

1 **Title: Estimating the effect of lipids on IGF axis and subsequent breast cancer risk**

2 **Short title:** Lipids, IGFs and breast cancer risk

3 Vanessa Y Tan^{1,2†}, Caroline J Bull^{1,2†}, Kalina M Biernacka³, Alexander Teumer⁴, Laura

4 Corbin^{1,2}, Tom Dudding^{1,12}, Eleanor Sanderson^{1,2}, Qibin Qi⁵, Robert C Kaplan^{5,6}, Jerome I

5 Rotter⁷, Nele Friedrich⁸, Uwe Völker⁹, Julia Mayerle^{10,11}, Claire M Perks³, Jeff MP Holly³,

6 Nicholas J Timpson^{1,2}

7 † Joint first

8 ¹Medical Research Council (MRC) Integrative Epidemiology Unit, Population Health
9 Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom.

10 ²Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United
11 Kingdom.

12 ³IGFs & Metabolic Endocrinology Group, School of Translational Health Sciences, Learning &
13 Research Building, Southmead Hospital, Bristol, United Kingdom.

14 ⁴Institute for Community Medicine, University Medicine Greifswald, Greifswald, 17475,
15 Germany.

16 ⁵Department of Epidemiology and Population Health, Albert Einstein College of Medicine,
17 Bronx, NY, 10461, USA.

18 ⁶Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle WA.

19 ⁷The Institute for Translational Genomics and Population Sciences, Department of Pediatrics,
20 The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance,
21 CA USA

22 ⁸Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald,
23 Greifswald, 17475, Germany.

24 ⁹Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald,
25 Greifswald, 17475, Germany.

26 ¹⁰Department of Medicine A, University Medicine Greifswald, Greifswald, 17475, Germany.

27 ¹¹Department of Medicine II, University Hospital, LMU Munich, Munich, 80539, Germany.

28 ¹²Bristol Dental School, University of Bristol, Bristol, United Kingdom.

29

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

33

34 **Abstract**

35 Circulating lipids have been associated with breast cancer (BCa). This association may, in
36 part, be due to an effect of lipids on insulin-like growth factors (IGFs), which have been
37 reliably associated with BCa. In two-sample Mendelian randomization (MR) analyses, we
38 found that low density lipoprotein (LDL-C) was associated with IGFBP-3 (beta:0.08 SD;
39 95%CI:0.02,0.15; $p=0.01$, per SD increase in LDL-C) and IGFBP-3 was associated with
40 postmenopausal BCa (OR:1.09; 95%CI:1.00,1.19; $p=0.05$, per SD increase in IGFBP-3). We
41 also found that triglycerides were associated with IGF-I (beta:-0.13SD; 95%CI:-0.22,-0.03, per
42 SD increase in triglycerides) and that IGF-I was associated with overall BCa
43 (OR:1.10;95%CI:1.02,1.18, per SD increase in IGF-I). Taken together, these results suggest
44 that IGFBP-3 may be a potential causal step between LDL-C and postmenopausal BCa and
45 IGF-I a potential causal for triglycerides. Our two-step MR results build on evidence linking
46 circulating lipids and IGFs with BCa, however, multivariable MR analyses are currently unable
47 to support this relationship due to weak instruments.

48

49 **Key words:** Breast cancer, Mendelian randomization, Insulin-like growth factors, lipids

50

51 **Main text words:**4281; **Abstract:**157; **References:**61; **Tables & Figures:** 10;

52 **Supplementary files:** 3; **Conflicts of interest:** None to declare

53

54

55

56

57

58

59 **Introduction**

60 Breast cancer (BCa) is one of the leading causes of cancer-related deaths worldwide^{1,2}.
61 Circulating lipids including low density lipoprotein (LDL-C), high density lipoprotein (HDL-C)
62 and triglycerides have been associated with known risk factors for BCa such as obesity^{3,4}.
63 However, evidence from observational studies investigating the relationship between lipid
64 profile and BCa risk has been conflicting^{5,6}. Statins competitively inhibit 3-hydroxy-3-
65 methylglutaryl-coenzyme A (*HMGCR*) to reduce circulating LDL-C and are commonly
66 prescribed for the prevention and treatment of cardiovascular disease, but have also been
67 hypothesised to have anti-cancer effects⁷. Most observational studies have found no
68 evidence of an association between statin use and BCa risk⁸. However, reverse causation,
69 confounding and other biases may limit the interpretation of these observational findings. The
70 associations, estimated from these observational studies, between statin therapy and breast
71 cancer are particularly problematic, due to selection and so called immortal time biases⁹.

72
73 Mendelian randomization (MR) is an analytical approach where genetic variants reliably
74 associated with risk factors of interest are used as “proxy” measures to investigate whether
75 the effect of an exposure on an outcome is likely to be causal¹⁰. A recent MR study reported a
76 positive association between LDL-C and HDL-C with overall or estrogen receptor (ER)-
77 positive BCa risk¹¹. It is becoming increasingly feasible to study drug effects in a MR
78 framework using genetic variants located in the genes of the drug’s protein targets¹². Genetic
79 variants in *HMGCR*, associated with increased LDL-C levels, have been found to associate
80 with increased overall and ER-positive BCa risk^{13,14}. The mechanism by which statins protect
81 against BCa and whether the effect differs by menopausal status remains to be elucidated.

82
83 Several randomised control trials have suggested that statin use is associated with alterations
84 in circulating levels of insulin-like growth factors (IGFs), suggesting that perturbation of
85 circulating lipids can alter levels of IGF axis traits¹⁵⁻¹⁷. The IGF axis includes IGF-I, IGF-II and
86 six IGF binding proteins (IGFBPs), which are important for cancer initiation and progression
87 through the regulation of cell growth, metabolism and survival¹⁸⁻²¹. A recent meta-analysis of
88 individual data from observational studies found evidence that IGF-I was associated with

89 increased BCa risk in both pre and postmenopausal women while IGFBP-3 was only
90 associated with increased BCa risk in postmenopausal women²². However, it is unclear if a
91 causal relationship between circulating IGFs and BCa exists, and whether estrogen receptor
92 or menopausal status is important.

93

94 The role of the IGF axis in metabolic regulation is well-established²³⁻²⁵. Intervention studies in
95 patients with growth hormone disorders support a link between IGF and lipid biology²⁶⁻²⁸.
96 Population-based studies examining the relationship between circulating IGFs and lipid profile
97 have yielded conflicting results²⁹⁻³¹, however, these studies may be limited by their cross-
98 sectional design. Hence, the direction of association and whether causation exists between
99 circulating lipids and IGFs is still unclear. Given previous evidence implicating lipids and IGFs
100 as potential modifiable risk factors for cancer³²⁻³⁴, there is motivation to assess the
101 relationship between IGFs and lipids to investigate the mechanisms underlying BCa risk.

102

103 We aimed to investigate circulating IGFs as potential intermediates between lipid profile and
104 BCa risk. We set out to examine the association between circulating lipid traits (LDL-C, HDL-
105 C and TG), IGFs (IGF-I and IGFBP-3) and BCa (stratified by menopausal and ER status)
106 using a two-step MR study design³⁵.

107

108 **Materials and Methods**

109 **Study design and data sources**

110 To estimate the causal relationship between lipids and BCa using MR **Figure 1,(i)**, we
111 obtained summary statistics from: i) a lipid GWAS conducted by the Global Lipids Genetic
112 Consortium (GLGC, n = 188,577)³⁶; ii) a BCa risk GWAS (overall (n=228,951) and stratified
113 by estrogen receptor status (n= 69,501 for ER positive BCa and n=21,468 for ER negative
114 BCa) conducted by the Breast Cancer Association Consortium (BCAC,³⁷; and iii) a BCa risk
115 GWAS (overall (n=230,954) and stratified by menopausal status (premenopausal BCa,
116 n=59,124 for premenopausal BCa and n=112,245 for postmenopausal BCa) conducted in UK
117 Biobank (see **Supplementary methods**).

118

119 To investigate the observational relationship between lipids and IGFs (**Figure 1, (ii)**), we used
120 individual-level data ($n = 3,812$) from the Study of Health in Pomerania (SHIP) where
121 circulating lipids and IGFs were measured (these data are described in detail in the
122 **Supplementary methods**).

123

124 To estimate the causal relationship between lipids and IGF-I or IGFBP-3 using MR (**Figure 1,**
125 (iii)), we obtained summary statistics from: i) the lipid GWAS conducted by the GLGC³⁶, as
126 described above and ii) the IGF GWAS conducted by the “IGF working group” of the
127 CHARGE consortium ($n = 30,884$)³⁸ (see **Supplementary methods**).

128

129 To estimate the causal relationship between circulating IGFs and BCa (**Figure 1,(iv)**), we
130 obtained summary statistics from: i) the IGF GWAS(29); ii) the BCa risk GWAS conducted by
131 BCAC³⁷ (overall and stratified by ER status); and iii) the BCa risk GWAS conducted in UKBB
132 (overall and stratified by menopausal status)(see **Supplementary methods**).

133

134 To assess the potential causal effect of changes in LDL-C levels due to lipid lowering drugs
135 such as statins on BCa and IGF levels (**Figure 1,(v)**), we constructed a genetic instrument to
136 proxy statin treatment which comprised of 5 SNPs (rs12916, rs10515198, rs12173076,
137 rs3857388, and rs7711235) within 100kb from the gene region for *HMGCR* and in low linkage
138 disequilibrium ($r^2 < 0.20$)³⁹. These SNPs were robustly associated ($p < 5 \times 10^{-8}$) with LDL-C in the
139 lipid GWAS conducted by GLGC³⁶. We obtained summary statistics for the genetic variants in
140 *HMGCR* from: i) the lipid GWAS [29]; ii) the IGF GWAS(29); iii) the BCa risk GWAS
141 conducted by BCAC³⁷; and iv) the BCa risk GWAS conducted in UKBB (see **Supplementary**
142 **methods**).

