Analysis of hyperlipidemia risk factors among pilots based on physical examination data: A study using a multilevel propensity score models

FEIFEI YU*, YI XIE* and JISHUN YANG

Naval Medical Center, Naval Medical University (Second Military Medical University), Shanghai 200433, P.R. China

Received November 30, 2023; Accepted May 23, 2024

DOI: 10.3892/etm.2024.12630

Abstract. Pilot tends to have a high prevalence of dyslipidemia. The present study aimed to identify key factors of pilot hyperlipidemia through thorough analysis of physical examination data, and to provide pilot-targeted health guidance to manage hyperlipidemia risks. The physical examination data of 1,253 pilot inpatients from January 2019 to June 2022, were evaluated and divided into two groups based on whether or not the pilot had hyperlipidemia. A total of three multivariate analysis models including logistic model, multilevel model and boosting propensity score were applied to find the risk factors of pilot hyperlipidemia. In the group of pilots with hyperlipidemia, four risk factors, including thrombin time, carbohydrate antigen 199, lymphocyte count and rheumatoid factor, were significantly different from pilots without hyperlipidemia, which might be positively associated with the incidence of hyperlipidemia. In future studies regarding pilots, whether hyperlipidemia is connected to abnormalities in thrombin time, carbohydrate antigen 199 and rheumatoid factor should be further explored. Based on the findings of the present study, pilot health management should be more refined and personalized, and attention should be paid to the risk factors of hyperlipidemia including diet and lifestyle.

Introduction

Hyperlipidemia is a prevalent medical condition defined by increased lipid levels in the blood, especially cholesterol and triglycerides, which is often connected with cardiovascular and cerebrovascular diseases such as coronary heart disease and cerebral atherosclerosis (1,2). One of the most prevalent metabolic disorders amongst pilots is hyperlipidemia (3). Previous research found that long-term hyperlipidemia, which is a risk factor for both pancreatitis and diabetes mellitus (1), might also increase the incidence of cerebrovascular disease (4). As it is closely related to the health of pilot, hyperlipidemia may impact flight safety and shorten the flying career of pilots (5), therefore it warrants further investigation.

Several researchers have suggested that the prevalence of dyslipidemia in pilots in China is substantially higher than the overall population. Specifically, the prevalence of dyslipidemia is reportedly 18.6% in the adult Chinese population (6), but is 20.79% amongst pilots in the Chinese civil aviation industry (7). This further increases to 22.7% in military pilots and crew (8). Yu *et al* (9) revealed that the prevalence of hyperlipidemia in pilots has increased annually, while the age of onset has decreased. For ~22 years, hyperlipidemia has been one of the most prevalent disease amongst pilots (10) with Dong *et al* (11) reporting that hyperlipidemia is the most common illness out of 270 other diseases in the pilot disease spectrum. Another study found that hyperlipidemia ranks in the top five in the spectrum of diseases among pilots and the incidence of marginal elevation of blood lipids including total cholesterol, triglyceride and low-density lipoprotein cholesterol (LDL-C) was 28% among 450 male flight personnel (7), which was significantly higher than the prevalence of the general population of the same age.

Polymorphisms in the 8th intron of the LPL gene and apolipoprotein E may affect the levels of blood lipids (12,13). However, several studies have found no significant difference between the genotypes of pilots and the general population (14-16). Therefore, it has been hypothesized that, instead of genetic factors, hyperlipidemia in pilots might be more closely related to environmental factors (17). Specifically, pilots are routinely subjected to the high intensity of flight training, physical exertion, frequent exposure to stress, high-load flights, which requires pilots to put in more operational skills and attention during the flight process, with increased heart rate and muscle tensing (18), and high-calorie diets (19). All these environmental factors may cause stress in the neural endocrine system and lead to metabolic disorders (20). Furthermore, researcher has shown that 98.9% of ~1,000 surveyed pilots have at least one negative lifestyle habit, such as smoking and drinking alcohol, which might also contribute to hyperlipidemia (21).

Correspondence to: Dr Jishun Yang, Naval Medical Center, Naval Medical University (Second Military Medical University), 800 Xiangyin Road, Yangpu, Shanghai 200433, P.R. China
E-mail: jasunyang@foxmail.com

*Contributed equally

Key words: pilot, hyperlipidemia, risk factors, propensity score, physical examination data

Although many studies have shown that the prevalence of pilot hyperlipidemia is high, there is no detailed report on the causes and characteristics of pilot hyperlipidemia. Previously, it has been hypothesized that this increased prevalence is associated with their unusual work environment (22), which involves high-intensity training, often leading to an increased alcohol intake and a high-fat diet. However, at present no complete supporting data or quantitative risk estimations have been carried out. The aim of the present study was to explore the association between physical examination indicators and the incidence of hyperlipidemia in pilots by using physical examination data.

Materials and methods

Data sources. The present study was based on pilot physical examination data, obtained from a sample of 1,253 pilots between January 2019 and June 2022 at the Naval Medical Center of Naval Medical University (Shanghai, China). The process of selection for the present study is shown in Fig. 1. The inclusion criteria was as follows: i) Pilots that were inpatients at the Naval Medical Center of Naval Medical University (Shanghai, China) from January 2019 to June 2022; and ii) pilots that underwent multiple physical examinations in the Naval Medical Center of Naval Medical University. The exclusion criteria included: i) Retired pilots; ii) pilots who had not undergone various tests such as blood tests and had only received physical therapy and traditional Chinese medicine treatment. The physical examination data were described and evaluated after being cleaned, and analysis results were presented in Tables I-III. Based on diagnostic information, 1,253 pilots were divided into two groups for the present study: i) Pilots with hyperlipidemia (n=14) whose disease diagnosis included hyperlipidemia [diagnosis code E78 in the International Statistical Classification of Diseases 10 (ICD-10) (23)]; and ii) pilots without hyperlipidemia (n=1,239). The diagnosis of hyperlipidemia was total cholesterol (TC) ≥6.2 mmol/l, triglyceride (TG) ≥2.3 mmol/l or LDL-C ≥4.2 mmol/l (23).

Variables. General conditions (such as height, weight, blood pressure and heart rate), blood routine examination, urine routine examination, blood biochemistry analysis, electrocardiogram and ultrasound or X-ray of each organ were the primary components of the physical examination.

