The Relationship Between Life's simple 7 and Generalized Anxiety Disorder: A Study Based on KNHANES and Machine Learning

Keywords
KNHANES, Machine Learning, Generalized Anxiety Disorder, Life's Simple 7

Abstract

Introduction
This study aimed to investigate the relationship between LS7 scores and GAD risk in a nationally representative Korean population.

Material and methods
This study included 12,366 participants from 2001, 2005, 2007–2023 cycles of the Korean National Health and Nutrition Examination Survey (KNHANES). Propensity score matching (1:1) balanced baseline characteristics between GAD and non-GAD groups. Logistic regression analysis examined the association between LS7 and GAD across three models. XGBoost combined with SHAP analysis identified key variables associated with GAD. Restricted cubic spline analysis explored potential nonlinear dose response relationships. Sensitivity analysis using complete case analysis was conducted to assess the robustness of the findings.

Results
After propensity score matching (N = 1,126), the LS7 score was significantly lower in the GAD group than in the Normal group (9.35 ± 1.77 vs. 9.65 ± 1.72, P = 0.004). Logistic regression demonstrated a significant inverse association between LS7 and GAD risk (Model 1: OR = 0.88, 95% CI: 0.82–0.95, P < 0.001), with a significant dose-response trend (P for trend = 0.002). SHAP analysis identified LS7 as a key contributing feature (mean SHAP value = 0.044). Restricted cubic spline analysis confirmed a significant overall inverse association (P for overall < 0.001) with no significant nonlinear relationship detected (P for nonlinear = 0.426), and a notable inflection point observed at an LS7 score of 10. Sensitivity analysis yielded consistent results.

Conclusions
A negative correlation exists between LS7 and generalized anxiety. Optimizing the risk factors covered by LS7 may help improve and prevent GAD.

The Relationship Between Life's simple 7 and Generalized Anxiety Disorder: A Study Based on KNHANES and Machine Learning

Background： Generalized Anxiety Disorder (GAD) is a prevalent mental health condition associated with significant personal and societal burden. Although Life's Simple 7 (LS7) is an established cardiovascular health metric, its association with GAD risk remains insufficiently studied.

Objective: This study aimed to investigate the relationship between LS7 scores and GAD risk in a nationally representative Korean population.

Methods: This study included 12,366 participants from the 2001, 2005, and 2007–2023 cycles of the Korean National Health and Nutrition Examination Survey (KNHANES). Propensity score matching (1:1) was employed to balance baseline characteristics between GAD and non-GAD groups. Logistic regression analysis was used to examine the association between LS7 and GAD across three models. The XGBoost algorithm combined with SHAP analysis was applied to identify key variables associated with GAD. Restricted cubic spline analysis was performed to explore potential nonlinear dose-response relationships. Sensitivity analysis using complete case analysis was conducted to assess the robustness of the findings.

Results: After propensity score matching (N = 1,126), the LS7 score was significantly lower in the GAD group than in the Normal group (9.35 ± 1.77 vs. 9.65 ± 1.72, $P = 0.004$). Logistic regression demonstrated a significant inverse association between LS7 and GAD risk (Model 1: OR = 0.88, 95% CI: 0.82–0.95, $P < 0.001$), with a significant dose-response trend (P for trend = 0.002). SHAP analysis identified LS7 as a key contributing feature (mean SHAP value = 0.044). Restricted cubic spline analysis confirmed a significant overall inverse association (P for overall < 0.001) with no significant nonlinear relationship detected (P for nonlinear = 0.426), and a notable inflection point observed at an LS7 score of 10. Sensitivity analysis yielded consistent results.

Conclusion: A negative correlation exists between LS7 and generalized anxiety. Optimizing the risk factors covered by LS7 may help improve and prevent GAD.

Keywords: Generalized Anxiety Disorder, Life's Simple 7, KNHANES, Machine Learning

1. Introduction

Generalized anxiety disorder (GAD) is a prevalent mental condition marked by chronic, unfocused apprehension and tension, often associated with symptoms like concentration difficulties, memory deficits, muscular tension, irritability, and sleep abnormalities [1][1]. Post-pandemic, as life pressures continue to mount, the prevalence of GAD has risen significantly. Epidemiological data indicates that the prevalence of GAD has reached as high as 7.3%. This situation not only severely compromises individuals' physical and mental health but also imposes a heavy burden on personal finances, family relationships, and other aspects of life [2]. Life's Simple 7 (LS7) is a health guideline developed by the American Heart Association based on cardiovascular health indicators, including four behavioral factors: Body Mass Index (BMI), physical activity, diet, and smoking. Initially designed to assess cardiovascular health, ongoing research has revealed associations between LS7 and non-cardiovascular conditions such as dementia, depression, and cognitive impairment [3]. Research demonstrates that higher LS7 scores effectively reduce dementia risk and slow cognitive decline. This effect may stem from enhanced social engagement, reduced depressive symptoms, and improved cognitive function [4].

Furthermore, numerous studies have identified associations between body mass index, serum cholesterol, blood pressure, and anxiety/depression [5][6][7]. However, to date, no comprehensive indicator integrating these influencing factors has been developed to investigate their relationship with anxiety. Therefore, this study utilizes data from the Korean National Health and Nutrition Examination Survey (KNHANES) to explore the association between LS7 and anxiety. The aim is to identify indicators that may influence anxiety risk factors and enhance anxiety prevention strategies.

2. Materials and Methods

2.1 Study Population

2.1.2 KNHANES

KNHANES is an ongoing, cross-sectional, nationally representative survey in Korea. For our study, we utilized KNHANES data from 2001, 2005, and 2007–2023, as these cycles contained complete GAD assessment data for participants [8]. KNHANES is a nationally representative cross-sectional survey conducted periodically by the Korea Disease Control and Prevention Agency (KDCA), assessing the health and nutritional status of the non-institutionalized Korean population through stratified, multistage probability sampling. It encompasses health interviews, health examinations, and nutritional surveys. The survey has been approved by the Institutional Review Board of the KDCA and reviewed by the Institutional Review Board of Seoul National University (Ethics No.: E2211/004-003). As this study involved only secondary analysis of publicly available de-identified data, no additional ethics approval was required.

From the initial cohort of 205,398 participants, we excluded cases for the following reasons: (1) Removal of data samples under 20 years of age (N=49,146). (2) Removal of samples lacking GAD data (N=139,776). (3) Removal of samples missing LS7 indicators and cutoff values (N=4,110). The final study cohort comprised 12,366 participants, including 563 GAD cases and 11,803 controls. The exclusion flowchart is shown in Figure 1.