143

144 **Genetic instruments for MR analyses**

145 The GLGC identified 185 SNPs ($r^2 < 0.2$) associated with at least HDL-C, LDL-C or TG
146 ($p < 5 \times 10^{-8}$)³⁶ (see **Supplementary Table 1** for list of SNPs). Due to the complex overlapping
147 nature of the lipid traits, genetic variants are commonly associated with more than one lipid
148 trait. To go some way to disentangling the roles of LDL-C, HDL-C and TG, we used

149 multivariable MR which was developed to estimate the causal effect of various correlated risk
150 factors when conditioned on one another in a single model^{40,41} (see **Supplementary**
151 **methods** for more details). For the multivariable MR analyses (**Figure 1**, (i)), all 185 SNPs
152 were included in the model. Univariable MR methods included SNPs associated with a given
153 lipid trait at a genome-wide level of significance ($p < 5 \times 10^{-8}$, 76 for LDL-C, 86 for HDL-C or 51
154 for TG). Sex-specific GWAS results were not available for lipid traits.

155

156 For the IGF MR analyses (**Figure 1**, (iii)&(iv)), we used 7 and 4 independent SNPs ($r^2 < 0.01$),
157 associated ($p < 5 \times 10^{-8}$) with IGF-I and IGFBP-3, respectively, as identified by a IGFs GWAS
158 (see **Supplementary Table 2** for list of SNPs)³⁸. IGF-II and IGFBP-2 were not investigated as
159 GWAS summary results are not currently available for these traits. For IGFBP-3, marked
160 differences in effect sizes by sex were found for 2 variants (with stronger associations in
161 women). As differences in effect size could potentially bias MR estimates⁴², analyses
162 estimating the causal relationship between IGF-I or IGFBP-3 and BCa were conducted using
163 female-specific effect estimates for IGF-I and IGFBP-3.

164

165 **Statistical analyses**

166 Multivariable MR was used to estimate the effect of LDL-C, HDL-C and TG on the odds of
167 BCa when conditioned on one another (**Figure 1**, (i)). As there were likely to be differences
168 between the causal estimates between BCAC and UKBB, a random-effects meta-analysis
169 was used to combine MR estimates using BCAC and UKBB summary statistics using the
170 'metan' package in Stata. Analyses stratified by ER status (in BCAC) and menopausal status
171 (in UKBB) were also conducted.

172

173 Observational associations between lipids and IGF-I or IGFBP-3 were assessed using linear
174 regression (**Figure 1**, (ii)). Fully adjusted models included age, sex, smoking status, body
175 mass index (BMI) and diabetes status. Associations of lipids and IGFs with potential
176 confounders were estimated using linear regression.

177

178 Bidirectional MR was performed to assess the causal effect of lipids on IGF-I and IGFBP-3
179 and vice versa (**Figure 1**, (iii)). SNP-exposure and SNP-outcome associations were combined
180 using the inverse variance weighted (IVW) method⁴³. MR-Egger⁴⁴, weighted median⁴⁵ and
181 multivariable MR^{40,41} were performed as sensitivity analyses.

182

183 To assess the causal effect of IGFs on BCa, SNP-exposure (SNP-IGF) and SNP-outcome
184 (SNP-BCa) associations were combined using inverse variance weighting (**Figure 1**, (iv))⁴³.
185 MR-Egger and weighted median were not conducted as these lack power for instruments with
186 a small number of SNPs⁴⁶. SNPs associated with IGF-I are also associated with other
187 components of the IGF pathway³⁸. To attempt to account for the overlapping associations for
188 these SNPs, we employed multivariable MR by regression of the SNP-BCa estimates on
189 SNP-IGF-I and SNP-IGFBP-3 estimates in a multivariable weighted regression model. A
190 random-effects meta-analysis was used to combine BCAC and UKBB estimates. Analyses
191 stratified by ER status (in BCAC) and menopausal status (in UKBB) were also conducted.

192

193 Multivariable MR⁴⁷ was conducted as an extension of the IVW method to test the hypothesis
194 that circulating IGF-I or IGFBP3 act as intermediate factors in any association between lipid
195 profile and BCa, identified by the analyses outlined above (**Figure 1**, (i-iv)). Two-sample
196 Conditional F-statistics⁴⁸ were estimated to predict instrument strength in the multivariable
197 model.

198

199 To estimate the causal relationship between statin use and BCa or IGF levels, we used a
200 random-effect linear regression model weighted by the inverse variance of the SNP-BCa
201 associations. The genetic correlation between the 5 SNPs in *HMGR* was incorporated into
202 the MR model using the weighted generalized linear regression method implemented in the
203 'MendelianRandomization' R package^{49,50}.

204

205 All MR analyses were performed using the MR-Base "TwoSampleMR" package⁵¹. All other
206 statistical analyses were performed using Stata version 14 (StataCorp, College Station,
207 Texas, USA) or R version 3.2.4.

208

209

210 **Results**

211 Re-assessing the causal relationship between circulating lipids with breast cancer

212 The relationship between circulating lipids and BCa in BCAC and UKBB was estimated using
213 multivariable MR. In BCAC, the odds ratio (OR) for overall BCa per SD increase in LDL-C,
214 HDL-C and TG was 1.06 (95%CI 1.01, 1.12; $p=0.02$), 1.07 (95%CI 1.01, 1.14; $p=0.03$) and
215 0.93 (95%CI: 0.86, 1.00; $p=0.06$), respectively (**Figure 2**). In UKBB, MR estimates of the
216 association between LDL-C, HDL-C, and TG and BCa were consistent with results in BCAC
217 (**Figure 2**). In a meta-analysis of BCAC and UKBB data, the OR for overall BCa per SD
218 increase in LDL-C, HDL-C and TG was 1.07 (95%CI 1.02, 1.12; $p=0.01$), 1.06 (95%CI 1.01,
219 1.10; $p=0.01$) and 0.94 (95%CI 0.89, 1.00; $p=0.05$), respectively (**Figure 2**).

220

221 Using data from BCAC, we extended this to the investigation of the relationship between
222 circulating lipids and BCa, stratified by ER status. The OR for ER-positive BCa per SD
223 increase in LDL-C, HDL-C and TG was 1.06 (95%CI 0.99, 1.12; $p=0.07$), 1.08 (95%CI 1.01,
224 1.15; $p=0.04$) and 0.93 (95%CI 0.86, 1.02; $p=0.11$), respectively. The OR for ER-negative
225 BCa per SD increase in LDL-C, HDL-C and TG was 1.05 (95%CI 0.98, 1.13; $p=0.16$), 1.08
226 (95%CI 0.99, 1.18; $p=0.09$) and 0.92 (95%CI 0.83, 1.02; $p=0.11$), respectively (**Table 1**).

227

228 Using data from UKBB, we investigated the relationship between circulating lipids and BCa,
229 stratified by menopausal status. We found evidence that LDL-C was associated with
230 increased odds of postmenopausal BCa (OR 1.14, 95%CI 1.04, 1.25; $p=0.004$), but there was
231 weak evidence for premenopausal BCa (OR 1.07, 95%CI 0.91, 1.26; $p=0.42$). There was little
232 evidence to suggest that HDL-C or TG associated with either premenopausal or
233 postmenopausal BCa (**Table 1**).

234

235 Observational relationship between circulating lipids and levels of IGF-I and IGFBP-3 in the

236 SHIP study

237 Study characteristics of the SHIP study are shown in **Supplementary Table 3**. The mean
238 (SD) IGF-I and IGFBP-3 levels were 142.1 (57.6) ng/ml and 1884.9 (503.5) ng/ml,
239 respectively. The mean (SD) LDL-C, HDL-C and TG levels were 3.57 (1.16) mmol/l, 1.45
240 (0.44) mmol/l and 1.82 (1.30) mmol/l, respectively. For IGF-I, a SD unit increase in LDL-C,
241 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 -
242 0.01, 0.05; $p=0.29$) and -0.16 (95%CI -0.19, -0.13; $p=2.26\times 10^{-22}$) SD unit change in IGF-I
243 levels, respectively. For IGFBP-3, a SD unit increase in LDL-C, HDL-C and TG was
244 associated with a -0.01 (95%CI -0.04, 0.02; $p=0.47$), 0.06 (95%CI 0.03, 0.10; $p=7.01\times 10^{-5}$)
245 and 0.03 (95%CI -0.003, 0.06; $p=0.08$) SD unit change in IGFBP-3 levels, respectively.
246 Circulating lipids and IGF traits were associated with potential confounders of a lipid- or IGF-
247 BCa relationship, including age, sex, smoking status, diabetes status and body mass index
248 (**Supplementary Table 4**). In the fully adjusted model, a SD unit increase in LDL-C, HDL-C
249 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;
250 $p=0.001$) and -0.06 (95%CI -0.09, -0.04; $p=1.5\times 10^{-5}$) SD unit change in IGF-I levels,
251 respectively. For IGFBP-3, a SD unit increase in LDL-C, HDL-C and TG was associated with
252 a 0.08 (95%CI 0.05, 0.11; $p=1.12\times 10^{-7}$), -0.02 (95%CI -0.05, 0.02; $p=0.30$) and 0.13 (95%CI
253 0.10, 0.16; $p=2.69\times 10^{-15}$) SD unit change in IGFBP-3 levels, respectively (**Table 2**).

254

255 Estimating the causal relationship between circulating lipids and IGFs

256 There was little evidence to suggest that LDL-C or HDL-C alter IGF-I levels. The MR estimate
257 of the causal relationship between TG and IGF-I was -0.13 SD units (95%CI -0.22, -0.03;
258 $p=0.01$) per SD unit increase in TG. Estimates obtained using MR-Egger regression,
259 weighted median and multivariable MR methods were of similar magnitude (**Table 3**). There
260 was little evidence for directional pleiotropy using Egger regression (Intercept: 0.007; se:
261 0.004; directionality $p=0.11$). In the reverse direction, there was little evidence to suggest that
262 IGF-I altered circulating LDL-C, HDL-C or TG (**Table 4**).

