DR. GEORGE HINDY (Orcid ID : 0000-0002-7257-9299)

Article type : Full Length

Running head: Cardiometabolic traits and osteoarthritis

Title: Cardiometabolic polygenic risk scores and osteoarthritis outcomes: a Mendelian randomization study from the Malmö Diet and Cancer Study and the UK Biobank

George Hindy MD, PhD[1,2], Kristina E. Åkesson MD, PhD[1], Olle Melander MD, PhD[1], Krishna G. Aragam MD, MS[2,3], Mary E. Haas PhD[2], Peter M. Nilsson MD, PhD[1], Umesh T. Kadam MBChB, PhD*[4], Marju Orho-Melander PhD*[1]

*Joint Senior Authors

1. Department of Clinical Sciences in Malmö, Lund University, Sweden

2. Program in Medical and Population Genetics and the Cardiovascular Disease Initiative, Broad Institute, Cambridge, Massachusetts, USA

3. Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

4. Diabetes Research Centre and Department of Health Sciences, University of Leicester, UK

Funding:

This work was supported by the European Research Council (consolidator grant nr 649021, Orho-Melander), the Swedish Research Council (Orho-Melander, Hindy) and the Region Skåne (Orho-Melander).

Correspondence to:

George Hindy, Department of Clinical Sciences in Malmö, Lund University, Jan Waldenströms gata 35, Clinical Research Center, 205 02 Malmö, Sweden. Email: George.Hindy@med.lu.se Telephone: +4640391248

ABSTRACT

Objective: To investigate the causal role of cardiometabolic risk factors in osteoarthritis (OA) using associated genetic variants.

Methods: We studied 27691 adults from the Malmö Diet and Cancer Study (MDCS) and replicated novel findings among 376435 adults from the UK Biobank. Trait-specific polygenic risk scores for LDL- and HDL-cholesterol (LDLC, HDLC), triglycerides (TG), BMI, fasting plasma glucose (FPG) and systolic blood pressure (SBP) were used to test the associations of genetically predicted elevations in each trait with incident OA-diagnosis (n=3559), OA-joint replacement (n=2780), or both (total-OA, n=4226) in Mendelian randomization (MR) analyses in MDCS, and with self-reported and/or hospital diagnosed OA (n=65213) in the UK Biobank. Multivariable MR, MR-Egger and weighted-median MR were used to adjust for potential pleiotropic biases.

Results: In MDCS, genetically predicted higher LDLC was associated with lower risk of OA-diagnosis (OR: 0.83; 95% CI: 0.73-0.95 per 1-SD) and total-OA (0.87; 0.78-0.98) which was supported by multivariable MR for OA-diagnosis (0.84; 0.75-0.95) and total-OA (0.87; 0.78-0.97), and by conventional two-sample MR for OA-diagnosis (0.86; 0.75-0.98). MR-Egger indicated no pleiotropic bias. Genetically predicted higher BMI was associated with

increased risk for OA-diagnosis (1.65; 1.14-2.41), while MR-Egger indicated pleiotropic bias and larger association with OA-diagnosis (3.25; 1.26-8.39), OA-joint replacement (3.81; 1.39-10.4) and total-OA (3.41; 1.43-8.15). No associations were observed between genetically predicted HDLC, TG, FPG or SBP and OA outcomes. The LDLC associations were replicated in the UK Biobank (0.95; 0.93-0.98).

Conclusion: Our MR study provides evidence for causal role of lower LDLC and higher BMI in OA.

It has been hypothesized that there are pathophysiologic links between cardiometabolic disease and osteoarthritis (OA) (1, 2). Several, studies have previously reported that cardiometabolic risk factors and diseases are associated with increased risk of OA (2, 3). Potential mechanisms could relate to systemic processes including cholesterol metabolism and associated inflammatory processes (2, 4). OA encompasses changes in articular, bone and cartilage structures (5) and the current clinical focus is on the modification of mechanical loading as a causal factor, treatment of psychosocial factors or replacement of intra-articular cartilage (6). Yet, studies of generalized OA suggest a potential role of systemic processes which raised the hypothesis that lipid metabolism disorders (7, 8) could be involved in pathogenic mechanisms of OA (9). Supporting evidence for this hypothesis includes shared mesenchymal origin of adipocytes and articular cells (10, 11), *in vitro* studies showing that higher lipid levels in the synovial fluid can induce arthritic changes (12), and experimental mouse models of atherosclerosis showing arthritic changes with high cholesterol diet (13, 14). Epidemiological studies have provided further support for the hypothesis indicating that cardiovascular disease and OA commonly co-occur (15), share similar risk factors (16, 17), and are both associated with higher mortality (18).

Observational studies, however, suffer from biases, mainly due to reverse causation and confounding, making it difficult to infer a causal role between cardiometabolic traits and OA. As genetic markers are distributed randomly at conception they can be used to infer causal relations between these traits and OA. This has been previously performed in a Mendelian randomization (MR) study using a genetic variant in the *FTO* locus previously associated with adiposity measured as body mass index (BMI), and more recently in the UK Biobank using polygenic risk scores (19). These studies have provided evidence supporting a causal role of higher BMI in increasing the risk of OA (20). Similarly, genetic variants associated with other cardiometabolic traits can be leveraged for deciphering their causal role in OA. However, for genetic variants to be used in MR studies they must reliably be associated with the exposure of interest, should exert their effect on the outcome solely through the exposure and should not associate with other factors that affect the outcome. Those assumptions are violated when variants exhibit horizontal pleiotropy, meaning that variants have effect(s) on disease outside of their effect on the exposure, and lead to bias in MR studies. As combinations of many genetic variants are needed to power most of MR studies, there is increased risk of violations through horizontal pleiotropy. Several methods have been developed to correct for such bias in the two-sample MR setting using exposure and outcome summary statistics for genetic associations obtained from independent samples.