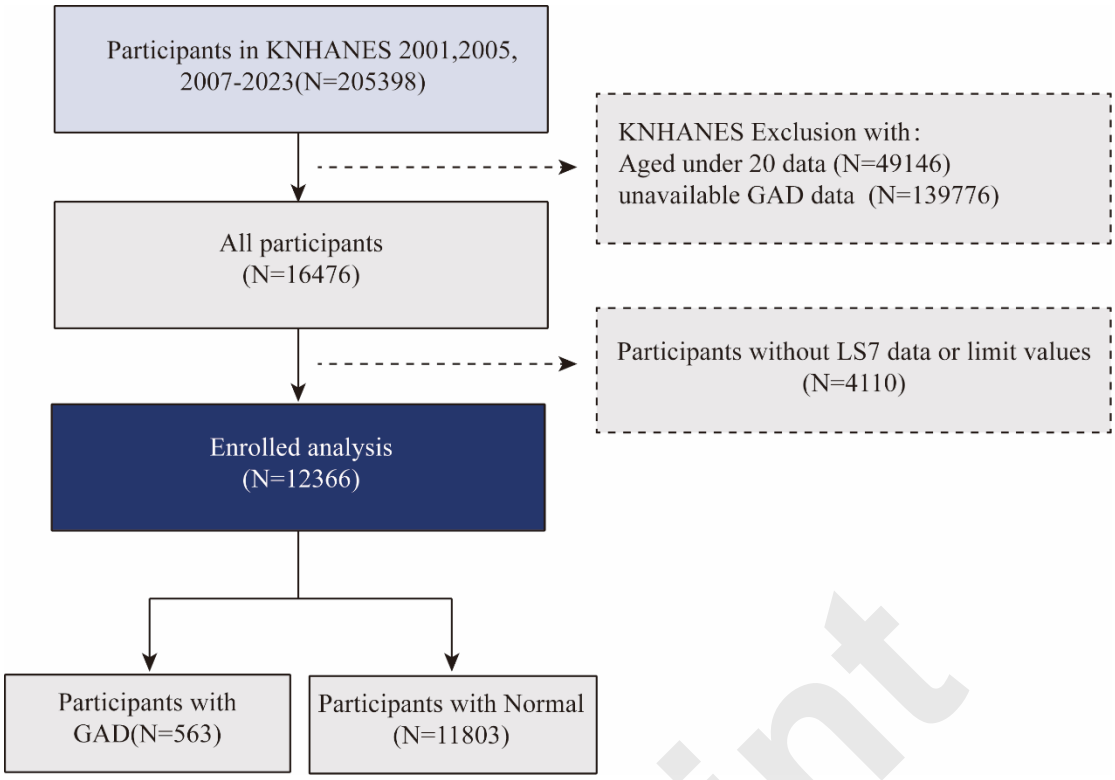


Fig. 1. Flowchart of the sample selection from the KNHANES

2.2 LS7 Indicator Measurement

LS7 comprises seven indicators: physical activity, dietary patterns, smoking habits, BMI, blood glucose, blood pressure, and cholesterol levels. Each indicator is categorized into three standards: ideal, moderate, and poor. Physical activity is measured by the frequency and duration of moderate-to-vigorous exercise (e.g., walking, running, swimming, yard work) over the past 30 days. Dietary patterns are assessed using the Healthy Eating Index-2015 (HEI-2015), which evaluates intake of 13 food groups. The HEI-2015 score is calculated by summing the daily intake scores for these 13 food groups [9]. The GAD-7 was used to assess generalized anxiety disorder symptoms. This seven-item self-report scale has been validated as a reliable anxiety screening tool and can measure anxiety severity. It is based on criteria from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition [10].

2.3. Diagnosis of GAD

GAD was diagnosed according to DSM-5 criteria, defined as excessive and uncontrollable anxiety persisting for at least 6 months, accompanied by at least three associated symptoms such as restlessness, fatigue, and difficulty concentrating, and not attributable to substances, other medical conditions, or other mental disorders [11]. Individuals with secondary anxiety due to medical conditions, substance use disorders, or who were pregnant or breastfeeding were excluded.

2.4 Covariates

Based on prior literature, several potential confounding factors were analyzed, including age, age group, gender, and household income relative to the poverty line. Age, age group, gender, and household income relative to the poverty line were determined using demographic variable files. Household income relative to the poverty line was categorized as low income (≤1.3), medium income (1.3–3.5), or high income (>3.5). Diabetes and hypertension were identified through

questionnaire surveys.

2.5 Statistical Methods

All data processing and analysis were performed using R software (version 4.1.1). Baseline characteristics were analyzed by describing continuous variables using mean and standard deviation (SD). Categorical variables were reported as sample counts and percentages. To examine differences in variable characteristics between GAD and Normal groups, we employed independent samples t-tests for continuous variables and chi-square tests for categorical variables, providing a comprehensive description of the entire population. For missing values in the data, if the proportion of missing data was less than 20%, multiple imputation was performed using the chained equations method with classification and regression trees (with the number of imputed datasets set to 5). Prior to imputation, observations with missing values in the outcome variable (GAD) and the primary exposure variable (LS7) were excluded. Variables with more than 20% missing data were also removed. The remaining covariates with missing proportions not exceeding 20% were imputed, including age group, gender, poverty income ratio (PIR), alanine aminotransferase (ALT), and white blood cell count, among other sociodemographic and biochemical indicators. Imputation was completed prior to propensity score matching to ensure complete covariate data for propensity score estimation.

Logistic regression was employed to analyze the potential association between LS7 and GAD. LS7 levels were incorporated into the regression model as both continuous and categorical variables. This facilitated the calculation of odds ratios (OR) and their corresponding 95% confidence intervals (95% CI). LS7 levels were introduced into the model as a continuous variable. The raw LS7 levels were divided into quartiles based on interquartile ranges and ranked from low to high. These quartiles were labeled as the first quartile (Q1), second quartile (Q2), third quartile (Q3), and fourth quartile (Q4). They were further evaluated as categorical variables. When LS7 levels were treated as categorical variables, the lowest quartile (Q1) served as the reference category. Three distinct statistical analysis groups were established: the Crude Model was unadjusted; Model 1 was adjusted for age group, gender, and poverty income ratio (PIR); Model 2 was adjusted for alanine aminotransferase (ALT) and white blood cell count. The three models represent parallel adjustment designs, each evaluating the influence of a different category of covariates. Restricted Cubic Splines (RCS) were then employed to examine whether a nonlinear relationship exists between LS7 and anxiety. Subgroup analyses were performed by age group and gender using multivariate logistic regression, with interaction effects between each subgroup and LS7 assessed by incorporating interaction terms into the model.

The XGBoost (Extreme Gradient Boosting) algorithm was applied as an exploratory tool to assess the relative contributions of various factors to GAD risk and identify key disease-associated features, rather than to develop a clinical predictive model[12][13]. XGBoost (Extreme Gradient Boosting) is an efficient Gradient Boosting Tree (GBT) algorithm belonging to ensemble learning. Based on decision tree models, XGBoost employs gradient boosting techniques to enhance predictive accuracy by integrating multiple weak learners (i.e., weak classifiers) into a single strong learner. We employed the XGBoost algorithm model to discuss the relative importance of indicators in influencing GAD. In this study, a learning rate of 0.08 was selected to balance model complexity and generalization capability. To prevent overfitting, the maximum depth was capped at 6, and the number of trees was set to 80% of the total sample size to achieve stable and accurate results.

Sensitivity analysis is a statistical method used to assess the robustness of research findings.

By altering key assumptions or parameters within a model, sensitivity analysis examines how sensitive the results are to these changes, thereby validating the reliability of conclusions. In this study, to ensure the robustness of the association between LS7 indicator levels and GAD, we employed a sensitivity analysis using direct deletion of missing values (NA values) rather than data imputation techniques. Statistical analysis was two-tailed, with $P < 0.05$ considered statistically significant. This study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cross-sectional studies, and the completed STROBE checklist is provided as supplementary material.

3. Results

3.1 Description of Baseline Characteristics of Patients

Baseline characteristics of the 1,126 participants, stratified by the presence of GAD, are presented in Table 1. The mean age was 48 ± 18 years, with 31.8% classified as old and 68.2% as young, and the cohort was 69.9% female. The distribution of participants across age groups (P=0.798), gender (P=0.649), and poverty-income ratio (PIR) categories (P=0.945) was balanced between the GAD (n=563) and Normal (n=563) groups, with no statistically significant differences observed. Similarly, no significant between-group differences were found for the mean levels of alanine aminotransferase (ALT) (22 ± 17 vs. 21 ± 18, p=0.405) or white blood cell count (6.03 ± 1.69 vs. 5.95 ± 1.71, P=0.412). However, a statistically significant difference was noted in the mean LS7 score, with the Normal group demonstrating a higher cardiovascular health score compared to the GAD group (9.65 ± 1.72 vs. 9.35 ± 1.77, P=0.004).

Table 1. Baseline characteristics of participants after propensity score matching.