263

264 The MR estimate of the causal relationship between LDL-C and IGFBP-3 was 0.08 SD units
265 (95%CI 0.02, 0.15; $p=0.01$) per SD unit increase in LDL. There was little evidence for an
266 association between HDL-C and IGFBP-3 (**Table 3**). The IVW estimate of the causal

267 relationship between TG and IGFBP-3 was 0.10 SD units (95%CI 0.01, 0.19; $p=0.03$) per SD
268 unit increase in TG. Estimates obtained using MR-Egger regression, weighted median and
269 multivariable MR methods were consistent with the IVW analyses for LDL-C and TG (**Table**
270 **3**). In the reverse direction, there was minimal evidence to suggest that IGFBP-3 altered
271 circulating LDL-C, HDL-C or TG (**Table 4**).

272

273 Estimating the causal relationship between circulating IGFs and breast cancer

274 Using data from BCAC, the OR for overall BCa per SD increase in IGF-I and IGFBP-3 (using
275 the IVW method) was 1.10 (95%CI 1.01, 1.20; $p=0.02$) and 1.00 (95%CI 0.96, 1.03; $p=0.95$),
276 respectively. Estimates from multivariable MR analyses were consistent with the IVW
277 analyses unadjusted for IGF-I (OR 1.12, 95%CI 1.03, 1.20; $p=0.01$) and IGFBP-3 (OR 0.98,
278 95%CI 0.95, 1.02; $p=0.40$). In UKBB, estimates for the effect of IGF-I on BCa were
279 comparable. For IGFBP-3, the OR for overall BCa per SD increase in IGFBP-3 was 1.08
280 (95%CI 1.01, 1.15; $p=0.03$) and 1.08 (95%CI 1.01, 1.16; $p=0.02$), estimated by IVW and
281 multivariable MR, respectively (**Table 5**). When results from BCAC and UKBB were meta-
282 analysed, the OR for overall BCa for IGF-I and IGFBP-3 was 1.10 (95%CI 1.02, 1.18; $p=0.01$)
283 and 1.02 (95%CI 0.99, 1.06; $p=0.21$), respectively, from IVW analyses (**Figure 3**).

284

285 We also investigated the relationship between IGF traits and BCa by ER-status. The OR for
286 ER-positive BCa per SD increase in IGF-I and IGFBP-3 (using the IVW method) was 1.10
287 (95%CI: 0.99, 1.20; $p=0.06$) and 1.00 (95%CI: 0.96, 1.04; $p=0.94$), respectively. The OR for
288 ER-negative BCa per SD increase in IGF-I and IGFBP-3 (using the IVW method) was 1.05
289 (95%CI: 0.90, 1.23; $p=0.53$) and 0.98 (95%CI: 0.92, 1.04; $p=0.43$), respectively. For IGF-I and
290 IGFBP-3, multivariable MR and IVW estimates were consistent for both ER-positive and ER-
291 negative BCa (**Table 5**).

292

293 Using data from UKBB, we investigated the relationship between IGF traits and BCa stratified
294 by menopausal status. The OR for postmenopausal BCa per SD increase in IGF-I and
295 IGFBP-3 (using the IVW method) was 0.97 (95%CI 0.78, 1.20; $p=0.77$) and 1.09 (95%CI
296 1.00, 1.19; $p=0.05$), respectively. The OR for premenopausal BCa per SD increase in IGF-I

297 and IGFBP-3 (using the IVW method) was 1.19 (95%CI 0.76, 1.87; $p=0.46$) and 0.93 (95%CI
298 0.77, 1.12; $p=0.45$), respectively. For IGF-I and IGFBP-3, multivariable MR and IVW
299 estimates were consistent for both pre- and post-menopausal BCa (**Table 5**).

300 Multivariable MR analyses to estimate whether IGFs are on the causal pathway between
301 lipids and BCa

302 As a sensitivity analysis, MVMR analyses were conducted to investigate if the effect of TG on
303 overall BCa was attenuated following adjustment for IGF-I. Using data from BCAC, the
304 multivariable MR OR for overall BCa per SD increase in TG, conditioned on IGF-I, was 0.92
305 (95%CI 0.86, 1.10; $p=0.03$). In UKBB, estimates for the effect of TG on overall BCa after
306 conditioning on IGF-I were comparable. When results from BCAC and UKBB were meta-
307 analysed, the multivariable MR OR for overall BCa per SD increase in TG, conditioned on
308 IGF-I, was 0.93 (95%CI 0.88, 0.99; $p=0.03$), which did not differ in size or magnitude when
309 compared to the IGF-I unadjusted model (OR 0.94; 95%CI 0.89, 1.00; $p=0.05$) (**Table 6**).

310

311 Using data from UKBB, MVMR analyses were conducted to investigate if the effect of LDL-C
312 on post-menopausal BCa was attenuated following adjustment for IGFBP-3. The multivariable
313 OR for postmenopausal BCa per SD increase in LDL-C, conditioned on IGFBP-3 was 1.11
314 (95%CI 1.02, 1.20; $p=0.02$), which did not differ by size and direction of effect when
315 compared to the IGFBP-3 unadjusted model (OR 1.13; 95%CI 1.04, 1.25; $p=0.004$) (**Table 6**).

316

317 Estimating the causal relationship between LDL-C raising genetic variants in *HMGCR* with
318 BCa and IGF levels.

319 LDL-C raising variants in *HMGCR* (which simulate activation of *HMGCR*) were associated
320 with increased odds of BCa in BCAC (OR 1.20, 95%CI 1.02, 1.40; $p=0.03$). In UKBB, the
321 estimated causal effect of LDL-C raising variants in *HMGCR* was consistent but less precise
322 (OR 1.34, 95%CI 0.90, 2.02; $p=0.15$). The meta-analysed OR for overall BCa in BCAC and
323 UKBB was 1.22 (95%CI 1.04, 1.43; $p=0.01$) (**Table 7**). Using data from BCAC, LDL-C raising
324 variants in *HMGCR* were associated with increased odds of ER-positive BCa (OR 1.21,
325 95%CI 1.00, 1.47; $p=0.05$) but not ER negative BCa (OR 1.20, 95%CI 0.90, 1.60; $p=0.23$).

326

327 Using data from UKBB, LDL-C raising variants in *HMGCR* were associated with increased
328 odds of postmenopausal BCa (OR 1.63, 95%CI 1.06, 2.49; $p=0.03$), but not premenopausal
329 BCa (OR 1.47, 95%CI 0.57, 3.75; $p=0.42$) (**Table 7**). There was weak evidence that statins
330 could reduce IGF-I (beta: 0.02, 95%CI -0.02, 0.43, $p=0.07$). There was little evidence that
331 genetic variants in *HMGCR* were associated with IGFBP-3 (**Table 7**).

332

333 **Discussion**

334 In this study, we aimed to estimate the effect of circulating lipids and IGF traits on BCa and to
335 investigate the role of circulating IGFs in the association between circulating lipids and BCa.
336 We used complementary approaches to investigate circulating IGFs as causal intermediates
337 between lipids and BCa. We found that TG was inversely associated with IGF-I and overall
338 BCa and that IGF-I is positively associated with overall BCa. We also found that LDL-C was
339 positively associated with IGFBP-3 and postmenopausal BCa and that IGFBP-3 is positively
340 associated with postmenopausal BCa. Two-step MR results supported a hypothetical
341 pathway between TG, IGF-I and overall BCa and also between LDL-C, IGFBP-3 and
342 postmenopausal BCa.

343

344 A recent two-sample MR study assessed the relationship between circulating lipids and BCa
345 in BCAC¹¹. In an attempt to account for pleiotropic effects, the authors conducted sensitivity
346 analyses by excluding pleiotropic variants. However, manual pruning of SNPs that are
347 considered pleiotropic may result in an instrument that does not account for the underlying
348 genetic architecture of the trait, resulting in findings that can be challenging to interpret⁵².
349 Given that genetic variants included in the model are associated with multiple lipid fractions,
350 the MR assumption that all genetic variants are valid instruments (i.e. not pleiotropic) is
351 unlikely to hold. In this investigation, we have employed a multivariable MR approach, which
352 allows genetic variants to have pleiotropic effects on other variables included in the model
353 and jointly estimates the independent causal effect of each risk factor on the outcome. Our
354 multivariable MR analysis found that TG is negatively associated with overall BCa, which was
355 not observed in previous MR studies^{11,14}. Care is needed in interpretation however as whilst
356 these may reflect real differences, they could also be because of differences in analytical

357 complexities. For example, MVMR model estimates the direct effect of each lipid conditional
358 on the other lipid traits which differ from the total effect of the lipids as estimated by
359 univariable MR analyses in the absence of pleiotropy.

360

361 Epidemiological evidence supports a relationship between lipids and IGFs¹⁵⁻¹⁷. In the context
362 of this relationship, we set out to assess the relationship between circulating lipids and IGFs
363 using a combination of observational and MR methods. Results from observational analyses
364 and MR analyses highlight that TG decreases IGF-I levels and that TG and LDL-C increases
365 IGFBP-3 levels. There was minimal evidence to suggest that circulating IGFs causally affect
366 lipid levels, although this finding may be limited by the availability of robust genetic
367 instruments for individual IGF traits⁵³.

