

1 **Association of genetic variants related to gluteofemoral versus abdominal fat**
2 **distribution with type 2 diabetes, coronary disease, and cardiovascular risk factors**

3
4 Luca A Lotta, MD, PhD,¹ Laura BL Wittemans, MSc,¹ Verena Zuber, PhD,² Isobel D
5 Stewart, PhD,¹ Stephen J Sharp, MSc,¹ Jian'an Luan, PhD,¹ Felix R Day, PhD,¹ Chen Li,
6 MSc,¹ Nicholas Bowker, MSc,¹ Lina Cai, MSc,¹ Emanuella De Lucia Rolfe, PhD,¹ Kay-Tee
7 Khaw, MB, ChB, MSc,³ John RB Perry, PhD,¹ Stephen O'Rahilly, MD,⁴ Robert A Scott,
8 PhD,¹ David B Savage, MD,⁴ Stephen Burgess, PhD,^{2,3} Nicholas J Wareham, MBBS, PhD,¹
9 Claudia Langenberg, MD, PhD.¹

10
11 1 MRC Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom

12 2 MRC Biostatistics Unit, University of Cambridge, Cambridge, United Kingdom

13 3 Department of Public Health and Primary Care, University of Cambridge, Cambridge,
14 United Kingdom

15 4 Metabolic Research Laboratories, Institute of Metabolic Science, University of Cambridge,
16 Cambridge, United Kingdom

17
18
19
20 Word Count (Abstract): 400

21 Word Count (Text): 4,503

22 Tables: 2

23 Figures: 2

24
25
26 Correspondence to:

27 Claudia Langenberg (claudia.langenberg@mrc-epid.cam.ac.uk)

28 Luca A Lotta (luca.lotta@mrc-epid.cam.ac.uk)

29 MRC Epidemiology Unit

30 University of Cambridge

31 Cambridge

32 CB20QQ

33 United Kingdom

34 Tel. +44 (0)1223 330315

35 Fax. +44 (0)1223 330316

36

37

38 **Abstract**

39

40 **Importance:** Body fat distribution, usually measured using waist-to-hip ratio (WHR), is an important
41 contributor to cardio-metabolic disease independent of body mass index (BMI). Whether mechanisms
42 that increase WHR via lower gluteofemoral (hip) or via higher abdominal (waist) fat distribution
43 affect cardio-metabolic risk is unknown.

44

45 **Objective:** To identify genetic variants associated with higher WHR specifically via lower
46 gluteofemoral or higher abdominal fat distribution and estimate their association with cardio-
47 metabolic risk.

48

49 **Design, Setting, and Participants:** Genome-wide association studies (GWAS) for WHR combined
50 data from the UK Biobank cohort and summary statistics from previous GWAS (data collection:
51 2006-2018). Specific polygenic scores for higher WHR via lower gluteofemoral or via higher
52 abdominal fat distribution were derived using WHR-associated genetic variants showing specific
53 association with hip or waist circumference. Associations of polygenic scores with outcomes were
54 estimated in three population-based cohorts, a case-cohort study and summary statistics from 6
55 GWAS (data collection: 1991-2018).

56

57 **Exposures:** Over 2.4 million common genetic variants (GWAS); polygenic scores for higher WHR
58 (follow-up analyses).

59

60 **Main outcomes and measures:** BMI-adjusted WHR and unadjusted WHR (GWAS); compartmental
61 fat mass measured by dual-energy X-ray absorptiometry (DEXA), systolic, diastolic blood pressure,
62 low-density lipoprotein cholesterol, triglycerides, fasting glucose, fasting insulin, type 2 diabetes and
63 coronary disease risk (follow-up analyses).

64

65 **Results:** Among 452,302 European-ancestry UK Biobank participants, mean age was 57 (SD=8)
66 years and mean WHR was 0.87 (SD=0.09). In genome-wide analyses, 202 independent genetic
67 variants were associated with higher BMI-adjusted WHR (N=660,648) and unadjusted WHR
68 (N=663,598). In DEXA analyses (N=18,330), the hip- and waist-specific polygenic scores for higher
69 WHR were specifically associated with lower gluteofemoral and higher abdominal fat, respectively.
70 In follow-up analyses (N=636,607), both polygenic scores were associated with higher blood
71 pressure, triglycerides and higher risk of diabetes (waist-specific score: odds ratio [OR], 1.57 [95%
72 CI, 1.34-1.83], absolute risk increase per 1000 participant-years [ARI], 4.4 [95% CI, 2.7-6.5], P<.001;
73 hip-specific score: OR, 2.54 [95% CI, 2.17-2.96], ARI, 12.0 [95% CI, 9.1-15.3], P<.001) and
74 coronary disease (waist-specific score: OR, 1.60 [95% CI, 1.39-1.84], ARI, 2.3 [95% CI, 1.5-3.3],
75 P<.001; hip-specific score: OR, 1.76 [95% CI, 1.53-2.02], ARI, 3.0 [95% CI, 2.1-4.0], P<.001), per 1
76 SD increase in BMI-adjusted WHR.

77

78 **Conclusions and Relevance:** Distinct genetic mechanisms may be linked to gluteofemoral and
79 abdominal fat distribution that are the basis for the calculation of the waist-to-hip ratio. If replicated in
80 additional diverse populations, these findings may have implications for risk assessment and treatment
81 of diabetes and coronary disease.

82 **Key points**

83

84 **Question:** Do genetic variants that are related to body fat distribution via lower levels of
85 gluteofemoral (hip) fat or via higher levels of abdominal (waist) fat show associations with
86 diabetes or coronary disease risk?