Variable	Overall	GAD	Normal	*p*-value

	N = 1,126	N = 563	N = 563	
Age	48 (18)	48 (18)	48 (18)	0.822
Age Group				0.798
Old	358 (32%)	177 (31%)	181 (32%)	
Young	768 (68%)	386 (69%)	382 (68%)	
Gender				0.649
Female	787 (70%)	397 (71%)	390 (69%)	
Male	339 (30%)	166 (29%)	173 (31%)	
PIR				0.945
≤ 1.3	370 (33%)	184 (33%)	186 (33%)	
1.3-3.5	745 (66%)	374 (66%)	371 (66%)	
> 3.5	11 (1.0%)	5 (0.9%)	6 (1.1%)	
ALT	21 (17)	22 (17)	21 (18)	0.405
White blood cell count	5.99 (1.70)	6.03 (1.69)	5.95 (1.71)	0.412
LS7	9.50 (1.75)	9.35 (1.77)	9.65 (1.72)	0.004

181

3.2 Relationship Between GAD and LS7

In the unadjusted model (Model 1), each unit increase in the continuous LS7 score was associated with a 10% lower risk of disease (OR 0.90, 95% CI 0.85 to 0.97, *P*=0.004). This association remained significant after adjusting for demographic factors (Model 2: OR 0.88, 95% CI 0.82 to 0.95, *P*<0.001) and for laboratory markers (Model 3: OR 0.91, 95% CI 0.84 to 0.97, *P*=0.007). When analyzed categorically, significant inverse associations were observed for the highest quartiles (Q3 and Q4) compared to the reference (Q1 [4,8)). In Model 2, the odds of disease were significantly lower for Q3 (OR 0.63, 95% CI 0.42 to 0.96, *P*=0.032) and Q4 (OR 0.60, 95% CI 0.40 to 0.90, *P*=0.014). The trend test was significant across all models (Crude Model: *P* = 0.007; Model 1: *P* = 0.002; Model 2: *P* = 0.013), which suggests that higher LS7 scores may help protect against GAD (Table 2).

Table 2. The association between LS7 exposure and GAD

	Crude Model			Model1			Model2		
Characteristic	**OR**	**95%CI**	***P*-value**	**OR**	**95%CI**	***P*-value**	**OR**	**95%CI**	***P*-value**
LS7									
LS7 continuous	0.90	(0.85, 0.97)	0.004	0.88	(0.82, 0.95)	<0.001	0.91	(0.84, 0.97)	0.007
LS7 quantile									
Q1(< 8)	Ref	Ref		Ref	Ref		Ref	Ref	
Q2(8-10)	0.93	(0.64, 1.37)	0.719	0.90	(0.61, 1.33)	0.587	0.95	(0.64, 1.40)	0.784
Q3(10-11)	0.68	(0.46, 1.02)	0.062	0.63	(0.42, 0.96)	0.032	0.69	(0.46, 1.05)	0.087
Q4(≥ 11)	0.67	(0.46, 0.98)	0.039	0.60	(0.40, 0.90)	0.014	0.68	(0.45, 1.03)	0.069
P for trend			0.007			0.002			0.013

Note: The Crude Model was non-adjusted. Model 1 was adjusted for age group, gender, and poverty income ratio (PIR). Model 2 was adjusted for alanine aminotransferase (ALT) and white blood cell count.

3.3 Machine Learning Analysis of GAD-Related Factors Using the XGBoost Model

To further identify key factors associated with GAD, the XGBoost algorithm was applied to rank variable importance using SHAP analysis. SHAP analysis revealed that triglycerides (mean SHAP value = 0.103), white blood cell count (mean SHAP value = 0.097), alanine aminotransferase (ALT; mean SHAP value = 0.050), and LS7 (mean SHAP value = 0.044) were the most influential features associated with GAD (Fig. 2A). Furthermore, SHAP dependency plots indicated that higher LS7 values were associated with a negative contribution to the predicted probability of GAD, suggesting a protective role of LS7 against GAD (Fig. 2B).

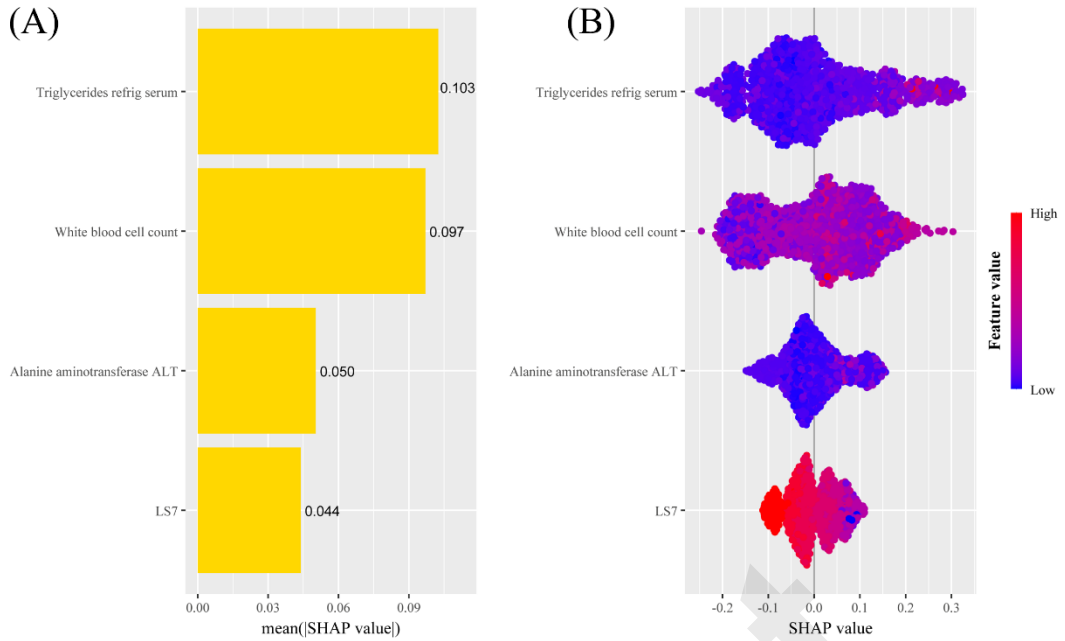


Fig 2. Clarification of machine learning models through SHAP summary and dependence plots

3.4 Stratified Statistics

Stratified results in Fig. 3A show that the GAD incidence rate gradually decreases across quartiles of LS7. The first quartile (Q1) had the highest incidence rate at 55.92%, while the fourth quartile (Q4) had the lowest at 45.87%. Fig. 3B displays GAD incidence stratified by age and PIR. Across all age groups, moderate PIR showed the highest GAD incidence, followed by low PIR. High PIR consistently exhibited the lowest GAD incidence across all age cohorts.

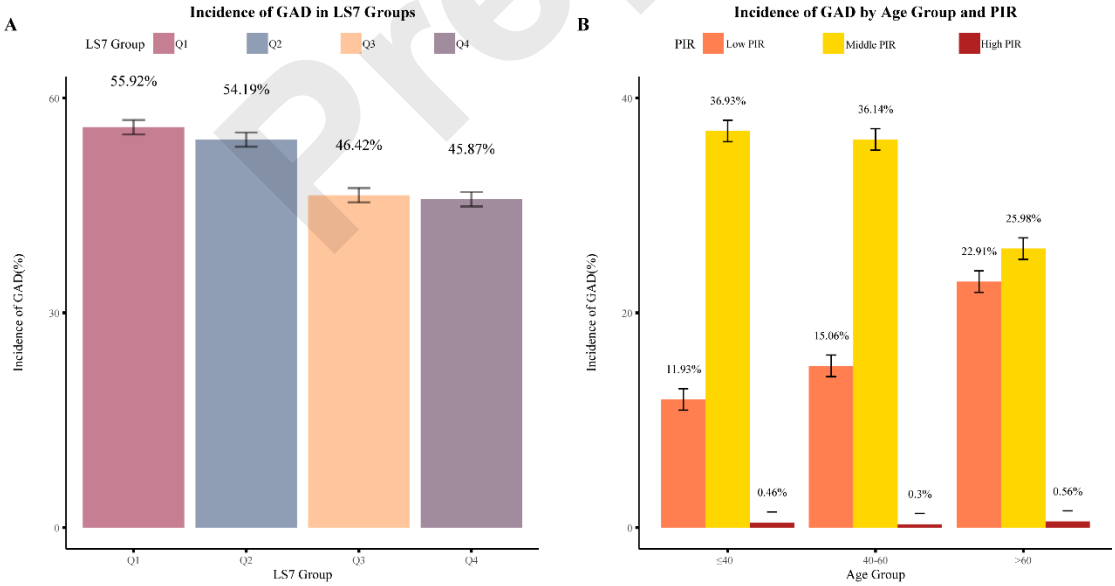


Fig. 3 (A) Bar graph depicting the incidence of GAD across different LS7 quartiles;(B) Comparative graph illustrating GAD incidence across age groups

3.5 RCS Curves for GAD and LS7 and Its Subgroups

Having confirmed a significant inverse association between LS7 scores and GAD risk

through logistic regression analysis, we further employed restricted cubic spline (RCS) analysis to explore whether this association followed a nonlinear dose-response pattern. RCS revealed a significant overall inverse association between LS7 scores and GAD risk in the total population (*P* for overall < 0.001), with no significant nonlinear relationship (*P* for nonlinear = 0.426) and a notable inflection point at an LS7 score of 10 (Fig. 4A). Subgroup analysis by gender indicated that this inverse association remained significant in females (*P* for overall < 0.001, *P* for nonlinear = 0.252), whereas no significant association was observed in males (*P* for overall = 0.121, *P* for nonlinear = 0.711), suggesting that the protective effect of higher LS7 scores against GAD may be more pronounced in females (Fig.4B).