There were 13 modules obtained from the electronic medical records of the pilots, including the basic information, admission record, first course of disease record, patient notification, vital sign record, surgery information, inspection information, examination information, diagnosis classification, diagnosis information, nutritional risk screening score, discharge record and discharge summary.

The basic information of patients, vital sign records, surgery information, examination information, diagnosis classification and diagnosis information are text-type variables or structured data, but needed to be de-duplicated, standardized or unified in dimensions. Therefore, in order to ensure the accuracy of data analysis, the data format was defined, kept consistent, duplicates were removed and variables were standardized. The additional modules contain web-based text data or unstructured data such as admission records, discharge summary, initial and daily progress notes, discharge records and surgical records.

The data relating to the basic information of patients was structured and cross-sectional and includes 44 variables, such as the number of hospitalizations, date of hospitalization, sex and marital status. The two modules of diagnosis information and diagnosis classification information are the matching modules. The diagnosis name is displayed in Chinese, and the associated ICD-10 code is provided for the diagnosis categorization. Both modules are structured, longitudinal data.

The data relating to laboratory inspection information is structured and longitudinal, and includes blood test results (including biochemical tests, immunological tests, routine blood tests and virology tests.) and clinical test results (including urine and stool tests).

The data relating to routine physical examination included the following detailed categories: Liver function, renal function, electrolyte levels, hepatitis B test, hepatitis C antibody test, treponema pallidum antibody test, human immunodeficiency virus and antigen antibody detection, blood cell analysis including whole blood C-reactive protein levels, urine analysis including urine sediment levels, stool routine with occult blood tests, prothrombin time measurement, prothrombin standard ratio, activated partial thrombin time, prothrombin time, fibrinogen concentration, levels of D-dimer, levels of fibrinogen degradation product, liver and kidney function testing, blood lipid levels and other tests such as gastrin testing, tumor markers tests, thyroid tests, pepsin testing, ABO and Rh blood group identification, and anti-rheumatoid lymphocyte testing.

The data relating to vital signs was structured and longitudinal, including stool frequency, height, weight, systolic blood pressure, diastolic blood pressure, respiration, heart rate and body temperature.

The surgery information was semi-longitudinal data, which included the type of surgery, the ICD code of the surgery, the name of the surgeon, the name of the assistant and the time of the surgery. However, it was challenging to extract and make use of the other information, such as discharge summaries, which were obtained from web page forms, which were in free text format, and were unstructured data.

Statistical analysis. The data were analyzed with Statistics Analysis System (version 9.4) and R language (version 4.2.2). For enumeration data, chi-square test or Fisher's exact test was used for comparison between groups, and for measurement data, two independent samples (unpaired) t-test was used for comparison between groups. In the single factor analysis, if $P<0.1$, the factor was incorporated into the multi-factor model for further modeling. All tests were two-sided and $P<0.05$ was considered to indicate a statistically significant difference.

Logistic regression model, multilevel model and boosting propensity score model were used to establish the multivariate analysis model to find the risk factors of pilot hyperlipidemia. The multilevel model, also termed the hierarchical model, can deal with multilevel statistical methods with a hierarchical structure. Some medical research data tend to be hierarchical or clustered, and multilevel models can be used to divide the residuals in the models to different levels (24). The individual random error is relatively accurate, so it can be estimated more

accurately than using the traditional regression model (25). In the present study, the resident region of pilots was used as a stratification factor.

Propensity score is a statistical method that has been widely used in observational studies in recent years (26). It calculates the conditional probability of an individual being assigned to an exposed group by combining covariates from the exposed and non-exposed groups. Individuals with similar values of exposed and non-exposed components were then matched or adjusted, weighted or stratified to estimate exposure effects. It uses the balanced distribution of covariates data to achieve the objective of unbiased estimation of exposure effects (27).

The boosting algorithm is a classification learning algorithm, which can complete correct classification or recognition of a dataset (28). As the boosting algorithm can classify and predict individuals, it can also be used to estimate propensity score (29). Compared with some traditional regression and classification methods, the boosting algorithm has the following advantages: i) It is a non-parametric classification or recognition method, which does not need to satisfy the assumption of log-linear parameters; ii) second, this method is a strong learning algorithm, which can use other machine learning algorithms, such as decision trees and support vector machines, as the basis for its iterations, and produce more accurate classification predictions and output propensity scores directly; and iii) if there is interaction between various variables in the data, or if the data is in a high-dimensional state, the method can still be used to predict effectively and obtain relatively accurate results (30).

Results

Demographic information. Among the 1,253 pilot participants, 393 were married (31%), while 855 were unmarried (68.24%), as shown in Table I. In terms of permanent residence, 352 (23.56%) of the pilots lived in Liaoning followed by 251 (20.08%) in Shandong and 243 (16.27%) in Shanghai, as shown in Fig. 2. Pilots were aged 19-57 years old, with a median age of 24.5 years old. The longest hospital stay was 77 days, the shortest was 1 day and the median hospital stay was 5 days.

Univariate analysis. Results of univariate analysis showed that age, apolipoprotein B, hepatitis B core antibody, thrombin time, lymphocyte count, cholesterol, squamous cell carcinoma antigen, rheumatoid factor, low-density lipoprotein, B27 human leukocyte antigen, homocysteine and apolipoprotein AI were different between patients with hyperlipidemia and without hyperlipidemia (Tables II and III).

Multivariate analysis

Multivariate logistic regression model. Multivariate logistic regression analysis was used to analyze the data of 1,253 pilots. As there was a variation in the physical examination items of each pilot, most of the examination items were crossed but not overlapped, thus the sample size for multiple factors analysis was small. A total of 947 patients had records including the data from the majority of the tests. Of these patients, there were only 14 patients with hyperlipidemia and 933 patients without hyperlipidemia. As shown in Table IV, carbohydrate antigen 199, thrombin time and age might be the risk factors of hyperlipidemia, which mean higher levels of these three markers might increase the risk of diagnosis of hyperlipidemia. Low-density lipoprotein and cholesterol are markers for the diagnosis of hyperlipidemia, an increase in the levels of these markers also represents an increased risk of hyperlipidemia. However, as a routine blood test, white blood cell count, although statistically significant, might not be used clinically as a marker for hyperlipidemia due to the fact

Table I. Summary of indicators for enumeration data.