In light of the conflicting evidence, we aimed to use previously identified cardiometabolic genetic variants, and MR approach, to test the hypothesis that each of the six cardiometabolic traits including LDL cholesterol (LDLC), HDL cholesterol (HDLC), triglycerides (TG), body mass index (BMI), fasting plasma glucose (FPG) and systolic blood pressure (SBP), play a causal role in the pathogenesis of OA.

MATERIALS AND METHODS

Study Population

The Malmö Diet and Cancer Study (MDCS) cohort recruited nearly 30000 adults aged 45 to 64 years from the population between 1991 and 1996. Baseline data with consent included anthropometrics, body composition, cardiovascular physiological measures, socio-economic and lifestyle factors, and medical history (21). The data on MDCS population was linked to clinical registers using the civil registration numbers which uniquely identify all inhabitants of Sweden. MDCS was approved by the Regional Ethical Committee at Lund University (LU 51-90) and the final provision of data to the research team was anonymized. We also used the UK Biobank, a large cohort of around 500000 adults aged 40 to 69, to replicate novel findings. We excluded samples with non-white ancestry, sex mismatches, excess heterozygosity, missingness and second degree relatives (22).

Baseline assessment in MDCS

Blood pressure was measured using a mercury-column sphygmomanometer after ten minutes of rest in the supine position. A balance-beam scale was used to measure weight (kilograms) with subjects wearing light clothing and no shoes. A fixed stadiometer was used to measure height (centimeters). BMI was measured as weight in kilograms divided by height in meters squared. Fasting serum lipids and fasting blood glucose were measured from blood samples drawn after an overnight fast. Fasting blood glucose was converted to FPG by multiplying the values by 1.13. Samples were analyzed by routine standard methods at the Department of Clinical Chemistry, Malmö University Hospital. Fasting blood measurements were only available in the MDC-CC. Apolipoprotein B (ApoB) and apolipoprotein A1 (ApoA1) were

measured in non-fasted plasma samples of the entire MDCS by Quest Diagnostics (San Juan Capistrano, CA, USA), blinded to case-control status, using an immunonephelometric assay run on the Siemens BNII (Siemens, Newark, DE, USA). The inter-assay variability was < 4.0% for both ApoA1 and ApoB.

Genetic measures

In MDCS, a MALDI-TOF mass spectrometer (Sequenom MassArray, Sequenom, San Diego, CA, USA) was used to genotype DNA samples using Sequenom reagents and protocols. Proxy SNPs were identified using SNAP version 2.2.2 when commercial primers were not available. SNPs that failed Sequenom genotyping were genotyped individually using TaqMan or KASPar allelic discrimination on an ABI 7900HT (Applied Biosystems, Life Technologies, Carlsbad, CA, USA), according to the manufacturer's instructions. Individuals with < 60% successfully genotyped SNPs and SNPs with a genotype success rate of <90% or deviation from Bonferroni-corrected Hardy-Weinberg Equilibrium in each set of SNPs of the specific traits were excluded. Using these exclusion criteria, genetic data was available for 26,435 individuals. Genotyped SNPs were selected based on previous genome-wide association studies (GWAS) and included 31 SNPs for BMI (23), 29 for SBP (24-26), 52 for LDLC (27, 28), 41 for HDLC (27), 26 for TG (27), and 15 for FPG (29). Trait-specific MDCS-weighted polygenic risk scores (PRSs) were created using the PLINK software (version 1.07).

In the UK Biobank, we extracted data on 185 lipid associated SNPs including 76 independent LDLC-associated SNPs, defined as SNPs showing associations with LDLC with p-value of less than $5 * 10^{-8}$, from the latest Global Lipids Genetics Consortium GWAS (28).

Additionally, gene-specific instruments for *HMGCR*, *PCSK9*, *LDLR* and *NPC1L1* were constructed by using independent LDLC associated SNPs in each locus, as described previously by Ference et al. (30, 31), to estimate the causal effect of LDLC through each of these genes. Gene-specific genotype data were not available in MDCS.

Cardiovascular and OA-related outcomes

Cardiovascular disease cases were defined as fatal or nonfatal myocardial infarction (MI) or stroke, or death due to ischemic heart disease, and linked using the Swedish Hospital Discharge Register, the Swedish Cause of Death Register, and the Stroke Register of Malmö. MI cases were defined using codes 410 and I21 (ICD9 and ICD10), and the stroke cases were defined using codes 430, 431, 434, and 436 of the ICD9, and codes I60, I61, I63, and I64 of the ICD10. Data from national registers provided by the Swedish Board on Health and Welfare (Swedish National Discharge Registers), and by Statistics Sweden, were used to define the three main study outcomes: (i) incident 'OA diagnosis' – a clinical record of either knee or hip OA, (ii) incident 'OA joint replacement'– a record of either a hip or knee joint arthroplasty surgery, and (iii) 'total OA' –either a diagnosis or joint replacement. 'Incident' outcome was defined as a new case occurring at any point in the follow-up time-period after baseline screening date (1996) up to the end of year 2014, which means there was nearly a 20-year mean follow up for the study participants. The case definitions for the outcomes were based on International Classification of Disease coding systems that covered the span of time period from 1970s to the current version (ICD-10). For osteoarthritis-arthroplasty the codes were 8423, 8424, 8428, 8433, NGB09, NGB19, NGB29, NGB39, NGB49 (Knee Arthroplasty) and 8410, 8411, NFB29, NFB39 and NFB49 (Hip Arthroplasty). Additionally, the participants with both joint replacements and a record of fractures (any cause) were

excluded from the OA outcomes sample, to ensure that the indication was likely to be OA-related. In MDCS, based on ICD10 codes, a total of 3559 individuals had OA diagnosis, 2780 had OA joint replacement, 4226 had total OA and 2113 had both OA diagnosis and OA joint replacement.