87

88 **Findings:** In genetic studies including up to 636,607 people, distinct polygenic risk scores for
89 increased waist-to-hip ratio via lower gluteofemoral or via higher abdominal fat distribution
90 were significantly associated with higher levels of cardio-metabolic risk factors and higher
91 risk for type 2 diabetes and coronary disease.

92

93 **Meaning:** Genetic mechanisms specifically linked to lower gluteofemoral or higher
94 abdominal fat distribution may independently contribute to the relationship between body
95 shape and cardio-metabolic risk.

96 **Introduction**

97 The distribution of body fat is associated with the propensity of overweight individuals
98 to manifest insulin resistance and its associated metabolic and cardiovascular complications.¹⁻

99 ⁵ The waist-to-hip ratio (WHR) is a widely-used, convenient and robustly validated indicator
100 of fat distribution and is linked to the risk of type 2 diabetes and coronary disease
101 independently of body mass index (BMI).¹⁻⁵ This observation has been used to infer that
102 accumulation of fat in the abdominal cavity is an independent causal contributor to cardio-
103 metabolic disease. Whilst many studies support this assertion and plausible mechanisms have
104 been proposed, the waist-to-hip ratio can also be increased by a reduction in its denominator,
105 the hip circumference. Evidence from several different forms of partial lipodystrophy^{6,7} and
106 functional studies of peripheral adipose storage compartments⁸⁻¹⁰ suggests that a primary
107 inability to expand gluteofemoral or hip fat can also underpin subsequent cardio-metabolic
108 disease risk. Emerging evidence from the analysis of common genetic variants associated
109 with greater insulin resistance but lower levels of hip fat suggests that similar mechanisms
110 may also be relevant to the general population.¹¹⁻¹⁴

111 In this study, large-scale human genetic data were used to investigate whether genetic
112 variants related to body fat distribution via lower levels of gluteofemoral (hip) fat or via
113 higher levels of abdominal (waist) fat are associated with type 2 diabetes or coronary disease
114 risk.

115 **Methods**

116 *Study design*

117 A multi-stage approach was adopted (**Table 1**). In Stage 1, genome-wide association
118 studies (GWAS) of waist-to-hip ratio with ($WHR_{\text{BMI-adjusted}}$) and without ($WHR_{\text{unadjusted}}$)
119 adjustment for BMI were performed to identify genetic variants associated with fat
120 distribution. Stage 1 included data from European ancestry participants of the UK Biobank
121 study and summary statistics from previously-published GWAS of the Genetic Investigation
122 of Anthropometric Traits (GIANT) consortium.¹⁵ In Stage 2, general, hip- and waist-specific
123 polygenic scores for higher WHR were derived using 202 genetic variants independently
124 associated with WHR in Stage 1. Stage 2 included data from European ancestry participants
125 of UK Biobank and summary statistics from GIANT.¹⁵ In Stage 3, associations of polygenic
126 scores with compartmental fat mass measured by dual-energy X-ray absorptiometry (DEXA)
127 were estimated in European ancestry participants from the UK Biobank, Fenland and EPIC-
128 Norfolk studies. In Stage 4, associations of polygenic scores with six cardio-metabolic risk
129 factors and with risk of type 2 diabetes and coronary artery disease were estimated using data
130 from European ancestry participants of UK Biobank, the EPIC-InterAct case-cohort study
131 and summary statistics from 6 previously-published GWAS. All studies were approved by
132 local institutional review boards and ethics committees and participants gave written
133 informed consent.

134

135 *Studies and participants*

136 UK Biobank (data collection: 2006-2018) is a prospective population-based cohort study
137 of people aged 40-69 years who were recruited in 2006-2010 from 22 centers located in
138 urban and rural areas across the United Kingdom.¹⁶

139 Fenland (data collection: 2005-2018) is a prospective population-based cohort study of

140 people born in 1950-1975 and recruited in 2005-2015 from outpatient primary care clinics in
141 Cambridge, Ely and Wisbech (United Kingdom).¹¹

142 EPIC-Norfolk (data collection: 1993-2018) is a prospective population-based cohort
143 study of individuals aged 40-79 and living in the Norfolk county (rural areas, market towns
144 and the city of Norwich) in the United Kingdom at recruitment from outpatient primary care
145 clinics in 1993-1997.¹⁷

146 EPIC-InterAct (data collection: 1991-2018) is a case-cohort study nested within the
147 European Prospective Investigation into Cancer and Nutrition (EPIC) study, a prospective
148 cohort study.¹⁸ EPIC study participants who developed type 2 diabetes after study baseline
149 constituted the incident case group of EPIC-InterAct and a randomly-selected group of
150 individuals free of diabetes at baseline constituted the subcohort.

151 Summary statistics from 11 GWAS published by research consortia between 2012 and
152 2015 were used in the different stages of the study (**eMethods 1 and eTable 1**). These
153 included genetic variant associations with BMI, $WHR_{BMI\text{-adjusted}}$, $WHR_{unadjusted}$, waist- and
154 hip-circumference from the GIANT consortium,^{15,19} associations with fasting glucose and
155 fasting insulin from the Meta-analyses of Glucose and Insulin-related Traits consortium
156 (MAGIC),^{20,21} associations with triglycerides and low-density lipoprotein cholesterol (LDL-
157 C) from the Global Lipid Genetic consortium (GLGC),²² associations with type 2 diabetes
158 from the Diabetes Genetics Replication and Meta-analysis (DIAGRAM) consortium²³ and
159 with coronary artery disease from the Coronary Artery Disease Genome-wide Replication
160 and Meta-analysis plus the Coronary Artery Disease Genetics consortium
161 (CARDIOGRAMplusC4D).²⁴ Data collection took place in 2012-2016.

