

# UCLA

## UCLA Previously Published Works

### Title

Investigation of the Interplay between Circulating Lipids and IGF-I and Relevance to Breast Cancer Risk: An Observational and Mendelian Randomization Study

### Permalink

<https://escholarship.org/uc/item/6jv896t8>

### Journal

Cancer Epidemiology Biomarkers & Prevention, 30(12)

### ISSN

1055-9965

### Authors

Tan, Vanessa Y  
Bull, Caroline J  
Biernacka, Kalina M  
[et al.](#)

### Publication Date

2021-12-01

### DOI

10.1158/1055-9965.epi-21-0315

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Published in final edited form as:

*Cancer Epidemiol Biomarkers Prev.* 2021 December 01; 30(12): 2207–2216.

doi:10.1158/1055-9965.EPI-21-0315.

## Investigation of the interplay between circulating lipids and IGF-I and relevance to breast cancer risk: an observational and Mendelian randomization study

Vanessa Y Tan<sup>#1,2</sup>, Caroline J Bull<sup>#1,2</sup>, Kalina M Biernacka<sup>3</sup>, Alexander Teumer<sup>4,5</sup>, Tom G Richardson<sup>1,2</sup>, Eleanor Sanderson<sup>1,2</sup>, Laura J Corbin<sup>1,2</sup>, Tom Dudding<sup>1,13</sup>, Qibin Qi<sup>6</sup>, Robert C Kaplan<sup>6,7</sup>, Jerome I Rotter<sup>8</sup>, Nele Friedrich<sup>9</sup>, Uwe Völker<sup>10</sup>, Julia Mayerle<sup>11,12</sup>, Claire M Perks<sup>3</sup>, Jeff M.P. Holly<sup>3</sup>, Nicholas J Timpson<sup>1,2</sup>

<sup>1</sup>Medical Research Council (MRC) Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

<sup>2</sup>Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

<sup>3</sup>IGFs & Metabolic Endocrinology Group, School of Translational Health Sciences, Learning & Research Building, Southmead Hospital, Bristol, United Kingdom

<sup>4</sup>Institute for Community Medicine, University Medicine Greifswald, Greifswald, 17475, Germany

<sup>5</sup>Department of Population Medicine and Lifestyle Diseases Prevention, Medical University of Bialystok, Bialystok, Poland

<sup>6</sup>Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, 10461, USA

<sup>7</sup>Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle WA

<sup>8</sup>The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA USA

<sup>9</sup>Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, 17475, Germany

<sup>10</sup>Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, 17475, Germany

<sup>11</sup>Department of Medicine A, University Medicine Greifswald, Greifswald, 17475, Germany

<sup>12</sup>Department of Medicine II, University Hospital, LMU Munich, Munich, 80539, Germany

<sup>13</sup>Bristol Dental School, University of Bristol, Bristol, United Kingdom

<sup>#</sup> These authors contributed equally to this work.

---

This work is licensed under a [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) International license.

Correspondence to: Nicholas J Timpson.

**Corresponding author:** Professor Nicholas J. Timpson, Address: MRC Integrative Epidemiology Unit, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, N.J.Timpson@bristol.ac.uk.

**Conflicts of interest** The authors declare no potential conflicts of interest

## Abstract

**Background**—Circulating lipids and insulin-like growth factor 1 (IGF-I) have been reliably associated with breast cancer (BCa). Observational studies suggest an interplay between lipids and IGF-I, however, whether these relationships are causal and if pathways from these phenotypes to BCa overlap is unclear.

**Methods**—Mendelian Randomisation (MR) was conducted to estimate the relationship between lipids or IGF-I and BCa risk using genetic summary statistics for lipids (low-density lipoprotein cholesterol, LDL-C; high-density lipoprotein cholesterol, HDL-C; triglycerides, TGs), IGF-I and BCa from GLGC/UKBB (N=239,119), CHARGE/UKBB (N=252,547) and BCAC (N=247,173), respectively. Cross-sectional observational and MR analyses were conducted to assess the bi-directional relationship between lipids and IGF-I in SHIP (N=3,812) and UKBB (N=422,389), and using genetic summary statistics from GLGC (N=188,577) and CHARGE/UKBB (N=469,872).

**Results**—In multivariable MR (MVMR) analyses, the OR for BCa per 1-SD increase in HDL-C and TG was 1.08 (95% CI: 1.04, 1.13) and 0.94 (95% CI: 0.89, 0.98), respectively. The OR for BCa per 1-SD increase in IGF-I was 1.09 (95% CI: 1.04, 1.15). MR analyses suggested a bi-directional TG-IGF-I relationship (TG-IGF-I beta per 1-SD: -0.13; 95% CI: -0.23, -0.04; and IGF-I-TG beta per 1-SD: -0.11; 95% CI: -0.18, -0.05). There was little evidence for a causal relationship between HDL-C and LDL-C with IGF-I. In MVMR analyses, associations of TG or IGF-I with BCa were robust to adjustment for IGF-I or TG, respectively.

**Conclusions**—Our findings suggest a causal role of HDL-C, TG and IGF-I in BCa. Observational and MR analyses support an interplay between IGF-I and TG, however, MVMR estimates suggest that TG and IGF-I may act independently to influence BCa.

**Impact**—Our findings should be considered in the development of prevention strategies for BCa, where interventions are known to modify circulating lipids and IGF-I.

## Keywords

Breast cancer; Mendelian randomization; Insulin-like growth factors; lipids

## Introduction

Breast cancer (BCa) is a leading cause of cancer-related death<sup>1,2</sup>, yet approximately 23% of cases in the UK are estimated to be preventable<sup>3</sup>. Circulating lipid and insulin-like growth factor (IGF) traits are frequently hypothesised to underlie the effect of modifiable factors such as obesity on cancer risk<sup>4</sup>, however, the extent to which circulating lipids and IGFs interact is unclear. It is necessary to determine the potential causal relationship between lipids, IGFs and BCa in order to prioritise intervention strategies for BCa prevention.

Observational studies investigating the relationship between lipid profile (low density lipoprotein cholesterol (LDL-C), high density lipoprotein cholesterol (HDL-C) and triglycerides (TG)) and BCa risk have found suggestive evidence that higher HDL-C and TG is associated with lower BCa risk<sup>5,6</sup>. Mendelian randomization (MR) studies, which use genetic variants as instruments for an exposure of interest, given their randomly allocated

and fixed nature<sup>7</sup>, support a causal role for HDL-C in reducing overall BCa risk, whereas effect estimates for LDL-C and TG have been less consistent<sup>8–11</sup>.

Insulin-like growth factor-1 (IGF-I) modulates cell growth, metabolism and survival, and is thought to be important in cancer initiation and progression<sup>12–14</sup>. Observationally, circulating IGF-I levels are associated with increased BCa risk in both pre and postmenopausal women<sup>13</sup> and recent MR estimates further support causality between circulating IGF-I and BCa risk<sup>15</sup>. Several randomised control trials have reported decreased circulating IGF-I levels following LDL-C lowering statin use, suggesting that perturbation of circulating lipids can alter levels of IGF-I in circulation<sup>16–18</sup>. Intervention studies in patients with growth hormone disorders also suggest that modification of IGF-I can alter circulating lipid levels<sup>19–21</sup>. Population-based studies examining the relationship between circulating IGF-I and lipid profile have yielded conflicting results<sup>22–24</sup>, however, these studies may be limited by their cross-sectional design. Hence, the direction of association and whether causation exists between circulating lipids and IGF-I is still unclear.

Given evidence implicating lipids and IGF-I as potential modifiable risk factors for BCa<sup>25–27</sup>, there is motivation to assess the bi-directional relationship between circulating lipids and IGF-I and to test the hypothesis that pathways from these phenotypes to BCa overlap. We set out to examine the causal relationships between circulating lipid traits, IGF-I and BCa using genetic (two-step<sup>28</sup> and multivariable MR (MVMR)) and cross-sectional observational study designs.