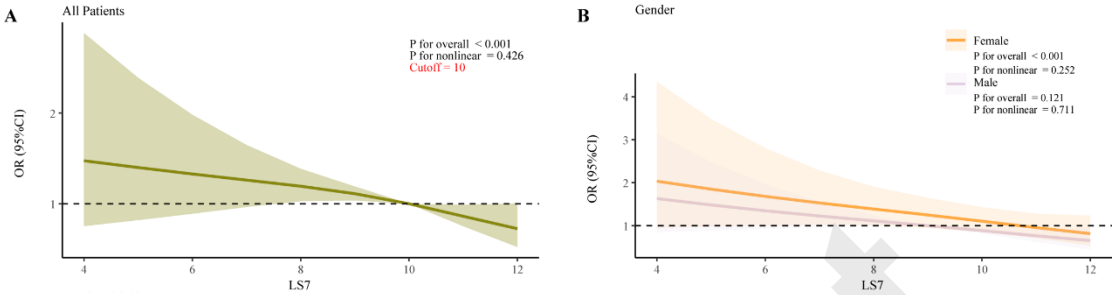


Fig. 4. Restricted cubic spline curves for the association between LS7 scores and GAD risk. (A) Overall population; (B) Stratified by gender (Female vs. Male);

3.6 Relationship Between LS7 and Subgroups Based on Baseline Characteristics

To examine whether the protective association between LS7 and GAD was consistent across different population subgroups, this study further analyzed the association between LS7 and GAD across subgroups defined by baseline characteristics using multivariate logistic regression (Fig. 5).In the overall analysis of 1,126 participants, a higher LS7 score was significantly associated with a lower risk of disease (OR 0.90, 95% CI 0.85-0.97, *P*=0.004). Subgroup analyses were conducted by age group and gender. The association was significant in the young group (OR 0.89, 95% CI 0.82-0.96, *P*=0.003) but not in the old group (OR 0.95, 95% CI 0.83-1.10, *P*=0.509), with no significant interaction by age (*P* for interaction = 0.401). Similarly, a significant inverse association was observed in females (OR 0.84, 95% CI 0.77-0.92, *P*<0.001) but not in males (OR 0.97, 95% CI 0.86-1.10, *P*=0.645). The interaction by gender approached but did not reach statistical significance (*P* for interaction = 0.058).

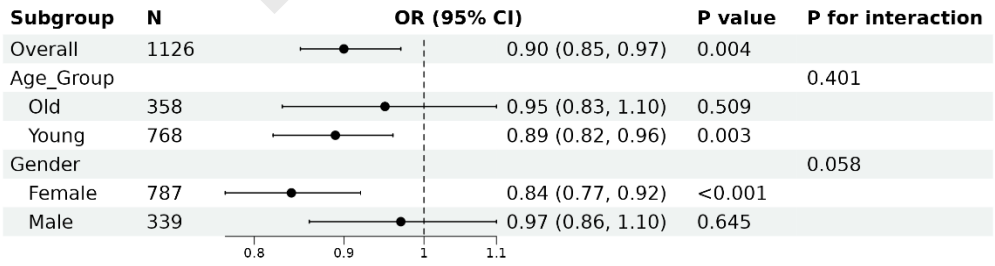


Fig. 5. Subgroup analyses of the relationship between LS7 Score and GAD.

3.7 Sensitivity Analysis

Sensitivity analysis using complete case analysis (excluding imputed data) yielded consistent results. The unadjusted model showed that for each one-unit increase in LS7 as a continuous

variable, the risk of GAD significantly decreased (OR = 0.91, 95% CI: 0.84–0.99, $P = 0.022$). After multivariable adjustment, this association remained significant (Model 1: OR = 0.90, 95% CI: 0.83–0.98, $P = 0.015$; Model 2: OR = 0.91, 95% CI: 0.83–0.99, $P = 0.032$). Quartile stratification analysis further confirmed this trend. Compared with Q1, participants in Q4 had a significantly lower risk of GAD in the Crude Model (OR = 0.59, 95% CI: 0.37–0.94, $P = 0.025$), Model 1 (OR = 0.55, 95% CI: 0.33–0.89, $P = 0.016$), and Model 2 (OR = 0.60, 95% CI: 0.36–0.99, $P = 0.045$). The trend test was significant across all models (Crude Model: $P = 0.011$; Model 1: $P = 0.007$; Model 2: $P = 0.018$) (Table S2). These findings indicate that the association between LS7 scores and GAD risk remains consistent across different analytical approaches, further supporting the robustness of the study conclusions.

4. Discussion

As a core comprehensive indicator for assessing cardiovascular health, the ideal state of LS7 not only correlates closely with improved anxiety symptoms but also demonstrates significant neuroprotective effects. Existing research indicates that when LS7 scores are at ideal or moderate levels, they show significant correlations with multiple brain structural parameters: specifically, a marked reduction in white matter hyperintensity volume ($\beta = -0.39$, $P < 0.001$) and a significant increase in total brain volume ($\beta = 0.14$, $P < 0.001$). This finding suggests that LS7 may exert its alleviating effect on anxiety-related neurodegeneration through the mediating mechanism of optimizing cerebrovascular health [14]. This is consistent with our finding that higher LS7 scores were significantly associated with reduced GAD risk (Model 1: OR = 0.886, 95% CI: 0.823–0.953), suggesting that optimization of cardiovascular health metrics may attenuate anxiety-related neurodegeneration through cerebrovascular pathways. Subsequent research demonstrates that behavioral intervention elements within the LS7 framework—such as consistent physical activity and nutritious food habits—are markedly correlated with the reduction of anxiety-like behaviors.

Another study revealed that high-intensity stress induces excessive activation and sensitization of serotonin (5-HT) neurons in the dorsal raphe nucleus (DRN) region of rodent brains, thereby triggering anxiety and depression symptoms associated with human stress. Appropriate physical exercise, however, inhibits 5-HT neuron activity by inducing neural circuit plasticity (enhancing 5-HT 1A auto receptor function and reducing 5-HT 2C receptor sensitivity), thereby attenuating anxiety symptoms induced by sharp increases in serotonin levels [15]. A bidirectional Mendelian randomization analysis examining the effects of physical activity, sedentary behavior, anxiety, depression, and well-being revealed that sedentary individuals had a 2.59-fold higher probability of anxiety compared to non-sedentary individuals, while regular physical activity promoted well-being [16]. These findings support the observed dose-response relationship between LS7 scores and GAD risk in our study (P for trend = 0.0024), wherein physical activity as a core LS7 component may contribute to GAD prevention through serotonergic pathway modulation. This effect may be attributed to post-exercise increases in β-endorphins, endocannabinoids, monoamines, and BDNF [17][18]. Recent updates to cardiovascular health frameworks further underscore the clinical relevance of our findings. The American Heart Association has expanded Life's Simple 7 to Life's Crucial 9, incorporating sleep health and social determinants of health as additional metrics. Notably, the ILEP-SMILE

algorithm has identified stress, which is strongly correlated with anxiety, as a critically important cardiovascular risk factor, suggesting that effective stress management may help maintain cardiovascular health [19]. Our findings that higher LS7 scores are inversely associated with GAD risk imply that jointly optimizing cardiovascular health behaviors and addressing anxiety and stress may confer synergistic cardioprotective benefits.