Indicators	n	Percentage
Marital status		
Unknown	3	0.24
Other	1	0.08
Widowed	1	0.08
Unmarried	855	68.24
Married	393	31.36
Total	1,253	100.00
Ethnicity		
Han	1,239	98.88
Manchu	6	0.48
Mongol	4	0.32
Tujia	1	0.08
Unknown	3	0.24
Total	1,253	100.00
Blood group (ABO)		
A	223	29.69
AB	76	10.12
B	198	26.36
O	254	33.82
Total	751	100.00
-	6	0.80
+	742	99.20
Total	748	100.00
Allergy		
Penicillin sodium for injection	12	1.03
76% compound meglumine diatrizoate injection	1	0.09
Azithromycin	2	0.17
Iodine	1	0.09
Compound sulfamethoxazole	1	0.09
Sulfonamides	5	0.41
Streptomycin	1	0.09
Tetanus antitoxin	1	0.09
Penicillin	16	1.37
Penicillin, tetanus antitoxin	1	0.09
Cephalosporin	6	0.53
None	1,119	95.97
Total	1,166	100.00

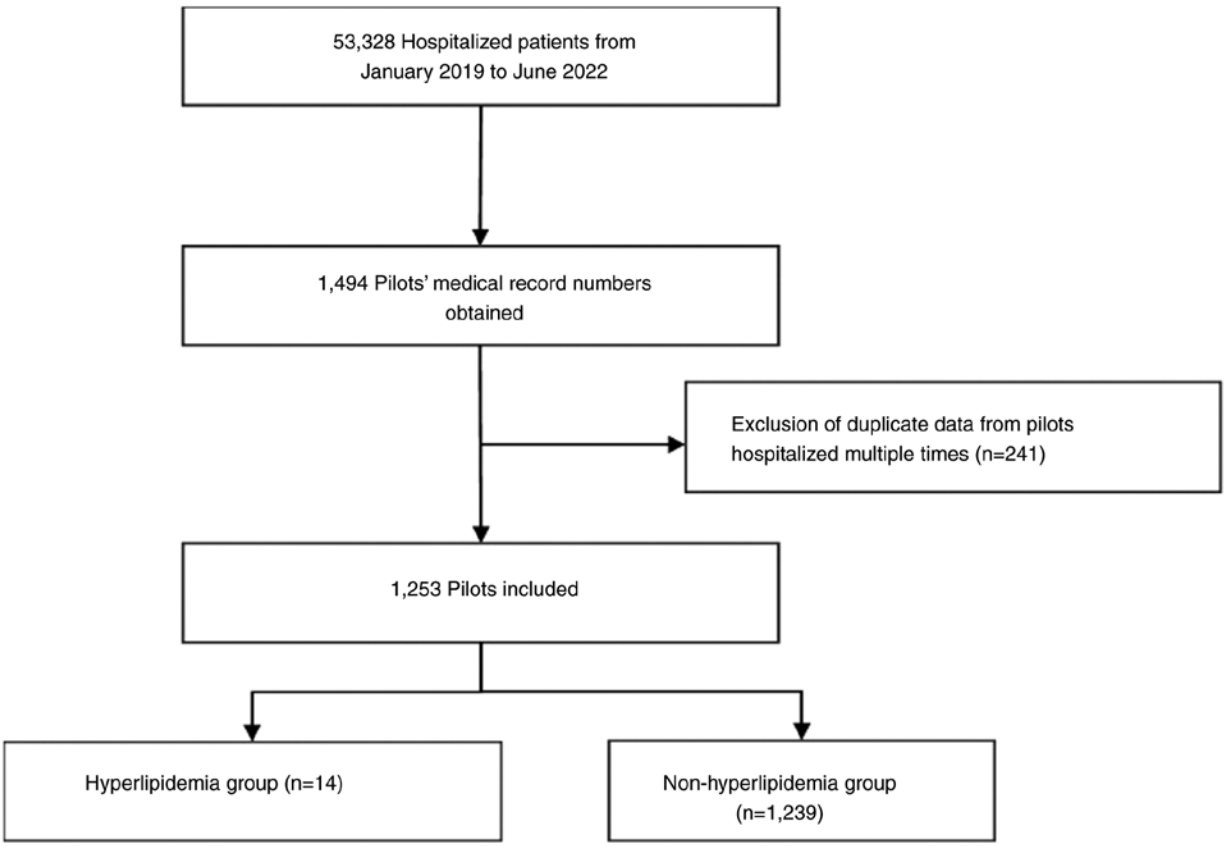

Figure 1. Flowchart of the selection process. ICD-10, international classification of diseases, version 10.