162 Detailed descriptions of study design, sources of data, and participants in each stage are
163 in **Tables 1-2, eMethods 1 and eTables 1-3**.

164

165 *Outcomes*

166 Outcomes of the study were WHR (Stage 1 and 2b), hip and waist circumference (Stage
167 2a), compartmental body fat masses (Stage 3), six cardio-metabolic risk factors (systolic and
168 diastolic blood pressure, fasting glucose, fasting insulin, triglycerides and LDL-C; Stage 4)
169 and two disease outcomes (type 2 diabetes and coronary disease; Stage 4).

170 Stage 1 and 2: WHR was defined as the ratio of the circumference of the waist to that of
171 the hip, both of which were estimated in cm using a Seca 200-cm tape measure. BMI-
172 adjusted WHR was obtained by calculating the residuals for a linear regression model of
173 WHR on age, sex and BMI.

174 Stage 3: compartmental fat masses were measured in grams by DEXA, a whole-body,
175 low-intensity X-ray scan that precisely quantifies fat mass in different body regions. In UK
176 Biobank, DEXA measures were obtained using a GE-Lunar iDXA instrument. In Fenland
177 and EPIC-Norfolk, DEXA scans were performed using a Lunar Prodigy advanced fan beam
178 scanner (GE Healthcare, Bedford, UK). Participants were scanned by trained operators using
179 standard imaging and positioning protocols. All the images were manually processed by one
180 trained researcher, who corrected DEXA demarcations according to a standardized procedure
181 as illustrated in **eFigure 1** and described in **eMethods 1**. In brief, the arm region included the
182 arm and shoulder area, the trunk region included the neck, chest, abdominal and pelvic areas.
183 The abdominal region was defined as the area between the ribs and the pelvis, and was
184 enclosed by the trunk region. The leg region included all of the area below the lines that form
185 the lower borders of the trunk. The gluteofemoral region included the hips and upper thighs,
186 and overlapped both leg and trunk regions. The upper demarcation of this region was below
187 the top of the iliac crest at a distance of 1.5 times the abdominal height. The DEXA
188 CoreScan® software (GE Healthcare, Bedford UK) was used to determine visceral
189 abdominal fat mass within the abdominal region.

190 Stage 4, risk factors: systolic and diastolic blood pressures were defined as the values of
191 arterial blood pressure in mmHg measured using an Omron monitor during the systolic and
192 diastolic phases of the heart cycle. Fasting insulin and fasting glucose were defined as the
193 values of insulin (log-transformed and expressed in log-pmol/L) in serum and glucose
194 (mmol/L) in whole blood measured in fasting state in non-diabetic individuals as previously
195 described.^{20,21} Triglycerides (log-transformed and expressed in log-mmol/L) and LDL-C
196 (mmol/L) levels in the circulation were measured using biochemical assays (triglycerides and
197 24% of LDL-C values in the GLGC study²²) or derived with the Friedewald formula (76% of
198 LDL-C values in the GLGC study²²) as previously described.²²

199 Stage 4, disease outcomes: for disease outcomes analyses in UK Biobank, binary
200 definitions of prevalent disease status and a case-control analytical design were used in line
201 with previous work.^{11,25,26} Definition of prevalent diabetes was consistent with validated
202 algorithms.²⁵ Participants were classified as cases of prevalent type 2 diabetes if they met the
203 following two criteria: (1) self-reported type 2 diabetes diagnosis or self-reported diabetes
204 medication at nurse interview or at digital questionnaire, or electronic health record
205 consistent with type 2 diabetes (International Statistical Classification of Diseases and
206 Related Health Problems version 10 [ICD-10] code E11); and (2) age at diagnosis >36 years
207 or use of oral anti-diabetic medications (to remove likely type 1 diabetes cases). Controls
208 were participants who (1) did not self-report a diagnosis of diabetes of any type, and (2) did
209 not take any diabetes medications, and (3) did not have an electronic health record of diabetes
210 of any type. In EPIC-InterAct, the outcome was incident type 2 diabetes. Incident type 2
211 diabetes case status was defined on the basis of evidence of type 2 diabetes from self-report,
212 primary care registers, drug registers (medication use), hospital record or mortality data.¹⁸
213 Incident type 2 diabetes cases were considered to be verified if evidence from a minimum of
214 two of these independent sources was present.¹⁸ Participants free from type 2 diabetes at

215 baseline were randomly selected from participating EPIC-study cohorts and constituted the
216 subcohort group of EPIC-InterAct. Participants with prevalent diabetes at study baseline were
217 excluded from EPIC-InterAct. In UK Biobank, prevalent coronary artery disease was defined
218 as either (1) myocardial infarction or coronary disease documented in the participant's
219 medical history at the time of enrolment by a trained nurse or (2) an electronic health record
220 of acute myocardial infarction or its complications (ICD-10 codes I21-I23). Controls were
221 participants who did not meet any of these criteria.