## Materials and Methods

### Study design

This study has four main components as outlined in Figure 1: (I) Estimation of the causal association between lipids and BCa using two-sample MR; (II) Estimation of the causal association between IGF-I and BCa using two-sample MR; (IIIA) Analysis of the observational association between lipids and IGF-I using individual level data from the Study of Health in Pomerania (SHIP) cohort and UK Biobank (UKBB); (IIIB) Estimation of the causal association between lipids and IGF-I using a bi-directional two-sample MR; (IV) MVMR analyses to estimate the independent causal effects of lipids and IGF-I on BCa

### Genome-wide association study and two-sample Mendelian randomization analyses

**Study populations and data sources—** Table 1 shows the data sources used for the two-sample MR analyses. To estimate the causal relationship between lipids or IGF-I with BCa using two sample MR analyses (Figure 1, Part I, II and IV), we used summary genome wide association study (GWAS) statistics from: a female-specific GWAS of circulating lipids conducted using individual-level data from UKBB<sup>29</sup> (under application #15825, N = 239,119), a female-specific meta-analysis of two IGF-I GWAS conducted by the IGF working group of the CHARGE consortium (N = 14,600)<sup>30</sup> and using individual-level data from UKBB<sup>29</sup> (under application #15825, N = 237,947) and a BCa GWAS conducted by Breast Cancer Association Consortium (BCAC, 133,384 breast cancer cases and 113,789 controls)<sup>31,32</sup>. Analyses stratified by estrogen receptor (ER) status were also conducted

(ER-positive, 69,501 cases and 105,974 controls and ER-negative, 21,468 cases and 105,974 controls). To estimate the bi-directional causal relationship between lipids and IGF-I (Figure 1, Part IIIB), we used summary GWAS statistics from: a sex-combined GWAS of circulating lipids (LDL-C, HDL-C and TG), conducted by Global lipids Genetic Consortium (GLGC, N = 188,577)<sup>33</sup>, a sex-combined IGF-I GWAS meta-analysis conducted by the IGF working group of the CHARGE consortium (N= 30,884) and UKBB (under application #15825, N= 438,988). Details on the UKBB, including geographical regions, recruitment processes and other characteristics have been described previously<sup>34</sup> (see Supplementary methods for details). All individual participant data used in this study were obtained from the UKBB study, who have obtained ethics approval from the Research Ethics Committee (REC; approval number: 11/NW/0382) and informed consent from all participants enrolled in UKBB.

### Lipids GWAS in UKBB

LDL-C, HDL-C and TG were measured using enzymatic selective protection, enzyme immunoinhibition and GPO/POD methods, respectively. The lipid measures were standardised using inverse rank normalisation such that the mean was 0 and standard deviation was 1. Given that the summary estimates from BCAC described BCa in females only, we conducted female-specific GWAS of LDL-C (N = 238,861), HDL-C (N = 217,373) and TG (N = 239,119) in UKBB female participants of European descent on K-means clustering of genetic ancestry data (K=4) after standard exclusions including withdrawn consent, mismatch between genetic and reported sex and putative sex chromosome aneuploidy<sup>35,36</sup>. We identified single nucleotide polymorphisms (SNPs) associated with LDL-C, HDL-C and TG using the BOLT-LMM (linear mixed model) software<sup>37</sup>. Analyses were adjusted for age and a binary variable denoting the genotyping chip individuals were allocated to in UKBB (the UKBB Axiom array or the UK BiLEVE array).

### IGF-I GWAS meta-analysis

IGF-I was measured in 468,384 individuals in UKBB using the chemiluminescent immunoassay (DiaSorin Ltd, UK). IGF-I measures were standardised using inverse rank normalisation such that the mean was 0 and standard deviation was 1. We conducted a sex-combined GWAS for IGF-I (N = 438,988) in UKBB participants of European descent using the same GWAS pipeline as the lipid GWAS as described above. Analyses were adjusted for age, sex and a binary variable denoting the genotyping chip individuals were allocated to in UKBB (the UKBB Axiom array or the UK BiLEVE array). SNP effect estimates and their standard errors from the IGF GWAS in UKBB were combined with those from the IGF-I GWAS (N = 30,884) conducted by IGF working group of the CHARGE consortium<sup>30</sup> by inverse-weighted meta-analysis using GWAMA<sup>38</sup>. Given that the summary estimates from BCAC and UKBB described BCa in females only, we also conducted a female-specific GWAS of IGF-I (N = 237,947) in UKBB using BOLT-LMM. Analyses were adjusted for age and a binary variable denoting the genotyping chip individuals were allocated to in UKBB (the UKBB Axiom array or the UK BiLEVE array). SNP effect estimates and their standard errors were then combined with those from the female-specific IGF-I GWAS conducted by the IGF working group of the CHARGE consortium (N = 14,600)<sup>30</sup>.

## Selection of genetic instruments for MR analyses

**Lipid instruments:** To investigate the causal association between lipids and BCa risk (Figure 1, Part I), we identified 135, 214 and 203 independent SNPs (clumped based on a linkage disequilibrium (LD)  $r^2 < 0.001$  and 1Mb window) associated with LDL-C, HDL-C and TG, respectively, at  $P < 5e^{-08}$  from the female-specific lipid GWAS conducted in UKBB (described above). Due to the complex overlapping nature of the lipid traits, genetic variants are commonly associated with more than one lipid trait. To disentangle the roles of LDL-C, HDL-C and TG, we also used MVMR which was developed to estimate the direct effect of various correlated risk factors when conditioned on one another in a single model<sup>39</sup> (see Supplementary methods for more details). For the MVMR methods, we included all female-specific GWAS-associated SNPs for LDL-C, HDL-C and TG in the model (Supplementary Table 1).

For the bi-directional MR analyses investigating the causal relationship between lipids and IGF-I (Figure 1, Part IIIB), we selected 76, 86 and 51 independent SNPs associated with LDL-C, HDL-C and TG at  $P < 5 \times 10^{-8}$  from the sex-combined lipid GWAS by GLGC ( $N = 188,577$ )<sup>33</sup>. As effect estimates taken from overlapping datasets can be biased in the direction of the null for two-sample MR analyses, we chose to use results from the sex-combined lipid GWAS conducted by GLGC and not UKBB for this analysis as the IGF-I instruments were derived from a IGF-I GWAS meta-analysis which included data from UK Biobank (female-specific lipid GWAS was not available from GLGC and thus not used for this analysis). For the MVMR methods, we selected 185 SNPs ( $r^2 < 0.2$ ) associated with LDL-C, HDL-C and TG ( $p < 5 \times 10^{-8}$ ) from the lipid GWAS by GLGC ( $N = 188,577$ )<sup>33</sup> (Supplementary Table 1).

**IGF-I instruments:** For the MR analyses investigating the causal association between IGF-I and BCa risk (Figure 1, Part II), we identified 278 independent (clumped based on a  $r^2 < 0.001$  within a 1Mb window) female-specific SNPs associated with the IGF-I at  $P < 5e^{-08}$  from the female-specific IGF-I GWAS meta-analysis (Supplementary Table 2) (described above).

To assess the causality and direction of association between lipids and IGF-I (Figure 1, Part IIIB), we identified 476 independent SNPs associated with IGF-I at the conventional GWAS threshold ( $p < 5e^{-08}$ ), within 1MB and at  $r^2 < 0.001$  from the sex-combined IGF-I GWAS meta-analysis (Supplementary Table 2) (described above). Sex-combined estimates for IGF-I were used as sex-specific lipid GWAS was not conducted by GLGC.