368

369 A meta-analysis of 17 studies and a case-control study showed that circulating IGF-I
370 increases odds of ER-positive BCa in pre- and postmenopausal women but not odds of ER-
371 negative BCa^{22,54}. This finding is concordant with results of our study, suggesting that
372 circulating IGF-I increases overall BCa, likely driven by ER-positive BCa. Our MR results
373 suggest that circulating IGF-I is more influential in ER-positive BCa. However, preclinical
374 evidence suggests that IGF signalling is mitogenic for both ER-positive and ER-negative
375 BCa⁵⁵. As we have used genetic variants that predict circulating levels of IGF-I, we cannot
376 rule out important aspects of tissue-specific regulation which may contribute to BCa.

377

378 Menopause causes substantial disturbances to lipid metabolism⁵⁶. Observational evidence
379 suggests that the association between lipids or IGFs and BCa depends on menopausal
380 status^{5,22}. Using data from UKBB, our MR results suggest that circulating LDL-C is more
381 influential in postmenopausal BCa compared to premenopausal BCa. In contrast to previous
382 studies⁵⁷⁻⁶⁰, the association between HDL-C and BCa did not differ by menopausal status in
383 our analysis. IGFBP-3 appeared to be more influential in post-menopausal BCa, consistent
384 with findings from a meta-analysis of observational studies²². However, confidence intervals
385 were overlapping for both LDL-C and IGFBP-3 estimates in premenopausal and
386 postmenopausal BCa analyses and the difference could be due to low statistical power in

387 subgroup analyses, in particular the small number of cases in the premenopausal analysis (n
388 = 1,781).

389

390 Our two-step MR results suggested a pathway between TG, IGF-I and overall BCa and a
391 pathway between LDL-C, IGFBP-3 and postmenopausal BCa. However, multivariable MR
392 analyses, wherein either IGF-I or IGFBP-3 were included as a covariate in the model, did not
393 attenuate the associations between TG and BCa or the associations between LDL-C and
394 postmenopausal BCa. This is counter to our interpretation of results, however our main
395 concern in this sensitivity analysis relates to instrument strength and the power of this
396 analysis. Instrument strength in a MVMR model can be assessed using a conditional F
397 statistic that summarises the relationship between genetic instruments and each exposure
398 conditional on the other exposures included in the MVMR model⁴⁸. We assessed the
399 instrument strength in the multivariable MR models and found that the conditional F statistics
400 for IGF-I and IGFBP-3 were low (**Table 6**). In addition, assessment of the Q-statistic for
401 heterogeneity found evidence of potential pleiotropy in the MVMR model (**Table 6**), however,
402 this could be due to presence of conditionally weak instruments for the IGF measures.

403

404 Given that our study found evidence that LDL-C increases both BCa and IGFBP-3 levels, it
405 was appropriate to assess how statins (the most commonly prescribed LDL-lowering therapy)
406 impact BCa and IGFBP-3 levels using LDL-C raising genetic variants in *HMGCR*. Increased
407 LDL-C as proxied by genetic variants in *HMGCR* was associated with increased overall BCa,
408 consistent with results from a previous MR study¹³. In our analysis the effect estimate for
409 overall BCa using the LDL-C raising genetic variants in *HMGCR* was larger than the effect
410 estimate using the 185 SNP genome-wide LDL-C score. It seems likely that the difference in
411 the results could reflect properties of the instruments employed. The larger number of genetic
412 variants used to proxy LDL-C levels could capture pathways beyond *HMGCR*, therefore
413 pleiotropy could influence the effect estimate. As the 5 SNPs in the *HMGCR* used in the statin
414 analysis are weakly correlated, the multiplicity of associations due to LD between SNPs could
415 bias the MR results⁶¹. However, we attempted to account for the LD between the SNPs in the
416 *HMGCR* instrument in our MR analyses by accounting for these^{49,50}. Differences in the results

417 using the *HMGCR* score and 185 SNP genome-wide LDL-C score could also be due to off-
418 target effects of statins (i.e. not via LDL-C). For example, a recent study by Kar et al¹³
419 suggested that BMI, rather than LDL-C, could be the true underlying mechanism responsible
420 for the protective association between genetic inhibition of *HMGCR* and BCa.

421

422 Randomised control trials have shown that statin use results in altered levels of IGF-I and
423 IGFBP-3¹⁵⁻¹⁷. Our study found weak evidence of an association between genetic variants in
424 *HMGCR* and IGF-I levels and minimal evidence of an association with IGFBP-3 levels. Failure
425 to detect an association between statins and IGFBP3 may be due to the relatively small size of
426 the IGF GWAS.

427

428 **Conclusion**

429 This study employed complementary observational and MR analyses to confirm that LDL-C
430 and HDL-C are positively associated with overall BCa and that TG is negatively associated
431 with overall BCa. Our study found that TG is negatively associated with IGF-I and that IGF-I is
432 positively associated with overall BCa, consistent with the TG-BCa association. We also
433 found that LDL-C is positively associated with IGFBP-3 and that IGFBP-3 is positively
434 associated with postmenopausal BCa, consistent with the LDL-C-postmenopausal BCa
435 association. Our two-step MR results build on evidence linking circulating lipids and IGFs with
436 BCa, however, multivariable MR analyses are currently unable to support this relationship due
437 to presence of conditionally weak instruments for the IGF measures.

438

439 **Funding**

440 This work was specifically supported by the Medical Research Council (MRC) Integrative
441 Epidemiology Unit (IEU) (MC_UU_12013/3) and by a Cancer Research UK Programme Grant
442 [The Integrative Cancer Epidemiology Programme, ICEP] (C18281/A19169). VT, KMB, CMP,
443 JMPH, NJT are supported by ICEP (C18281/A19169). NJT is a Wellcome Trust Investigator
444 (202802/Z/16/Z) and works within the University of Bristol NIHR Biomedical Research Centre
445 (BRC). LJC is supported by NJT's Wellcome Trust Investigator grant (202802/Z/16/Z). NJT
446 and LJC work in the MRC IEU at the University of Bristol which is supported by the MRC

447 (MC_UU_00011) and the University of Bristol. TD received support from Wellcome (grant ref
448 201268/Z/16/Z) and is now funded by the NIHR as an Academic Clinical Fellow. QQ is
449 supported by a Scientist Development Award (K01HL129892) from the NHLBI. JR is
450 supported in part by the National Center for Advancing Translational Sciences, CTSI grant
451 UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease
452 Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes
453 Endocrinology Research Center. The views expressed are those of the authors and not
454 necessarily those of the NHS, the NIHR, or the Department of Health. This publication is the
455 work of the authors who will serve as guarantors for the contents of this paper.

456 **Acknowledgements**

457 Quality Control filtering of the UK Biobank data was conducted by R.Mitchell, G.Hemani,
458 T.Dudding, L.Paternoster as described in the published protocol
459 (doi:10.5523/bris.3074krb6t2frj29yh2b03x3wxj). SHIP is part of the Community Medicine
460 Research net of the University of Greifswald, Germany, which is funded by the Federal
461 Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the
462 Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-
463 West Pomerania, and the network 'Greifswald Approach to Individualized Medicine
464 (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A).

465

466 **Tables and Figures**

467 Figure 1. Flow diagram of study design.

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

470 Figure 3. Estimates of the causal relationship between circulating IGF traits and overall breast
471 cancer in BCAC and UK Biobank.

472 Table 1. Estimates of the causal relationship between lipids and breast cancer stratified by
473 ER status or menopausal status.

474 Table 2. Beta estimates of SD unit change in IGF-I and IGFBP-3 per SD unit increase in HDL-
475 C, LDL-C or TG based on observational analyses in SHIP.

476 Table 3. Beta estimates of SD unit change in IGF-I and IGFBP-3 per SD unit increase in HDL-

477 C, LDL-C or TG based on two-sample and multivariable Mendelian Randomisation analyses.
478 Table 4. Beta estimates of SD unit change in LDL-C, HDL-C and triglycerides per SD unit
479 increase in IGF-I or IGFBP-3 based on two-sample and multivariable Mendelian
480 randomization analyses.
481 Table 5. Estimates of the causal relationship between IGF-I or IGFBP-3 and breast cancer
482 risk stratified by ER or menopausal status.
483 Table 6. Multivariable MR analysis of the association of LDL-C and TG with breast cancer
484 conditioned on IGFs.
485
486 Table 7. Estimates of the causal relationship between LDL-C raising variants in HMGCR with
487 breast cancer or IGF levels.
488 Supplementary Table 1. Genetic instruments used to construct genetic risk scores for lipid
489 traits used in multivariable MR analyses
490 Supplementary Table 2. Genetic instruments used to construct genetic risk scores for IGF-I
491 and IGFBP-3
492 Supplementary Table 3. Study characteristics of the SHIP cohort
493 Supplementary Table 4. Association of IGF axis and lipids with potential confounders in the
494 SHIP study

495

496 **References**

- 497 1 Torre, L. A. *et al.* Global cancer statistics, 2012. *CA Cancer J Clin* **65**, 87-
498 108, doi:10.3322/caac.21262 (2015).
- 499 2 Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2012. *CA Cancer J*
500 *Clin* **62**, 10-29, doi:10.3322/caac.20138 (2012).
- 501 3 Garcia-Estevez, L. & Moreno-Bueno, G. Updating the role of obesity and
502 cholesterol in breast cancer. *Breast Cancer Res* **21**, 35,
503 doi:10.1186/s13058-019-1124-1 (2019).
- 504 4 Blucher, C. & Stadler, S. C. Obesity and Breast Cancer: Current Insights on
505 the Role of Fatty Acids and Lipid Metabolism in Promoting Breast Cancer
506 Growth and Progression. *Front Endocrinol (Lausanne)* **8**, 293,
507 doi:10.3389/fendo.2017.00293 (2017).
- 508 5 Touvier, M. *et al.* Cholesterol and breast cancer risk: a systematic review
509 and meta-analysis of prospective studies. *Br J Nutr* **114**, 347-357,
510 doi:10.1017/S000711451500183X (2015).
- 511 6 Ni, H., Liu, H. & Gao, R. Serum Lipids and Breast Cancer Risk: A Meta-
512 Analysis of Prospective Cohort Studies. *PLoS one* **10**, e0142669,
513 doi:10.1371/journal.pone.0142669 (2015).
- 514 7 Chan, K. K., Oza, A. M. & Siu, L. L. The statins as anticancer agents. *Clin*
515 *Cancer Res* **9**, 10-19 (2003).