222

223 *Statistical analysis*

224 Stage 1: in UK Biobank, GWAS analyses were performed using BOLT-LMM,²⁷ which
225 fits linear mixed-models accounting for relatedness between individuals using a genomic
226 kinship matrix.^{27,28} An inverse-variance weighted, fixed-effect meta-analysis of results from
227 UK Biobank and GIANT was performed using METAL.²⁹ This study focused on 2,446,094
228 common genetic variants in autosomal chromosomes (i.e. not X or Y chromosome) with
229 minor allele frequency $\geq 0.5\%$ captured in both UK Biobank and GIANT. Restriction to
230 European ancestry individuals, use of linear mixed-models (UK Biobank) and adjustment for
231 genetic principal components and genomic inflation factor (GIANT) were used to minimize
232 type I error. Quality measures of genuine genetic association signal versus possible
233 confounding by population stratification or relatedness included the mean χ^2 statistic, the
234 linkage-disequilibrium score (LDSC) regression intercept and its attenuation ratio (**eMethods**
235 **2**), as recommended for genetic studies of this size using linear mixed model estimates.²⁸
236 Values of LDSC-regression intercept below 1.5 and an attenuation ratio statistic (a measure
237 of proportionality between LDSC-regression intercept and χ^2 statistic calculated as: [LDSC
238 intercept – 1] / [mean χ^2 statistic – 1]) equal to or below 0.08 are consistent with optimal
239 control of genetic confounding.²⁸ Genetic variants were taken forward to Stage 2 if they were

240 associated with both $WHR_{BMI-adjusted}$ and $WHR_{unadjusted}$ at the conventional genome-wide level
241 of statistical significance³⁰ ($P < 5 \times 10^{-08}$ in each analysis). The use of both BMI-adjusted and
242 unadjusted results prevented the inclusion of variants associated with higher WHR via
243 collider bias³¹ or via a primary association with higher BMI. A forward-selection process was
244 used to select independent genetic variants for Stage 2. At each iteration, the genetic variant
245 with the lowest P-value for $WHR_{BMI-adjusted}$ was selected, while genetic variants within
246 1,000,000 base pairs either side of that genetic variant were discarded from further iterations.
247 The resulting list of genetic variants was further filtered on the basis of pairwise linkage
248 disequilibrium such that the final list of independent genetic variants had no or negligible
249 correlation (pairwise $R^2 < .05$). Full details about genetic analyses are in **eMethods 2**.

250 Stage 2: polygenic scores capturing genetic predisposition to higher WHR were derived
251 by combining the 202 independent genetic variants from Stage 1 (or subsets of the 202
252 variants as described below), weighted by their association with $WHR_{BMI-adjusted}$ in Stage 1. A
253 general polygenic score for higher WHR was derived by combining all 202 genetic variants.
254 A waist-specific polygenic score capturing genetic predisposition to higher WHR via higher
255 abdominal fat was derived by combining 36 variants specifically associated with waist
256 ($P < .00025$, a Bonferroni correction for 202 genetic variants) but not with hip circumference
257 ($P > .20$, an arbitrary threshold). A hip-specific polygenic score capturing genetic
258 predisposition to higher WHR via lower gluteofemoral fat was derived by combining 22
259 variants specifically associated with hip ($P < .00025$) but not with waist circumference ($P > .50$,
260 a stricter arbitrary threshold which was necessary because of residual associations with waist
261 circumference of a polygenic score initially derived using $P > .20$, **eMethods 3**). A fourth
262 polygenic score was derived by combining 144 genetic variants not included in the waist- or
263 hip-specific polygenic scores.

264 The statistical performance of these polygenic scores was assessed by estimating the

265 proportion of the variance in $WHR_{BMI-adjusted}$ accounted for by the score (variance explained)
266 and by the F-statistic (**eMethods 4**). The F-statistic is a measure of the ability of the
267 polygenic score to predict the independent variable ($WHR_{BMI-adjusted}$). Values of F-statistic
268 above 10 have been considered to provide evidence of a statistically-robust polygenic
269 score.^{26,32} Statistical power calculations for the association with disease outcomes were also
270 performed (**eMethods 4 and eFigure 2**).

271 Stage 3 and 4: associations of polygenic scores with DEXA phenotypes, cardio-
272 metabolic risk factors and outcomes were estimated in each study separately and results were
273 combined using fixed-effect inverse-variance weighted meta-analysis. In individual-level
274 data analyses, polygenic scores were calculated for each study participant by adding the
275 number of copies of each contributing genetic variant weighted by its association estimate in
276 SD units of $WHR_{BMI-adjusted}$ per allele from Stage 1. Association of polygenic scores with
277 outcomes were estimated using linear, logistic or Cox regression models as appropriate for
278 outcome type and study design. Regression models were adjusted for age, sex and genetic
279 principal components or a genomic kinship matrix to minimize genetic confounding. In UK
280 Biobank disease outcomes analyses, prevalent disease status was defined as a binary variable
281 and logistic regression was used to estimate the odds ratio of disease per 1 SD increase in
282 $WHR_{BMI-adjusted}$ due to a given polygenic score. In EPIC-InterAct, Cox regression weighted
283 for case-cohort design was used to estimate the hazard ratio of incident type 2 diabetes per 1
284 SD increase in $WHR_{BMI-adjusted}$ due to a given polygenic score. In summary statistics analyses,
285 estimates equivalent to those of individual-level analyses were obtained using inverse-
286 variance weighted meta-analysis of the association of each genetic variant in the polygenic
287 score with the outcome, divided by the association of that genetic variant with WHR_{BMI-}
288 $adjusted$.³³ These analytical approaches assume normal distributions for polygenic scores and
289 continuous outcomes. They also assume a linear relationship of the polygenic score with

290 continuous outcomes (linear regression), or with the log-odds of binary outcomes (logistic
291 regression), or with the log-hazard of incident disease (Cox regression). All of these
292 assumptions were largely met in this study (**eMethods 5, eTable 4 and eFigures 3-6**). Meta-
293 analyses of log-odds ratios and log-hazard ratios of disease assumed that these estimates are
294 similar, an assumption which was shown to be reasonable in a sensitivity analysis conducted
295 in EPIC-InterAct (**eMethods 5 and eFigure 7**).

