



OPEN

## The association between lipid biomarkers and osteoarthritis based on the National Health and Nutrition Examination Survey and Mendelian randomization study

Guoxin Huang<sup>1,4</sup>, Xian Zhong<sup>2,4</sup>, Meiling Zhang<sup>3</sup>, Ming Xu<sup>2✉</sup>, Bin Pei<sup>1✉</sup> & Da Qian<sup>2✉</sup>

To explore the association between lipid markers and osteoarthritis (OA). First, the National Health and Nutrition Examination Survey (NHANES) database was used to screen participants with lipid markers, OA and relevant covariates, and logistic regression was used to analyze the association between lipid markers and OA; Then, under the theoretical framework of Mendelian randomization (MR), two-sample MR was performed using GWAS data of lipid markers and OA to explore the causal association between the two, which was analyzed by inverse variance weighting (IVW) method. Heterogeneity test, sensitivity analysis and pleiotropy analysis were also performed. The NHANES database screened a total of 3706 participants, of whom 836 had OA and 2870 did not have OA. When lipid markers were used as continuous variables, multivariate logistic results showed an association between HDL, LDL and OA (HDL, OR (95%):1.01 (1.00, 1.01); LDL, OR (95%):1.00 (0.99, 1.00)). When lipid markers were used as categorical variables, multivariate logistic results showed the fourth quartile result of 0.713 (0.513, 0.992) for LDL relative to the first quartile. In MR study, the results of the IVW method for TG, TL, HDL and LDL showed OR (95% CI) of 1.06 (0.97–1.16), 0.95 (0.85–1.06), 0.94 (0.86–1.02) and 0.89 (0.80–0.998) with P-values of 0.21, 0.37, 0.013, 0.046. The heterogeneity tests and multiplicity analyses showed P-values greater than 0.05, and sensitivity analyses showed no abnormal single nucleotide polymorphisms. Through NHANES database and MR analyses, LDL was found to be a protective factor for OA, while HDL still needs further study. Our results provide new biomarkers for preventive and therapeutic strategies for OA.

Degenerative joint diseases such as osteoarthritis (OA), the clinical manifestations are characterized by joint pain, restricted joint movement and bone rubbing sounds, affect patients' quality of life seriously<sup>1</sup>. Globally, the age-standardized point prevalence and the annual incidence rate of OA in 2017 were 3754.2<sup>2</sup> and 181.2 per 100,000 and has a wide range of factors such as age, obesity, inflammation, etc<sup>3</sup>. In recent years, studies have identified lipid biomarkers that may be associated with the development of OA, including triglycerides (TG), total cholesterol (TC), high-density lipoprotein (HDL), and low-density lipoprotein (LDL)<sup>4</sup>. For example, Puenpatom et al. found differences between TG and HDL in OA and non-OA populations<sup>5</sup>. Moreover, Garcia-Gil et al. found that levels of TG were associated with hand OA, whereas TC and LDL were not associated with hand OA<sup>6</sup>. Though these studies have explored the association between lipid biomarkers and OA, the association between the two is still unclear.

<sup>1</sup>Department of Evidence-Based Medicine Center, Xiangyang No.1 People's Hospital, Hubei University of Medicine, Xiangyang 441000, China. <sup>2</sup>Department of Burn and Plastic Surgery-Hand Surgery, Changshu Hospital Affiliated to Soochow University, Changshu No.1 People's Hospital, Changshu 215500, China. <sup>3</sup>Department of the second ward of Orthopedic, Xiangyang No.1 People's Hospital, Hubei University of Medicine, Xiangyang 441000, China. <sup>4</sup>These authors contributed equally: Guoxin Huang and Xian Zhong. ✉email: drxuming@126.com; binpei@hbm.u.edu.cn; drqianda@hotmail.com

Existing methods for studying causal analysis contain propensity score matching (PSM), inverse probability weighting (IPW) and Mendelian randomization. PSM is mainly used to deal with selective bias in observational data, and IPW is more suitable for dealing with missing data or non-response. Among these methods, MR provides a way to estimate causal effects based on genetic information, which can better control for potential confounders and thus reduce bias and uncertainty. MR is a statistical method for analysing the causal relationship between exposure factors and outcomes by using instrumental variables (IV)<sup>7</sup>. George published the first MR article in 2003, suggesting that MR can provide insight into environmental determinants of disease and formalizing a research framework and study design for MR<sup>8</sup>. MR, which is based on whole-genome sequencing data, is effective in reducing bias, similar to RCT studies, and has been widely used in studies of causal relationships between exposure factors and outcomes<sup>9</sup>. Mendelian randomisation is one of the more widely used causal inference methods in epidemiology in recent years<sup>10</sup>.

The National Health and Nutrition Examination Survey (NHANES) is a population-based cross-sectional survey that is designed to collect information on the health and nutrition of the US population, containing demographic data and lifestyle and health and nutrition status information on participants, and has been used extensively in the study of morbidity factors<sup>11,12</sup>. In this study we screened participants with lipid markers from NHANES database and firstly combined cross-sectional study and MR study to explore the association between them from two dimensions, which makes the conclusions more reliable.

## Methods and materials

This study firstly examined the association between lipid biomarkers and OA using data from the NHANES. Then, in the framework of MR analysis, we performed two-sample MR based on published data from genome-wide association studies to further assess the causal association between lipid biomarker and OA.