Not only exercise, but healthy eating can also effectively alleviate anxiety symptoms.Within the LS7 framework, dietary quality is one of the seven core components. Poor dietary patterns, particularly those rich in refined carbohydrates, have been shown to elevate C-reactive protein levels [20] and decrease brain-derived neurotrophic factor levels [21] , causing neuronal oxidative stress degeneration, which may contribute to GAD development [22]. Another study indicates that nutrients such as B vitamins, vitamin C, magnesium, and zinc can alleviate metabolic disorders and abnormal hormone secretion caused by long-term chronic stress. They do this by protecting the integrity and stability of neuronal membrane structures through the production of neurotransmitters like serotonin, norepinephrine, and dopamine, thereby reducing the risk of anxiety disorders [23]. These mechanisms may partly explain the inverse association between LS7 scores and GAD risk observed in our study, as dietary optimization within the LS7 framework could attenuate neuroinflammation and support neurotransmitter homeostasis, thereby reducing GAD susceptibility.

At the population level, ideal LS7 scores correlate with lower cerebral small vessel disease (CSVD) risk (cOR=0.73). Research indicates that white matter hyperintensity lesions caused by CSVD disrupt cortical and subcortical networks regulating emotional control, inducing anxiety, depression, apathy, impulsivity, and other emotional disorders in CSVD patients [24]. This is consistent with our finding that higher LS7 scores were associated with significantly reduced GAD risk (Q4 vs Q1: OR = 0.604, 95% CI: 0.399–0.909), suggesting that LS7-mediated reduction in CSVD burden may represent one pathway through which cardiovascular health optimization attenuates GAD susceptibility. Healthy lifestyle habits can reduce the incidence of cerebrovascular disease, thereby indirectly lowering anxiety prevalence.Under normal conditions, only minimal amounts of the pro-inflammatory cytokine IL-6 can cross the BBB. However, elevated IL-6 levels in cerebrospinal fluid have been observed in patients with severe anxiety and depression, a finding associated with increased BBB permeability [25]. Menard et al. observed significantly reduced Claudin-5 protein levels in the BBB of anxiety-prone rats compared to normal rats [26]. Furthermore, IL-6 decreases Claudin-5-related mRNA concentrations, inhibits Claudin-5 protein synthesis, increases BBB permeability, and induces anxiety and depressive symptoms [27]. Notably, several LS7 components—including physical activity, healthy diet, and blood pressure control—have been shown to maintain BBB integrity by suppressing IL-6-mediated inflammatory cascades, which may contribute to the protective association between higher LS7 scores and reduced GAD risk observed in our study

Beyond this, optimizing medical indicators (such as blood pressure and blood glucose) demonstrates more pronounced preventive effects against brain inflammation and blood-brain barrier disruption. Research indicates that elevated blood glucose concentrations can induce systemic low-grade inflammatory mechanisms, leading to increased blood-brain barrier permeability. High glucose levels may also impair blood-brain barrier function by increasing paracellular permeability through metabolic overload and reducing trans endothelial resistance in human brain microvascular endothelial cells [28]. Within the LS7 framework, blood glucose

control is a core component; our finding that higher LS7 scores were associated with reduced GAD risk (Crude Model: OR = 0.905, 95% CI: 0.845–0.968) may partly reflect the neuroprotective effects of optimized glycemic control on BBB integrity. Elevated cholesterol levels also adversely affect BBB maintenance. Consistent with prior findings, after 30 days of high-cholesterol diet in mice, increased plasma cholesterol levels were accompanied by BBB disruption and heightened permeability in the prefrontal cortex and hippocampus [29]. Other studies report that high-cholesterol diets activate glial cells, mediate inflammatory cytokine expression, and induce more intense neuroinflammation [30]. This neuroinflammatory mechanism may further explain why triglycerides emerged as the most influential feature in our SHAP analysis (mean |SHAP value| = 0.103), underscoring the importance of lipid management within the LS7 framework in reducing GAD susceptibility.

Additionally, studies indicate that LS7 scores correlate with the activity of the renin-angiotensin-aldosterone system (RAAS). Higher LS7 scores are associated with reduced RAAS activity and decreased levels of inflammatory markers IL-6 and CRP [31]. This is directly relevant to our findings, as higher LS7 scores were associated with significantly reduced GAD risk across all three models (Model 1: OR = 0.886; Model 2: OR = 0.889), suggesting that RAAS suppression and reduced systemic inflammation may represent key mechanistic pathways underlying this protective association. Current research widely acknowledges the association between depression and anxiety with heightened RAAS axis function, which is linked to neuroinflammation, stress, and oxidative stress induction [32]. Ang II was previously identified as a prohypertensive factor in brain regions associated with cardiovascular function, and recent studies have implicated it in motor activity, anxiety, learning, and memory [33].Chronic neuroinflammation triggers microglial activation and neural dysfunction in the brain. This leads to reduced levels of brain-derived neurotrophic factor (BDNF), exacerbates nitric oxide stress, and alters monoamine levels, thereby increasing the risk of psychiatric disorders [34].

This evidence collectively demonstrates that LS7 exerts anxiolytic effects through a multi-target mechanism (modulation of neuroinflammation, cerebrovascular protection, and metabolic optimization).

Limitations in medical record data, particularly incompleteness, may impact assessment. Future prospective, multicenter cohort studies are warranted to provide clearer guidance for clinical management.

5.Funding

This work was funded by the Science and Technology Research Project of Henan Province (grant 252102310469) and the Scientific Research Project for the "Double First-Class" Construction of Traditional Chinese Medicine in Henan Province (grant HSRP-DFCTCM-2023-1-09). The funders played no role in the study design, data collection, data analysis, data interpretation, writing of the report, or decision to submit the article for publication.

6. Ethical Approval

This study analyzed de-identified information downloaded from a public database (KNHANES). All participants provided written informed consent prior to study participation.

7.Conflict of interest

The authors declare no conflict of interest.

References

[1] Munir S, Takov V. Generalized Anxiety Disorder. 2022 Oct 17. In: StatPearls [Internet].

Treasure Island (FL): StatPearls Publishing; 2025 Jan.

[2] Mishra AK, Varma AR. A Comprehensive Review of the Generalized Anxiety Disorder. Cureus. 2023 Sep 28;15(9):e46115.

[3] Wang L, Chang G, Cai S, Zou X, Qin M, Tan Y. The association of Life's Simple 7 and infertility among U.S. women. Front Endocrinol (Lausanne). 2024 Feb 1;15:1288289.

[4] Ogunmoroti O, Allen NB, Cushman M, Michos ED, Rundek T, Rana JS, Blankstein R, Blumenthal RS, Blaha MJ, Veledar E, Nasir K. Association Between Life's Simple 7 and Noncardiovascular Disease: The Multi-Ethnic Study of Atherosclerosis. J Am Heart Assoc. 2016 Oct 20;5(10):e003954.

[5] Chen L, Liu Q, Xu F, Wang F, Luo S, An X, Chen J, Tang N, Jiang X, Liang X. Effect of physical activity on anxiety, depression and obesity index in children and adolescents with obesity: A meta-analysis. J Affect Disord. 2024 Jun 1;354:275-285..

[6] Zou L, Tian Y, Wang Y, Chen D, Lu X, Zeng Z, Chen Z, Lin C, Liang Y. High-cholesterol diet promotes depression- and anxiety-like behaviors in mice by impact gut microbe and neuroinflammation. J Affect Disord. 2023 Apr 14;327:425-438.

[7] Bigalke JA, Durocher JJ, Greenlund IM, Keller-Ross M, Carter JR. Blood pressure and muscle sympathetic nerve activity are associated with trait anxiety in humans. Am J Physiol Heart Circ Physiol. 2023 Apr 1;324(4):H494-H503.

[8] Kweon S, Kim Y, Jang MJ, Kim Y, Kim K, Choi S, Chun C, Khang YH, Oh K. Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES). Int J Epidemiol. 2014 Feb;43(1):69-77.

[9] Krebs-Smith SM, Pannucci TE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, Wilson MM, Reedy J. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet. 2018 Sep;118(9):1591-1602.

[10] Toussaint A, Hüsing P, Gumz A, Wingenfeld K, Härter M, Schramm E, Löwe B. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). J Affect Disord. 2020 Mar 15;265:395-401.

[11] First MB. Diagnostic and statistical manual of mental disorders, 5th edition, and clinical utility. J Nerv Ment Dis. 2013 Sep;201(9):727-9.

[12] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems. 2017;30.