- 516 8 Islam, M. M. *et al.* Exploring association between statin use and breast
517 cancer risk: an updated meta-analysis. *Arch Gynecol Obstet* **296**, 1043-
518 1053, doi:10.1007/s00404-017-4533-3 (2017).
- 519 9 Suissa, S. Immortal time bias in observational studies of drug effects.
520 *Pharmacoepidemiol Drug Saf* **16**, 241-249, doi:10.1002/pds.1357 (2007).
- 521 10 Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors
522 for causal inference in epidemiological studies. *Human molecular genetics*
523 **23**, R89-98, doi:10.1093/hmg/ddu328 (2014).
- 524 11 Nowak, C. & Arnlov, J. A Mendelian randomization study of the effects of
525 blood lipids on breast cancer risk. *Nature communications* **9**, 3957,
526 doi:10.1038/s41467-018-06467-9 (2018).
- 527 12 Walker, V. M., Davey Smith, G., Davies, N. M. & Martin, R. M. Mendelian
528 randomization: a novel approach for the prediction of adverse drug
529 events and drug repurposing opportunities. *International journal of*
530 *epidemiology* **46**, 2078-2089, doi:10.1093/ije/dyx207 (2017).
- 531 13 Kar, S. P. *et al.* Body mass index and the association between low-density
532 lipoprotein cholesterol as predicted by HMGCR genetic variants and
533 breast cancer risk. *International journal of epidemiology*,
534 doi:10.1093/ije/dyz047 (2019).
- 535 14 Orho-Melander, M. *et al.* Blood lipid genetic scores, the HMGCR gene and
536 cancer risk: a Mendelian randomization study. *International journal of*
537 *epidemiology* **47**, 495-505, doi:10.1093/ije/dyx237 (2018).
- 538 15 Bergen, K., Brismar, K. & Tehrani, S. High-dose atorvastatin is associated
539 with lower IGF-1 levels in patients with type 1 diabetes. *Growth hormone*
540 *& IGF research : official journal of the Growth Hormone Research Society*
541 *and the International IGF Research Society* **29**, 78-82,
542 doi:10.1016/j.ghir.2016.06.001 (2016).
- 543 16 Narayanan, R. P. *et al.* Atorvastatin administration is associated with
544 dose-related changes in IGF bioavailability. *European journal of*
545 *endocrinology* **168**, 543-548, doi:10.1530/EJE-12-0844 (2013).
- 546 17 Szkodzinski, J. *et al.* Effect of HMG-CoA (3-hydroxy-3-methylglutaryl-CoA)
547 reductase inhibitors on the concentration of insulin-like growth factor-1
548 (IGF-1) in hypercholesterolemic patients. *Pharmacological reports : PR* **61**,
549 654-664 (2009).
- 550 18 Perks, C. M. & Holly, J. M. IGF binding proteins (IGFBPs) and regulation of
551 breast cancer biology. *J Mammary Gland Biol Neoplasia* **13**, 455-469,
552 doi:10.1007/s10911-008-9106-4 (2008).
- 553 19 Perks, C. M. & Holly, J. M. Insulin-like growth factor binding proteins
554 (IGFBPs) in breast cancer. *J Mammary Gland Biol Neoplasia* **5**, 75-84
555 (2000).
- 556 20 Biernacka, K. M., Perks, C. M. & Holly, J. M. Role of the IGF axis in prostate
557 cancer. *Minerva Endocrinol* **37**, 173-185 (2012).
- 558 21 Pollak, M. The insulin and insulin-like growth factor receptor family in
559 neoplasia: an update. *Nat Rev Cancer* **12**, 159-169, doi:10.1038/nrc3215
560 (2012).
- 561 22 Endogenous, H. *et al.* Insulin-like growth factor 1 (IGF1), IGF binding
562 protein 3 (IGFBP3), and breast cancer risk: pooled individual data
563 analysis of 17 prospective studies. *Lancet Oncol* **11**, 530-542,
564 doi:10.1016/S1470-2045(10)70095-4 (2010).

- 565 23 Clemmons, D. R. Metabolic actions of insulin-like growth factor-I in
566 normal physiology and diabetes. *Endocrinology and metabolism clinics of*
567 *North America* **41**, 425-443, vii-viii, doi:10.1016/j.ecl.2012.04.017 (2012).
- 568 24 Kim, H. S. Role of insulin-like growth factor binding protein-3 in glucose
569 and lipid metabolism. *Ann Pediatr Endocrinol Metab* **18**, 9-12,
570 doi:10.6065/apem.2013.18.1.9 (2013).
- 571 25 Russo, V. C., Azar, W. J., Yau, S. W., Sabin, M. A. & Werther, G. A. IGFBP-2:
572 The dark horse in metabolism and cancer. *Cytokine Growth Factor Rev* **26**,
573 329-346, doi:10.1016/j.cytogfr.2014.12.001 (2015).
- 574 26 Mauras, N., Martinez, V., Rini, A. & Guevara-Aguirre, J. Recombinant
575 human insulin-like growth factor I has significant anabolic effects in
576 adults with growth hormone receptor deficiency: studies on protein,
577 glucose, and lipid metabolism. *J Clin Endocrinol Metab* **85**, 3036-3042,
578 doi:10.1210/jcem.85.9.6772 (2000).
- 579 27 Mauras, N. *et al.* Insulin-like growth factor I and growth hormone (GH)
580 treatment in GH-deficient humans: differential effects on protein, glucose,
581 lipid, and calcium metabolism. *J Clin Endocrinol Metab* **85**, 1686-1694,
582 doi:10.1210/jcem.85.4.6541 (2000).
- 583 28 Mauras, N. & Haymond, M. W. Metabolic effects of recombinant human
584 insulin-like growth factor-I in humans: comparison with recombinant
585 human growth hormone. *Pediatr Nephrol* **10**, 318-323 (1996).
- 586 29 Friedrich, N. *et al.* Cross-sectional and longitudinal associations between
587 insulin-like growth factor I and metabolic syndrome: a general population
588 study in German adults. *Diabetes/metabolism research and reviews* **29**,
589 452-462, doi:10.1002/dmrr.2412 (2013).
- 590 30 Lam, C. S. *et al.* Circulating insulin-like growth factor-1 and its binding
591 protein-3: metabolic and genetic correlates in the community.
592 *Arteriosclerosis, thrombosis, and vascular biology* **30**, 1479-1484,
593 doi:10.1161/ATVBAHA.110.203943 (2010).
- 594 31 Eggert, M. L. *et al.* Cross-sectional and longitudinal relation of IGF1 and
595 IGF-binding protein 3 with lipid metabolism. *European journal of*
596 *endocrinology* **171**, 9-19, doi:10.1530/EJE-13-1017 (2014).
- 597 32 Melvin, J. C., Holmberg, L., Rohrmann, S., Loda, M. & Van Hemelrijck, M.
598 Serum lipid profiles and cancer risk in the context of obesity: four meta-
599 analyses. *J Cancer Epidemiol* **2013**, 823849, doi:10.1155/2013/823849
600 (2013).
- 601 33 Chandler, P. D. *et al.* Lipid biomarkers and long-term risk of cancer in the
602 Women's Health Study. *Am J Clin Nutr* **103**, 1397-1407,
603 doi:10.3945/ajcn.115.124321 (2016).
- 604 34 Pollak, M. Insulin and insulin-like growth factor signalling in neoplasia.
605 *Nat Rev Cancer* **8**, 915-928, doi:10.1038/nrc2536 (2008).
- 606 35 Relton, C. L. & Davey Smith, G. Two-step epigenetic Mendelian
607 randomization: a strategy for establishing the causal role of epigenetic
608 processes in pathways to disease. *International journal of epidemiology*
609 **41**, 161-176, doi:10.1093/ije/dyr233 (2012).
- 610 36 Global Lipids Genetics, C. *et al.* Discovery and refinement of loci associated
611 with lipid levels. *Nature genetics* **45**, 1274-1283, doi:10.1038/ng.2797
612 (2013).