### Study population in NHANES

In the cross-sectional study, data from a total of nine cycles of the NHANES database from 2003 to 2020 were used. Data on the presence or absence of osteoarthritis were obtained from survey documentation data. The data were first used to identify participants without osteoarthritis according to "Doctor ever said you had arthritis" and then to identify participants with osteoarthritis according to "Which type of arthritis". Please see the NHANES database for details of the populations that can be measured. ([https://www.cdc.gov/Nchs/Nhanes/2003-2004/MCQ\\_C.htm#Component\\_Description](https://www.cdc.gov/Nchs/Nhanes/2003-2004/MCQ_C.htm#Component_Description)). Informed consent was obtained from all subjects in the National Health and Nutrition Examination Survey. Information for the assessment of OA was obtained through questionnaires and self-reports<sup>13,14</sup>. In a study conducted by March et al. the agreement between self-reported OA and clinically well-defined OA was 85%, indicating that OA can be accurately diagnosed in the majority of case reports<sup>15</sup>.

### Lipid biomarkers in NHANES

Lipid biomarkers include TG, TC, HDL and LDL. Lipid Biomarker tests are performed on a fasted state and serum samples are processed, stored and transported to the partner laboratory service for analysis. TG, TC, HDL, LDL units are all in mg/dl<sup>16</sup>. The measurement of TC and TG was performed with enzymatic assays, and the measurement of HDL was performed with immunoassays. LDL was calculated according to the Friedewald equation<sup>17</sup>. Detailed specimen collection and processing instructions are discussed in the NHANES Laboratory/Medical Technologists Procedures Manual. The NHANES quality control and quality assurance protocols meet the 1988 Clinical Laboratory Improvement Act mandates.

### Covariates used in NHANES

Refer to previous studies<sup>18–20</sup>, variables that may confound the association between lipid biomarkers and OA were collected. Information regarding demographics and lifestyle factors was collected by questionnaire, including age (years), sex (female, male), race/ethnicity (White, Mexican, Black, other), education (under high school, high school or equivalent, above high school), marital status (married, living with partner, separated, divorced, widowed, never married), smoking (never, former, now), alcohol intake (never, former, heavy, mild, moderate), cancer (no, yes), use of atherosclerosis conditions (no, unknown, yes). Health examination was performed in the mobile centers. Body mass index (BMI, kg/m<sup>2</sup>). Diabetes was defined as achieving a fasting glucose level of 126 mg/dL or reporting a previous diagnosis. Hypertension was defined as resting blood pressure (BP) persistently at or above 140/90 mmHg or reporting a previous diagnosis. Poverty is an index based on family income and federally defined poverty thresholds, reflecting income related to family needs<sup>21</sup>. Poverty ranges from 0 (no income) to 5 (greater than or equal to five times the federal poverty level)<sup>22</sup>. Physical activity is measured in MET (Metabolic equivalent, MET). MET (metabolic equivalent, MET) is the oxygen consumption required to maintain resting metabolism<sup>23</sup>. Based on energy expenditure in a quiet sitting position, it is a commonly used indicator to express the relative level of energy metabolism in various activities. The recommended MET values for work-related vigorous exercise, moderate exercise, walking or cycling, strenuous exercise, and moderate exercise were 8.0, 4.0, 4.0, 8.0, and 4.0, respectively. For each activity, physical activity was calculated in MET-min per week by multiplying the number of days by the average duration times the recommended MET and summing the resulting values to obtain an estimate of total physical activity<sup>24</sup>.

### Sources of MR

The MR study used two-sample MR. Exposure factors were TG, TC, HDL, LDL. GWAS data for TG, HDL and LDL are from the paper published by Richardson et al. in 2020<sup>25</sup>. TC data from the paper Borges CM et al. 2020<sup>25</sup>. Outcome factors for OA were derived from hospital-confirmed OA, containing 50,508 Sample sizes and 15,845,511 Single Nucleotide Polymorphisms (SNPs) sequenced by Zengini E et al. in 2018<sup>26</sup>. Informed consent

was obtained from all subjects in the original genome-wide association studies. Detailed information is shown in Table 1.

### Statistical analysis

The NHANES data analysis refers to the NHANES statistical tutorial and follows its complex multi-stage probability sampling with weighting of the sample. The weighting variable was chosen as  $wtsaf2yr$ , calculated as  $1/9 * wtsaf2yr$ , and all analyses were performed under complex weighting. Continuous variables in normal distribution should be described as mean  $\pm$  standard deviation (SD) or else reported as median (Range). Variance homogeneous and normal distributed continuous variables could be compared by student t-test, otherwise, the Mann–Whitney *U*-test or Kruskal–Wallis H-test should be used. Count data were statistically described by rates, and Poisson regression or Negative binomial regression was used for comparison between groups. Weight logistic regression models were used to test the associations of TG, TC, HDL, LDL with OA. All covariates were using the lowest quartile as the reference. Model 1 is weight logistic regression, the independent variable is each lipid biomarker and the dependent variable is OA; Model 2 was adjusted for age, sex, and race/ethnicity; Model 3 was further adjusted for age, sex, race/ethnicity, BMI, marital status, education, poverty, smoking, alcohol intake, hypertension, diabetes, cancer, physical activity and atherosclerosis. To better explore the association between lipid biomarkers and OA, logistic regression was conducted to explore the association with OA when lipids are used as quartiles.