[13] Liang P, Wu J. Association between nutritional risk and CRP levels in patients with acute exacerbation of chronic obstructive pulmonary disease. Front Med (Lausanne). 2025 Jun 24;12:1611981.

[14] Acosta JN, Both CP, Rivier C, Szejko N, Leasure AC, Gill TM, Payabvash S, Sheth KN, Falcone GJ. Analysis of Clinical Traits Associated With Cardiovascular Health, Genomic Profiles, and Neuroimaging Markers of Brain Health in Adults Without Stroke or Dementia. JAMA Netw Open. 2022 May 2;5(5):e2215328.

[15] Greenwood BN, Fleshner M. Exercise, stress resistance, and central serotonergic systems. Exerc Sport Sci Rev. 2011 Jul;39(3):140-9.
[16] Casanova F, O'Loughlin J, Karageorgiou V, Beaumont RN, Bowden J, Wood AR, Tyrrell J. Effects of physical activity and sedentary time on depression, anxiety and well-being: a bidirectional Mendelian randomisation study. BMC Med. 2023 Dec 18;21(1):501.
[17] Heyman E, Gamelin FX, Goekint M, Piscitelli F, Roelands B, Leclair E, Di Marzo V, Meeusen R. Intense exercise increases circulating endocannabinoid and BDNF levels in humans--possible implications for reward and depression. Psychoneuroendocrinology. 2012 Jun;37(6):844-51.
[18] Archer T, Josefsson T, Lindwall M. Effects of physical exercise on depressive symptoms and biomarkers in depression. CNS Neurol Disord Drug Targets. 2014;13(10):1640-53.
[19] Banach M, Fogacci F, Atanasov AG, Stoian AP, Jóźwiak J, Bytyçi I, Pećin I, Katsiki N, Mazidi M, Gouni-Berthold I, Penson PE, Gaita D, Cicero AFG; International Lipid Expert Panel (ILEP). A 360° perspective on cardiovascular prevention: the International Lipid Expert Panel SiMple tIps for the heaLthy hEart (ILEP-SMILE). Arch Med Sci. 2025 May 29;21(3):711-718.
[20] Maes M. Major depression and activation of the inflammatory response system. Adv Exp Med Biol. 1999;461:25-46.
[21] Molteni R, Barnard RJ, Ying Z, Roberts CK, Gómez-Pinilla F. A high-fat, refined sugar diet reduces hippocampal brain-derived neurotrophic factor, neuronal plasticity, and learning. Neuroscience. 2002;112(4):803-14.
[22] Shirayama Y, Chen AC, Nakagawa S, Russell DS, Duman RS. Brain-derived neurotrophic factor produces antidepressant effects in behavioral models of depression. J Neurosci. 2002 Apr 15;22(8):3251-61.
[23] Horrobin D, Bennett C. Phospholipid metabolism and the pathophysiology of psychiatric and neurological disorders. In: Peet M, Glen I, Horrobin D, eds. Phospholipid Spectrum Disorders in Psychiatry and Neurology. 2nd ed. Carnforth, UK: Maius Press; 2003:3–47.
[24] Xie JZ, Zhang Y, Li SH, Wei H, Yu HL, Zhou QZ, Wei LY, Ke D, Wang Q, Yang Y, Wang JZ. P301S-hTau acetylates KEAP1 to trigger synaptic toxicity via inhibiting NRF2/ARE pathway: A novel mechanism underlying hTau-induced synaptic toxicities. Clin Transl Med. 2022 Aug;12(8):e1003.
[25] Lindqvist D, Hall S, Surova Y, Nielsen HM, Janelidze S, Brundin L, Hansson O. Cerebrospinal fluid inflammatory markers in Parkinson's disease--associations with depression, fatigue, and cognitive impairment. Brain Behav Immun. 2013 Oct;33:183-9.
[26] Menard C, Pfau ML, Hodes GE, Kana V, Wang VX, Bouchard S, Takahashi A, Flanigan ME, Aleyasin H, LeClair KB, Janssen WG, Labonté B, Parise EM, Lorsch ZS, Golden SA, Heshmati M, Tamminga C, Turecki G, Campbell M, Fayad ZA, Tang CY, Merad M, Russo SJ. Social stress induces neurovascular pathology promoting depression. Nat Neurosci. 2017 Dec;20(12):1752-1760.

[27] Cohen SS, Min M, Cummings EE, Chen X, Sadowska GB, Sharma S, Stonestreet BS. Effects of interleukin-6 on the expression of tight junction proteins in isolated cerebral microvessels from yearling and adult sheep. Neuroimmunomodulation. 2013;20(5):264-73.

[28] Jayaraj RL, Azimullah S, Beiram R. Diabetes as a risk factor for Alzheimer's disease in the Middle East and its shared pathological mediators. Saudi J Biol Sci. 2020 Feb;27(2):736-750.

[29] de Oliveira J, Engel DF, de Paula GC, Dos Santos DB, Lopes JB, Farina M, Moreira ELG, de Bem AF. High Cholesterol Diet Exacerbates Blood-Brain Barrier Disruption in LDLr-/- Mice: Impact on Cognitive Function. J Alzheimers Dis. 2020;78(1):97-115.

[30] Zou L, Tian Y, Wang Y, Chen D, Lu X, Zeng Z, Chen Z, Lin C, Liang Y. High-cholesterol diet promotes depression- and anxiety-like behaviors in mice by impact gut microbe and neuroinflammation. J Affect Disord. 2023 Apr 14;327:425-438.

[31] Yuan YE, Haas AV, Williams GH, Taylor H, Seely EW, Adler GK. Association Between Life's Simple 7 and Biomarkers of Cardiovascular Disease: Aldosterone, Interleukin-6, C-Reactive Protein. J Am Heart Assoc. 2023;12(10):e028718.

[32] Gong S, Deng F. Renin-angiotensin system: The underlying mechanisms and promising therapeutical target for depression and anxiety. Front Immunol. 2023 Jan 24;13:1053136.

[33] Saavedra JM, Ando H, Armando I, Baiardi G, Bregonzio C, Juorio A, Macova M. Anti-stress and anti-anxiety effects of centrally acting angiotensin II AT1 receptor antagonists. Regul Pept. 2005 Jun 30;128(3):227-38.

[34] Leonard BE. Inflammation and depression: a causal or coincidental link to the pathophysiology? Acta Neuropsychiatr. 2018 Feb;30(1):1-16.