- 613 37 Michailidou, K. *et al.* Association analysis identifies 65 new breast cancer
614 risk loci. *Nature* **551**, 92-94, doi:10.1038/nature24284 (2017).
- 615 38 Teumer, A. *et al.* Genomewide meta-analysis identifies loci associated
616 with IGF-I and IGFBP-3 levels with impact on age-related traits. *Aging cell*
617 **15**, 811-824, doi:10.1111/accel.12490 (2016).
- 618 39 Yarmolinsky, J. *et al.* Association Between Genetically Proxied Inhibition
619 of HMG-CoA Reductase and Epithelial Ovarian Cancer. *JAMA* **323**, 646-
620 655, doi:10.1001/jama.2020.0150 (2020).
- 621 40 Burgess, S., Dudbridge, F. & Thompson, S. G. Re: "Multivariable Mendelian
622 randomization: the use of pleiotropic genetic variants to estimate causal
623 effects". *American journal of epidemiology* **181**, 290-291,
624 doi:10.1093/aje/kwv017 (2015).
- 625 41 Burgess, S. & Thompson, S. G. Multivariable Mendelian randomization: the
626 use of pleiotropic genetic variants to estimate causal effects. *American*
627 *journal of epidemiology* **181**, 251-260, doi:10.1093/aje/kwu283 (2015).
- 628 42 Lawlor, D. A. Commentary: Two-sample Mendelian randomization:
629 opportunities and challenges. *International journal of epidemiology* **45**,
630 908-915, doi:10.1093/ije/dyw127 (2016).
- 631 43 Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization
632 analysis with multiple genetic variants using summarized data. *Genetic*
633 *epidemiology* **37**, 658-665, doi:10.1002/gepi.21758 (2013).
- 634 44 Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with
635 invalid instruments: effect estimation and bias detection through Egger
636 regression. *International journal of epidemiology* **44**, 512-525,
637 doi:10.1093/ije/dyv080 (2015).
- 638 45 Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent
639 Estimation in Mendelian Randomization with Some Invalid Instruments
640 Using a Weighted Median Estimator. *Genetic epidemiology* **40**, 304-314,
641 doi:10.1002/gepi.21965 (2016).
- 642 46 Burgess, S. *et al.* Using published data in Mendelian randomization: a
643 blueprint for efficient identification of causal risk factors. *Eur J Epidemiol*
644 **30**, 543-552, doi:10.1007/s10654-015-0011-z (2015).
- 645 47 Sanderson, E., Davey Smith, G., Windmeijer, F. & Bowden, J. An
646 examination of multivariable Mendelian randomization in the single-
647 sample and two-sample summary data settings. *International journal of*
648 *epidemiology* **48**, 713-727, doi:10.1093/ije/dyy262 (2019).
- 649 48 Sanderson, E., Spiller, W., Bowden, J. . Testing and Correcting for Weak
650 and Pleiotropic Instruments in Two-Sample Multivariable Mendelian
651 Randomisation. *bioRxiv*, doi:<https://doi.org/10.1101/2020.04.02.021980>
652 (2020).
- 653 49 Yavorska, O. O. & Burgess, S. MendelianRandomization: an R package for
654 performing Mendelian randomization analyses using summarized data.
655 *International journal of epidemiology* **46**, 1734-1739,
656 doi:10.1093/ije/dyx034 (2017).
- 657 50 Burgess, S., Dudbridge, F. & Thompson, S. G. Combining information on
658 multiple instrumental variables in Mendelian randomization: comparison
659 of allele score and summarized data methods. *Stat Med* **35**, 1880-1906,
660 doi:10.1002/sim.6835 (2016).

- 661 51 Hemani, G. *et al.* The MR-Base platform supports systematic causal
662 inference across the human phenome. *Elife* **7**, doi:10.7554/eLife.34408
663 (2018).
- 664 52 Holmes, M. V., Ala-Korpela, M. & Smith, G. D. Mendelian randomization in
665 cardiometabolic disease: challenges in evaluating causality. *Nat Rev*
666 *Cardiol* **14**, 577-590, doi:10.1038/nrcardio.2017.78 (2017).
- 667 53 Bonilla, C. *et al.* Assessing the role of insulin-like growth factors and
668 binding proteins in prostate cancer using Mendelian randomization:
669 Genetic variants as instruments for circulating levels. *International*
670 *journal of cancer* **139**, 1520-1533, doi:10.1002/ijc.30206 (2016).
- 671 54 Kaaks, R. *et al.* Insulin-like growth factor I and risk of breast cancer by age
672 and hormone receptor status-A prospective study within the EPIC cohort.
673 *International journal of cancer* **134**, 2683-2690, doi:10.1002/ijc.28589
674 (2014).
- 675 55 Surmacz, E. & Bartucci, M. Role of estrogen receptor alpha in modulating
676 IGF-I receptor signaling and function in breast cancer. *J Exp Clin Cancer*
677 *Res* **23**, 385-394 (2004).
- 678 56 Auro, K. *et al.* A metabolic view on menopause and ageing. *Nature*
679 *communications* **5**, 4708, doi:10.1038/ncomms5708 (2014).
- 680 57 Agnoli, C. *et al.* Metabolic syndrome and postmenopausal breast cancer in
681 the ORDET cohort: a nested case-control study. *Nutr Metab Cardiovasc Dis*
682 **20**, 41-48, doi:10.1016/j.numecd.2009.02.006 (2010).
- 683 58 Furberg, A. S., Veierod, M. B., Wilsgaard, T., Bernstein, L. & Thune, I. Serum
684 high-density lipoprotein cholesterol, metabolic profile, and breast cancer
685 risk. *J Natl Cancer Inst* **96**, 1152-1160, doi:10.1093/jnci/djh216 (2004).
- 686 59 Kucharska-Newton, A. M. *et al.* HDL-cholesterol and incidence of breast
687 cancer in the ARIC cohort study. *Ann Epidemiol* **18**, 671-677,
688 doi:10.1016/j.annepidem.2008.06.006 (2008).
- 689 60 Moorman, P. G. *et al.* Association between high-density lipoprotein
690 cholesterol and breast cancer varies by menopausal status. *Cancer*
691 *epidemiology, biomarkers & prevention : a publication of the American*
692 *Association for Cancer Research, cosponsored by the American Society of*
693 *Preventive Oncology* **7**, 483-488 (1998).
- 694 61 Swerdlow, D. I. *et al.* Selecting instruments for Mendelian randomization
695 in the wake of genome-wide association studies. *International journal of*
696 *epidemiology* **45**, 1600-1616, doi:10.1093/ije/dyw088 (2016).
- 697

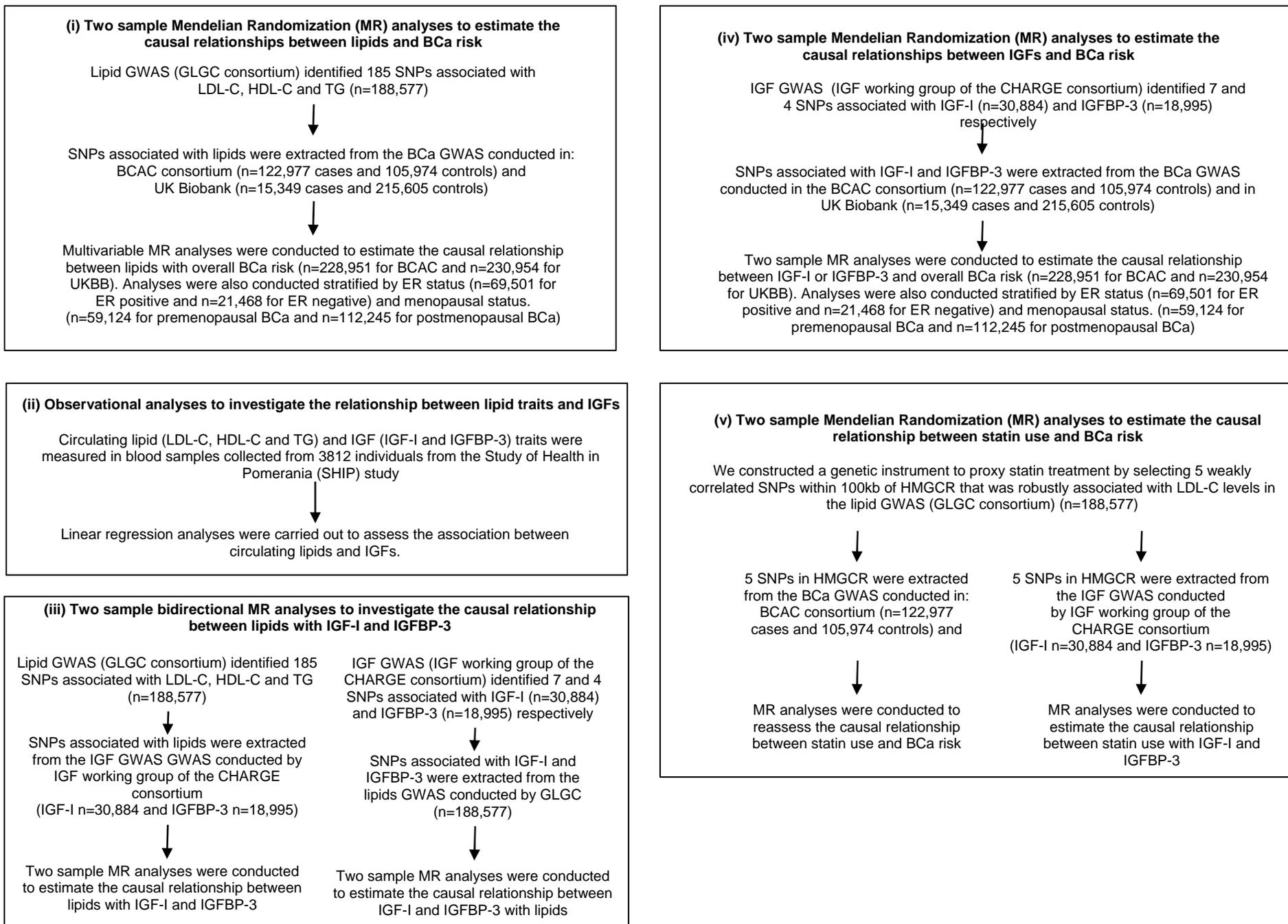


Figure 1. Flow diagram of study design.