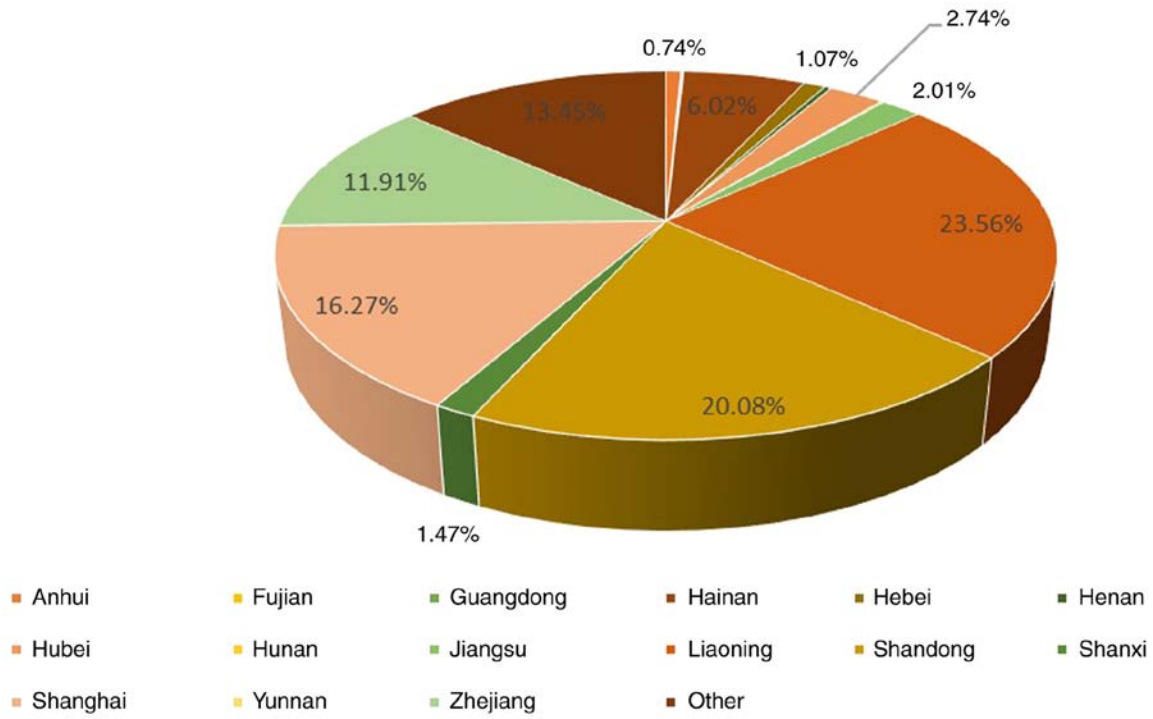

Figure 2. Permanent residence distribution of the pilots.

that white blood cell count is usually used to determine the inflammatory response and could be elevated in patients for a number of reasons including infection (31).

A high-fat diet and age are widely recognized factors related to hyperlipidemia (32). However in the present study, two more relatively new risk factors, the carbohydrate

Table II. T-test analysis of the risk factors of hyperlipidemia.

	T-test			95% confidence interval	
Indicators	P-value	Mean difference	Standard error	The lower limit	Ceiling
Age	<0.001	-4.150	0.586	-5.385	-2.915
Neutrophil percentage	0.097	-3.492	2.104	-7.620	0.635
Apolipoprotein B	0.001	0.164	0.049	0.069	0.260
Hepatitis B core antibody	<0.001	-0.719	0.066	-0.849	-0.588
Hepatitis B E antibody	0.495	-0.075	0.110	-0.292	0.141
Width of platelet distribution	0.124	-0.233	0.152	-0.531	0.064
Platelets	0.088	19.569	11.477	-2.949	42.087
Carbohydrate antigen 50	0.151	1.035	0.721	-0.380	2.450
Carbohydrate antigen 199	0.086	2.679	1.560	-0.382	5.740
Eosinophil granulocyte	0.301	0.052	0.049	-0.053	0.158
Mean platelet volume	0.074	-0.540	0.302	-1.132	0.053
Thrombin time	<0.001	0.901	0.237	0.437	1.365
Uric acid	0.670	7.043	16.531	-25.391	39.477
Immunoglobulin M	0.207	33.391	26.384	-18.522	85.303
Immunoglobulin G	0.085	282.172	163.075	-38.689	603.034
Lymphocyte count	0.019	0.346	0.147	0.057	0.635
Lymphocyte percentage	0.091	3.411	2.013	-0.540	7.361
Streptolysin	0.151	-35.402	24.663	-83.799	12.995
Antithyroglobulin	0.664	-7.948	18.291	-43.836	27.940
Thyroid peroxidase antibodies	0.607	-12.242	23.786	-58.913	34.429
α-fetoprotein	0.095	-0.564	0.337	-1.224	0.097
Red blood cells	0.077	1.489	0.841	-0.161	3.140
Aspartate transaminase	0.148	-3.423	2.364	-8.062	1.215
Alanine transaminase	0.125	-4.787	3.121	-10.910	1.336
Low-density lipoprotein	<0.001	0.768	0.194	0.386	1.149
Cholesterol	0.019	0.482	0.205	0.080	0.885
Large platelet ratio	0.086	-3.592	2.087	-7.686	0.503
Complement C4	0.366	-3.014	3.327	-9.560	3.532
Complement C3	0.179	-15.401	11.423	-37.877	7.076
White blood cells	0.011	2.174	0.854	0.499	3.850

antigen 199 and thrombin time, were identified from the aforementioned multifactor analysis. Although, whether these two relatively new risk factors are causally related to the pathogenesis of hyperlipidemia needs to be verified by further research.

Multilevel model analysis. The physical examination data of 1,253 pilots were further analyzed using a multilevel model. The behavioral patterns of the pilots may fit into certain residential clusters due to different regional cultures and eating customs, thus the resident region of each pilot was taken as a horizontal factor. As shown in Figure 2, the majority of pilots resided in the following provinces or regions, which were divided into 8 categories based on their locations: i) Anhui, Henan and Hubei (categorized as 1); ii) Hebei and Beijing (categorized as 2), Hainan (categorized as 3); iv) Guangxi, Fujian and Guangdong (categorized as 4); v) Jiangsu, Zhejiang and Shanghai (categorized as 5); vi) Liaoning (categorized as 6); vii) Shandong and Shanxi (categorized as 7); and viii) Yunnan and unknown regions (categorized as 8).

The resident regions of the pilots were categorized as aforementioned, and the spatial model was then applied to check whether the data have a multilevel structure. The results showed that the estimated intercept value of the second level was -5.444 and P=0.001. The results showed that the data of the present study have obvious hierarchical structure. After which, the data were analyzed by multilevel model analysis, and the results were shown in Table V. After taking into account the level factor of the pilot resident region, more relevant factors were derived from the results of the aforementioned multilevel model analysis for the diagnosis of hyperlipidemia, including thrombin time, carbohydrate antigen 199, lymphocyte count and rheumatoid factor. The results of thrombin time and carbohydrate antigen 199 were consistent with the results of the general logistic regression. However, lymphocyte count and rheumatoid factor may require further exploring.

Table III. Chi-square test (or Fisher's exact test) analysis of the risk factors of hyperlipidemia.