MR was used to explore the causal relationship between lipid biomarkers and OA. All IV were selected using the same criteria. Exposure factors with genome-wide significance parameters were set to  $P < 5 \times 10^{-8}$ , the linkage disequilibrium parameter ( $r^2$ ) parameter was set to 0.001, and the genetic distance was set to 10 MB to screen for IV with no linkage effects. Association between lipid biomarker and OA was assessed using an inverse variance weighting (IVW) method as the main statistical method, theory has been described in previous studies<sup>27,28</sup>. Heterogeneity was examined using the IVW method and the MR-Egger method. Sensitivity analysis was performed using the leave-one-out method. Pleiotropy analysis was performed using the Egger-intercept method. Finally, the strength of association of the genetic instruments for each putative risk factor was quantified by the *F* statistic ( $F = \beta^2/se^2$ ) for all SNPs, to assess the power of the SNPs<sup>29</sup>. All statistical analyses were performed using R software (Version 4.1.2; <http://www.R-project.org>, R Foundation for Statistical Computing, TUNA Team, Tsinghua University), the "nhanesR" package for NHANES data analysis and the "TwoSampleMR" package for MR analysis.

### Ethical approval

Informed consent was obtained from all subjects in the original genome-wide association studies and National Health and Nutrition Examination Survey, which were approved by NCHS Ethics Review Board. Therefore, per the guidelines of the XYZ Institutional Review Board, IRB assessment was not necessary.

## Results

### Lipid biomarkers and OA in NHANES

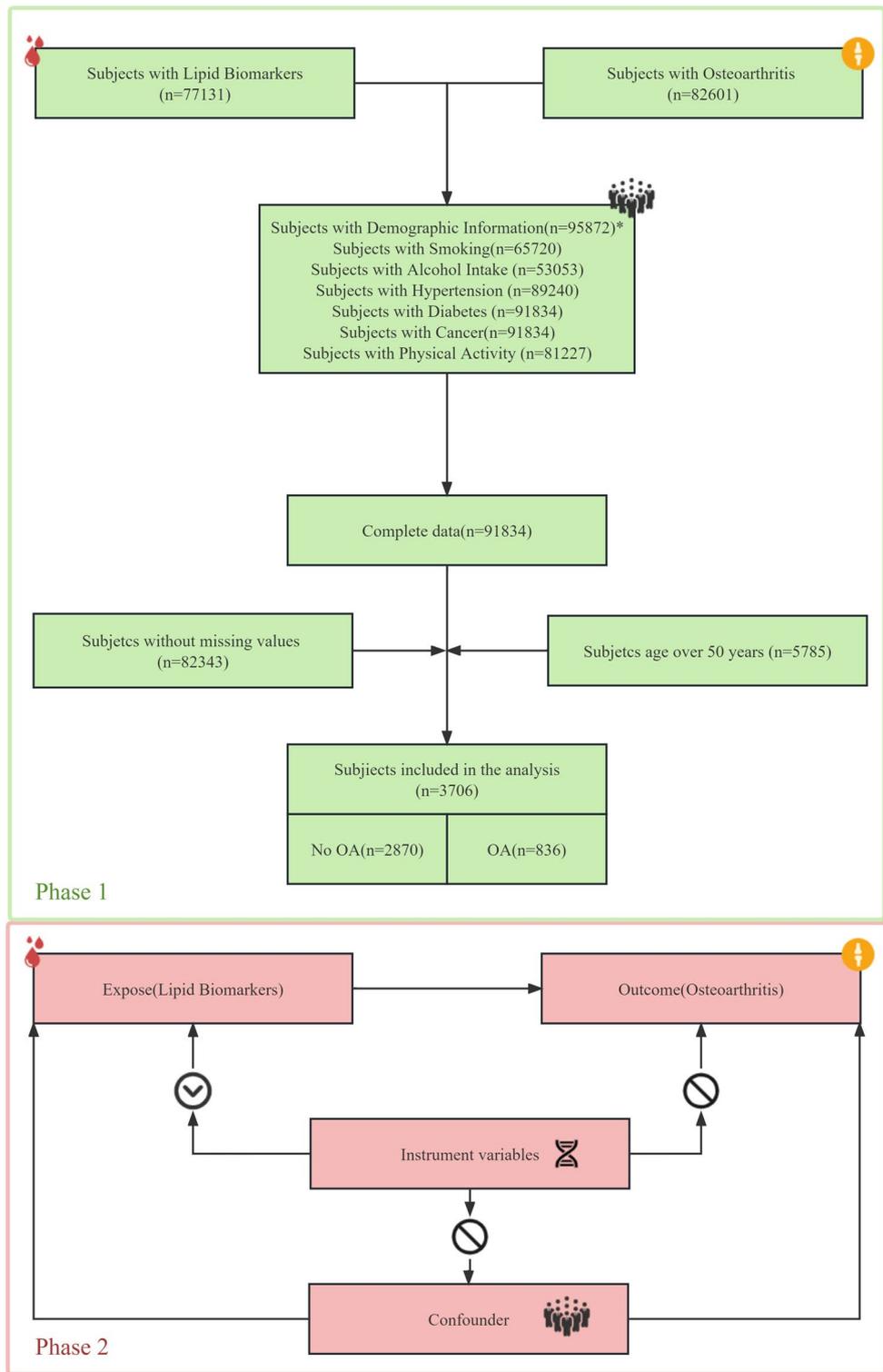
Between 2003 and 2020, 82,601 participants were evaluated for OA and 77,131 participants were available for lipid biomarkers results, resulting in a total of 91,834 participants after combining all covariates. After removing participants with missing values, 9492 participants remained. 3706 participants were older than 50 years, of whom 836 had OA and 2870 did not have OA. After weighting, this represents 40,802,041 participants. The detailed process is shown in Fig. 1. The two groups were grouped according to whether they had OA or not, and there was a statistical difference ( $P < 0.05$ ) between the two groups in age, BMI, race/ethnicity, marital status, alcohol intake, hypertension, cancer, atherosclerosis, HDL, LDL in Table 2. The results of univariate logistic showed that the OR (95% CI) for TG, TC, HDL and LDL were 1.00 (1.00, 1.00), 1.00 (1.00, 1.00), 1.01 (1.00, 1.01) and 0.99 (0.99, 1.00), respectively, with P values of 0.12, 0.08, 0.02 and  $< 0.001$ . The results of multifactorial logistic (model 3) showed that the OR (95% CI) for TG, TC, HDL, and LDL were 1.00 (1.00, 1.00), 1.00 (1.00, 1.00), 1.01 (1.00, 1.01), and 1.00 (0.99, 1.00), respectively, and the p-value was 0.62, 0.37, 0.049, and 0.049, respectively. The detailed results are shown in Table 3. When lipid biomarkers were divided into quartiles, the detailed results of the logistics results are shown in Table 3.