296 In Stage 3 and 4, associations with continuous outcomes were expressed in standardized
297 or clinical units of outcome per 1 SD increase in $WHR_{BMI-adjusted}$ (corresponding to 0.056 ratio
298 units of age-, sex- and BMI-residualized WHR in UK Biobank) due to a given polygenic
299 score (**eMethods 5 and eTable 5**). Associations with disease outcomes were expressed as
300 odds ratios (OR) for outcome per 1 SD increase in $WHR_{BMI-adjusted}$ due to a given polygenic
301 score. Absolute risk increases (ARI) for disease outcomes were estimated using the estimated
302 ORs and the incidence of type 2 diabetes or coronary disease in the United States (**eMethods**
303 **5**). The threshold of statistical significance for association with DEXA phenotypes was
304 $P < .0016$ ($0.05/32 = 0.0016$, Bonferroni correction for 8 outcomes and 4 polygenic scores), that
305 for association with cardio-metabolic risk factors was $P < .0021$ ($0.05/24 = 0.0021$, Bonferroni
306 correction for 6 outcomes and 4 polygenic scores), and that for association with type 2
307 diabetes and coronary disease was $P < .0063$ ($0.05/8 = 0.0063$, Bonferroni correction for 2
308 outcomes and 4 polygenic scores). All reported P-values were from 2-tailed statistical tests.

309 In addition to deriving specific polygenic scores, the independent association of
310 gluteofemoral or abdominal fat distribution with outcomes was studied using multivariable
311 genetic association analyses adjusting for either of these two components of body fat
312 distribution (**eMethods 6 and eFigure 8**). Adjusting for abdominal fat distribution measures
313 was used as a way of estimating the residual association of the polygenic score with
314 outcomes via gluteofemoral fat distribution, while adjusting for gluteofemoral fat distribution

315 measures as a way of estimating the residual association via abdominal fat distribution
316 (**eFigure 8**). To obtain adjusted association estimates, multivariable weighted regression
317 models were fitted in which the association of the 202-variant general polygenic score
318 (exposure) with cardio-metabolic risk factors or diseases (outcomes) was estimated while
319 adjusting for a polygenic score comprising the same 202 genetic variants but weighted for
320 measures of abdominal fat distribution or measures of gluteofemoral fat distribution
321 (covariates).³⁴ A detailed description of these analysis methods and their assumptions is in
322 **eMethods 6 and eFigures 8-9**. This method was also used to conduct a post hoc exploratory
323 analysis of the association of the hip-specific polygenic score with cardio-metabolic disease
324 outcomes after adjusting for visceral abdominal fat mass estimates.

325 Six different secondary or sensitivity analyses were conducted to estimate the association
326 of polygenic scores with other phenotypes including high-density lipoprotein cholesterol
327 (HDL-C), triglyceride/HDL-C ratio, height, and non-diabetic hyperglycemia, and to assess
328 the robustness of the main analysis to associations with height, sex-specific associations, or
329 the possibility of false positive associations in Stage 1 or Stage 2 (**eMethods 7**).

330 Statistical analyses were performed using STATA v14.2 (StataCorp, College Station,
331 Texas 77845 USA), R v3.2.2 (The R Foundation for Statistical Computing), BOLT-LMM
332 v2.3.2^{27,28} and METAL v2011-03-25.²⁹

333 **Results**

334 *Genetic predisposition to higher WHR via lower gluteofemoral or via higher abdominal fat*

335 Among 452,302 European ancestry participants of UK Biobank, mean age was 57
336 (SD=8) years, women were 245,351 (54%) and mean WHR was 0.87 (SD=0.09; **Table 2**). In
337 genome-wide association analyses of $\text{WHR}_{\text{BMI-adjusted}}$ (N=660,648, mean $\chi^2=2.50$, LDSC-
338 regression intercept, 1.098 [95% CI, 1.063, 1.134], attenuation ratio, 0.07 [95% CI, 0.04,
339 0.09]) and $\text{WHR}_{\text{unadjusted}}$ (N=663,598, mean $\chi^2=2.68$, LDSC-regression intercept, 1.096 [95%
340 CI, 1.064, 1.129], attenuation ratio, 0.06 [95% CI, 0.04, 0.08]) there was evidence of optimal
341 control for genetic confounding (**eMethods 2, eFigures 10-11**). A total of 202 independent
342 genetic variants were associated with both $\text{WHR}_{\text{BMI-adjusted}}$ and $\text{WHR}_{\text{unadjusted}}$ ($P < 5 \times 10^{-08}$ in
343 each analysis; **eTable 6, eFigures 12-13**). These 202 genetic variants were used to derive
344 polygenic scores for higher WHR (**Table 1**). The 202-variant general score (variance in
345 $\text{WHR}_{\text{BMI-adjusted}}$ explained by score in UK Biobank=3.4%, F-statistic=12,231), 22-variant hip-
346 specific score (variance explained=0.4%, F-statistic=1,550), 36-variant waist-specific score
347 (variance explained=0.4%, F-statistic=1,444), and 144-variant general score (variance
348 explained=2.6%, F-statistic=9,177) were statistically robust polygenic scores for $\text{WHR}_{\text{BMI-}}$
349 adjusted (**eMethods 4 and eFigure 2**).