In the UK Biobank the OA definition was based on self-reported medical history and/or ICD9 and ICD10 codes in hospital-based medical records. A total of 65213 individuals had either self-reported or medical record-based OA, 36128 had self-reported OA, 43744 had medical-record based OA and 14659 had both.

Statistical analysis

In this study, we performed several different MR analyses. In MDCS, we performed i) one-sample conventional and multivariable MR using SNP-exposure and SNP-outcome data from MDCS ii) two-sample conventional and multivariable MR using SNP-exposure data from publicly available GWAS databases and SNP-outcome data from MDCS (32, 33). For significant associations, we continued with further methods to control for pleiotropy. We used iii) two-sample MR Egger analyses and estimation of Egger intercept which reflects total horizontal pleiotropy (34); and iv) two-sample weighted median MR which can provide accurate estimates given that at least 50% of the variants are valid instruments (35). For LDLC, we performed replication analyses in the UK Biobank using two-sample conventional MR, and correcting for pleiotropy by multivariable MR, MR Egger and weighted median MR. In MDCS, prevalent OA cases were excluded and controls were defined as all individuals who did not develop OA during the follow-up period. In the UK Biobank, cases

were defined as both prevalent or incident OA and controls were defined as all individuals free from OA.

In order to achieve normality of distribution and comparability between the studied traits, all of the studied traits (LDLC, HDLC, TG, BMI, SBP, FPG) were natural log-transformed and converted to z-scores in MDCS. Linear regression was used to study the association between the weighted PRSs and their respective traits. To estimate the un-confounded causal effect of the trait on incident cases, we performed a conventional one-sample MR analyses employing a two-stage regression. The predicted fitted values from the linear regression of the traits by their respective PRSs were used as the predictor variables for osteoarthritis outcomes in logistic regression models adjusted for age and sex.

The PRSs included pleiotropic SNPs, which may have biased our results. Excluding these SNPs would have largely weakened the PRSs as instrumental variables and therefore, we used a previously described and applied inverse-variance weighted multivariable MR approach (36). In this approach, the SNP-outcome β coefficients obtained from logistic regression of 168 independent SNPs (r^2 <0.2) on incident OA were included as outcome variables in a multivariable model with SNP-LDLC, SNP-HDLC, SNP-TG, SNP-BMI, SNP-FPG, and SNP-SBP β coefficients obtained from the linear regression of the same SNPs on each of these cardio-metabolic traits in MDCS. This multivariable model was weighted by the inverse-variance of the SNP-outcome association and the intercept was fixed to zero. The multivariable MR model corrects for pleiotropic bias across the studied traits and not for potential pleiotropy with other traits not included in the model (32, 37). To correct for possible bias from total or cardiovascular mortality as competing risks, we used the Fine and Gray proportional subdistribution hazard models and age as the underlying time variable in sensitivity analyses.

We followed-up significant findings using summary level data for the same set of SNPs in the most recent publicly available GWAS data, except for blood pressure (28, 38, 39) which was not publicly available, to perform two-sample MR analyses using inverse-variance weighted MR or conventional MR. This is a regression of SNP-outcome β coefficients on each instrumental SNP-exposure β coefficients weighted by the inverse-variance of the SNP-outcome associations and the intercept fixed to zero (33). To detect and correct for known and unknown pleiotropic bias, multivariable MR and Egger MR (MR-Egger) were performed. MR-Egger is similar to the inverse-variance weighted MR but the intercept is left unconstrained and represents the average unbalanced pleiotropy by each instrument. It provides valid estimates even if all SNPs are invalid instruments. However, it is limited by the untestable InSIDE (Instrument Strength Independent of Direct Effect) assumption that the distribution of the pleiotropic effects of the SNPs on the outcome are independent of their associations with the exposure (34). We additionally performed weighted-median MR which provides a causal estimate given that at least 50% of variants are valid instruments (35). Finally, we performed two-sample conventional MR association analysis between a 76-SNP LDLC GWAS-weighted (28) PRS and OA in the UK Biobank including 65213 OA cases and 311222 controls. Restricted conventional MR analyses were also performed using a 16-SNP LDLC-specific GWAS-weighted PRS, defined as LDLC-associated SNPs that remained after excluding those associated with HDLC and/or TG, using a p-value of 0.05. Then, we performed sensitivity MR analyses including MR-Egger and weighted median MR using the 76 LDLC-associated SNPs, and multivariable MR using 185 LDLC, HDLC and/or TG associated SNPs (28). Finally, we performed conventional MR analyses between LDLC and OA using GWAS-weighted gene-specific LDLC PRSs in each of the *HMGCR*, *PCSK9*, *LDLR* and *NPC1L1* genes.

STATA SE 13.1 (StataCorp LP) and R (version 3.3) were used for the statistical analyses. For the two-sample MR analyses, we used the "MendelianRandomization" package in R (40).

RESULTS

Incident OA endpoints in MDCS

From the total sample of 27691 MDCS participants at baseline, by the end of the mean follow-up time of 17.4 years, 19350 (70%) participants were alive, 8091 (29.2%) had died and 250 (0.9%) had emigrated. During the follow-up, 3559 (12.9%) had an OA diagnosis, 2780 (10%) OA joint replacement and 4226 (15.3%) had total OA (Table 1). The variances explained by PRSs were 7.3% for LDLC, 5.7% for HDLC, 4.3% for TG, 0.8% for BMI, 2.3% for FPG and 0.5% for SBP in MDCS.