### Causal association between lipid biomarkers and OA in MR

The same statistical process was used for the causal analysis of lipid biomarkers and OA. The IVW results for TG showed an OR (95% CI) of 1.059 (0.969 to 1.157), TC showed an OR (95% CI) of 0.950 (0.851 to 1.061), HDL showed an OR (95% CI) of 0.936 (0.858 to 1.021), LDL showed an OR (95% CI) of 0.892 (0.797–0.998). The

Expose/outcome	Year	Population	Sex	Sample size	Number of SNPs	Author	GWAS ID	PMID
TG	2020	European	Males and females	441,016	12,321,875	Richardson Tom	ieu-b-111	32203549
TC	2020	European	Males and females	115,078	12,321,875	Borges CM	met-d-Total_C	NA
HDL	2020	European	Males and females	403,943	12,321,875	Richardson Tom	ieu-b-109	32203549
LDL	2020	European	Males and females	440,546	12,321,875	Richardson Tom	ieu-b-110	32203549
OA	2018	European	Males and females	50,508	15,845,511	Zengini E	ebi-a-GCST005814	29559693

**Table 1.** Detailed information on expose/outcome factor.



**Figure 1.** Study design overview.

F-values are all greater than 10. Heterogeneity tests, sensitivity analysis, and pleiotropy analysis were all negative. According to the three assumptions of MR, there is a causal relationship between LDL and OA, and there is no causal relationship between TG, TC, HDL and OA. Detailed results are shown in Table 4.

Characteristics	Total	No-OA	OA	P value
Sex, n (%)				
Female	1993 (49.76)	1676 (53.74)	317 (38.42)	Ref
Male	1713 (50.24)	1194 (46.26)	519 (61.58)	<0.0001
Age (year), mean (SD)	61.80 (0.19)	60.81 (0.21)	64.61 (0.36)	<0.0001
BMI, mean (SD)	28.42 (0.13)	27.94 (0.14)	29.78 (0.26)	<0.0001
Race/ethnicity, n (%)				
White	1946 (79.11)	1382 (76.61)	564 (86.24)	Ref
Mexican	462 ( 3.88)	396 (4.47)	66 (2.19)	<0.0001
Black	624 ( 7.65)	521 (8.61)	103 (4.90)	<0.0001
Other	674 ( 9.37)	571 (10.31)	103 (6.68)	<0.0001
Marital status, n (%)				
Married	2282 (66.44)	1778 (66.57)	504 (66.05)	Ref
Living with partner	133 ( 3.22)	108 (3.48)	25 (2.49)	0.431
Separated	108 ( 1.61)	96 (1.94)	12 (0.67)	0.019
Divorced	510 (13.48)	397 (13.64)	113 (13.01)	0.975
Widowed	425 ( 9.16)	290 ( 7.96)	135 (12.58)	<0.0001
Never married	248 ( 6.09)	201 (6.41)	47 (5.19)	0.316
Education, n (%)				
Under high school	768 (11.11)	637 (11.51)	131 (9.96)	Ref
High school or equivalent	837 (23.19)	636 (23.68)	201 (21.80)	0.002
Above high school	2101 (65.70)	1597 (64.81)	504 (68.24)	<0.0001
Poverty, mean (SD)	3.51 (0.05)	3.52 (0.05)	3.45 (0.10)	0.43
Smoking, n (%)				
Never	1849 (50.86)	1455 (52.02)	394 (47.54)	Ref
Former	1312 (35.04)	982 (33.55)	330 (39.31)	0.026
Now	545 (14.10)	433 (14.44)	112 (13.15)	0.735
Alcohol intake, n (%)				
Never	488 (10.48)	379 (10.34)	109 (10.88)	Ref
Former	732 (15.24)	572 (15.48)	160 (14.57)	0.862
Mild	1625 (48.59)	1223 (47.07)	402 (52.92)	0.344
Moderate	463 (14.56)	358 (14.79)	105 (13.88)	0.912
Heavy	398 (11.13)	338 (12.32)	60 ( 7.74)	0.015
Hypertension, n (%)				
No	1549 (46.28)	1274 (49.29)	275 (37.68)	Ref
Yes	2157 (53.72)	1596 (50.71)	561 (62.32)	<0.0001
Diabetes, n (%)				
No	3091 (87.65)	2405 (88.31)	686 (85.75)	Ref
Yes	615 (12.35)	465 (11.69)	150 (14.25)	0.295
Cancer, n (%)				
No	3154 (83.75)	2485 (85.47)	669 (78.85)	Ref
Yes	552 (16.25)	385 (14.53)	167 (21.15)	<0.0001
Physical activity (MET/week), mean (SD)	3103.84 (101.91)	3149.90 (116.79)	2972.47 (192.50)	0.42
Atherosclerosis, n (%)				
No	974 (24.99)	883 (30.07)	91 (10.51)	Ref
Unknown	2728 (74.93)	1984 (69.85)	744 (89.43)	<0.0001
Yes	4 (0.08)	3 (0.08)	1 (0.06)	0.328
TG (mg/dl), mean (SD)	123.50 (1.65)	122.10 (1.90)	127.51 (3.02)	0.12
TC (mg/dl), mean (SD)	201.47 (0.91)	202.39 (1.04)	198.84 (1.72)	0.07
HDL (mg/dl), mean (SD)	57.43 (0.48)	56.90 (0.55)	58.95 (0.75)	0.02
LDL (mg/dl), mean (SD)	119.33 (0.79)	121.07 (0.88)	114.39 (1.53)	<0.001