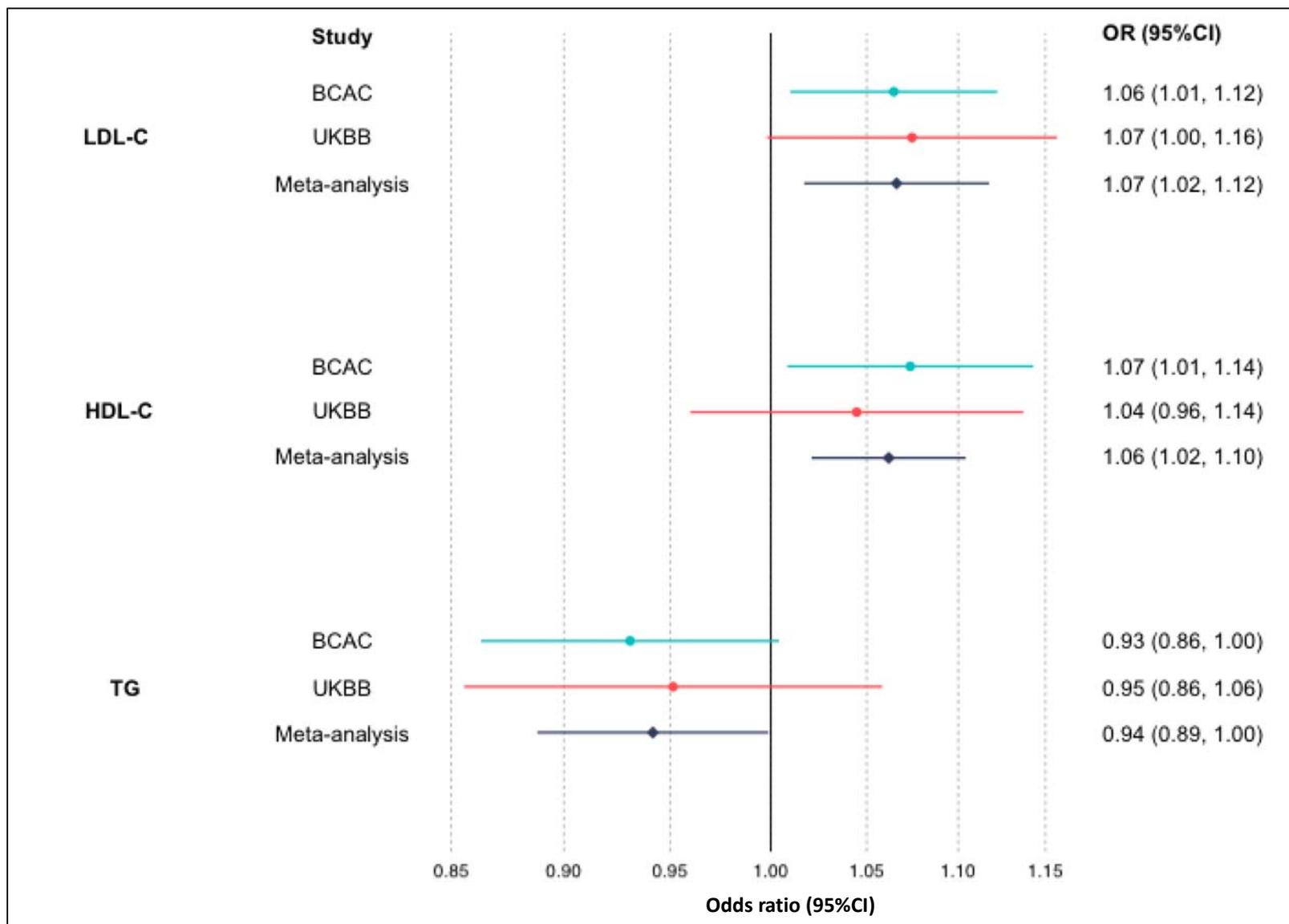


Figure 2. Estimates of the causal relationship between circulating lipid traits and breast cancer risk in BCAC and UK Biobank.

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=122,977 cases and 105,974 controls) and UK Biobank (UKBB) (n=15,349 cases and 215,605 controls). Squares represent point estimates from individual analyses. Horizontal lines represent the 95% confidence intervals. Pooled estimates (represented by diamonds) were obtained using random-effects meta-analysis.

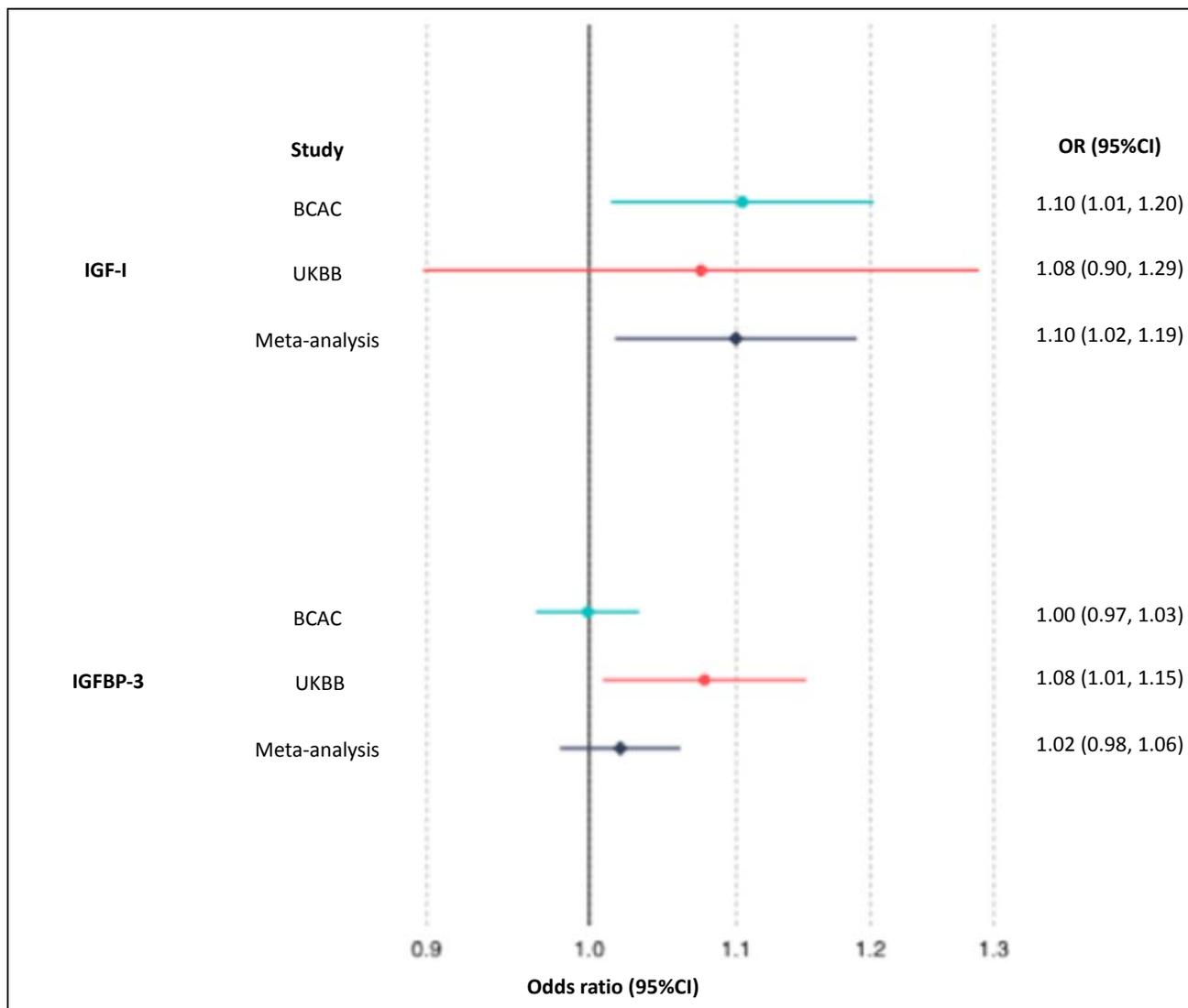


Figure 3. Estimates of the causal relationship between circulating IGF traits and breast cancer risk in BCAC and UK Biobank.

The forest plot shows the estimate of the causal effect of IGF-I and IGFBP-3 on breast cancer risk using summary data from the breast cancer association consortium (BCAC) (n=122,977 cases and 105,974 controls) and UK Biobank (UKBB) (n=15,349 cases and 215,605 controls). Squares represent point estimates from individual analyses. Horizontal lines represent the 95% confidence intervals. Pooled estimates (represented by diamonds) were obtained using random effects meta-analysis.

Table 1. Estimates of the causal relationship between lipids and breast cancer stratified by ER status or menopausal status.

Exp	Consortium	Outcome	N	OR ^a (95% CI)	p
LDL-C	BCAC	ER positive BCa	69,501	1.06 (0.99, 1.12)	0.07
	BCAC	ER negative BCa	21,468	1.05 (0.98, 1.13)	0.16
HDL-C	BCAC	ER positive BCa	69,501	1.08 (1.01, 1.15)	0.04
	BCAC	ER negative BCa	21,468	1.08 (0.99, 1.18)	0.09
TG	BCAC	ER positive BCa	69,501	0.93 (0.86, 1.02)	0.11
	BCAC	ER negative BCa	21,468	0.92 (0.83, 1.02)	0.11
LDL-C	UKBB	Premenopause BCa	59,124	1.07 (0.91, 1.26)	0.42
	UKBB	Postmenopause BCa	112,245	1.14 (1.04, 1.25)	0.004
HDL-C	UKBB	Premenopause BCa	59,124	1.07 (0.89, 1.30)	0.46
	UKBB	Postmenopause BCa	112,245	1.02 (0.93, 1.13)	0.64
TG	UKBB	Premenopause BCa	59,124	1.02 (0.81, 1.30)	0.84
	UKBB	Postmenopause BCa	112,245	0.91 (0.80, 1.03)	0.14

OR, odds ratio; 95% CI, 95% confidence interval; p, p-value; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; TG, triglycerides; MR, Mendelian randomization; BCa, Breast cancer; BCAC, breast cancer association consortium; UKBB, UK Biobank; ER, estrogen receptor.

^aAssociations are per 1 SD unit increase in LDL-C, HDL-C or TG from multivariable MR analyses

Table 2. Beta estimates of SD unit change in IGF-I and IGFBP-3 per SD unit increase in HDL-C, LDL-C or TG based on observational analyses in SHIP.