**Statistical analyses**—We examined the association of lipids and IGF-I with overall, ER positive and ER negative BCa using SNP estimates from the female-specific GWAS of lipids, IGF-I and BCa (Figure 1, Part I and II). Details of the SNPs included in each analysis, and proxies used, are provided in Supplementary Table 1 and 2. Summary statistics were harmonised using the `harmonise_data` function within the `TwoSampleMR` R package<sup>40</sup>. All GWAS were assumed to be coded on the forward strand and harmonization was confirmed as consistent using option 2 of the “action” argument. Univariable causal estimates were combined using the inverse-variance weighted (IVW)

method<sup>41</sup>. We performed the following sensitivity analyses, each robust to some form of potential unbalanced horizontal pleiotropy: 1) MR-Egger regression method<sup>42</sup> to test overall directional pleiotropy and provide a valid causal estimate, taking into account the presence of pleiotropy and; 2) weighted median method<sup>43</sup> which provides a consistent estimate of causal effect if at least 50% of the information in the analysis comes from variants that are valid instrumental variables. Due to the complex overlapping nature of the lipid traits, we also performed multivariable IVW MR and MVMR Egger analyses to disentangle the roles of LDL-C, HDL-C and TG in BCa. As each of these sensitivity analyses make differing pleiotropy assumptions, consistency of causal effect estimates was interpreted to strengthen conclusions.

We examined the bi-directional relationship between lipid and IGF-I (Figure 1, Part IIIB) using SNP estimates from the sex-combined GWAS of lipids conducted by GLGC and IGF-I GWAS meta-analysis (described above) and the MR models described above. We performed the MR Steiger directionality test<sup>44</sup> to determine whether the observed observations were directionally causal based on the variance explained by the genetic instruments in the exposure and outcome and tests if the variance in the outcome is less than the exposure. We also performed LD score regression to look at the genetic correlation between lipids and IGF-I<sup>45</sup>.

MVMR was conducted as an extension of the IVW method to test the hypothesis that circulating IGF-I may act as an intermediate factor in any reported association between circulating lipids and BCa, or vice versa (Figure 1, Part IV). For the MVMR analyses, we fitted a model with LDL-C, HDL-C, TG and IGF-I. Two-sample conditional F-statistics were estimated to provide some assessment of instrument strength of each exposure when accounting for the prediction of other exposures in the multivariable model using the MVMR R package by Sanderson et al (<http://github.com/WSpiller/MVMR>)<sup>46,47</sup>.

In each instance, MR estimates are interpreted as the change in outcome per SD unit change in the exposure. Estimates for BCa outcomes reflect odds ratios (OR). All MR analyses were performed using the MR-Base “TwoSampleMR” package<sup>40</sup>. All other statistical analyses were performed using Stata version 14 (StataCorp, College Station, Texas, USA) or R version 3.2.4.

## Observational analyses

### Study Populations

**Study of Health in Pomerania (SHIP) Participants:** Observational analyses (Figure 1, Part IIIA) of the relationship between lipids and IGF-I were examined in a cross-sectional study within SHIP<sup>48</sup>, a population-based project conducted in Northeast Germany (see Supplemental Methods for details). All participants underwent standardized medical examination, blood sampling and extensive computer-aided personal interview. Data on socio-demographic characteristics and medical histories were collected. The present study includes unrelated individuals with both lipid and IGF-I measurements (N = 3812; these data are described in detail in the Supplementary methods). All participants gave written informed consent and the study conformed to the principles of the Declaration of Helsinki

as reflected by an a priori approval of the Ethics Committee of the University of Greifswald (Greifswald, Germany).

**UK Biobank participants:** Replication analyses to investigate the observational relationship between lipids and IGF-I were examined in a cross-sectional study within UKBB<sup>29</sup> (under application # 16009) (described above and in the Supplementary methods). We included individuals with both lipid and IGF-I measurements in the present study.

**Statistical analysis**—Observational associations between lipids and IGF-I were assessed in the SHIP and UKBB using linear regression. Fully adjusted models included age, sex, smoking status, body mass index (BMI) and diabetes status. Associations of lipids and IGFs with potential confounders were estimated using linear regression.

## Results

### Part I: Two-sample MR analysis to estimate the causal effect of lipids on BCa

The univariable IVW analyses found little evidence that LDL-C was associated with BCa (OR: 1.01; 95%CI: 0.97, 1.06;  $p=0.59$ ). There was evidence that HDL-C was associated with increased odds of overall BCa (OR: 1.08; 95%CI: 1.04, 1.13;  $p=0.0002$ ) and TG was associated with decreased odds of overall BCa (OR: 0.94; 95%CI: 0.89, 0.98;  $p=0.01$ ) (Figure 2). Estimates of all causal associations between lipids and overall BCa are shown in Figure 2. Sensitivity analyses using methods that take into account potential genetic pleiotropy did not result in substantive changes in the estimates.

When assessed together using MVMR, the estimated causal odds ratio from multivariable IVW for LDL-C, HDL-C and TG were 1.04 (95%CI: 1.00, 1.07;  $p=0.05$ ), 1.07 (95%CI: 1.03, 1.11;  $p=0.0002$ ) and 0.95 (95%CI: 0.91, 0.99;  $p=0.01$ ), respectively (Figure 2). The estimated causal odds ratios from the MVMR-Egger analyses was similar to the multivariable IVW analyses for LDL-C, HDL-C and TG, with little evidence of directional pleiotropy (LDL-C intercept: 0.001; se: 0.001;  $p=0.26$ ; HDL-C intercept: -0.001; se: 0.001;  $p=0.31$ ; TG intercept: 0.001, se: 0.001;  $p=0.28$ ).

We also investigated the relationship between circulating lipids and BCa, stratified by ER status (Supplementary Figure 1). There was little evidence that LDL-C was associated with associated with ER-positive or ER-negative breast cancer from univariable IVW and multivariable IVW analyses. For HDL-C, the estimated causal odds ratios from univariable IVW (OR: 1.07; 95%CI: 1.01, 1.12;  $p=0.01$ ) and multivariable IVW analyses (OR: 1.06; 95%CI: 1.01, 1.10;  $p=0.01$ ) had similar direction and magnitude of association, with both analyses suggesting that HDL-C increases odds of ER-positive BCa. There was also evidence from both univariable IVW (OR: 1.10; 95%CI: 1.04, 1.17;  $p=0.002$ ) and multivariable IVW analyses (OR: 1.08; 95%CI: 1.03, 1.14;  $p=0.004$ ) that HDL-C increases odds of ER-negative BCa risk. For TG, the estimated causal odds ratios from univariable IVW (OR: 0.94; 95%CI: 0.89, 0.99;  $p=0.03$ ) and multivariable IVW analyses (OR: 0.94; 95%CI: 0.90, 0.99;  $p=0.01$ ) were consistent, with both analyses suggesting that TG decreases odds of ER-positive BCa. However, for TG, there was evidence of directional pleiotropy from MVMR-Egger analyses (intercept: 0.002; se: 0.001;  $p=0.01$ ). Sensitivity



analyses using univariable MR-Egger, weighted median and MVMR-Egger did not result in substantive changes in the estimates.

## Part 2: Two-sample MR analysis to estimate the causal effect of IGF-I on BCa

IGF-I was associated with increased odds of overall BCa (IVW OR: 1.09; 95% CI 1.04, 1.15;  $p=0.001$ ) from IVW analyses. Estimates from MR-Egger and weighted median analyses were consistent with the IVW estimates (Table 2).

Using data from BCAC, we investigated the relationship between IGF-I and BCa by ER-status. The odds of ER-positive BCa and ER-negative BCa was 1.09 (95% CI: 1.02, 1.15;  $p=0.01$ ) and 1.04 (95% CI: 0.96, 1.12;  $p=0.37$ ), respectively, from IVW analyses (Table 2).

## Part IIIA: The observational association between lipids and IGFs

Study characteristics of the SHIP and UKBB study are shown in Supplementary Table 3. In SHIP, the mean (SD) LDL-C, HDL-C and TG levels were 3.57 (1.16) mmol/l, 1.45 (0.44) mmol/l and 1.82 (1.30) mmol/l, respectively. The mean (SD) IGF-I was 142.1 (57.6) ng/ml.