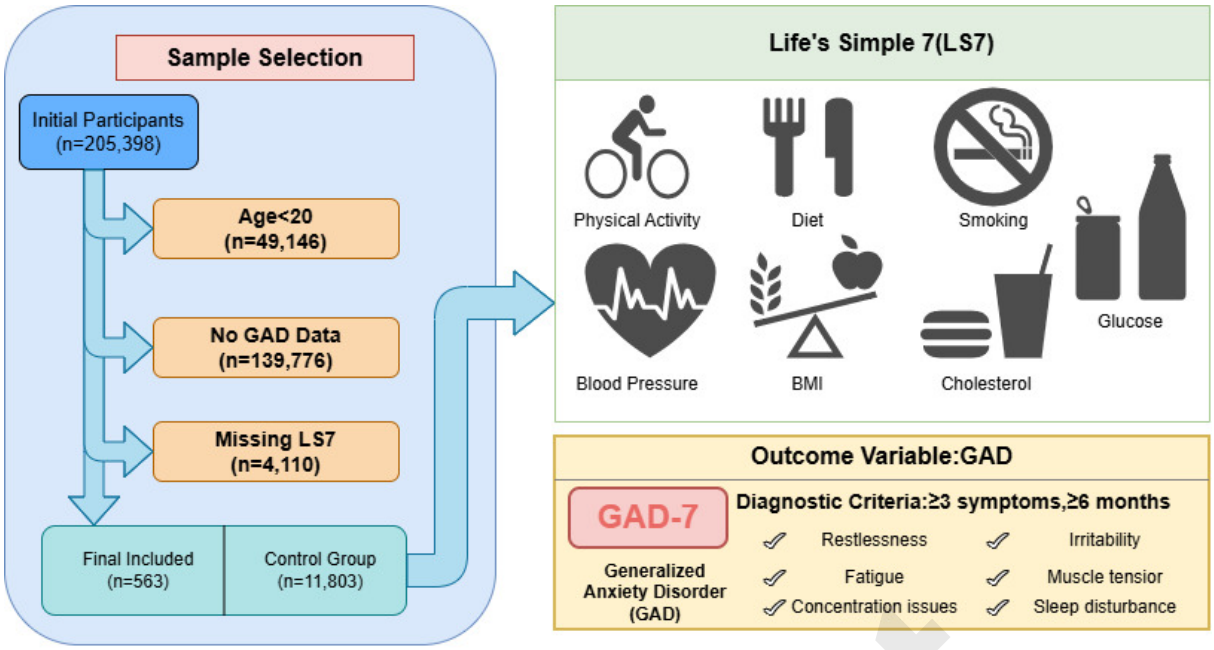

Sample Selection
Initial Participants (n=205,398)
Age<20 (n=49,146)
No GAD Data (n=139,776)
Missing LS7 (n=4,110)
Final Included (n=563)
Control Group (n=11,803)
Life's Simple 7(LS7)
Physical Activity
Diet
Smoking
Glucose
Blood Pressure
BMI
Cholesterol
Outcome Variable:GAD
GAD-7
Generalized Anxiety Disorder (GAD)
Diagnostic Criteria:≥3 symptoms,≥6 months
Restlessness
Irritability
Fatigue
Muscle tension
Concentration issues
Sleep disturbance

Table 1. Baseline characteristics of participants after propensity score matching.

Variable	Overall N = 1,126	GAD N = 563	Normal N = 563	*p*-value
Age	48 (18)	48 (18)	48 (18)	0.822
Age Group				0.798
Old	358 (32%)	177 (31%)	181 (32%)	
Young	768 (68%)	386 (69%)	382 (68%)	
Gender				0.649
Female	787 (70%)	397 (71%)	390 (69%)	
Male	339 (30%)	166 (29%)	173 (31%)	
PIR				0.945
≤ 1.3	370 (33%)	184 (33%)	186 (33%)	
1.3-3.5	745 (66%)	374 (66%)	371 (66%)	
> 3.5	11 (1.0%)	5 (0.9%)	6 (1.1%)	
ALT	21 (17)	22 (17)	21 (18)	0.405
White blood cell count	5.99 (1.70)	6.03 (1.69)	5.95 (1.71)	0.412
LS7	9.50 (1.75)	9.35 (1.77)	9.65 (1.72)	0.004

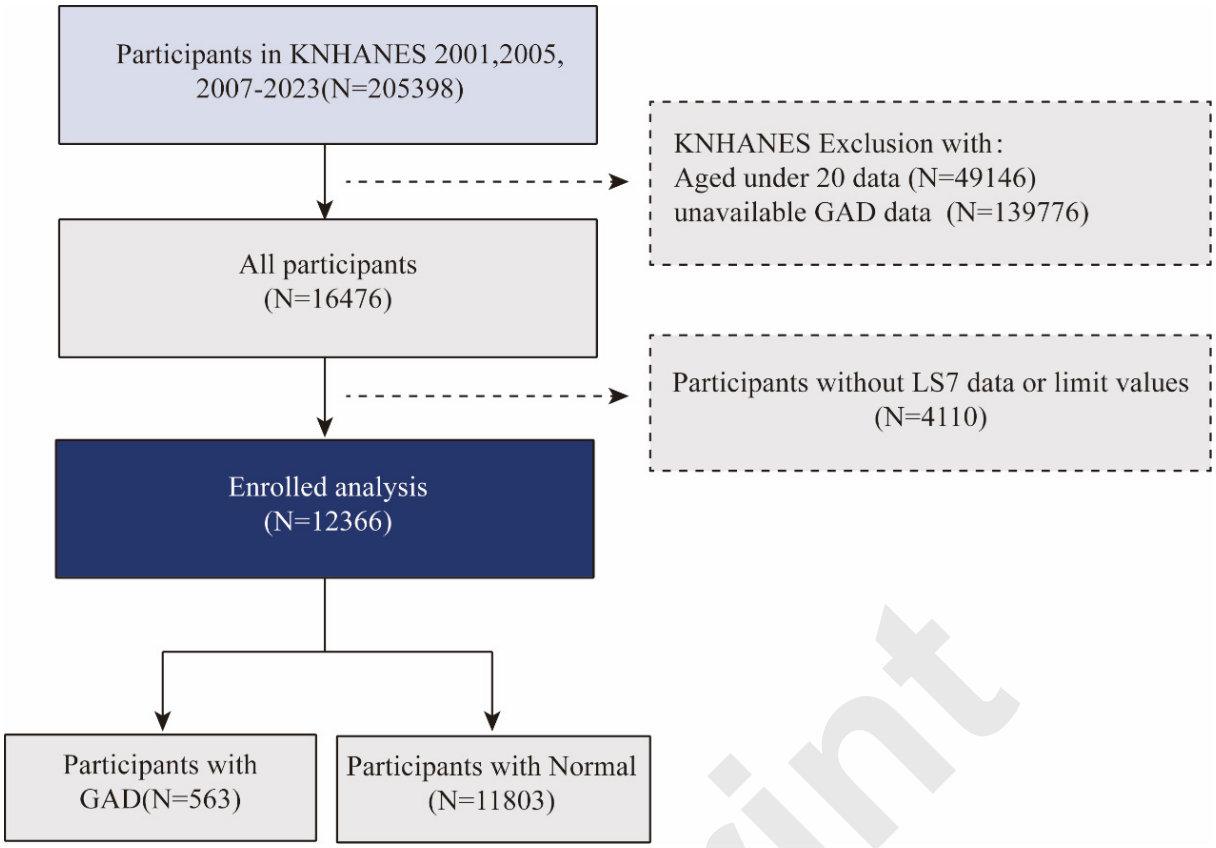
Table 2. The association between LS7 exposure and GAD

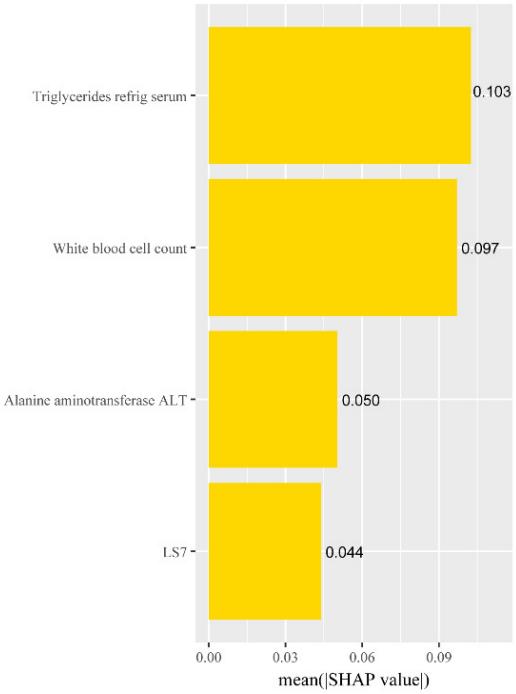
	Crude Model			Model1			Model2		
Characteristic	**OR**	**95%CI**	***P*-value**	**OR**	**95%CI**	***P*-value**	**OR**	**95%CI**	***P*-value**
LS7									
LS7 continuous	0.90	(0.85, 0.97)	0.004	0.88	(0.82, 0.95)	<0.001	0.91	(0.84, 0.97)	0.007
LS7 quantile									
Q1(< 8)	Ref	Ref		Ref	Ref		Ref	Ref	
Q2(8-10)	0.93	(0.64, 1.37)	0.719	0.90	(0.61, 1.33)	0.587	0.95	(0.64, 1.40)	0.784
Q3(10-11)	0.68	(0.46, 1.02)	0.062	0.63	(0.42, 0.96)	0.032	0.69	(0.46, 1.05)	0.087
Q4(≥ 11)	0.67	(0.46, 0.98)	0.039	0.60	(0.40, 0.90)	0.014	0.68	(0.45, 1.03)	0.069
P for trend			0.007			0.002			0.013