Conventional and multivariable one-sample MR in MDCS

The associations of PRS-predicted elevations of cardiometabolic traits with OA outcomes in MDCS using conventional MR are shown in Figure 1. Genetically predicted elevations in LDLC were associated with lower risk of OA diagnosis (OR: 0.83; 95% CI: 0.73-0.95) and total OA (0.87; 0.78-0.98) but not with OA joint replacement (0.94; 0.81-1.08). Genetically predicted elevations in BMI were associated with increased risk for OA diagnosis (1.65; 1.14-2.41) and with a trend for increased risk for total OA (1.42; 1.00-2.02) but not with OA joint replacement (1.31; 0.86-2.00). Genetically predicted elevations in SBP were associated with a reduced risk of OA diagnosis (0.54; 0.35-0.83), OA joint replacement (0.43; 0.26-0.70) and total OA (0.58; 0.39-0.88). Multivariable MR adjusted for pleiotropy across the six

cardiometabolic traits supported the observed associations of BMI and LDLC with OA. However, the association between genetically elevated SBP and OA outcomes was lost when adjusting for pleiotropy (Supplementary Figure 1). Sensitivity analyses that accounted for competing risks showed that the association between LDLC and OA diagnosis was not likely to be biased by total mortality (Subdistribution Hazard Ratio (SHR): 0.83; 0.74-0.94) or cardiovascular mortality (SHR: 0.84; 0.74-0.94) as competing risks (Supplementary Table 1).

Two-sample multivariable-, MR-Egger- and weighted median MR analysis to correct for pleiotropy using SNP-exposure associations from publicly available GWAS databases for SNP-outcome associations in MDCS

Using published GWAS summary-level data for cardiometabolic traits, we performed two-sample MR analyses for LDLC and BMI with OA outcomes in MDCS. Conventional MR indicated an inverse association between LDLC and OA diagnosis (0.86; 0.75-0.98). Multivariable MR adjusting for estimates for HDLC, TG, BMI, and FPG indicated lower risk of OA diagnosis by genetically elevated LDLC (0.85; 0.73-0.98). Similar analyses indicated no association with OA joint replacement. The MR-Egger intercept indicated no pleiotropic bias (p-intercept=0.51) and the MR-Egger estimate was consistent with conventional and multivariable MR, yet not significant (p=0.41). Weighted-median MR indicated an inverse association between LDLC and total OA (0.82; 0.68-0.99) and a tendency with OA diagnosis (0.83; 0.68-1.02) (Table 2).

Conventional MR analyses showed a tendency for increased risk of OA diagnosis by genetically elevated BMI (1.41; 0.96-2.08) but not for other OA outcomes. Multivariable MR showed significantly increased risk of OA diagnosis and total OA by elevated BMI (1.39;

1.05-1.86 and 1.32; 1.01-1.72). The MR-Egger intercept; however, indicated that the BMI SNPs likely exhibited unbalanced pleiotropy in their association with all three OA outcomes (p-intercept= 0.06 for OA diagnosis, and 0.01 for both OA joint replacement and total OA). The MR-Egger estimates suggested around 3.3-3.8 fold increased risk for OA diagnosis (3.25; 1.26-8.39), joint replacement (3.81; 1.39-10.4), and total OA (3.41; 1.43-8.15) by 1-SD higher BMI. Higher BMI was associated with both OA joint replacement and total OA using weighted-median MR (2.01; 1.19-3.21 and 2.07; 1.33-3.21, respectively) (Table 2).

Two-sample MR analyses in the UK Biobank to replicate the association between LDLC and OA

Finally, we performed two-sample MR analyses in the UK Biobank. A 76-SNP LDLC instrument was used in more than 65213 OA cases and 311222 controls in the UK Biobank. Using the 76-SNP GWAS-weighted PRS, each 1-SD elevated LDLC was associated with lower risk of OA (0.95; 0.93-0.98). Restricting the PRS to 16 LDLC specific SNPs indicated a larger magnitude of decreased risk of OA (0.86; 0.79-0.95). Gene-specific LDLC PRSs analyses showed inverse association between the *LDLR* gene specific PRS and OA (0.92; 0.88-0.97). While the *HMGCR* gene specific PRS showed a tendency for lower OA risk (0.92; 0.90-1.01), the *NPC1L1*- gene specific PRS suggested increased risk of OA risk (1.12, 0.98-1.27). Finally, sensitivity two-sample MR-Egger, multivariable and weighted-median MR analyses with SNP-data from public GWAS databases indicated inverse associations between LDLC and OA with odds ratios ranging from 0.91 to 0.95 (Figure 2) and the MR-Egger intercept did not suggest evidence of pleiotropy. Similar results were obtained when self-reported and medical-record based OA were analyzed separately (Supplementary Figures 2-3).

DISCUSSION

Our study has three separate key findings. First, our results suggest an inverse causal role between LDLC and OA in large Swedish and British cohorts, which was not explained after adjustment for the genetically predicted levels of other cardiometabolic traits. Second, our findings support earlier evidence of a direct causal role between BMI and OA (19, 20). Third, there was a lack of association between genetically predicted levels of other cardiometabolic factors and OA.