Exposure	Outcome	N	Unadjusted		Adjusted ^a	
			Beta ^b (95%CI)	p	Beta ^b (95%CI)	p
LDL-C	IGF-I	3812	-0.11 (-0.14, -0.08)	2x10 ⁻¹²	0.03 (0.004, 0.06)	0.03
HDL-C	IGF-I	3812	0.02 (-0.01, 0.05)	0.29	-0.05 (-0.08, -0.02)	0.001
TG	IGF-I	3812	-0.16 (-0.19, -0.13)	2x10 ⁻²²	-0.06 (-0.09, -0.04)	1.5x10 ⁻⁵
LDL-C	IGFBP-3	3812	-0.01 (-0.04, 0.02)	0.47	0.08 (0.05, 0.11)	1.12x10 ⁻⁷
HDL-C	IGFBP-3	3812	0.06 (0.03, 0.10)	7.01x10 ⁻⁵	-0.02 (-0.05, 0.02)	0.30
TG	IGFBP-3	3812	0.03 (-0.003, 0.06)	0.08	0.13 (0.10, 0.16)	2.69x10 ⁻¹⁵

N, sample size; Beta, regression coefficient; 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 1; IGFBP-3, insulin-like growth factor binding protein 3

^aAnalyses were adjusted for age, gender, smoking status, body mass index and diabetes status

^bBeta refers to the s.d unit change in IGF-I or IGFBP-3 levels per 1 s.d unit decrease in HDL-C, LDL-C or TG.

Table 3. Beta estimates of SD unit change in IGF-I and IGFBP-3 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					
		IVW Beta ^a (95% CI)	<i>p</i>	Weighted Median Beta ^a (95% CI)	<i>p</i>	MR Egger regression Beta ^a (95% CI)	<i>p</i>	Multivariable MR Beta ^a (95% CI)	<i>p</i>
LDL-C	IGF-I	0.01 (-0.04, 0.07)	0.64	0.02 (-0.05, 0.09)	0.64	0.001 (-0.09, 0.10)	0.99	0.04 (-0.02, 0.10)	0.21
HDL-C	IGF-I	0.01 (-0.05, 0.07)	0.80	-0.01 (-0.08, 0.07)	0.87	0.03 (-0.08, 0.13)	0.62	-0.02 (-0.08, 0.04)	0.56
TG	IGF-I	-0.13 (-0.22, -0.03)	0.01	-0.04 (-0.13, 0.06)	0.46	-0.23 (-0.39, -0.07)	0.01	-0.12 (-0.18, -0.06)	<0.001
LDL-C	IGFBP-3	0.08 (0.02, 0.15)	0.01	0.05 (-0.03, 0.12)	0.26	0.09 (-0.02, 0.20)	0.10	0.08 (0.02, 0.14)	0.01
HDL-C	IGFBP-3	-0.06 (-0.13, 0.01)	0.08	-0.05 (-0.17, 0.06)	0.36	-0.05 (-0.17, 0.06)	0.36	-0.02 (-0.08, 0.04)	0.61
TG	IGFBP-3	0.10 (0.01, 0.19)	0.03	0.08 (-0.03, 0.19)	0.14	0.13 (-0.02, 0.27)	0.10	0.06 (0.001, 0.12)	0.04

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

^aBeta refers to the s.d unit change in IGF-I or IGFBP-3 levels per 1 s.d 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 or IGFBP-3 based on two-sample and multivariable Mendelian randomization analyses.

Exposure	Outcome	Main analysis		Sensitivity analysis	
		IVW Beta ^a (95% CI)	p	Multivariable MR Beta ^a (95% CI)	p
IGF-I	LDL-C	-0.07 (-0.21, 0.08)	0.37	-0.04 (-0.18, 0.10)	0.55
IGF-I	HDL-C	-0.02 (-0.13, 0.08)	0.65	-0.01 (-0.11, 0.09)	0.77
IGF-I	TG	-0.39 (-1.00, 0.22)	0.21	-0.35 (-0.90, 0.20)	0.20
IGFBP-3	LDL-C	-0.01 (-0.04, 0.02)	0.60	-0.01 (-0.07, 0.05)	0.81
IGFBP-3	HDL-C	-0.002 (-0.02, 0.02)	0.87	-0.01 (-0.05, 0.03)	0.69
IGFBP-3	TG	-0.02 (-0.05, 0.02)	0.33	0.02 (-0.22, 0.26)	0.90

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

^aBeta refers to the s.d unit change in HDL-C, LDL-C, TG or TC per SD unit change in IGF-I and IGFBP-3 levels.

Table 5. Estimates of the causal relationship between IGF-I or IGFBP-3 and breast cancer stratified by ER or menopausal status.

Exposure	Consortium	Outcome	N	Main analysis		Sensitivity analysis	
				IVW		Multivariable MR	
				OR ^a (95%CI)	p	OR ^a (95%CI)	p
IGF-I	BCAC	Overall BCa	228,951	1.10 (1.01, 1.20)	0.02	1.12 (1.03, 1.20)	0.01
	BCAC	ER positive BCa	69,501	1.10 (0.99, 1.20)	0.06	1.12 (1.02, 1.22)	0.02
	BCAC	ER negative BCa	21,468	1.05 (0.90, 1.23)	0.53	1.04 (0.90, 1.21)	0.59
IGFBP-3	BCAC	Overall BCa	228,951	1.00 (0.97, 1.03)	0.95	0.98 (0.95, 1.02)	0.40
	BCAC	ER positive BCa	69,501	1.00 (0.96, 1.04)	0.94	0.99 (0.95, 1.03)	0.52
	BCAC	ER negative BCa	21,468	0.98 (0.92, 1.04)	0.43	0.97 (0.91, 1.03)	0.30
IGF-I	UKBB	Overall BCa	230,954	1.08 (0.90, 1.29)	0.43	1.02 (0.87, 1.19)	0.83
	UKBB	Premenopausal BCa	59,124	1.19 (0.76, 1.87)	0.46	1.27 (0.82, 1.99)	0.29
	UKBB	Postmenopausal BCa	112,245	0.97 (0.78, 1.20)	0.77	0.92 (0.75, 1.12)	0.40
IGFBP-3	UKBB	Overall BCa	230,954	1.08 (1.01, 1.15)	0.03	1.08 (1.01, 1.16)	0.02
	UKBB	Premenopausal BCa	59,124	0.93 (0.77, 1.12)	0.45	0.93 (0.76, 1.13)	0.45
	UKBB	Postmenopausal BCa	112,245	1.09 (1.00, 1.19)	0.05	1.11 (1.01, 1.21)	0.02

IVW, inverse variance weighted; OR, odds ratio; 95% CI, 95% confidence interval; p, p-value; se, standard error; MR, Mendelian randomization; UKBB, UK Biobank; IGF-I, insulin-like growth factor I; IGFBP-3, insulin-like growth factor binding protein 3.

^aAssociations are per 1 SD unit increase in IGF-I or IGFBP-3.

Table 6. Multivariable MR analysis of the association of LDL-C and TG with breast cancer conditioned on IGFs.

Exposure	Consortium	Outcome	MVMR Beta (95%CI)	<i>p</i>	F statistic for MVMR	Q statistic for heterogeneity
TG ^a IGF-I	BCAC	Overall BCa	0.92 (0.86, 1.10)	0.03	17.80 2.38	Q=579.42; <i>p</i> = 4.1x10 ⁻⁴⁶
TG ^a <i>p</i>	UKBB	Overall BCa	0.95 (0.86, 1.05)	0.25	31.22 2.42	Q=327.29; <i>p</i> =1.74x10 ⁻⁹
TG ^a IGF-I	Meta-analysis	Overall BCa	0.93 (0.88, 0.99)	0.03		
LDL-C ^b IGFBP-3		Postmenopausal BCa	1.11 (1.02, 1.20)	0.02	46.74 5.34	Q=274.59; <i>p</i> =4.71x10 ⁻⁵

OR, odds ratio; 95%CI, 95% confidence interval; *p*, *p*-value; LDL-C, low density lipoprotein cholesterol; MVMR, Multivariable Mendelian randomization; BCa, Breast cancer; IGFBP-3, insulin-like growth factor 3.

^aAssociations are per 1 SD unit increase in TG when conditioned on IGF-I.

^bAssociations are per 1 SD unit increase in LDL-C when conditioned on IGFBP-3.

Table 7. Estimates of the causal relationship between LDL-C raising variants in *HMGCR* with breast cancer or IGF levels.

Exposure	Consortium	Outcome	N	OR ^a (95% CI)	p	
LDL-C raising genetic variants in <i>HMGCR</i>	BCAC	Overall BCa	228,951	1.20 (1.02, 1.40)	0.03	
	UKBB	Overall BCa	230,954	1.34 (0.90, 2.02)	0.15	
	Meta-analysis	Overall BCa	459,905	1.22 (1.04, 1.43)	0.01	
	BCAC	ER positive BCa	69,501	1.21 (1.00, 1.47)	0.05	
	BCAC	ER negative BCa	21,468	1.20 (0.90, 1.60)	0.23	
	UKBB	Premenopausal BCa	59,124	1.47 (0.57, 3.75)	0.42	
	UKBB	Postmenopausal BCa	112,245	1.63 (1.06, 2.49)	0.03	
	IGF working group of CHARGE consortium		IGF-I levels	30,884	0.20 (-0.02, 0.43)	0.07
			IGFBP-3 levels	18,995	0.03 (-0.24, 0.30)	0.82

OR, odds ratio; 95% CI, 95% confidence interval; p, p-value; LDL-C, low density lipoprotein cholesterol; *HMGCR*, HMG-CoA reductase; BCa, breast cancer; BCAC, breast cancer association consortium; UKBB, UK Biobank; IGF-I, insulin-like growth factor 1; IGFBP-3, insulin-like growth factor binding protein 3.

^aAssociations are per 1 SD unit increase in LDL-C as proxied by genetic variation in *HMGCR*.