The observational associations between circulating lipids and IGF-I using data from SHIP are shown in Supplementary Table 4. In the unadjusted analyses, a SD unit increase in LDL-C, HDL-C and TG was associated with a -0.11 (95% CI -0.14, -0.08;  $p=2.27\times 10^{-12}$ ), 0.02 (95% CI -0.01, 0.05;  $p=0.29$ ), -0.16 (95% CI -0.19, -0.13;  $p=2.26\times 10^{-22}$ ) SD unit change in IGF-I levels, respectively. Circulating lipids and IGF-I were associated with potential confounders of a lipid-or IGF-BCa relationship, including age, sex, smoking status, diabetes status and body mass index (Supplementary Table 5). In the fully adjusted model, a SD unit increase in LDL-C, HDL-C and TG was associated with a 0.03 (95% CI 0.004, 0.06;  $p=0.03$ ), -0.05 (95% CI -0.08, -0.02;  $p=0.001$ ) and -0.06 (95% CI -0.09, -0.04;  $p=1.5\times 10^{-5}$ ) SD unit change in IGF-I levels, respectively (Supplementary Table 4).

We undertook observational analyses using data from UKBB (Supplementary Table 4). In unadjusted analyses in UKBB, a SD unit increase in LDL-C, HDL-C and TG was associated with a -0.01 (95% CI -0.014, -0.006;  $p=2.08\times 10^{-19}$ ), -0.03 (95% CI -0.032, -0.028;  $p=8.9\times 10^{-102}$ ) and -0.05 (95% CI -0.054, -0.046;  $p=9.37\times 10^{-204}$ ) SD unit change in IGF-I levels, respectively. The association between LDL-C and TG with IGF-I is directionally consistent but smaller in magnitude compared to the analyses in SHIP. In the fully adjusted model, a SD unit increase in LDL-C, HDL-C and TG was associated with a -0.001 (95% CI -0.005, 0.003;  $p=0.57$ ), -0.04 (95% CI -0.044, -0.036;  $p=1.88\times 10^{-112}$ ) and -0.01 (95% CI -0.014, -0.006;  $p=5.05\times 10^{-20}$ ) SD unit change in IGF-I levels, respectively. For the adjusted analyses, the association between HDL-C and TG with IGF-I is directionally consistent but smaller in magnitude compared to the adjusted analyses in SHIP.

## Part IIIB: Two-sample MR analysis to estimate the bi-directional causal association between lipids and IGF-I

We estimated the causal effect of LDL-C, HDL-C and TG on IGF-I levels using two-sample MR. There was weak evidence that LDL-C or HDL-C affect levels of IGF-I (Table 3).

The univariable IVW analyses suggested that a SD unit increase in TG (approximately

81.8mg/dL) is associated with a -0.13 (95%CI: -0.23, -0.04;  $p=0.01$ ) SD unit change in IGF-I. Estimates from sensitivity analyses using methods that take into account potential pleiotropy were in the same direction but differed in magnitude to univariable IVW estimates. Estimates from multivariable IVW (-0.12; 95%CI: -0.20, -0.05;  $p=0.002$ ) and MVMR-Egger methods (-0.17; 95%CI: -0.26, -0.08;  $p=0.0003$ ) were consistent with the univariable IVW estimates, with weak evidence of directional pleiotropy (intercept: 0.003; se: 0.001;  $p=0.054$ ) (Table 3). The MR Steiger directionality test suggested that the observed association was directionally causal (Supplementary Table 6).

In the reverse direction (estimation of the causal effect of IGF-I on lipids), there was little evidence to suggest that IGF-I levels impact circulating LDL-C or HDL-C (Table 4). A SD unit increase in IGF-I (approximately 49.76ng/ml) was associated with a -0.11 (95%CI: -0.18, -0.05;  $p=0.001$ ) SD unit change in TG, using the univariable IVW method. This estimate was larger in magnitude in the univariable MR-Egger analyses (-0.28; 95%CI: -0.42, -0.15;  $p=0.00004$ ), with evidence of directional pleiotropy (intercept: 0.01; se:0.002;  $p=0.004$ ). The estimate attenuated towards the null in the weighted median analyses (-0.02; 95%CI: -0.07, 0.02;  $p=0.34$ ). The MR Steiger directionality test suggested that the observed association was directionally causal (Supplementary Table 6).

Supplementary Table 7 shows the estimated genetic correlations between IGF-I with lipid traits. There was little evidence of genetic correlation between IGF-I with the circulating lipids except for HDL-C although the genetic correlation was low (genetic correlation (rG): 0.04; se: 0.02;  $p=0.05$ ).

#### Part IV: MVMR analyses to estimate the direct effect of lipids/IGF-I on BCa

Multivariable IVW analyses were conducted to investigate whether the effect of lipids on overall BCa was attenuated following adjustment for IGF-I, or vice versa (Table 5). Using data from BCAC, the MVMR OR for overall BCa per SD increase in TG, conditioned on HDL-C, LDL-C and IGF-I, was 0.95 (95%CI: 0.92, 0.99;  $p=0.03$ ) (Table 5), which was comparable to the IGF-I unadjusted model (OR: 0.95; 95%CI: 0.91, 0.99;  $p=0.01$ ) (Figure 2). The MVMR OR for overall BCa per SD increase in IGF-I, conditioned on LDL-C, HDL-C and TG, was 1.09 (95%CI: 1.05, 1.14;  $p=0.0001$ ) (Table 5), which was comparable to the lipid unadjusted model (OR: 1.09; 95%CI: 1.04, 1.15;  $p=0.001$ ) (Table 2). We assessed likely instrument strength in the MVMR models and found that the conditional F statistics for LDL-C, HDL-C and TG and IGF-I were 52.87, 40.83, 36.48 and 45.14, respectively, suggesting sufficient instrument strength (Table 5). Evaluation of horizontal pleiotropy using a modified form of Cochran's Q-statistic with respect to the differences in MVMR estimates across the instruments found evidence of potential pleiotropy in the MVMR model ( $p<4.31\times 10^{-135}$ ).

## Discussion

In this study, we explored the interplay between circulating lipids and IGF-I and the relevance to BCa risk. Using two-sample MR, there was strong evidence that HDL-C is positively associated with BCa risk, whereas TG is negatively associated with BCa risk. Results from observational and MR analyses highlight that TG decreases IGF-I and

that IGF-I decreases TG; providing evidence of a bi-directional relationship between TG and IGF-I. The LD score regression analysis contributed towards evidence of a causal relationship between TG and IGF-I, as opposed to shared heritability. Findings from our observational and MR analyses point to an interplay between TG and IGF-I, however, MVMR estimates suggest effects of TG and IGF-I on BCa are independent.

The effects of lipids on BCa risk have been investigated by several MR studies<sup>8–11</sup>, which found consistent evidence in support of HDL-C playing a causal role in the aetiology of BCa. Conversely, evidence in support of the causal role of LDL-C and TG on BCa risk has been less consistent. In agreement with two previous MR studies<sup>8,11</sup>, our study provides evidence that HDL-C is positively associated with overall and ER-positive BCa whereas TG is negatively associated with overall BCa and ER-positive BCa. Findings from our study build on prior investigations through the inclusion of a large female-specific GWAS of lipid traits in UKBB, thus expanding the number of robust female-specific genetic instruments for each lipid trait. As the BCa outcome was defined in females, we chose to use female-specific instruments for the lipid traits as sex-specific effects have been observed for these exposures<sup>49</sup>. This is in contrast to previous lipid-BCa MR studies<sup>8–11</sup> which had used non-sex specific lipid instruments, resulting in reduced precision in their estimates in comparison to our results.