Our LDLC findings are in contrast to previous observational studies, which have suggested elevated serum cholesterol levels as a risk factor for OA (1, 17, 41). Previously, it has been hypothesized that lipid accumulation in the cartilage might trigger OA development (1) and high cholesterol diet has been shown to induce arthritic changes in experimental mouse models of atherosclerosis (13, 14). However, our findings now provide an alternative and opposite hypothesis that higher LDLC levels may be associated with a lower risk of OA.

The key implication of our findings relates to the potential role of LDLC lowering drugs in the pathogenesis of OA. Statins are key and widespread drugs for LDLC lowering in cardiovascular disease management and observational studies have provided conflicting results on the relationship between statins and OA. While some epidemiological studies have indicated that statin use is associated with a lower risk (42, 43), others have observed increased risk (44, 45), or lack of association (46, 47). The key limitations of the previous evidence are that the studies, with few exceptions, have been either small or have used OA joint replacement as the outcome. Further, observational associations are prone to biases due to confounding and reverse causation, which can be bypassed by using genetic variants as instruments. Our study is to the best of our knowledge the first MR study to examine causal associations of genetically predicted LDLC with incident OA endpoints. The association

between lower LDLC instrumented by the LDLC PRSs and *LDLR* gene PRS, and tendency for association between lower LDLC through *HMGCR* gene (encoding statin target), and higher OA risk contests most previous evidence and suggests that statins may increase the risk of OA. The linked implication is that future research needs to specifically investigate the effects of statins on OA with radiographic definitions to demonstrate pathologic progression or non-progression.

Our findings also support direct causality between BMI and OA (19, 20). The higher prevalence of the metabolic syndrome among patients with OA and the lower age of OA onset among individuals with the metabolic syndrome has led to the hypothesis that OA could be part of a systemic metabolic disorder characterized by obesity, dyslipidemia (based on HDLC and/or triglyceride levels), dysglycemia and hypertension (16, 48, 49). The only component of the metabolic syndrome that indicated causal link to OA in our study was BMI. The effect estimates using the different MR methods were strikingly different. While the conventional MR analysis in MDCS indicated around 65% increased risk of OA diagnosis with each 1-SD genetically higher BMI, the MR-Egger analysis using GWAS summary data indicated that the BMI genetic instrument exhibited pleiotropic bias and thus this analysis, correcting for pleiotropy, may have provided a more accurate estimate. It showed that each 1-SD of genetically predicted BMI increased the risk for all OA outcomes with odds ratios ranging from 3.25-3.81.

Previous evidence has suggested that hyperglycemia or diabetes status is associated with an increased risk of OA (1, 50), but we did not observe association between FPG and OA in our study. This result is concordant with recent observations in the UK Biobank, where no association was found between genetic risk for type 2 diabetes and OA (19). It is possible that hyperglycemia may still exert a causal but weak association on OA, as lack of evidence could

be a consequence of the rather weak instrument for FPG and insufficient statistical power. Regardless, it is intriguing that earlier MR studies have shown clear evidence that genetic lowering of LDLC increases the risk of type 2 diabetes (30, 51), similar to what was observed in our study for OA.

In the initial conventional MR analyses we observed an inverse association between genetically elevated SBP and lower risk of OA, which lost significance in multivariable MR analyses. The genetic instrument for SBP was the weakest, explaining only 0.5% of the variance in baseline SBP, which also means that a causal association with OA cannot be excluded pending stronger genetic instruments for blood pressure.

The main strength of our study is the use of different MR analyses to analyze the causal role of cardiometabolic traits in OA. We have used both one-sample and two-sample strategies that account for biases mainly due to pleiotropy. Importantly, we were able to replicate the rather controversial finding, that lower LDLC increases the risk of OA, in the large UK Biobank using up-to-date summary-level genetic data for lipid traits. However, the MR estimate for OA using the 76-SNP GWAS-weighted LDLC PRS was smaller in the UK Biobank as compared to the MR estimate using the 52-SNP LDLC PRS in MDCS. It can be speculated that this could be due to misclassification in the self-reported OA phenotype in the UK Biobank. Nevertheless, we must acknowledge some limitations in our study. First, we cannot confirm negative associations and thus we cannot rule out possible weaker causal associations of HDLC, TG, FPG and SBP with OA outcomes, as the genetic instruments and the used sample size might be underpowered to detect them. Second, we used two different OA end-points, one based on clinical diagnosis and the other on the surgical intervention of arthroplasty. Clinical diagnosis provides the phenotype for the presentation of symptomatic OA, whatever the joint location, whereas joint arthroplasty provides the phenotype for

intervention for either severe large joint pain or radiographic joint destruction (52). This approach enabled the ascertainment of the most symptomatic OA cases, yet with the caveat that the study-defined phenotype is heterogeneous and prone to other non-random factors that influence indications for joint replacement (53). Third, we did not correct for multiple testing. However, considering a Bonferroni corrected p-value of 0.0028, the LDLC-OA findings (P = 0.004 from multivariable MR) can be considered as borderline significant, and replication in UK Biobank further strengthened this observation. In addition, prior evidence of a causal role of BMI in OA motivated the analyses of BMI as a positive control. Fourth, it is important to note that our results are likely relevant to hip and knee OA, as they are the most prevalent diagnoses in our cases, and thus may not necessarily be true for OA in other joints as hand. Finally, misclassification of OA outcomes in our samples may have biased our estimates towards the null and could at least partially explain some of the null findings.

In conclusion, our study suggests causal associations between lower LDLC and increased risk of OA. This evidence challenges the current perspectives from epidemiological studies and indicates that future investigations need to focus on the mechanisms linking lower LDLC with OA pathogenesis to potentially identify therapeutic targets, and on investigating how the widespread use of LDLC lowering drugs may impact the pathogenesis of OA and related outcomes.