350 In 18,330 people with DEXA compartmental fat measures, all polygenic scores for
351 higher WHR were associated with a higher abdominal-to-gluteofemoral fat mass ratio, a
352 refined measure of body fat distribution, but were associated with different patterns of
353 compartmental fat mass distribution (**Figure 1, eFigures 14-15**). The general 202-variant and
354 144-variant polygenic scores were associated with higher visceral abdominal and lower
355 gluteofemoral fat mass (**Figure 1A, eFigure 15**). The waist-specific polygenic score for
356 higher WHR was associated with higher abdominal fat mass, but not with gluteofemoral or
357 leg fat mass (**Figure 1B**). The hip-specific polygenic score for higher WHR was associated

358 with lower gluteofemoral and leg fat mass, but did not show statistically-significant
359 associations with abdominal fat mass (**Figure 1B**). Participants with higher values of the hip-
360 specific polygenic score had numerically higher visceral abdominal fat mass, but the
361 difference was not statistically significant when accounting for multiple tests (**Figure 1B**).

362

363 *Associations with cardio-metabolic risk factors and disease outcomes*

364 In 636,607 people, the 202-variant polygenic score for higher WHR was associated with
365 higher odds of type 2 diabetes and coronary artery disease and an unfavorable cardio-
366 metabolic risk profile (**eFigure 16**), consistent with previous studies of ~50 genetic
367 variants.^{15,26,35} In secondary analyses, there were associations with lower HDL-C, higher
368 triglyceride/HDL-C ratio and higher odds of non-diabetic hyperglycemia (**eMethods 7 and**
369 **eTables 7-8**). Associations with cardio-metabolic disease outcomes were similar in men and
370 women with no evidence of sex-interaction ($P_{\text{interaction}}$ for type 2 diabetes=0.19; $P_{\text{interaction}}$ for
371 coronary artery disease=0.80; **eTable 9**).

372 Both hip-specific and waist-specific polygenic scores for higher WHR were associated
373 with higher systolic, diastolic blood pressure and triglycerides (**Figure 2A**), with similar
374 association estimates for a 1 SD increase in $\text{WHR}_{\text{BMI-adjusted}}$. While the hip-specific polygenic
375 score was associated with higher fasting insulin and higher LDL-C, the waist-specific
376 polygenic score did not have statistically-significant associations with these traits (**Figure**
377 **2A**). Both the hip-specific and the waist-specific polygenic scores were associated with
378 higher odds of type 2 diabetes and coronary disease (**Figure 2B**), similarly in men and
379 women (**eTable 9**). The hip-specific polygenic score had a statistically larger association
380 estimate for diabetes than the waist-specific polygenic score per 1 SD increase in $\text{WHR}_{\text{BMI-}}$
381 adjusted (OR, 2.54 [95% CI, 2.17-2.96] vs 1.57 [1.34-1.83]; ARI, 12.0 [95% CI, 9.1-15.3] vs
382 4.4 [95% CI, 2.7-6.5] cases per 1000 participant-years; $P_{\text{heterogeneity}} < .001$; **Figure 2B**). In a

383 post-hoc multivariable analysis adjusting for visceral abdominal fat mass estimates, the hip-
384 specific polygenic score showed a statistically-significant association with higher odds of
385 type 2 diabetes and coronary disease (OR for diabetes per 1 SD increase in $WHR_{BMI-adjusted}$
386 due to the hip-specific polygenic score, 2.84 [95% CI, 1.98-4.08], ARI, 14.4 [95% CI, 7.6-24]
387 cases per 1000 participant-years, $P < .001$; OR for coronary disease, 1.74 [95% CI, 1.35-2.25],
388 ARI, 2.9 [95% CI, 1.4-4.9] cases per 1000 participant-years, $P < .001$). The 144-variant
389 polygenic score showed associations with risk factors and disease outcomes similar to those
390 observed for the 202-variant general polygenic score (**eFigure 15**). Sensitivity analyses
391 supported the robustness of the main analysis to sex-specific associations, associations with
392 height, or the possibility of false positive associations in Stage 1 or Stage 2 (**eMethods 7**,
393 **eTables 9-11**).

394 In multivariable analyses adjusting for hip circumference estimates, the 202-variant
395 polygenic score had a pattern of association with compartmental fat mass, cardio-metabolic
396 risk factors and disease outcomes which was similar to that of the waist-specific polygenic
397 score (**eFigure 8D and eFigure 17**). The 202-variant polygenic score remained associated
398 with higher risk of type 2 diabetes and coronary disease even when adjusting for hip
399 circumference and leg fat mass in the same model (**eTable 12**).

400 In multivariable analyses adjusting for waist circumference estimates, the 202-variant
401 polygenic score had a pattern of association with compartmental fat mass, cardio-metabolic
402 risk factors and disease outcomes which was similar to that of the hip-specific polygenic
403 score (**eFigure 8C and eFigure 17**). The 202-variant polygenic score remained associated
404 with higher risk of type 2 diabetes and coronary disease even when adjusting for waist
405 circumference and visceral abdominal fat mass in the same model (**eTable 12**).

406 In multivariable analyses adjusting for both waist and hip circumference estimates, the
407 202-variant polygenic score was not associated with risk of type 2 diabetes or coronary

408 disease (**eFigure 8B and eTable 12**).