Given the high degree of inter-relatedness amongst the lipid traits, genetic variants associated with one lipid trait will also be associated with other lipid traits. The main challenge in investigating the association between lipids and BCa is addressing the potential for horizontal pleiotropy (violation of the exclusion restriction assumption) which can confound the MR estimates<sup>50</sup>. With this in mind, we applied MVMR to simultaneously estimate the direct causal effect of various correlated lipid traits on BCa when conditioned on one another in a single model. A MR study on lipid and BCa risk using MVMR was recently published<sup>10</sup>; the authors included potential confounders such as body mass index and age of menarche in their MVMR model and their primary findings was that both HDL-C and LDL-C were associated with increased BCa risk. We did not adjust for potential confounders in our MVMR model in order to maintain instrument strength (as indicated by the conditional F statistic), however, we conducted MVMR-Egger analyses which can provide reliable evidence regarding causation even in the presence confounding through unbalanced horizontal pleiotropy. Indeed, results from our MVMR-Egger analyses were consistent with our multivariable IVW analyses.

We investigated the relationship between IGF-I and BCa using female-specific instruments from a meta-analysis of two major IGF-I GWAS from UKBB and CHARGE. Our study found evidence that IGF-I increases overall BCa, likely driven by ER-positive BCa which is concordant with results from previous observational and MR studies<sup>15,51</sup>. Given that SNPs associated with IGF-I are also associated with other components of the IGF axis, there is a possibility that our results could be biased due to pleiotropy with other components of the IGF axis<sup>52</sup>. GWAS analyses for other components of the IGF axis have been conducted<sup>30,53</sup>. However, due to a lack of robust genetic instruments for these individual IGF traits, we were unable to use MVMR to investigate the direct effect of IGF-I conditioned on the other components of the IGF axis. Preclinical evidence suggests that IGF signalling is

mitogenic for both ER-positive and ER-negative BCa<sup>54</sup>. In contrast, our MR results suggest that circulating IGF-I is more influential in ER-positive BCa compared to ER-negative BCa. However, our MR analyses for ER-negative BCa could lack statistical power due to the smaller sample size compared to ER-positive BCa. As we have used genetic variants that predict circulating levels of IGF-I, we cannot rule out important aspects of tissue-specific regulation which may contribute to BCa.

Our finding that IGF-I associates with decreased TG is consistent with findings from previous observational<sup>24</sup>, interventional<sup>17,55–57</sup> and candidate gene studies<sup>58,59</sup>. IGF-I can inhibit growth hormone secretion by negative feedback in the growth hormone (GH)-IGF axis. It has been suggested that the effect of circulating IGF-I on circulating TG levels is most likely due its effect on growth hormone or insulin secretion<sup>56,58,60</sup>. To the best of our best knowledge, this is the first study suggesting that TG levels causally reduce IGF-I levels. The mechanisms by which circulating TGs affect circulating IGF-I levels remain to be elucidated. The liver is the main source of circulating IGF-I (accounting for ~75% of circulating IGF-I) and many variables are known to control IGF-I synthesis and secretion, including nutrient intake, insulin and growth hormone levels<sup>61,62</sup>. It is possible that TGs could affect the hepatic synthesis of IGF-I.

Our observational and MR estimates support a bi-directional relationship between TG and IGF-I and suggest that the pathways from these phenotypes to BCa overlap. In the context of this, we applied MVMR approaches to investigate the direct effects of TG and IGF-I on BCa, independently of each other. The associations between TG or IGF-I with BCa from the MVMR analyses (wherein LDL-C, HDL-C, TG and IGF-I were included in the model) was not attenuated when compared to the univariable MR associations. This lack of attenuation is not likely due to weak instrument bias, which is a common problem in MVMR, however, we cannot rule out directional pleiotropy as a source of bias as a modified form of Cochran's Q-statistic found evidence of potential horizontal pleiotropy in the MVMR model. To this end, further work is required to elucidate other potential modifiable risk factors that drive the putative causal relationship between lipids or IGF-I with BCa.

Our study has several limitations. First, despite the large sample size of the UK Biobank, this cohort is not representative of the general population due to the recruitment of generally healthier individuals with higher socioeconomic status. Hence, our findings might be prone to selection bias and cannot be generalised to the UK population. Second, our study was focused primarily on individuals of European ancestry. Although population homogeneity eliminates population admixture as a potential confounder in our analyses, the findings drawn from this study might not be generalisable to non-European populations.

## Conclusion

In conclusion, our findings highlight a causal role for HDL-C, TG and IGF-I in BCa risk. Observational and bi-directional MR analyses support an interplay between IGF-I and TG, however, results from MVMR analyses suggest that TG and IGF-I may act independently to influence BCa. These relationships should be considered in the development of prevention strategies for BCa, where interventions are known to modify circulating lipid and IGF traits

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

This work was specifically supported by the Medical Research Council (MRC) Integrative Epidemiology Unit (IEU) (MC\_UU\_12013/3) and by a Cancer Research UK Programme Grant [The Integrative Cancer Epidemiology Programme, ICEP] (C18281/A29019). This research was funded in part by the Wellcome Trust (202802/Z/16/Z). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. VT, KMB, CMP, JMPH, NJT are supported by ICEP (C18281/A19169). NJT is a Wellcome Trust Investigator (202802/Z/16/Z) and works within the University of Bristol NIHR Biomedical Research Centre (BRC). LJC is supported by NJT's Wellcome Trust Investigator grant (202802/Z/16/Z). NJT and LJC work in the MRC IEU at the University of Bristol which is supported by the MRC (MC\_UU\_00011) and the University of Bristol. TD received support from Wellcome (grant ref 201268/Z/16/Z) and is now funded by the NIHR as an Academic Clinical Fellow. QQ is supported by a Scientist Development Award (K01HL129892) from the NHLBI. RK is supported by the National Heart, Lung and Blood Institute of the National Institutes of Health (HL105756). JR is supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI\_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). This publication is the work of the authors who will serve as guarantors for the contents of this paper. Quality Control filtering of the UK Biobank data was conducted by R.Mitchell, G.Hemani, T.Dudding, L.Paternoster as described in the published protocol([doi:10.5523/bris.3074krb6t2fj29yh2b03x3wxj](https://doi.org/10.5523/bris.3074krb6t2fj29yh2b03x3wxj)).

## Data Availability

Results of the female-specific lipid GWAS and the IGF-I GWAS will be made available to download from the OpenGWAS database (<http://gwas.mrcieu.ac.uk>).