**Table 2.** Weighted selected characteristics of study population in female and male, NHANES (Weighted N = 40,804,313).

Characteristics	Model	Continuous		Q1	Q2		Q3		Q4	
		OR (95% CI)	P value		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
TG	Model 1	1.00 (1.00,1.00)	0.115	Ref	0.998 (0.729,1.367)	0.992	1.270 (0.967,1.669)	0.085	1.226 (0.918,1.637)	0.167
	Model 2	1.00 (1.00,1.00)	0.051	Ref	0.952 (0.690,1.314)	0.952	1.210 (0.898,1.630)	1.21	1.269 (0.927,1.738)	1.269
	Model 3	1.00 (1.00,1.00)	0.619	Ref	0.920 (0.654,1.293)	0.627	1.063 (0.776,1.457)	0.7	1.033 (0.741,1.440)	0.847
TC	Model 1	1.00 (1.00,1.00)	0.077	Ref	0.930 (0.723,1.197)	0.572	0.875 (0.695,1.102)	0.255	0.771 (0.586,1.015)	0.063
	Model 2	1.00 (0.99,1.00)	0.007	Ref	0.974 (0.747,1.269)	0.842	0.849 (0.663,1.087)	0.192	<b>0.679 (0.508,0.907)</b>	<b>0.009</b>
	Model 3	1.00 (1.00,1.00)	0.374	Ref	1.046 (0.804,1.361)	0.735	0.975 (0.746,1.273)	0.849	0.860 (0.628,1.177)	0.342
HDL	Model 1	1.01 (1.00,1.01)	0.015	Ref	1.135 (0.834,1.545)	0.418	1.128 (0.821,1.552)	0.454	1.385 (1.027,1.870)	0.033
	Model 2	1.00 (0.99,1.00)	0.538	Ref	0.952 (0.688,1.316)	0.763	0.842 (0.607,1.168)	0.3	0.908 (0.654,1.262)	0.564
	Model 3	1.01 (1.00,1.01)	<b>0.049</b>	Ref	1.062 (0.759,1.487)	0.721	1.062 (0.747,1.510)	0.735	1.409 (0.953,2.083)	0.085
LDL	Model 1	0.99 (0.99,1.00)	< <b>0.001</b>	Ref	0.904 (0.670,1.220)	0.506	<b>0.672 (0.511,0.884)</b>	<b>0.005</b>	<b>0.579 (0.429,0.780)</b>	< <b>0.001</b>
	Model 2	0.99 (0.99,1.00)	< <b>0.001</b>	Ref	0.971 (0.704,1.340)	0.857	<b>0.726 (0.542,0.973)</b>	<b>0.033</b>	<b>0.582 (0.423,0.800)</b>	<b>0.001</b>
	Model 3	1.00 (0.99,1.00)	<b>0.049</b>	Ref	1.043 (0.759,1.432)	0.795	0.751 (0.547,1.033)	0.077	<b>0.713 (0.513,0.992)</b>	<b>0.044</b>

**Table 3.** Weighted ORs (95% CIs) of the associations between lipid biomarkers and OA. Model 1: no adjusted; Model 2: adjusted for age, sex, and race/ethnicity; Model 3: adjusted for all the covariates. TG: Q1 [18,77], Q2 (77,108], Q3 (108,153], Q4 (153,399]; TC: Q1 [79,171], Q2 (171,198], Q3 (198,226], Q4 (226,463]; HDL: Q1 [21, 44], Q2 (44,53], Q3 (53,66], Q4 (66,173]; LDL: Q1 [15,94], Q2 (94,116], Q3 (116,141], Q4(141,375]. TG triglycerides, TC total cholesterol, HDL high density lipoprotein, LDL low density lipoprotein. Significant values are in bold.

Exposure	SNPs	Method	OR(95%CI)	$P_{\text{Effect}}$	$P_{\text{Heterogeneity}}$	$P_{\text{Intercept}}$
TG	215	IVW	1.059 (0.969–1.157)	0.209	0.862	0.893
TC	48	IVW	0.950 (0.851–1.061)	0.366	0.761	0.929
HDL	247	IVW	0.936 (0.856–1.021)	0.133	0.971	0.243
LDL	126	IVW	0.892 (0.797–0.998)	0.046	0.676	0.864

**Table 4.** Mendelian randomization analysis of the main results. TG triglycerides, TC total cholesterol, HDL high density lipoprotein, LDL low density lipoprotein, IVW inverse variance weighting.