Indicators	Hyperlipidemia group	Non-hyperlipidemia group	P-value
Squamous cell carcinoma antigen			0.047
High	11	1,167	
Normal (not detected)	3	72	
Rheumatoid factor			0.018
High	2	17	
Normal (not detected)	12	1,222	
Low-density lipoprotein			<0.001
Low	0	584	
High	11	306	
Normal	3	349	
High-density lipoprotein			0.091
Low	1	267	
High	1	298	
Normal (not detected)	12	674	
Alkaline phosphatase			0.093
Low	0	52	
High	1	6	
Normal (not detected)	13	1,181	
Globulin			0.120
Low	3	215	
High	1	12	
Normal (not detected)	10	1,012	
B27 human leukocyte antigen			0.002
Weak positive	0	1	
Positive	1	22	
Negative	13	703	
Not tested	0	513	
Homocysteine			0.001
High	4	44	
Normal (not detected)	10	1,195	
Apolipoprotein AI			0.049
High	0	2	
Low	3	666	
Normal (not detected)	11	571	

Boosting propensity score model. After calculating the propensity score by the boosting algorithm, the effect values of the covariates were estimated by inverse probability weighting, and the results were shown in Table VI. The results were consistent with the results of the multilevel model analysis, and more variable differences (lymphocyte count and rheumatoid factor) could be obtained in comparison with the general logistic regression model. The estimation of effect values was also more precise and closer to the results of the multilevel model analysis.

Discussion

Carbohydrate antigen 199 is widely distributed in normal glandular epithelial cells, and exists as salivary mucin in serum under physiological conditions (33). When inflammation and obstruction occur in the glandular epithelial cells, the fluid in the lumen cannot be released smoothly, and the vascular permeability increase may lead to an increase in serum carbohydrate antigen 199 levels (34). Previous studies have shown that the increase of carbohydrate antigen 199 is related to inflammation or metabolic diseases (35,36). In addition, the decrease of metabolic clearance rate in blood may also cause the increase of carbohydrate antigen 199 level (37). Secondary hyperlipidemia is often induced by metabolic disorders (38), such as diabetes, obesity and hypothyroidism. Therefore, there may be an interaction between carbohydrate antigen 199 and hyperlipidemia.

Previous studies have shown that thrombin time is closely related to the degree of cardiac cell damage (39). After the surface of the arteriosclerosis plaque is damaged, platelets released from the body will aggregate and adhere to the collagen and microfibers under the intima of blood vessels in patients with cardiovascular disease (CVD) (40). In severe cases, persistent activation of the coagulation system and hypercoagulation may lead to myocardial anoxia and metabolites may not be effectively eliminated (41). As a chronic disease with abnormal lipid metabolism, hyperlipidemia is closely related to the occurrence of CVDs (42). Thus, there might be some interactions between thrombin time and lipid metabolism.

It has been noted that patients with active rheumatoid arthritis may have higher serum lipid levels, particularly cholesterol (43). Rheumatic activity may influence lipid distribution and result in lipid metabolism disorders in patient with rheumatoid arthritis (44). Some researchers have demonstrated that the serum TG and TC levels of patients with rheumatoid arthritis were not higher than those of the general population before they were diagnosed with rheumatoid arthritis (45). Another study revealed that the predominant dyslipidemia-related characteristics in patients with rheumatoid arthritis were low TC, LDL-C and HDL-C (46). However, a case-control study demonstrated that patients with rheumatoid arthritis have a greater burden of coronary atherosclerosis at their first angiogram that is independent of traditional cardiovascular risk factors (47). At the same time, hyperlipidemia represented one of the major cardiovascular risk factors that (47) affect children and adolescents early in life (48). Further studies and specific investigations are still necessary to explore the potential relationship between blood lipids and rheumatoid factors.

Table IV. Results of logistic regression analysis on the risk factors of pilot hyperlipidemia.

Variable	β-value	Standard error	Wald	Degree of freedom	P-value	Odds ratio	95% confidence interval	
							Lower limit	Upper limit
Carbohydrate antigen 199	0.055	0.024	5.168	1	0.023	1.056	1.008	1.107
Thrombin time	1.442	0.353	16.639	1	<0.001	4.228	2.115	8.451
Squamous cell carcinoma antigen	0.095	0.072	1.772	1	0.183	1.100	0.956	1.266
Low-density lipoprotein	5.213	1.417	13.534	1	<0.001	183.591	11.422	2950.824
Cholesterol	-3.970	1.397	8.073	1	0.004	0.019	0.001	0.292
White blood cells	0.209	0.088	5.723	1	0.017	1.233	1.039	1.464
Age	-0.195	0.081	5.778	1	0.016	0.822	0.701	0.965

Table V. Results of multilevel model analysis on the risk factors of pilot hyperlipidemia.

Effect	Estimate	Standard error	Degree of freedom	T-value	P-value
Constant term	-26.439	5.550	7	-4.76	0.002
Thrombin time	0.979	0.274	1,482	3.57	<0.001
Carbohydrate antigen 199	0.053	0.027	1,482	1.96	0.050
Lymphocyte	1.248	0.450	1,482	2.77	0.006
Rheumatoid factor	3.037	0.940	6	3.23	0.018

Table VI. Boosting propensity score model analysis of pilot hyperlipidemia risk factors.

Parameter	Coefficient	Standard error	Wald card square	P-value	Odds ratio	95% confidence interval	
						Lower limit	Upper limit
The constant term	-38.896	1.926	407.868	<0.001	-	-	-
Thrombin time	1.464	0.096	233.435	<0.001	4.322	3.582	5.215
Carbohydrate antigen 199	0.040	0.017	5.459	0.020	1.041	1.007	1.077
Lymphocyte	1.514	0.175	75.025	<0.001	4.544	3.226	6.400
Rheumatoid factor	0.914	0.293	9.733	0.002	6.215	1.972	19.583

One possible explanation is that the increase of inflammatory factors in serum may affect the function of adipose tissue and liver (49). It was also reported that the structure and function of vascular endothelium changed noticeably during the active phase of rheumatoid arthritis, and endothelial dysfunction may affect lipid metabolism (50). It is also speculated that disease activity in patients with rheumatoid arthritis may be associated with lipid metabolism, and abnormal lipid metabolism could aggravate inflammatory reaction (51,52). Therefore, more clinical trials and cohort studies are needed to clarify the causal relationship between rheumatoid factor and hyperlipidemia.

Recently, hyperlipidemia has been identified as an inflammation-related disease (53), and research has observed that during the pathological development of atherosclerosis-related CVDs, monocytes enter the endothelium after inflammation-induced overexpression of endothelial cells (54). As already established, lymphocytes are the protagonist of inflammatory and immune responses (55). A research study showed that interleukin-38, mainly secreted by B lymphocyte, infiltrated peripheral blood mononuclear cells or keratinocytes, has a significant function in the development of hyperlipidemia and its related CVDs (40). On the other hand, the stimulation of interleukin-36, which is also a lymphokine, could significantly inhibit the anti-hyperlipidemic effects

of the interleukin-36 receptor and affected the process of hyperlipidemia and the related atherosclerosis (56). Therefore, the association between inflammatory response with the development of hyperlipidemia and CVDs is well established.