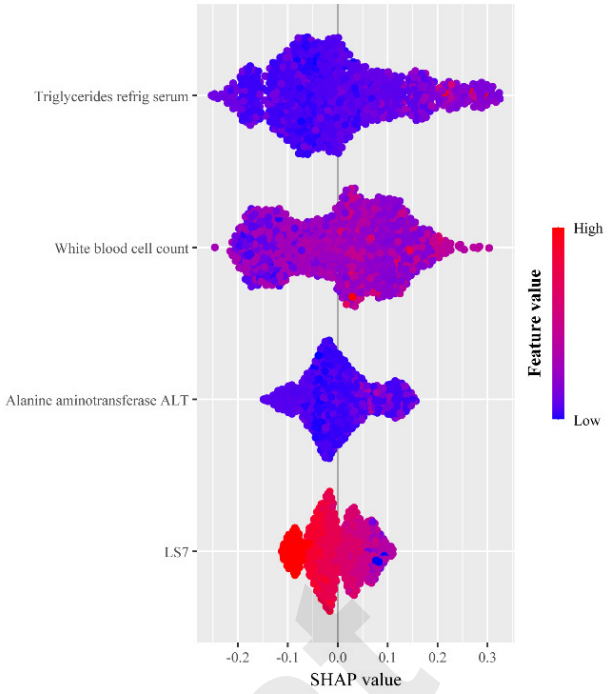
Note: The Crude Model was non-adjusted ; The Model 1 was adjusted by Age、Gender、PIR; The Model 2 was adjusted by ALT、 White blood cell count

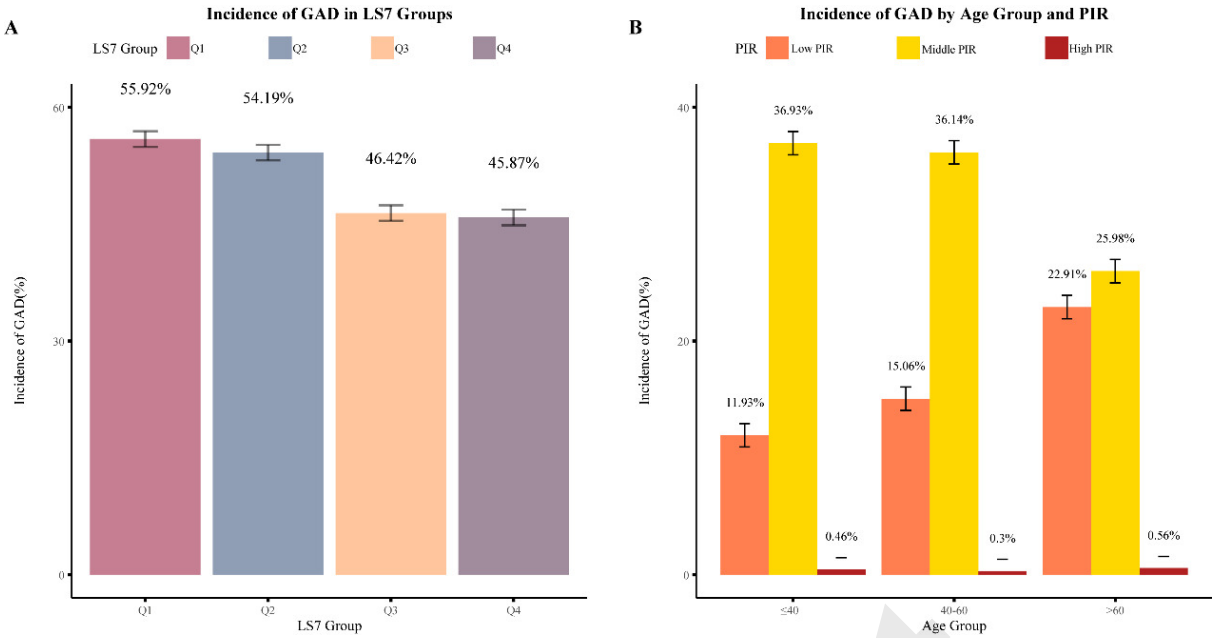
Table S2. Univariate logistic regression analysis of the association between LS7 exposure and GAD (Sensitivity analysis)

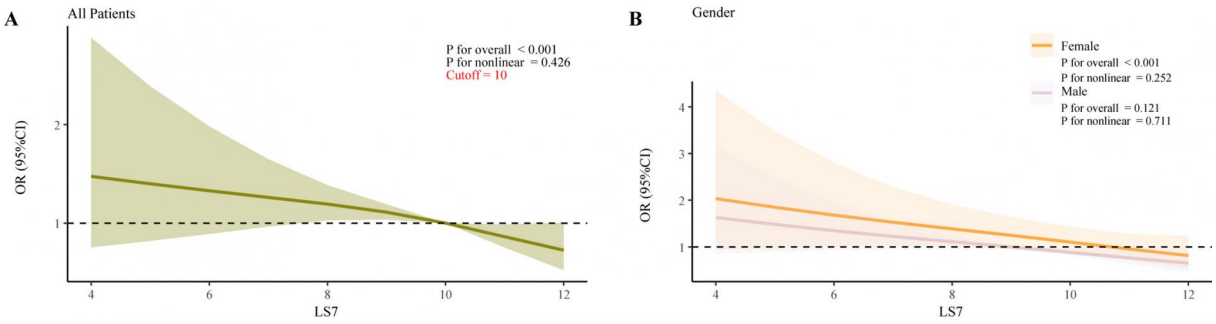
	Crude Model			Model1			Model2		
Characteristic	**OR**	**95%CI**	**P-value**	**OR**	**95%CI**	**P-value**	**OR**	**95%CI**	**P-value**
LS7									
LS7 continuous	0.91	(0.84, 0.99)	0.022	0.90	(0.83, 0.98)	0.015	0.91	(0.83, 0.99)	0.032
LS7 quantile									
Q1(< 8)	Ref	Ref		Ref	Ref		Ref	Ref	
Q2(8-10)	0.80	(0.51, 1.27)	0.347	0.78	(0.48, 1.24)	0.293	0.83	(0.51, 1.34)	0.443
Q3(10-11)	0.61	(0.38, 0.99)	0.048	0.58	(0.35, 0.96)	0.034	0.62	(0.37, 1.03)	0.063
Q4(≥ 11)	0.59	(0.37, 0.94)	0.025	0.55	(0.33, 0.89)	0.016	0.60	(0.36, 0.99)	0.045
p for trend			0.011			0.007			0.018


Participants in KNHANES 2001,2005, 2007-2023(N=205398)
KNHANES Exclusion with:
Aged under 20 data (N=49146)
unavailable GAD data (N=139776)
All participants (N=16476)
Participants without LS7 data or limit values (N=4110)
Enrolled analysis (N=12366)
Participants with GAD(N=563)
Participants with Normal (N=11803)


Triglycerides refrig serum
White blood cell count
Alanine aminotransferase ALT
LS7
0.103
0.097
0.050
0.044
0.00
0.03
0.06
0.09
mean(|SHAP value|)


Triglycerides refrig serum
White blood cell count
Alanine aminotransferase ALT
LS7
-0.2
-0.1
0.0
0.1
0.2
0.3
SHAP value
Feature value
High
Low


A
Incidence of GAD in LS7 Groups
LS7 Group
Q1
Q2
Q3
Q4
55.92%
54.19%
46.42%
45.87%
Incidence of GAD(%)
LS7 Group
B
Incidence of GAD by Age Group and PIR
PIR
Low PIR
Middle PIR
High PIR
11.93%
36.93%
0.46%
15.06%
36.14%
0.3%
22.91%
25.98%
0.56%
Incidence of GAD(%)
Age Group
≤40
40-60
>60


A
All Patients
P for overall < 0.001
P for nonlinear = 0.426
Cutoff = 10
OR (95%CI)
LS7
B
Gender
Female
P for overall < 0.001
P for nonlinear = 0.252
Male
P for overall = 0.121
P for nonlinear = 0.711
OR (95%CI)
LS7

Subgroup	N	OR (95% CI)	P value	P for interaction
Overall	1126	0.90 (0.85, 0.97)	0.004	
Age_Group				0.401
Old	358	0.95 (0.83, 1.10)	0.509	
Young	768	0.89 (0.82, 0.96)	0.003	
Gender				0.058
Female	787	0.84 (0.77, 0.92)	<0.001	
Male	339	0.97 (0.86, 1.10)	0.645