## Discussion

OA is a common clinical degenerative condition. With an increasing proportion of obese and old people, OA brings a huge economic burden to society<sup>30</sup>. Early detection of risk factors for osteoarthritis can help in the prevention and treatment of the disease. In our study, we combined a cross-sectional study and MR to explore the relationship between lipid biomarkers and OA. The logistic regression results showed no association between TG, TC, and OA, but an association between HDL, LDL and OA. MR used a two-sample MR method, and the results showed no causal association between TG, TC, HDL and OA, but a causal association between LDL and OA (IVW results showed an OR value of 0.892 (0.797–0.998),  $P$ -value = 0.046).

By summarizing previous research, we found that our results are more reliable. We combined a cross-sectional study and MR to explore the relationship between lipid biomarkers and OA. By combining these two methods, we can effectively combine the advantages and disadvantages of the two methods. Based on the above results, we have found that HDL is a protective factor for OA, which should be paid enough attention and has some guiding sense for the clinic. At the same time, the control of lipid biomarkers should be strengthened to help the prevention and treatment of OA.

TG is the most abundant and most productive energy substance in the body and has been found to be closely associated with diseases such as coronary heart disease and diabetes<sup>31</sup>. Previous studies on the relationship between TG and OA are unclear; Zhou et al. found that the prevalence and incidence of knee OA increased by 9% and 5% respectively for a one-unit increase in TG, respectively<sup>32</sup>; Puenpatom et al. found high TG in people with OA in comparison to those without OA (47% vs. 32%)<sup>5</sup>; Askaria et al. found an association between TG and OA<sup>33</sup>. In contrast, our study did not find an association between TG and OA in the cross-sectional study, the same conclusion as that found by Zhang et al.<sup>34</sup>, who found no difference in TG between the OA and healthy groups, and Hindy et al. found no association between TG and OA<sup>35</sup>. To confirm the results of the cross-sectional study, further analysis that utilised MR showed that there was no causal relationship between TG and OA. Previously, Funck-Brentano et al. also used MR but found no causal relationship between TG and OA<sup>36</sup>, and Zengini et al. also found no causal relationship between TG and OA<sup>26</sup>. Combining the cross-sectional results with those of MR reveals that there is no relationship between TG and OA.

TC is a lipid-like substance found in blood lipoproteins and essential for cells. Some previous studies have examined the association between TG and OA<sup>37</sup>. Singh et al. found that the OA group had a higher proportion of high TC (32% vs. 24%) than did the control group<sup>38</sup>; Abdurhman et al. found that high levels of TC were associated with OA<sup>22</sup>; Zhang et al. found that levels of TC were higher in the OA group in comparison to the

healthy group<sup>34</sup>. Meanwhile, Schwage et al. found no association between TC and OA, while Chingford found no association between TC and the incidence of hand OA<sup>39</sup>. In contrast, our study first used a cross-sectional study to find no association between TC and OA, and then combined it with MR to find no causal relationship between the two, thus supporting the conclusion that there is no association between TC and OA.

HDL is an anti-atherosclerotic lipoprotein that is synthesised mainly in the liver and transports cholesterol from extra-hepatic tissues to the liver for metabolism<sup>39</sup>. Our study showed HDL as a risk factor for OA in a cross-sectional study, which is the same as the findings of some of the previous studies. Pan et al. found an association between reduced HDL and a loss of medial tibial cartilage volume<sup>40</sup>, while Askaria et al. found an association between HDL and OA<sup>33</sup>; Zhang et al. found reduced levels of HDL in the OA group in comparison to the healthy group<sup>34</sup>; Puenpatom et al. found lower HDL in people with OA than in those without OA (44% vs. 38%)<sup>5</sup>. However, MR analysis shows no causal link between HDL and OA, which is the same finding as that of Hindy et al. and Schwage et al.<sup>35,39</sup>, who found no observed association between HDL and OA in their observational studies, while Funck-Brentano et al. used MR to find no causal relationship between HDL and OA<sup>36</sup>. The results of the cross-sectional study contradict the results of Mendelian randomisation, and further studies are needed to clarify the relationship between HDL and OA.

LDL is a cholesterol-rich lipoprotein. Kruisbergen et al. found that LDL activation of circulating monocytes was a risk factor for OA<sup>41</sup>; Oliviero et al. found higher levels of serum LDL in patients with OA in comparison to controls<sup>42</sup>; Mishra et al. found higher LDL in the OA group than in the control group<sup>43</sup>. However, the results of this type of study are contrary to the results of the present study. In the cross-sectional study, the logistic regression results showed that the OR for LDL was less than 1 and that the *P*-value was less than 0.05, whereby suggesting a association between LDL and OA, while the MR results showed an OR value (95% CI) of 0.892 (0.797–0.998). A heterogeneity test, sensitivity analysis, and pleiotropy analysis all showed negative results, which suggested a causal relationship between LDL and OA. The relationship between LDL and OA was demonstrated at two levels. Previously, George Hindy, E. Gill, Wang et al. using Mendelian randomisation, all found LDL to be a protective factor in OA<sup>44,45</sup>, consistent with the results of the present study, and suggested a corresponding possible mechanism by which LDL reduces APOA1 levels and serum amyloid A-induced arthritic inflammation in human primary chondrocytes and fibroblast-like synoviocytes.

However, there are still some shortcomings in this study. Due to the limitation of the data source, it is not possible to further analyse the type of OA, such as osteoarthritis of the knee, osteoarthritis of the hip, etc.; The OA data in the NHANES database is derived from questionnaires of patients' recollections, and there may be a certain recollection bias; Although the MR method was adopted in this study to investigate the causality of the two, but MR's prerequisite is the existence of a linear relationship between the two, if not then MR is not applicable. Although we have combined cross-sectional studies and MR, prospective cohort data are still needed for verifying this, and basic experiments can be performed to explore the role of lipid markers in the development of OA.