Using the platform of artificial intelligence data analysis technology, the digitized data can be used for the screening of relevant risk factors or causes of disease, and formulate individualized health guidelines based on their vital signs, behaviors, lifestyle and personal preference (57). Individual intervention could be used to reduce the prevalence of hyperlipidemia, and reduce or delay the occurrence and development of serious complications such as cerebrovascular disease by altering health-related behaviors (58).

Feasible approaches for lowering the prevalence of hyperlipidemia include health promotion, health education and targeted medical treatment (59). Targeted health intervention may be achieved by gathering and recording comprehensive data on the demographics, living habits, lifestyle, occupational history, disease history, family history, treatment history, clinical characteristics, laboratory tests, imaging tests, functional tests, quality of life tests, health status, genetic factors, physical fitness and life trajectory of individuals (60).

The results of the present study indicate that in order for pilots to keep blood lipid levels healthy, they should undergo regular physical examinations to monitor the blood lipid indicators, and measurements of carbohydrate antigen 199, rheumatoid factor and thrombin time. Weight loss, smoking cessation, healthy diet and maintaining physical and mental well-being are also beneficial for controlling further deterioration of blood lipid levels (61,62). On the other hand, pilot management departments should implement more health promotion and education programs to benefit pilots (63).

It should be noted that in the present study, the diagnostic criteria for hyperlipidemia were more stringent than those in a previous study (64). In the present study, if TC ≥6.2 mmol/l, TG ≥2.3 mmol/l or LDL-C ≥4.2 mmol/l, it was considered as hyperlipemia (65). However, most previous studies used TC ≥5.2 mmol/l, TG ≥1.7 mmol/l or LDL-C ≥3.4 mmol/l as the diagnostic criteria for hyperlipidemia, which might be considered as dyslipidemia (63,66). This resulted in a relatively low incidence of hyperlipidemia in the present study.

The main limitations of the present study were that it was difficult to efficiently extract useful information from the large amount of unstructured text in the electronic medical record system, and required a significant amount of manpower and material resources. The majority of the acquired variables had missing values due to the varied contents of the physical examination, which was not conducive to analysis. In addition, the resource data of diagnoses was from the front page of the medical records, which were mainly based on the knowledge of physicians and their ability to diagnose all the conditions of the patients. If the doctors did not notice the abnormal lipid parameters of the patients or did not include all diagnoses on the front page of the medical record, there might be bias of the data. However, this was considered unlikely.

In conclusion, to the best of our knowledge, this is the first time a large amount of data about pilots was structurally categorized and thoroughly analyzed for the first, and multiple advanced statistical models were carried out to obtain more meaningful results. Several related risk factors of hyperlipidemia, including carbohydrate antigen 199, thrombin time, lymphocytes and rheumatoid factor were found in the present study, and might provide the basis for further research.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The data generated in the present study are not publicly available due to confidentiality restrictions, but may be requested from the corresponding author.

Authors' contributions

FY designed the study, performed the data analysis, interpreted the results, and wrote and revised the manuscript; YX performed data collection and drew the plot. JY reviewed and checked the references and tables, and performed the data collation and validation. All authors read and approved the final version of the manuscript. YX and JY confirm the authenticity of all the raw data.

Ethics approval and consent to participate

The present study complied with the ethical guidelines of the Declaration of Helsinki and was approved by the Medical Research Ethics Committee of the Naval Medical Center (Shanghai, China; approval no. 2019010101). Oral informed consent for participation was obtained from the patients.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Ewald N, Hardt PD and Kloer HU: Severe hypertriglyceridemia and pancreatitis: presentation and management. Curr Opin Lipidol 20: 497-504, 2009.
2. Han Y, Jiang X, Qin Y, Zhao Y, Zhang G and Liu C: Correction: A cross-sectional study exploring the relationship between the dietary inflammatory index and hyperlipidemia based on the national health and nutrition examination survey (2005-2018). Lipids Health Dis 22: 161, 2023.
3. Malpica D: Metabolic syndrome, hyperlipidemias, and associated clinical markers among military airmen. Aerosp Med Hum Perform 94: 604-609, 2023.
4. Nelson RH: Hyperlipidemia as a risk factor for cardiovascular disease. Prim Care 40: 195-211, 2013.
5. Alaminos-Torres A, Martínez-Álvarez JR, López-Ejeda N and Marrodán-Serrano MD: Atherogenic risk, anthropometry, diet and physical activity in a sample of Spanish commercial airline pilots. Int J Environ Res Public Health 19: 4128, 2022.

