality, and multicollinearity were examined separately for LDL and HDL cholesterol, adjusting for age, sex, and alcohol consumption. Given the very large sample size, formal

normality tests were intentionally avoided, as they are known to be overly sensitive; instead, visual diagnostics guided model adequacy.
Overall, the pre-analysis confirms that the data structure and statistical assumptions are appropriate for testing the proposed hypotheses: (H1) decreasing likelihood of current smoking with increasing age, (H2) lower body weight among current smokers compared to non-smokers and former smokers, and (H3) higher LDL and lower HDL cholesterol levels among smokers. The analyses are explicitly observational and do not permit causal inference, but they are well-suited to identifying robust population-level associations between smoking behavior and key physiological indicators in a Korean adult population.

Results:

H1
An ordinal logistic regression was conducted to predict smoking status (1 = never smoked, 2 = former smoker, 3 = current smoker) from age and sex. Age was modeled using a natural spline function (df = 3) to account for potential non-linear effects, and sex was included as a covariate. A likelihood-ratio test comparing the linear age model with the spline model indicated that the non-linear specification provided a significantly better fit, $\Delta\chi^2(2) = 8{,}729$, $p < .001$. The proportional odds assumption was met, as indicated by a non-significant nominal test.
Results showed a significant non-linear association between age and smoking status. The first spline component was negatively associated with smoking status, $\beta = -0.192$, SE = 0.011, $z = -18.20$, $p < .001$, whereas the second spline component was not statistically significant, $\beta = 0.035$, SE = 0.026, $z = 1.33$, $p = .183$. The third spline component showed a strong negative association, $\beta = -1.481$, SE = 0.017, $z = -86.66$, $p < .001$, confirming a pronounced non-linear age effect.
Sex was a highly significant predictor, with males having substantially higher odds of belonging to higher smoking categories compared to females, $\beta = 3.520$, SE = 0.007, $z = 508.48$, $p < .001$. The estimated threshold parameters were 3.009 for the transition from never smokers to former smokers and 4.247 for the transition from former smokers to current smokers.
Visual inspection of model-based predicted probabilities indicated a non-linear age pattern, with the highest probability of current smoking observed around 40 years of age, followed by a gradual decline at older ages. This pattern was more pronounced among males.

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 emerged as 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 also significantly associated with body weight, showing both a significant linear trend, $b = 0.37$, SE = 0.02, $p < .001$, and a significant quadratic trend, $b = -0.75$, SE = 0.02, $p < .001$, indicating a non-linear relationship across smoking categories. Former smokers exhibited higher body weight compared to both never smokers and current smokers.
Effect sizes were examined using partial eta squared (η^2_p) to assess practical relevance. Sex showed a large effect on body weight ($\eta^2_p = .34$), whereas age showed a small-to-moderate effect ($\eta^2_p = .06$). Smoking status exhibited a very small effect ($\eta^2_p = .001$). Although all predictors reached statistical significance due to the large sample size, only sex demonstrated a practically meaningful association with body weight.

H3
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 was observed, $V = .041$, $F(4, 1{,}982{,}588) = 10{,}325$, $p < .001$. Significant multivariate effects were also found for age, $V = .021$, $F(2, 991{,}293) = 10{,}868$, $p < .001$, and sex, $V = .056$, $F(2, 991{,}293) = 29{,}604$, $p < .001$.
Follow-up linear regression analyses were conducted separately for HDL and LDL cholesterol, additionally adjusting for alcohol consumption (DRK_YN).

For HDL cholesterol, smoking status was significantly associated with HDL levels, $F(2, 991{,}293) = 1{,}388$, $p < .001$. 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. Increasing age was associated with lower HDL cholesterol, $b = -0.12$, $SE < 0.01$, $p < .001$, and males had substantially lower HDL levels than females, $b = -10.06$, $SE = 0.04$, $p < .001$. Alcohol consumption was associated with higher HDL cholesterol, $b = 4.75$, $SE = 0.03$, $p < .001$. The model explained 12.9% of the variance in HDL cholesterol ($R^2 = .129$).
For LDL cholesterol, smoking status was also significantly associated with LDL levels, $F(2, 991{,}293) = 6.07$, $p = .002$, and remained significant after Holm correction ($p = .007$). 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, $b = -3.03$, $SE = 0.08$, $p < .001$. The model explained a very small proportion of variance in LDL cholesterol ($R^2 = .003$).
Overall, smoking status showed a statistically significant association with lipid profile at both the multivariate and univariate levels. However, the association was substantially stronger for HDL cholesterol than for LDL cholesterol. Visualization suggested sex-specific patterns, with smoking status being more strongly related to HDL cholesterol among women than among men, whereas corresponding differences in LDL cholesterol were comparatively small.