## Conclusion

In summary, our study used cross-sectional studies and MR to demonstrate the relationship between lipid biomarkers and OA. LDL is a protective factor for OA. No relationship exists between TG, TC and OA, while HDL still needs to be proved by further studies. Our findings provide new biomarkers for preventive and therapeutic strategies for OA, but further studies on the underlying mechanisms are still needed.

## Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 12 September 2023; Accepted: 6 January 2024

Published online: 16 January 2024

## References

- Hunter, D. J. Osteoarthritis preface. *Rheum. Dis. Clin. N. Am.* <https://doi.org/10.1016/j.rdc.2012.11.005> (2013).
- Safiri, S. et al. Global, regional and national burden of osteoarthritis 1990–2017: A systematic analysis of the Global Burden of Disease Study 2017. *Ann. Rheum. Dis.* **79**, 819–828 (2020).
- Sun, X. et al. Osteoarthritis in the middle-aged and elderly in China: Prevalence and Influencing Factors. *Int. J. Environ. Res. Public Health* **16**, 4701 (2019).
- Zarrrouk, A. et al. Lipid biomarkers in Alzheimer's disease. *Curr. Alzheimer Res.* **15**, 303–312 (2018).
- Puenpatom, R. A. & Victor, T. W. Increased prevalence of metabolic syndrome in individuals with osteoarthritis: An analysis of NHANES III data. *Postgrad. Med.* **121**, 9–20 (2009).
- Garcia-Gil, M. et al. Serum lipid levels and risk of hand osteoarthritis: The Chingford prospective cohort study. *Sci. Rep.* <https://doi.org/10.1038/s41598-017-03317-4> (2017).
- Skrivankova, V. W. et al. Strengthening the reporting of observational studies in epidemiology using Mendelian randomization: The STROBE-MR statement. *JAMA J. Am. Med. Assoc.* **326**, 1614–1621 (2021).
- Smith, G. D. & Ebrahim, S. "Mendelian randomization": Can genetic epidemiology contribute to understanding environmental determinants of disease?. *Int. J. Epidemiol.* **32**, 1–22 (2003).
- Yuan, S. et al. Genetic liability to rheumatoid arthritis in relation to coronary artery disease and stroke risk. *Arthritis Rheumatol.* **74**, 1638–1647 (2022).
- Birney, E. Mendelian randomization. *Cold Spring Harb. Perspect. Med.* **12**, a041302 (2022).
- Wang, M. et al. Coffee consumption and prostate cancer risk: Results from National Health and Nutrition Examination Survey 1999–2010 and Mendelian randomization analyses. *Nutrients* **13**, 2317 (2021).
- Wang, X., Seo, Y. A. & Park, S. K. Serum selenium and non-alcoholic fatty liver disease (NAFLD) in US adults: National Health and Nutrition Examination Survey (NHANES) 2011–2016. *Environ. Res.* **197**, 111190 (2021).