6. Wang D, Chen J, Cai W, Chen Y, Jiang C and Yu X: Hyperlipidemia and intestinal microflora distributions among flying personnel. Prev Med 34: 665-671, 2022.
7. Liu P, Bai S, Chen X, Mu H, Wang R, Li F and Du P: Risk factors analysis and health management countermeasures of dyslipidemia of flight personnel in 2020 annual physical examination. In: Man-Machine-Environment System Engineering: Proceedings of the 21st International Conference on MMESE. Long S and Dhillon BS (eds). Springer, Singapore, pp88-93, 2022.
8. Maculewicz E, Pabin A, Kowalczuk K, Dziuda Ł and Białek A: Endogenous risk factors of cardiovascular diseases (CVDs) in military professionals with a special emphasis on military pilots. J Clin Med 11: 4314, 2022.
9. Yu L, Zhang Y and Wang L: Analysis of hyperlipidemia in fighter pilots. Chin J Convalescent Med 21: 166-167, 2012.
10. Bin-Bin S, Xiao-Wei WU, Chun-Lei Z, *et al*: Disease spectrum of hospitalized helicopter pilots. Mil Med J S China, 2012.
11. Dong Y, Chen H, Ma J, *et al*: Investigation and analysis of diseases of high-performance fighter pilots in different regions. J Third Mil Med Univ 22: 2313-2314, 2015.
12. Smith RC, Segman RH, Golcer-Dubner T, Pavlov V and Lerer B: Allelic variation in ApoC3, ApoA5 and LPL genes and first and second generation antipsychotic effects on serum lipids in patients with schizophrenia. Pharmacogenomics J 8: 228-236, 2008.
13. Bays HE, Kirkpatrick C, Maki KC, Toth PP, Morgan RT, Tondt J, Christensen SM, Dixon D and Jacobson TA: Obesity, dyslipidemia, and cardiovascular disease: A joint expert review from the obesity medicine association and the national lipid association 2024. Obes Pillars 10: 100108, 2024.
14. Huang H, Liu J, Feng Y and Chen W: The distribution of apolipoprotein E gene polymorphism in Chinese civil aircrews, and a possible risk factor to their overweight and dyslipidemia is cumulative flight time. Clin Chim Acta 416: 36-40, 2013.
15. Qing C, Hong-Jin L and Tao C: Effect of low density lipoprotein receptor gene nco site polymorphism on blood lipid in pilots. J Prev Med Chin People's Liberation Army 26: 242-245, 2008.
16. Cai Q, Liu H J, Chen T, *et al*: Relationship between Hind III site polymorphism of lipoprotein lipase gene and serum lipid level in pilots. Med J National Defending Forces S China 20: 723-726, 2010.
17. Qing C, Hong-Jin L and Tao C: Relationship between lipoprotein gene Ser447Ter site polymorphism and serum lipid in pilots. Clin J Med Off 39: 209-211, 2011.
18. Jiang JY, Sun YC, Zhan X and Yan CQ: Study on evaluating pilot workload based on muluti-source physiological data fusion. Chin J Ergon 29, 2023.
19. Liu T, Qiu B, Zheng J, Zhang C, Qi Y, Fan J, Li L and Gao J: Prevalence of and trends for dyslipidemia among pilots from one airline in China. Int J Travel Med Glob Healt 7: 18-22, 2019.
20. Kong L, Zhao H, Wang F, Zhang R, Yao X, Zuo R, Li J, Xu J, Qian Y, Kang Q and Fan C: Endocrine modulation of brain-skeleton axis driven by neural stem cell-derived perilipin 5 in the lipid metabolism homeostasis for bone regeneration. Mol Ther 31: 1293-1312. 2023.
21. Choi YY and Kim KY: Effects of physical examination and diet consultation on serum cholesterol and health-behavior in the Korean pilots employed in commercial airline. Ind Health 51: 603-611, 2013.
22. Xiao D, Li H, Wang X, Wang B, Yan Y and Men K: Prevalence of disease spectrum and sick leave time associated with illness in helicopter pilots. Aviat Space Environ Med 84: 234-236, 2013.
23. World Health Organization (WHO): International statistical classification of diseases and related health problems (ICD-10) in occupational health. WHO, Geneva, 1999. https://iris.who.int/bitstream/handle/10665/66100/WHO_SDE_OEH_99.11.pdf?sequence=1. Accessed May 26, 2024.
24. Atkins DC: Using multilevel models to analyze couple and family treatment data: basic and advanced issues. J Fam Psychol 19: 98-110, 2005.
25. Stawski RS: Multilevel analysis: An introduction to basic and advanced multilevel modeling (2nd edition). Struct Equ Model 20: 541-550, 2013.
26. Eilat-Tsanani S, Ernst P and Suissa S: Real-world effectiveness of single-inhaler triple therapy for COPD: Impact of diabetes comorbidity. COPD 21: 2327345, 2024.
27. Rosenbaum PR and Rubin DB: The central role of the propoensity score in observational studies for causal effects. Biometrika 70: 41-55, 1983.
28. Meir R and Rätsch G: An introduction to boosting and leveraging. In: Advanced Lectures on Machine Learning. Mendelson S and Smola A (eds). Springer, Berlin, Heidelberg, pp119-184, 2003.
29. McCaffrey DF, Ridgeway G and Morral AR: Propensity score estimation with boosted regression for evaluating causal effects in observational studies. Psychol Methods 9: 403-425, 2004.
30. Lee BK, Lessler J and Stuart EA: Improving propensity score weighting using machine learning. Stat Med 29: 337-346, 2010.
31. George EL and Panos A: Does a high WBC count signal infection? Nursing 35: 20-2, 2005.
32. Chen JS, Xie PF and Feng H: The role of exercise in improving hyperlipidemia-renal injuries induced by a high-fat diet: A literature review. PeerJ 11: e15435, 2023.
33. Pendry SD, Singhal N, Neo EL, Foreman D and Winter JM: Elevation of the tumor marker CA19-9 in a pancreatic cancer survivor with benign prostatic hyperplasia: A clinical case report. Clin Case Rep 12: e8929, 2024.
34. Liu S and Yuan H: The further evaluation of the clinical significance of serum CA199 and its diagnostic value of hepatobiliary and pancreatic diseases. J Med Theor Prac (Chin) 22: 1417-1419, 2009.
35. Bao SQ, Li FB and Jiang X: Correlation between carbohydrate antigen 199 and glycemic control in patients with type 2 diabetes mellitus. Chin Med J (Engl) 132: 984-986, 2019.
36. Gul K, Nas S, Ozdemir D, Gumus M, Ersoy R and Cakir B: CA 19-9 level in patients with type 2 diabetes mellitus and its relation to the metabolic control and microvascular complications. Am J Med Sci 341: 28-32, 2011.
37. Adachi M, Sekine T, Umemoto A, Tsukikawa S, Imai K and Yachi A: Mechanism of clearance of circulating CA19-9 in rats. Tumour Biol 11: 51-58, 1990.
38. Al-Tonsi AA, Abdel-Gayoum AA and Saad M: The secondary dyslipidemia and deranged serum phosphate concentration in thyroid disorders. Exp Mol Pathol 76: 182-187, 2004.
39. Xu M, Li Y, Zhao W, Song X, Gan G, Li B and Zhou X: Association between admission prothrombin time activity and hospital readmission in heart failure: A retrospective study. Clin Chim Acta 548: 117463, 2023.
40. Lai M, Peng H, Wu X, Chen X, Wang B and Su X: IL-38 in modulating hyperlipidemia and its related cardiovascular diseases. Int Immunopharmacol 108: 108876, 2022.
41. Xu DM, Li Q, Yi JX, Cai XJ, Xie L, Fang W, Qiu JF, Xu CW, He CL, Xu XR, *et al*: Investigation of lymphocyte subsets in peripheral blood of patients with dyslipidemia. Int J Gen Med 14: 5573-5579, 2021.
42. Robinson GA, Peng J, Pineda-Torra I, Ciurtin C and Jury EC: Metabolomics defines complex patterns of dyslipidaemia in juvenile-SLE patients associated with inflammation and potential cardiovascular disease risk. Metabolites 12: 3, 2021.
43. Georgiadis AN, Papavasiliou EC, Lourida ES, Alamanos Y, Kostara C, Tselepis AD and Drosos AA: Atherogenic lipid profile is a feature characteristic of patients with early rheumatoid arthritis: Effect of early treatment-a prospective, controlled study. Arthritis Res Ther 8: R82, 2006.
44. Yan J, Yang S, Han L, Ba X, Shen P, Lin W, Li T, Zhang R, Huang Y, Huang Y, *et al*: Dyslipidemia in rheumatoid arthritis: The possible mechanisms. Front Immunol 14: 1254753, 2023.
45. Myasoedova E, Crowson CS, Kremers HM, Fitz-Gibbon PD, Therneau TM and Gabriel SE: Total cholesterol and LDL levels decrease before rheumatoid arthritis. Ann Rheum Dis 69: 1310-1314, 2010.
46. Dursunoğlu D, Evrengül H, Polat B, Tanrıverdi H, Çobankara V, Kaftan A and Kılıç M: Lp(a) lipoprotein and lipids in patients with rheumatoid arthritis: Serum levels and relationship to inflammation. Rheumatol Int 25: 241-245, 2005.
47. Warrington KJ, Kent PD, Frye RL, Lymp JF, Kopecky SL, Goronzy JJ and Weyand CM: Rheumatoid arthritis is an independent risk factor for multi-vessel coronary artery disease: A case control study. Arthritis Res Ther 7: R984-R991. 2005.
48. Mainieri F, La Bella S and Chiarelli F: Hyperlipidemia and cardiovascular risk in children and adolescents. Biomedicines 11: 809, 2023.
49. Wisse BE: The inflammatory syndrome: The role of adipose tissue cytokines in metabolic disorders linked to obesity. J Am Soc Nephrol 15: 2792-2800, 2004.
50. Peyronnel C, Totoson P, Martin H and Demougeot C: Relevance of circulating markers of endothelial activation for cardiovascular risk assessment in rheumatoid arthritis: A narrative review. Life Sci 314: 121264, 2023.