## References

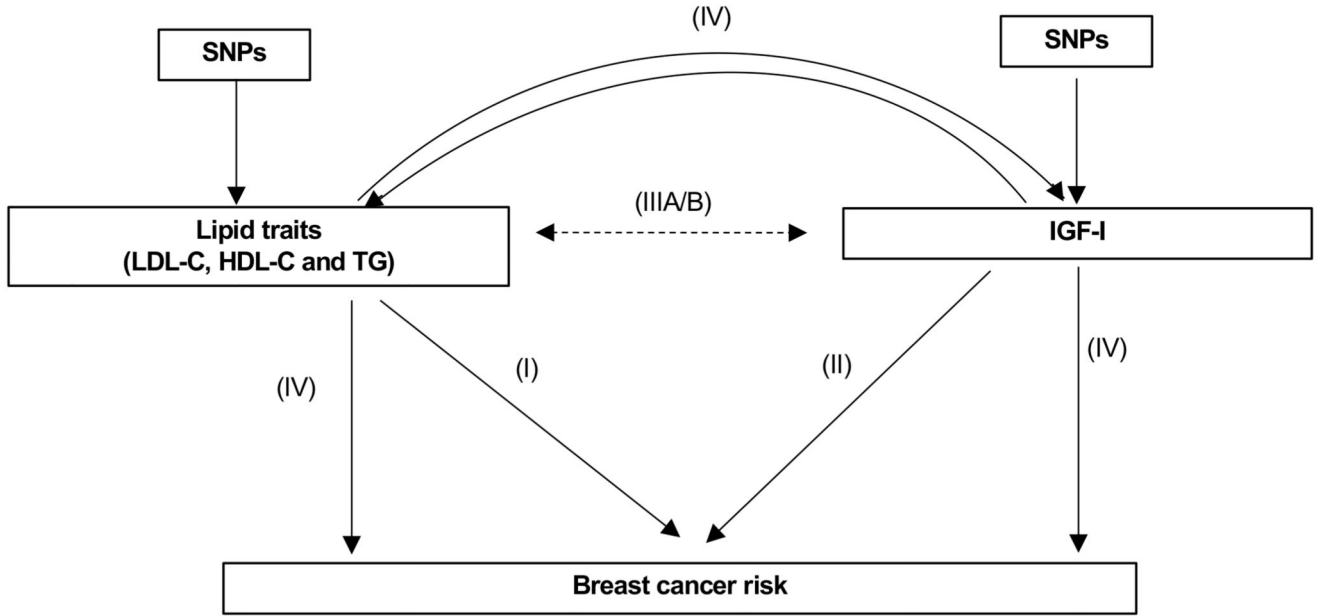
1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015
2. Siegel RL, Miller KD, Jemal A, Cancer Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016; 66 (1) :7–30. DOI: 10.3322/caac.21332statistics [PubMed: 26742998]
3. Brown KF, Rungay H, Dunlop C, et al. The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015. *Br J Cancer.* 2018
4. Avgerinos KI, Spyrou N, Mantzoros CS, Dalamaga M. Obesity and cancer risk: Emerging biological mechanisms and perspectives. *Metabolism.* 2019
5. Touvier M, Fassier P, His M, et al. Cholesterol and breast cancer risk: A systematic review and meta-analysis of prospective studies. *Br J Nutr.* 2015
6. Ni H, Liu H, Gao R. Serum lipids and breast cancer risk: A meta-Analysis of prospective cohort studies. *PLoS One.* 2015
7. Smith GD, Ebrahim S. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003
8. Beeghly-Fadiel A, Khankari NK, Delahanty RJ, et al. A Mendelian randomization analysis of circulating lipid traits and breast cancer risk. *Int J Epidemiol.* 2019
9. Nowak C, Ärnlöv J. A Mendelian randomization study of the effects of blood lipids on breast cancer risk. *Nat Commun.* 2018
10. Johnson KE, Siewert KM, Klarin D, et al. The relationship between circulating lipids and breast cancer risk: A Mendelian randomization study. *PLoS Med.* 2020

11. Orho-Melander M, Hindy G, Borgquist S, et al. Blood lipid genetic scores, the HMGCR gene and cancer risk: A Mendelian randomization study. *Int J Epidemiol*. 2018
12. Perks CM, Holly JMP. IGF binding proteins (IGFBPs) and regulation of breast cancer biology. *J Mammary Gland Biol Neoplasia*. 2008
13. Pollak M. The insulin and insulin-like growth factor receptor family in neoplasia: An update. *Nat Rev Cancer*. 2012
14. Perks CM, Holly JMP. Insulin-like growth factor binding proteins (IGFBPs) in breast cancer. *J Mammary Gland Biol Neoplasia*. 2000
15. Murphy N, Knuppel A, Papadimitriou N, et al. Insulin-like growth factor-1, insulin-like growth factor-binding protein-3, and breast cancer risk: observational and Mendelian randomization analyses with ~430 000 women. *Ann Oncol*. 2020
16. Szkodziński J, Romanowski W, Hudzik B, et al. Effect of HMG-CoA (3-hydroxy-3-methylglutaryl-CoA) reductase inhibitors on the concentration of insulin-like growth factor-1 (IGF-1) in hypercholesterolemic patients. *Pharmacol Reports*. 2009
17. Bergen K, Brismar K, Tehrani S. High-dose atorvastatin is associated with lower IGF-1 levels in patients with type 1 diabetes. *Growth Horm IGF Res*. 2016
18. Narayanan RP, Gittins M, Siddals KW, et al. Atorvastatin administration is associated with dose-related changes in IGF bioavailability. *Eur J Endocrinol*. 2013
19. Mauras N, Haymond MW. Metabolic effects of recombinant human insulin-like growth factor-I in humans: Comparison with recombinant human growth hormone. *Pediatr Nephrol*. 1996
20. Mauras N, Martinez V, Rini A, Guevara-Aguirre J. Recombinant Human Insulin-Like Growth Factor I Has Significant Anabolic Effects in Adults with Growth Hormone Receptor Deficiency: Studies on Protein, Glucose, and Lipid Metabolism 1. *J Clin Endocrinol Metab*. 2000
21. Mauras N, O'Brien KO, Welch S, et al. Insulin-Like Growth Factor I and Growth Hormone (GH) Treatment in GH-Deficient Humans: Differential Effects on Protein, Glucose, Lipid, and Calcium Metabolism 1. *J Clin Endocrinol Metab*. 2000
22. Eggert ML, Wallaschofski H, Grotevendt A, et al. Cross-sectional and longitudinal relation of IGF1 and IGF-binding protein 3 with lipid metabolism. *Eur J Endocrinol*. 2014
23. Friedrich N, Nauck M, Schipf S, Völzke H, Brabant G, Wallaschofski H. Cross-sectional and longitudinal associations between insulin-like growth factor I and metabolic syndrome: A general population study in German adults. *Diabetes Metab Res Rev*. 2013
24. Lam CSP, Chen MH, Lacey SM, et al. Circulating insulin-like growth factor-1 and its binding protein-3: Metabolic and genetic correlates in the community. *Arterioscler Thromb Vasc Biol*. 2010
25. Melvin JC, Holmberg L, Rohrmann S, Loda M, Van Hemelrijck M. Serum lipid profiles and cancer risk in the context of obesity: Four meta-analyses. *J Cancer Epidemiol*. 2013
26. Chandler PD, Song Y, Lin J, et al. Lipid biomarkers and long-term risk of cancer in the Women's Health Study. *Am J Clin Nutr*. 2016
27. Pollak M. Insulin and insulin-like growth factor signalling in neoplasia (*Nature Reviews Cancer* (2008) 8, (915-928)). *Nat Rev Cancer*. 2009
28. Relton CL, Davey Smith G. Two-step epigenetic mendelian randomization: A strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol*. 2012
29. Allen NE, Sudlow C, Peakman T, Collins R. UK biobank data: Come and get it. *Sci Transl Med*. 2014
30. Teumer A, Qi Q, Nethander M, et al. Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits. *Aging Cell*. 2016
31. Michailidou K, Lindström S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017
32. Zhang H, Ahearn TU, Lecarpentier J, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat Genet*. 2020
33. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013

34. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med.* 2015
35. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. *Nat Protoc.* 2010
36. Mitchell, R, Hemani, G, Dudding, T, Paternoster, L. UK Biobank Genetic Data: MRC-IEU Quality Control. 2017.
37. Loh PR, Tucker G, Bulik-Sullivan BK, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet.* 2015
38. Mägi R, Morris AP. GWAMA: Software for genome-wide association meta-analysis. *BMC Bioinformatics.* 2010
39. Burgess S, Thompson SG. Multivariable Mendelian randomization: The use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol.* 2015
40. Hemani G, Zheng J, Elsworth B, et al. The MR-base platform supports systematic causal inference across the human phenome. *Elife.* 2018
41. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol.* 2013
42. Bowden J, Smith GD, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015
43. 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
44. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 2017
45. Zheng J, Erzurumluoglu AM, Elsworth BL, et al. LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics.* 2017
46. Sanderson E, Spiller W, Bowden J. Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomisation. *bioRxiv.* 2020
47. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol.* 2019
48. Völzke H, Alte D, Schmidt CO, et al. Cohort profile: The study of health in Pomerania. *Int J Epidemiol.* 2011
49. Palmisano BT, Zhu L, Eckel RH, Stafford JM. Sex differences in lipid and lipoprotein metabolism. *Mol Metab.* 2018
50. Labrecque J, Swanson SA. Understanding the Assumptions Underlying Instrumental Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. *Curr Epidemiol Reports.* 2018
51. Key TJ, Appleby PN, Reeves GK, et al. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: Pooled individual data analysis of 17 prospective studies. *Lancet Oncol.* 2010
52. Bonilla C, Lewis SJ, Rowlands MA, et al. Assessing the role of insulin-like growth factors and binding proteins in prostate cancer using Mendelian randomization: Genetic variants as instruments for circulating levels. *Int J Cancer.* 2016
53. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. *Nature.* 2018
54. Surmacz E, Bartucci M. Role of estrogen receptor  $\alpha$  in modulating IGF-I receptor signaling and function in breast cancer. *J Exp Clin Cancer Res.* 2004
55. Agnoli C, Berrino F, Abagnato CA, et al. Metabolic syndrome and postmenopausal breast cancer in the ORDET cohort: A nested case-control study. *Nutr Metab Cardiovasc Dis.* 2010
56. Bianda T, Zapf J, Froesch ER, Schmid C. Insulin-like growth factor-I lowers fasting and postprandial triglyceride levels without affecting chylomicron clearance in healthy men. *J Intern Med.* 1999