Acknowledgements

We thank the participants of the Malmö Diet and Cancer Study and the UK Biobank. Analyses in the UK Biobank were conducted via application 7089 via a protocol approved by the Partners HealthCare Institutional Review Board.

Author Contributions

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Drs. Hindy and Orho-Melander had full access to all of the data and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Hindy, Åkesson, Melander, Kadam, Orho-Melander

Acquisition of data. Hindy, Melander, Aragam, Haas, Nilsson, Kadam, Orho-Melander

Analysis and Interpretation of Data. Hindy, Åkesson, Melander, Nilsson, Kadam, Orho-Melander

REFERENCES

1. Zhuo Q, Yang W, Chen J, Wang Y. Metabolic syndrome meets osteoarthritis. Nat Rev Rheumatol. 2012;8(12):729-37.
2. Tall AR, Yvan-Charvet L. Cholesterol, inflammation and innate immunity. Nat Rev Immunol. 2015;15(2):104-16.
3. Kadam UT, Holmberg A, Blagojevic M, Nilsson PM, Akesson K. Risk factors for cardiovascular disease and future osteoarthritis-related arthroplasty: a population-based cohort study in men and women from Malmo, Sweden. Scand J Rheumatol. 2011;40(6):478-85.
4. Aspden RM, Scheven BA, Hutchison JD. Osteoarthritis as a systemic disorder including stromal cell differentiation and lipid metabolism. Lancet. 2001;357(9262):1118-20.
5. Brandt KD, Doherty M, Lohmander S. Osteoarthritis. Oxford ; New York: Oxford University Press; 2003.
6. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthritis Cartilage. 2007;15(9):981-1000.
7. Thijssen E, van Caam A, van der Kraan PM. Obesity and osteoarthritis, more than just wear and tear: pivotal roles for inflamed adipose tissue and dyslipidaemia in obesity-induced osteoarthritis. Rheumatology (Oxford). 2015;54(4):588-600.

8. Dahaghin S, Bierma-Zeinstra SM, Koes BW, Hazes JM, Pols HA. Do metabolic factors add to the effect of overweight on hand osteoarthritis? The Rotterdam Study. Ann Rheum Dis. 2007;66(7):916-20.
9. Conaghan PG, Vanharanta H, Dieppe PA. Is progressive osteoarthritis an atheromatous vascular disease? Ann Rheum Dis. 2005;64(11):1539-41.
10. Pittenger MF, Mackay AM, Beck SC, Jaiswal RK, Douglas R, Mosca JD, et al. Multilineage potential of adult human mesenchymal stem cells. Science. 1999;284(5411):143-7.
11. Diascro DD, Jr., Vogel RL, Johnson TE, Witherup KM, Pitzenberger SM, Rutledge SJ, et al. High fatty acid content in rabbit serum is responsible for the differentiation of osteoblasts into adipocyte-like cells. J Bone Miner Res. 1998;13(1):96-106.
12. Lippiello L, Walsh T, Fienhold M. The association of lipid abnormalities with tissue pathology in human osteoarthritic articular cartilage. Metabolism. 1991;40(6):571-6.
13. de Munter W, Blom AB, Helsen MM, Walgreen B, van der Kraan PM, Joosten LA, et al. Cholesterol accumulation caused by low density lipoprotein receptor deficiency or a cholesterol-rich diet results in ectopic bone formation during experimental osteoarthritis. Arthritis Res Ther. 2013;15(6):R178.
14. Gierman LM, Kuhnast S, Koudijs A, Pieterman EJ, Kloppenburg M, van Osch GJ, et al. Osteoarthritis development is induced by increased dietary cholesterol and can be inhibited by atorvastatin in APOE*3Leiden.CETP mice--a translational model for atherosclerosis. Ann Rheum Dis. 2014;73(5):921-7.
15. Kadam UT, Jordan K, Croft PR. Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales. Ann Rheum Dis. 2004;63(4):408-14.
16. Singh G, Miller JD, Lee FH, Pettitt D, Russell MW. 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 Manag Care. 2002;8(15 Suppl):S383-91.
17. Sturmer T, Sun Y, Sauerland S, Zeissig I, Gunther KP, Puhl W, et al. Serum cholesterol and osteoarthritis. The baseline examination of the Ulm Osteoarthritis Study. J Rheumatol. 1998;25(9):1827-32.
18. Nuesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Juni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. BMJ. 2011;342:d1165.
19. Zengini E, Hatzikotoulas K, Tachmazidou I, Steinberg J, Hartwig FP, Southam L, et al. Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis. Nat Genet. 2018;50(4):549-58.
20. Panoutsopoulou K, Metrustry S, Doherty SA, Laslett LL, Maciewicz RA, Hart DJ, et al. The effect of FTO variation on increased osteoarthritis risk is mediated through body mass index: a Mendelian randomisation study. Ann Rheum Dis. 2014;73(12):2082-6.
21. Manjer J, Carlsson S, Elmstahl S, Gullberg B, Janzon L, Lindstrom M, et al. The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. Eur J Cancer Prev. 2001;10(6):489-99.
22. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. bioRxiv. 2017.
23. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010;42(11):937-48.
24. International Consortium for Blood Pressure Genome-Wide Association S, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011;478(7367):103-9.

25. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet. 2009;41(6):666-76.
26. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet. 2011;43(10):1005-11.
27. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010;466(7307):707-13.
28. Global Lipids Genetics C, Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45(11):1274-83.
29. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet. 2010;42(2):105-16.
30. Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, et al. Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. N Engl J Med. 2016;375(22):2144-53.
31. Ference BA, Majeed F, Penumetcha R, Flack JM, Brook RD. Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or both: a 2 x 2 factorial Mendelian randomization study. J Am Coll Cardiol. 2015;65(15):1552-61.
32. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. Am J Epidemiol. 2015;181(4):251-60.
33. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658-65.
34. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512-25.
35. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40(4):304-14.
36. Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. Nat Genet. 2013;45(11):1345-52.
37. Burgess S, Dudbridge F, Thompson SG. Re: "Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects". Am J Epidemiol. 2015;181(4):290-1.
38. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-206.
39. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet. 2012;44(6):659-69.
40. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017.
41. Hart DJ, Doyle DV, Spector TD. Association between metabolic factors and knee osteoarthritis in women: the Chingford Study. J Rheumatol. 1995;22(6):1118-23.
42. Kadam UT, Blagojevic M, Belcher J. Statin use and clinical osteoarthritis in the general population: a longitudinal study. J Gen Intern Med. 2013;28(7):943-9.
43. Clockaerts S, Van Osch GJ, Bastiaansen-Jenniskens YM, Verhaar JA, Van Glabbeek F, Van Meurs JB, et al. Statin use is associated with reduced incidence and progression of knee osteoarthritis in the Rotterdam study. Ann Rheum Dis. 2012;71(5):642-7.

44. Beattie MS, Lane NE, Hung YY, Nevitt MC. Association of statin use and development and progression of hip osteoarthritis in elderly women. J Rheumatol. 2005;32(1):106-10.
45. Mansi IA, Mortensen EM, Pugh MJ, Wegner M, Frei CR. Incidence of musculoskeletal and neoplastic diseases in patients on statin therapy: results of a retrospective cohort analysis. Am J Med Sci. 2013;345(5):343-8.
46. Riddle DL, Moxley G, Dumenci L. Associations between statin use and changes in pain, function and structural progression: a longitudinal study of persons with knee osteoarthritis. Ann Rheum Dis. 2013;72(2):196-203.
47. Michaelsson K, Lohmander LS, Turkiewicz A, Wolk A, Nilsson P, Englund M. Association between statin use and consultation or surgery for osteoarthritis of the hip or knee: a pooled analysis of four cohort studies. Osteoarthritis Cartilage. 2017.
48. Puenpatom RA, Victor TW. Increased prevalence of metabolic syndrome in individuals with osteoarthritis: an analysis of NHANES III data. Postgrad Med. 2009;121(6):9-20.
49. Engstrom G, Gerhardsson de Verdier M, Rollof J, Nilsson PM, Lohmander LS. C-reactive protein, metabolic syndrome and incidence of severe hip and knee osteoarthritis. A population-based cohort study. Osteoarthritis Cartilage. 2009;17(2):168-73.
50. Schett G, Kleyer A, Perricone C, Sahinbegovic E, Iagnocco A, Zwerina J, et al. Diabetes is an independent predictor for severe osteoarthritis: results from a longitudinal cohort study. Diabetes Care. 2013;36(2):403-9.
51. White J, Swerdlow DI, Preiss D, Fairhurst-Hunter Z, Keating BJ, Asselbergs FW, et al. Association of Lipid Fractions With Risks for Coronary Artery Disease and Diabetes. JAMA Cardiol. 2016;1(6):692-9.
52. Hernandez-Vaquero D, Fernandez-Carreira JM. Relationship between radiological grading and clinical status in knee osteoarthritis. A multicentric study. BMC Musculoskelet Disord. 2012;13:194.
53. Ibrahim SA, Siminoff LA, Burant CJ, Kwoh CK. Understanding ethnic differences in the utilization of joint replacement for osteoarthritis: the role of patient-level factors. Med Care. 2002;40(1 Suppl):I44-51.

FIGURE LEGENDS

Figure 1. One-sample conventional Mendelian randomization analyses for genetically predicted elevations in cardiometabolic traits and osteoarthritis (OA) outcomes in the Malmö Diet and Cancer Study (MDCS). Odds Ratios (OR) for OA outcomes are shown per 1-SD of genetically elevated levels of cardiometabolic traits using respective polygenic risk scores (PRS). Fitted values of each trait by its respective PRS were used as predictors of incident OA outcomes in MDCS in a two-stage least square regression analysis. Genetically higher LDL cholesterol (LDLC) was associated with lower risk of OA diagnosis and total OA, genetically higher body mass index (BMI) was associated with higher risk of OA-

diagnosis and genetically higher systolic blood pressure (SBP) was associated with lower risk of total OA outcomes. HDL cholesterol (HDLC), triglycerides (TG) and fasting plasma glucose (FPG) did not associate with OA outcomes.

Figure 2. Association between genetic predisposition for elevated LDL cholesterol (LDLC) and osteoarthritis (OA) in the UK Biobank. Odds ratios (OR) for OA using two-sample Mendelian randomization (MR) analyses in the UK Biobank. Conventional MR estimates were obtained by analyzing LDLC polygenic risk scores (PRS) created from GWAS identified and gene-specific single nucleotide polymorphisms (SNPs) in relation to OA in the UK biobank weighted by SNP-LDLC associations in the Global Lipids Genetics Consortium (28). Genes for the gene-specific analyses (*HMGCR*, *PCSK9*, *LDLR*, *NPC1L1*) were mainly selected due to them encoding for LDLC lowering targets. Two-sample sensitivity MR analyses were performed using MR-Egger (34), weighted-median (35) and multivariable MR (32) using summary SNP-exposure data from the Global Lipids Genetic Consortium (28) and summary SNP-outcome data from the UK Biobank. MR indicated that genetically elevated LDLC decreases the risk of OA. Gene-specific analyses indicated lower risk of OA by *LDLR*-mediated higher LDLC. A similar trend was observed with *HMGCR* instrument, although did not reach statistical significance.