51. Ferreira HB, Melo T, Paiva A and Domingues MDR: Insights in the role of lipids, oxidative stress and inflammation in rheumatoid arthritis unveiled by new trends in lipidomic investigations. Antioxidants (Basel) 10: 45, 2021.
52. Bonacina F, Pirillo A, Catapano AL and Norata GD: HDL in immune-inflammatory responses: Implications beyond cardiovascular diseases. Cells 10: 1061, 2021.
53. Kreiner FF, Kraaijenhof JM, von Herrath M, Hovingh GKK and von Scholten BJ: Interleukin 6 in diabetes, chronic kidney disease, and cardiovascular disease: Mechanisms and therapeutic perspectives. Expert Rev Clin Immunol 18: 377-389, 2022.
54. Leclercq G, Servera LA, Danilin S, Challier J, Steinhoff N, Bossen C, Odermatt A, Nicolini V, Umaña P, Klein C, *et al*: Dissecting the mechanism of cytokine release induced by T-cell engagers highlights the contribution of neutrophils. Oncoimmunology 11: 2039432, 2022.
55. Ou Q, Power R and Griffin MD: Revisiting regulatory T cells as modulators of innate immune response and inflammatory diseases. Front Immunol 14: 1287465, 2023.
56. Bonfigli AR, Protic O, Olivieri F, Montesanto A, Malatesta G, Di Pillo R and Antonicelli R: Effects of a novel nutraceutical combination (BruMeChol™) in subjects with mild hypercholesterolemia: Study protocol of a randomized, double-blind, controlled trial. Trials 21: 616, 2020.
57. Alanazi A: Clinicians' views on using artificial intelligence in healthcare: Opportunities, challenges, and beyond. Cureus 15: e45255, 2023.
58. Lalo R, Zekja I and Kamberi F: Association of cardiovascular disease risk and health-related behaviors in stroke patients. Int J Environ Res Public Health 20: 3693, 2023.
59. Michos ED, McEvoy JW and Blumenthal RS: Lipid management for the prevention of atherosclerotic cardiovascular disease. N Engl J Med 381: 1557-1567, 2019.
60. Damen JA, Hooft L, Schuit E, Debray TP, Collins GS, Tzoulaki I, Lassale CM, Siontis GC, Chiocchia V, Roberts C, *et al*: Prediction models for cardiovascular disease risk in the general population: Systematic review. BMJ 353: i2416, 2016.
61. Ornish D, Scherwitz LW, Billings JH, Brown SE, Gould KL, Merritt TA, Sparler S, Armstrong WT, Ports TA, Kirkeeide RL, *et al*: Intensive lifestyle changes for reversal of coronary heart disease. JAMA 280: 2001-2007, 1998.
62. Kreisberg RA and Oberman A: Medical management of hyperlipidemia/dyslipidemia. J Clin Endocrinol Metab 88: 2445-2461, 2003.
63. Liu T, Qiu B, Zhang C, Deng M, Liang Z and Qi Y: Health-related quality of life in pilots of a Chinese commercial airline. Arch Environ Occup Health 76: 511-517, 2021.
64. Garg A and Radhakrishnan S: Pediatric hyperlipidemia. Indian Heart J 76 (Suppl 1): S104-S107, 2024.
65. Joint Committee on the Chinese Guidelines for Lipid Management: Chinese guidelines for lipid management (2023). Zhonghua Xin Xue Guan Bing Za Zhi 51: 221-255, 2023 (In Chinese).
66. Bernier AV, Sentner CP, Correia CE, Theriaque DW, Shuster JJ, Smit GPA and Weinstein DA: Hyperlipidemia in glycogen storage disease type III: Effect of age and metabolic control. J Inherit Metab Dis 31: 729-732, 2008.