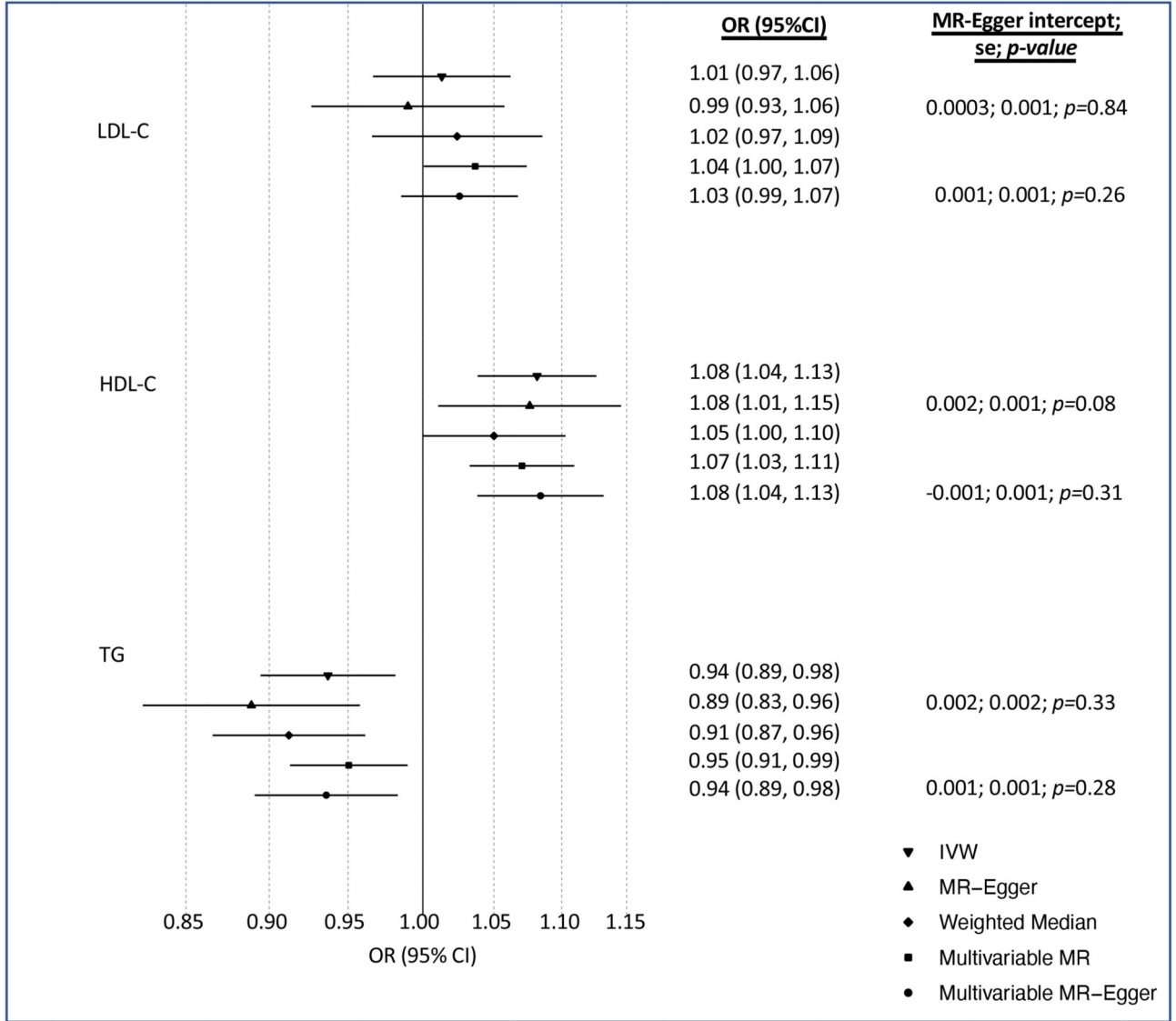
57. Christ ER, Carroll PV, Albany E, et al. Effect of IGF-I therapy on VLDL apolipoprotein B100 metabolism in type 1 diabetes mellitus. *Am J Physiol – Endocrinol Metab.* 2002
58. Nielsen EMD, Hansen L, Lajer M, et al. A common polymorphism in the promoter of the IGF-I gene associates with increased fasting serum triglyceride levels in glucose-tolerant subjects. *Clin Biochem.* 2004
59. Wang WC, Chiu YF, Chung RH, et al. IGF1 gene is associated with triglyceride levels in subjects with family history of hypertension from the SAPPHiRE and TWB projects. *Int J Med Sci.* 2018
60. Froesch ER, Hussain M. Recombinant human insulin-like growth factor-I: a therapeutic challenge for diabetes mellitus. *Diabetologia.* 1994
61. Clemmons DR. Metabolic Actions of Insulin-Like Growth Factor-I in Normal Physiology and Diabetes. *Endocrinol Metab Clin North Am.* 2012
62. Underwood LE, Thissen JP, Lemozy S, Ketelslegers JM, Clemmons DR. Hormonal and nutritional regulation of IGF-I and its binding proteins. *Horm Res Paediatr.* 1994





**Figure 1. Overview of methods employed**

I) Two-sample MR analysis of the effect of lipid traits on breast cancer risk; II) Two-sample MR analysis of the effect of IGF-I on breast cancer risk; III) Analysis of the observational association between lipid traits and IGF-I; and IV) Multivariable MR to estimate the direct effect of IGF-I or TG on BCa conditioned on each other. The solid lines represent analyses using female-specific instruments and dashed lines represent analyses using non-sex specific instruments.



**Figure 2. Estimates of the causal relationship between circulating lipid traits and overall breast cancer in BCAC and UK Biobank.**

The forest plot shows the estimate of the causal effect of LDL-C, HDL-C and TG on overall breast cancer risk using summary data from the breast cancer association consortium (BCAC) (n=133,384 cases and 113,789 controls). Circles represent point estimates from individual analyses. Horizontal lines represent the 95% confidence intervals.