13. Davis, A. M., King, L. K., Stanaitis, I. & Hawker, G. A. Fundamentals of osteoarthritis: Outcome evaluation with patient-reported measures and functional tests. *Osteoarthr. Cartil.* **30**, 775–785 (2022).
14. Liang, Z. *et al.* Association between pyrethroid exposure and osteoarthritis: A national population-based cross-sectional study in the US. *BMC Public Health* <https://doi.org/10.1186/s12889-023-16225-2> (2023).
15. March, L. M., Schwarz, J. M., Carfrae, B. H. & Bagge, E. Clinical validation of self-reported osteoarthritis. *Osteoarthr. Cartil.* **6**, 87–93 (1998).
16. Fan, Y. *et al.* Serum albumin mediates the effect of multiple per- and polyfluoroalkyl substances on serum lipid levels. *Environ. Pollut.* **266**, 115138 (2020).
17. Nie, Q., Wang, C. & Zhou, L. Association between serum selenium levels and lipids among people with and without diabetes. *Nutrients* **15**, 3190 (2023).
18. Ekroos, K. *et al.* Lipid-based biomarkers for CVD, COPD, and aging—A translational perspective. *Prog. Lipid Res.* **78**, 101030 (2020).
19. Kan, B. *et al.* Association between lipid biomarkers and osteoporosis: A cross-sectional study. *BMC Musculoskelet. Disord.* <https://doi.org/10.1186/s12891-021-04643-5> (2021).
20. Wise, B. L. *et al.* Osteoporotic fractures men Mros G: Frailty and hip osteoarthritis in men in the MrOS cohort. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **69**, 602–608 (2014).
21. Chen, Z., He, J., Chen, L., Wu, X. & Yu, X. Association between the nickel exposure and lipid profiles in general population from NHANES. *Environ. Sci. Pollut. Res.* **29**, 66383–66388 (2022).
22. Abramoff, B. & Caldera, F. E. Osteoarthritis pathology, diagnosis, and treatment options. *Med. Clin. N. Am.* **104**, 293 (2020).
23. Tian, X. *et al.* Physical activity reduces the role of blood cadmium on depression: A cross-sectional analysis with NHANES data. *Environ. Pollut.* **304**, 119211 (2022).
24. Feng, X. *et al.* Association between physical activity and kidney stones based on dose-response analyses using restricted cubic splines. *Eur. J. Public Health* **30**, 1206–1211 (2020).
25. Richardson, T. G. *et al.* Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS Med.* **17**, e1003062 (2020).
26. Zengini, E. *et al.* Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis. *Nat. Genet.* **50**, 549 (2018).
27. Hong, W. *et al.* Gut microbiome causal impacts on the prognosis of breast cancer: A Mendelian randomization study. *BMC Genom.* <https://doi.org/10.1186/s12864-023-09608-7> (2023).
28. Carter, P. *et al.* Coffee consumption and cancer risk: A Mendelian randomisation study. *Clin. Nutr.* **41**, 2113–2123 (2022).
29. Pierce, B. L. & Burgess, S. Efficient design for Mendelian randomization studies: Subsample and 2-sample instrumental variable estimators. *Am. J. Epidemiol.* **178**, 1177–1184 (2013).
30. Victoria Castell, M. *et al.* Osteoarthritis and frailty in elderly individuals across six European countries: Results from the European Project on OsteoArthritis (EPOSA). *BMC Musculoskelet. Disord.* <https://doi.org/10.1186/s12891-015-0807-8> (2015).
31. Budoff, M. Triglycerides and triglyceride-rich lipoproteins in the causal pathway of cardiovascular disease. *Am. J. Cardiol.* **118**, 138–145 (2016).
32. Zhou, M. *et al.* The cross-sectional and longitudinal effect of hyperlipidemia on knee osteoarthritis: Results from the Dongfeng-Tongji cohort in China. *Sci. Rep.* <https://doi.org/10.1038/s41598-017-10158-8> (2017).
33. Askari, A. *et al.* Relationship between metabolic syndrome and osteoarthritis: The FASA osteoarthritis study. *Diabetes Metab. Syndr.* **11**(Suppl 2), S827–S832 (2017).
34. Zhang, K. *et al.* High-density lipoprotein cholesterol and apolipoprotein A1 in synovial fluid: Potential predictors of disease severity of primary knee osteoarthritis. *Cartilage* **13**, 1465S–1473S (2021).
35. Hindy, G. *et al.* Cardiometabolic polygenic risk scores and osteoarthritis outcomes: A Mendelian randomization study using data from the malmo diet and cancer study and the UK Biobank. *Arthritis Rheumatol.* **71**, 925–934 (2019).
36. Funck-Brentano, T., Nethander, M., Moverare-Skrkt, S., Richette, P. & Ohlsson, C. Causal Factors for Knee, Hip, and Hand Osteoarthritis: A Mendelian Randomization Study in the UK Biobank. *ARTHRITIS & RHEUMATOLOGY* **71**, 1634–1641 (2019).
37. Soska, V. Cholesterol measurement and current guidelines. *Vnitřní lékařství* **68**, 54–57 (2022).
38. Singh, G., Miller, J. D., Lee, F. H., Pettitt, D. & Russell, M. W. Prevalence of cardiovascular disease risk factors among US adults with self-reported osteoarthritis: Data from the Third National Health and Nutrition Examination Survey. *Am. J. Managed Care* **8**, S383–391 (2002).
39. Schwager, J. L. *et al.* Multicenter osteoarthritis study G: Association of serum low-density lipoprotein, high-density lipoprotein, and total cholesterol with development of knee osteoarthritis. *Arthritis Care Res.* **74**, 274–280 (2022).
40. Pan, F., Tian, J., Mattap, S. M., Cicuttini, F. & Jones, G. Association between metabolic syndrome and knee structural change on MRI. *Rheumatology* **59**, 185–193 (2020).
41. Kruisbergen, N. N. L., van Gemert, Y., Blom, A. B., van den Bosch, M. H. J. & van Lent, P. L. E. M. Activation of circulating monocytes by low-density lipoprotein—a risk factor for osteoarthritis?. *Rheumatology* **62**, 42–51 (2022).
42. Oliviero, F. *et al.* A comparative study of serum and synovial fluid lipoprotein levels in patients with various arthritides. *Clin. Chim. Acta* **413**, 303–307 (2012).
43. Mishra, R. *et al.* A comparative analysis of serological parameters and oxidative stress in osteoarthritis and rheumatoid arthritis. *Rheumatol. Int.* **32**, 2377–2382 (2012).
44. Gill, D., Karhunen, V., Malik, R., Dichgans, M. & Sofat, N. Cardiometabolic traits mediating the effect of education on osteoarthritis risk: A Mendelian randomization study. *Osteoarthr. Cartil.* **29**, 365–371 (2021).
45. Wang, Z. *et al.* Effect of blood lipids and lipid-lowering therapies on osteoarthritis risk: A Mendelian randomization study. *Front. Med.* <https://doi.org/10.3389/fmed.2022.990569> (2022).

## Acknowledgements

We thank the NHANES database, Richardson et al., Borges CM et al. for providing statistics data for the analyses. We thank to Zhang Jing (Shanghai Tongren Hospital) for his work on the NHANES database. His outstanding work, nhanesR package and webpage, makes it easier for us to explore NHANES database.

## Author contributions

B.P., D.Q., M.X. and G.H. participated in study design. B.P., G.H. and X.Z. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. G.H., X.Z. and M.Z. analyzed the data and wrote the paper. B.P., D.Q. and M.X. jointly supervised the study. All authors read and approved the submitted version.

## Funding

This study was funded by Suzhou Youth Science and Technology Program (KJXW2021067), Suzhou Science and Technology Development Program (SKYD2022049), Changshu Science and Technology Development Program (CS202235), Changshu Health and Health Commission Science and Technology Program (CSWS202109), Key Disciplines of Changshu (CSZDXK202303).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.X., B.P. or D.Q.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024