* LDLC PRS was created using SNPs that were previously associated with LDL cholesterol at GWAS level ($p < 5 \times 10^{-8}$) in the Global Lipids Genetics Consortium (28).

† LDLC PRS was created using SNPs that were previously associated with LDLC at GWAS significant level ($p < 5 \times 10^{-8}$) and not associated with either HDL cholesterol and/or TG ($p > 0.05$) in the Global Lipids Genetics Consortium (28).

Table 1: Characteristics of the Malmö Diet and Cancer Study by osteoarthritis (OA) outcomes

	OA diagnosis	OA joint replacement	Total OA*	Non-OA†
	n=3559	**n=2780**	**n=4226**	**n=23465**
Age, n = 27691	58.5±7.3[#]	60.2±7.3[#]	59.2±7.4[#]	57.8±7.6
Men, n= 10,916 (39.4%)	1183(33.2)[#]	875(31.5)[#]	1371(32.4)[#]	9545(40.7)
Women, n= 16,775 (60.6%)	2376(66.8)[#]	1905(68.5)[#]	2855(67.6)[#]	13920(59.3)
Baseline measures				
BMI (kg/m^2), n=27649	27.3 ± 4.3[#]	26.9 ± 4.3[#]	27.0 ± 4.3[#]	25.5 ± 3.8
LDLC (mmol/L), n=5137	4.2 ± 1.0	4.3 ± 1.0	4.2 ± 1.0	4.2 ± 1.0
ApoA (mg/dL), n=27022	158 ± 28	159 ± 28	158 ± 28	156 ± 28
ApoB (mg/dL), n=27018	107 ± 25	108 ± 25	108 ± 26	107 ± 26
SBP (mmHg), n= 27648	141 ± 19	143 ± 20	142 ± 20	141 ± 20
DBP (mmHg), n=27646	86 ± 10	86 ± 10	86 ± 10	86 ± 10
Never Smokers, n=10461	1466 (41.2)[#]	1168 (42.0)[#]	1736 (41.1)[#]	8725 (37.2)
Former Smokers, n=9375	1276 (35.9)[#]	958 (34.5)[#]	1471 (34.8)[#]	7904 (33.7)
Current Smokers, n=7843	815 (22.9)[#]	654 (23.5)[#]	1017 (24.1)[#]	6826 (29.1)
Myocardial Infarction, n=538	59 (1.7)	47(1.7)	73 (1.7)	465 (2.0)
Cardiovascular Disease, n=837	79 (2.2)[#]	64 (2.3)[#]	101 (2.4)[#]	736 (3.1)
Diabetes n=1209	144 (4.0)	120 (4.3)	186 (4.4)	1023 (4.4)

Data is presented as mean ± SD or count (%)

*Total OA refers to participants with incident OA-diagnosis, OA-joint replacement or both.

†Non-OA refers to participants without incident OA-diagnosis or OA-joint replacement

[#]Denotes statistical significance when compared with the rest of the population ($p<0.05$) after adjustment for age and sex

Table 2. Two-sample Mendelian randomization (MR) for LDL cholesterol (LDLC) and Body Mass Index (BMI) with osteoarthritis (OA) endpoints in the Malmö Diet and Cancer Study

		LDL cholesterol (LDLC)		Body Mass Index (BMI)	
MR method	**OA endpoint**	**OR (95% CI)**	**p-value**	**OR (95% CI)**	**p-value**
Conventional MR*	OA diagnosis	0.86 (0.75-0.98)	0.029	1.41 (0.96-2.08)	0.089
	OA joint replacement	1.01 (0.85-1.20)	0.910	1.19 (0.77-1.82)	0.442
	Total OA	0.93 (0.81-1.06)	0.281	1.24 (0.86-1.79)	0.266
Multivariable MR†	OA diagnosis	0.85 (0.73-0.98)	0.022	1.39 (1.05-1.86)	0.025
	OA joint replacement	1.02 (0.87-1.19)	0.853	1.35 (0.98-1.86)	0.069
	Total OA	0.91 (0.80-1.04)	0.178	1.32 (1.01-1.72)	0.048
MR-Egger‡	OA diagnosis	0.91 (0.73-1.13)	0.405	3.25 (1.26-8.39)	0.015
	OA joint replacement	0.97 (0.73-1.28)	0.807	3.81 (1.39-10.4)	0.009
	Total OA	0.94 (0.75-1.18)	0.583	3.41 (1.43-8.15)	0.006
Egger intercept‡	OA diagnosis	1.00 (0.98-1.01)	0.505	0.97 (0.93-1.00)	0.062
	OA joint replacement	1.00 (0.99-1.02)	0.689	0.96 (0.92-0.99)	0.014
	Total OA	1.00 (0.98-1.01)	0.882	0.96 (0.93-0.99)	0.013
Weighted-median MR§	OA diagnosis	0.83 (0.68-1.02)	0.078	1.39 (0.87-2.24)	0.172
	OA joint replacement	0.90 (0.72-1.13)	0.354	2.01 (1.19-3.39)	0.009
	Total OA	0.82 (0.68-0.99)	0.045	2.07 (1.33-3.21)	0.001

* Inverse-variance weighted MR provides estimates without correcting for pleiotropy (33)

† Inverse-variance weighted multivariable MR provides estimates after correcting for pleiotropy with other cardiometabolic traits (32)

‡ Egger MR. Egger intercept reflects total horizontal pleiotropy (34)

§ Weighted-median MR provides accurate estimates given that at least 50% of variants are valid instrumental variable (35)