409 **Discussion**

410 This large study identified distinct genetic variants associated with a higher WHR via
411 specific associations with lower gluteofemoral or higher abdominal fat distribution. Both
412 these distinct sets of genetic variants were associated with higher levels of cardio-metabolic
413 risk factors and a higher risk of type 2 diabetes and coronary disease. While this study
414 supports the theory that an enhanced accumulation of fat in the abdominal cavity may be a
415 cause of cardiovascular and metabolic disease, it also provides novel evidence of a possible
416 independent role of the relative inability to expand the gluteofemoral fat compartment.

417 Previous studies of ~50 genomic regions associated with BMI-adjusted WHR¹⁵ have
418 shown an association between genetic predisposition to higher WHR and higher risk of
419 cardio-metabolic disease,^{26,35} mirroring the well-established BMI-independent association of
420 a higher WHR with incident cardiovascular and metabolic disease in large-scale
421 observational studies.^{2,3} While these results have been widely interpreted as supportive of the
422 role of abdominal fat deposition in cardio-metabolic risk independent of overall adiposity, the
423 etiologic contribution of lower levels of gluteofemoral and peripheral fat to these associations
424 has not been considered.

425 The results of this study support the hypothesis that an impaired ability to preferentially
426 deposit excess calories in the gluteofemoral fat compartment leads to higher cardio-metabolic
427 risk in the general population. This is consistent with observations in severe forms of partial
428 lipodystrophy^{6,7} and with the emerging evidence of a shared genetic background between
429 extreme lipodystrophies and fat distribution in the general population.¹¹ This large human
430 genetic study adds to a growing body of evidence linking gluteofemoral and subcutaneous
431 adipose tissue biology with a favorable metabolic profile.⁸⁻¹⁰ The hip-specific polygenic score
432 for higher WHR was not significantly associated with measures of central fat in DEXA
433 analyses and, in a post hoc analysis, its association with cardio-metabolic disease outcomes

434 was independent of visceral abdominal fat mass. These associations may perhaps reflect the
435 secondary deposition within ectopic fat depots, such as liver, cardiac and skeletal muscle and
436 pancreas, of excess calories that cannot be accommodated in gluteofemoral fat.^{36,37}

437 It has been hypothesized that the association between fat distribution and cardio-
438 metabolic risk is due to an enhanced deposition of intra-abdominal fat generating a molecular
439 milieu that fosters abdominal organ insulin resistance.³⁸ The results of this study support a
440 role of abdominal fat distribution, but they also suggest that impaired gluteofemoral fat
441 distribution may contribute to the relationship between body shape and cardio-metabolic
442 health outcomes.

443

444 **Limitations**

445 This study has several limitations. First, as this is an observational study, it cannot
446 establish causality. Second, the discovery and characterization of genetic variants was
447 conducted in a large dataset but was limited to individuals of European ancestry. While the
448 genetic determinants of anthropometric phenotypes may be partly shared across different
449 ethnicities,^{15,39,40} further investigations in other populations and ethnicities will be required
450 for a complete understanding of the genetic relationships between body shape and cardio-
451 metabolic risk. Third, this study was largely based on population-based cohorts, the
452 participants of which are usually healthier than the general population, and used analytical
453 approaches that deliberately minimize the influence of outliers, in this case people with
454 extreme fat distribution. Genetic studies in people with extreme fat distribution may help
455 broaden understanding of the genetic basis of this risk factor. Fourth, while disease case
456 definitions were based on widely-adopted criteria, misclassification of cases/controls cannot
457 be excluded, which would bias association estimates towards the null. Fifth, absolute risk
458 increase estimates are based on incidence rates and odds ratios calculated in different

459 populations and therefore assume that these populations are similar. Sixth, P-value thresholds
460 used to exclude associations with the other component of fat distribution for genetic variants
461 included in waist- or hip-specific polygenic scores were arbitrarily chosen, but are more
462 stringent than traditionally used cutoffs (e.g. $P > .05$) and polygenic score results were
463 confirmed by multivariable genetic analyses which were independent of such thresholds.
464 Seventh, this analysis focused on common genetic variants captured in both UK Biobank and
465 GIANT and, by design, did not investigate the role of rare genetic variation or of other
466 variants captured by dense imputation in UK Biobank. Eighth, there was a statistically-
467 significant difference in the association of hip- versus waist-specific polygenic scores with
468 diabetes risk, with greater estimated magnitude of association for the hip-specific polygenic
469 score. However, given that the difference in absolute risk was small, this observation does not
470 necessarily represent a strong signal of mechanistic difference or differential clinical
471 importance in the relationship between the gluteofemoral versus abdominal components of
472 fat distribution and diabetes risk.

473

474 **Conclusions**

475 Distinct genetic mechanisms may be linked to gluteofemoral and abdominal fat
476 distribution that are the basis for the calculation of the waist-to-hip ratio. If replicated in
477 additional diverse populations, these findings may have implications for risk assessment and
478 treatment of diabetes and coronary disease.

479 **Acknowledgement**

480

481 Data access: L.A.L had full access to all the data in the study and takes responsibility for the integrity
482 of the data and the accuracy of the data analysis.

483 Acknowledgement: This research has been conducted using the UK Biobank resource. This study has
484 been conducting using data from the EPIC-InterAct, Fenland and EPIC-Norfolk studies. The authors
485 gratefully acknowledge the help of the MRC Epidemiology Unit Support Teams, including Field,
486 Laboratory and Data Management Teams.