**Table 1**  
**Data sources used for two-sample Mendelian randomization (MR) analyses**

Exposure/Outcome	Consortium or cohort study	Participants (N)	Part I: Two-sample MR analysis of the causal effect of lipids on BCa	Part II: Two-sample MR analysis of the causal effect of IGFs on BCa	Part IIIB: The bi-directional causal association between lipids and IGFs	Part IV: Multivariable MR analyses to estimate the direct effect of IGF-I or IG on BCa
<b>Insulin-like growth factor 1 (IGF-1)</b>	IGF working group of the CHARGE consortium	30,884 (sex-combined GWAS) <sup>a</sup>			X	
		14,600 (female-specific GWAS) <sup>b</sup>		X		X
	UKBB	438,988 (sex-combined GWAS) <sup>a</sup>			X	
<b>Lipid traits (LDL-C, HDL-C and TG)</b>		237,947 (female-specific GWAS) <sup>b</sup>		X		X
	Global lipids Genetics consortium (GLGC)	188,577 (sex-combined GWAS)			X	
	UKBB	LDL-C:238,861; HDL-C:217,373; TG:239,119 (female-specific GWAS)	X			X
<b>Breast cancer (BCa)</b>	Breast cancer association consortium (BCAC)	Overall: 133,384 cases and 113,789 controls	X	X		X
		ER-positive: 69,501 cases and 105,974 controls	X		X	
		ER-negative: 21,468 cases and 105,974 controls	X		X	

LDL-C, low density lipoprotein cholesterol; HDL-C, high density lipoprotein cholesterol; TG, triglycerides; UKBB, UK Biobank; BCa, breast cancer; N, sample size; MR, Mendelian Randomization

<sup>a</sup>Sex-combined IGF-I GWAS conducted by IGF working group of the CHARGE consortium was combined with those from the sex-combined IGF-I GWAS conducted using individual level data from UKBB

<sup>b</sup>Female-specific IGF-I GWAS conducted by IGF working group of the CHARGE consortium was combined with those from the female-specific IGF-I GWAS conducted using individual level data from UKBB

**Table 2**  
**Estimates of the causal relationship between IGF-I and breast cancer**

Exposure	Consortium	Outcome	N	Main analysis			Sensitivity analyses			MR-Egger intercept (se; <i>p</i> value)
				IVW		Weighted Median	MR-Egger			
				OR <sup>a</sup> (95% CI)	<i>p</i>		OR <sup>a</sup> (95% CI)	<i>p</i>		
IGF-I	BCAC	Overall	247,173	1.09 (1.04, 1.15)	0.001	1.08 (1.02, 1.14)	0.01	1.09 (0.97, 1.21)	0.14	0.0002 (0.002; <i>p</i> =0.91)
		BCa								
IGF-I	BCAC	ER positive	69,501	1.09 (1.02, 1.15)	0.01	1.07 (0.99, 1.15)	0.07	1.10 (0.97, 1.25)	0.12	-0.001 (0.002; <i>p</i> =0.77)
		BCa								
IGF-I	BCAC	ER negative	21,468	1.04 (0.96, 1.12)	0.37	1.04 (0.93, 1.17)	0.46	1.13 (0.96, 1.33)	0.16	-0.003 (0.003; <i>p</i> =0.26)
		BCa								

IVW, inverse variance weighted; OR, odds ratio; 95% CI, 95% confidence interval; *p*, *p*-value; se, standard error; MR, Mendelian randomization; IGF-I, insulin-like growth factor I; N, sample size

<sup>a</sup> Associations are per 1 SD unit increase in IGF-I

**Table 3**  
**Beta estimates of SD unit change in IGF-I per SD unit increase in HDL-C, LDL-C or TG based on two-sample and multivariable Mendelian randomization analyses.**

Exposure	Outcome	Main analysis			Sensitivity analyses							
		Univariable IVW	<i>p</i>	Univariable weighted Median	Univariable MR Egger regression	<i>p</i>	Univariable MR Egger intercept (se; <i>p</i> )	Multivariable MR Beta <sup>a</sup> (95% CI)	<i>p</i>	Multivariable Egger Beta <sup>a</sup> (95% CI)	<i>p</i>	Multivariable MR Egger intercept (se; <i>p</i> )
LDL-C	IGF-I	0.01 (-0.04, 0.06)	0.64	0.02 (-0.001, 0.03)	-0.01 (-0.09, 0.06)	0.73	0.002 (0.002; <i>p</i> =0.41)	0.03 (-0.02, 0.08)	0.28	-0.02 (-0.08, 0.05)	0.58	0.003 (0.001; <i>p</i> =0.02)
HDL-C	IGF-I	0.02 (-0.04, 0.08)	0.52	0.004 (-0.02, 0.02)	0.04 (-0.05, 0.14)	0.38	-0.001 (0.002; <i>p</i> =0.53)	-0.02 (-0.08, 0.05)	0.62	-0.03 (-0.10, 0.05)	0.48	0.001 (0.001; <i>p</i> =0.59)
TG	IGF-I	-0.13 (-0.23, -0.04)	0.01	-0.03 (-0.05, -0.001)	-0.24 (-0.39, -0.08)	0.004	0.01 (0.004; <i>p</i> =0.10)	-0.12 (-0.20, -0.05)	0.002	-0.17 (-0.26, -0.08)	0.0003	0.003 (0.001, <i>p</i> =0.054)

N, sample size; Beta, regression coefficient; IVW, inverse variance weighted; 95% CI, 95% confidence interval; *p*, *p*-value; MR, mendelian randomization; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; TG, triglycerides; IGF-I, insulin-like growth factor I

<sup>a</sup>Beta refers to the SD unit change in IGF-I levels per 1 SD unit increase in HDL-C, LDL-C or TG.

**Table 4**  
**Beta estimates of SD unit change in LDL-C, HDL-C and triglycerides per SD unit increase in IGF-I based on two-sample and multivariable Mendelian randomization analyses.**

Exposure	Outcome	Main Analysis		Sensitivity Analyses			MR Egger Intercept (se; p value)
		IVW Beta <sup>a</sup> (95% CI)	p	Weighted Median Beta <sup>a</sup> (95% CI)	p	MR Egger regression Beta <sup>a</sup> (95% CI)	
IGF-I	LDL-C	-0.06 (-0.14, 0.01)	0.09	-0.02 (-0.08, 0.03)	0.40	-0.06 (-0.21, 0.08)	0.39 (0.002; <i>p</i> =0.99)
	HDL-C	0.01 (-0.03, 0.06)	0.59	-0.02 (-0.06, 0.03)	0.48	0.01 (-0.08, 0.09)	0.0002 (0.001; <i>p</i> =0.86)
IGF-I	TG	-0.11 (-0.18, -0.05)	0.001	-0.02 (-0.07, 0.02)	0.34	-0.28 (-0.42, -0.15)	0.01 (0.002; <i>p</i> =0.004)

N, sample size; Beta, regression coefficient; IVW, inverse variance weighted; 95% CI, 95% confidence interval; *p*, *p*-value; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; TG, triglycerides; IGF-I, insulin-like growth factor I

<sup>a</sup>Beta refers to the SD unit change in HDL-C, LDL-C, TG or TC per SD unit change in IGF-I levels

**Table 5**  
**Multivariable MR analysis of the direct effect of lipids/IGF-I on Breast cancer.**

Exposure	Outcome	MVMR Beta (95%CI)	p	Conditional F statistic for MVMR	Q statistic for heterogeneity
LDL-C <sup>a</sup>	Overall BCa	1.04 (1.00, 1.08)	0.05	52.87	Q=2221.84 p=4.31×10 <sup>-135</sup>
HDL-C <sup>b</sup>	Overall BCa	1.07 (1.03, 1.11)	0.0005	40.83	
TG <sup>c</sup>	Overall BCa	0.95 (0.92, 0.99)	0.03	36.48	
IGF-I <sup>d</sup>	Overall BCa	1.09 (1.05, 1.14)	0.0001	45.14	

OR, odds ratio; 95%CI, 95% confidence interval; p, p-value; LDL-C, low density lipoprotein cholesterol; HDL-C, high density lipoprotein cholesterol; TG, triglycerides; MVMR, Multivariable Mendelian randomization; BCa, Breast cancer;

<sup>a</sup> Associations are per 1 SD unit increase in LDL-C when conditioned on IGF-I, HDL-C and TG.

<sup>b</sup> Associations are per 1 SD unit increase in HDL-C when conditioned on IGF-I, LDL-C and TG.

<sup>c</sup> Associations are per 1 SD unit increase in TG when conditioned on IGF-I, HDL-C and LDL-C

<sup>d</sup> Associations are per 1 SD unit increase in IGF-I when conditioned on TG, HDL-C and LDL-C