Discussion:

This study demonstrates that smoking behavior varies nonlinearly across the lifespan, with the highest probability of current smoking observed in mid-adulthood and a marked decline thereafter. Sex emerged as the strongest determinant of smoking status, body weight, and lipid profiles, consistently exceeding the magnitude of smoking-related effects. While smoking status was statistically associated with body weight and LDL cholesterol, these effects were negligible in practical terms, emphasizing the limited clinical relevance of these associations. In contrast, smoking status showed a robust and meaningful association with HDL cholesterol, supporting its role as a key behavioral determinant of lipid-related cardiovascular risk.

Limitations:

Several limitations should be considered. First, the cross-sectional design precludes causal inference, particularly regarding age-related changes in smoking behavior and health outcomes. Second, smoking status was assessed categorically without information on smoking intensity, duration, or cessation history, which may obscure dose–response relationships. Third, despite extensive covariate adjustment, residual confounding by unmeasured lifestyle factors (e.g., diet, physical activity) cannot be excluded. Finally, the very large sample size increases statistical power such that trivial effects achieve statistical significance, necessitating cautious interpretation with emphasis on effect sizes rather than p-values.

References:

1. Akbartabartoori, M., Lean, M. E. J., & Hankey, C. R. (2005). Relationships between cigarette smoking, body size and body shape. International Journal of Obesity, 29(12), 1473–1479.
2. Albanes, D., Jones, D. Y., & Micozzi, M. S. (1987). Associations between smoking and body weight in the US population: analysis of NHANES II. American Journal of Public Health, 77(4), 439–444.
3. Bobo, F. T., Thanasekaran, P., Joice, A., Yadecha, B., & Alebel, A. (2018). Prevalence of tobacco smoking and factors associated with the initiation of smoking among university students in Ethiopia: A cross-sectional study. BMC Public Health, 18(1), 132.
4. Li, X. X., Zhao, Y., Huang, L. X., Xu, H. X., Liu, X. Y., & Yang, J. J. (2018). Effects of smoking and alcohol consumption on lipid profile in male adults in northwest rural China. Public Health, 157, 7–13.

5. Mammas, I. N., Bertsias, G. K., & Linardakis, M. (2003). Cigarette smoking, alcohol consumption, and serum lipid profile among medical students in Greece. European Journal of Public Health, 13(3), 278–282.
6. Phillips, N. R., Havel, R. J., & Kane, J. P. (1981). Levels and interrelationships of serum and lipoprotein cholesterol and triglycerides: Association with adiposity and consumption of ethanol and tobacco. Arteriosclerosis, 1(1), 13–24.
7. Plurphanswat, N., & Rodu, B. (2014). The association of smoking and demographic characteristics on body mass index and obesity among adults in the US, 1999–2012. BMC Obesity, 1(1), 18.
8. Sneve, M., & Jorde, R. (2008). A cross-sectional study on the relationship between body mass index and smoking, and longitudinal changes in body mass index in relation to change in smoking status: The Tromsø Study. Scandinavian Journal of Public Health, 36(4), 397–403.
9. Wakabayashi, I. (2008). Associations of alcohol drinking and cigarette smoking with serum lipid levels in healthy middle-aged men. Alcohol and Alcoholism, 43(3), 274–280.
10. World Health Organization. (2023). WHO global report on trends in prevalence of tobacco use 2000–2025 (5th ed.). Geneva, Switzerland: WHO Press.
11. Zaid, M., Miura, K., Okayama, A., & Nakagawa, H. (2018). Associations of high-density lipoprotein particle and cholesterol with alcohol intake, smoking, and BMI: The INTERLIPID study. Circulation Journal, 82(10), 2501–2510.