487 Funding: This study was funded by the United Kingdom’s Medical Research Council through grants
488 MC_UU_12015/1, MC_PC_13046, MC_PC_13048 and MR/L00002/1. This work was supported by
489 the MRC Metabolic Diseases Unit (MC_UU_12012/5) and the Cambridge NIHR Biomedical
490 Research Centre and EU/EFPIA Innovative Medicines Initiative Joint Undertaking (EMIF grant:
491 115372). EPIC-InterAct Study funding: funding for the InterAct project was provided by the EU FP6
492 program (grant number LSHM_CT_2006_037197). D.B.S. and S.O’R. are supported by the
493 Wellcome Trust (WT107064 and WT095515 respectively) the MRC Metabolic Disease Unit, the
494 National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre and the NIHR
495 Rare Disease Translational Research Collaboration.

496 Role of Sponsor: The funding bodies had no role in the design or conduct of the study; collection,
497 management, analysis or interpretation of the data; preparation, review or approval of the manuscript
498 or the decision to submit the manuscript for publication.

499 Conflict of Interest Disclosures: R. A. S. is an employee and shareholder of GlaxoSmithKline Plc.
500 (GSK). S.O’R. reports personal fees from CVMED TASAP Pfizer Advisory Board, personal fees
501 from AstraZeneca CVMED iMed External Scientific Panel, personal fees from MedImmune
502 Cardiovascular & Metabolic Disease (CVMED) iMed Advisory Board, personal fees from ERX
503 Pharmaceuticals Scientific Advisory Board, outside the submitted work. The other authors report no
504 conflict of interest relative to this study.

505

506

507 **Tables**

508

509 **Table 1. Summary of the study design.**

510

Stage and aim	Independent variables	Outcome variables	Outcome data sources	Statistical significance
Stage 1: Genetic discovery Identify genetic variants associated with fat distribution	~2.4 million common genetic variants genome-wide	BMI-adjusted WHR (N=660,648) and unadjusted WHR (N=663,598)	UK Biobank; GIANT (summary statistics)	$P < 5 \times 10^{-08}$ in each analysis
Stage 2a: Derivation of polygenic scores for higher WHR^a Select genetic variants into polygenic scores for higher WHR capturing different components of fat distribution	202 independent genetic variants from Stage 1	Hip (N=664,446) and waist (N=683,549) circumference	UK Biobank; GIANT (summary statistics)	Hip- or waist specific WHR-associated genetic variant: $P < .00025$ for association with either hip or waist and at least $P > 0.2$ for association with the other
Stage 2b: Polygenic score performance Assess polygenic scores performance using variance explained and F-statistic	Four polygenic scores for higher WHR ^a	BMI-adjusted WHR (N=350,721) ^b	UK Biobank	F-statistic >10
Stage 3: Polygenic score validation Association of polygenic scores for higher WHR with detailed compartmental fat distribution measures	Polygenic scores for higher WHR from Stage 2b	Arm, trunk, abdominal, abdominal visceral, abdominal subcutaneous, gluteofemoral, leg fat mass and abdominal/gluteofemoral fat mass ratio measured by DEXA (N=18,330)	Fenland; EPIC-Norfolk; UK Biobank	$P < .0016$
Stage 4: Cardio-metabolic risk association Association of polygenic scores for higher WHR with cardiovascular risk factors and disease outcomes	Polygenic scores for higher WHR from Stage 2b	Risk factors: systolic (N=451,402), diastolic (N=451,415) blood pressure; fasting insulin (N=108,557), fasting glucose (N=133,010); triglycerides (N=188,577), LDL-C (N=188,577) Outcomes: type 2 diabetes (69,677 cases, 551,081 controls), coronary disease (85,358 cases, 551,249 controls)	Risk factors: UK Biobank; MAGIC (summary statistics); GLGC (summary statistics) Disease outcomes: UK Biobank; EPIC-InterAct; DIAGRAM (summary statistics); CARDIoGRAMplusC4D (summary statistics)	$P < .0021$ for risk factors $P < .0063$ for disease outcomes

511 Abbreviations: WHR, waist-to-hip ratio; BMI, body mass index; DEXA, dual-energy X-ray absorptiometry; LDL-C, low-density lipoprotein cholesterol. Studies
512 participating in each stage are described in details in the Methods section, Table 2, eMethods 1 and eTables 1-3.

513 a The four polygenic scores included: (1) general polygenic score for higher WHR including all 202 independent genetic variants from Stage 1; (2) waist-specific
514 polygenic score for higher WHR including 36 genetic variants associated with waist but not hip in Stage 2a; (3) hip-specific polygenic score for higher WHR
515 including 22 genetic variants associated with hip but not waist in Stage 2a; (4) general polygenic score for higher WHR including 144 genetic variants not
516 included in the waist-specific or hip-specific polygenic scores.
517 b Variance explained was estimated using linear regression models in unrelated European ancestry participants of UK Biobank.¹⁶

518 **Table 2. Participants of UK Biobank included in this study.**

519

Study	UK Biobank
Country	United Kingdom
Genotyping chip	Affymetrix UK BILEVE and UK Biobank Axiom arrays
Imputation panel	Haplotype Reference Consortium r1.1
Participants, N	452,302
Female sex, N (%)	245,351 (54)
Male sex, N (%)	206,951 (46)
Age at baseline, mean years (SD)	57 (8)
Age at baseline in women, mean years (SD)	57 (8)
Age at baseline in men, mean years (SD)	57 (8)
Currently smoking, N (%)	47,036 (10)
Currently smoking in women, N (%)	21,867 (9)
Currently smoking in men, N (%)	25,165 (12)
BMI, mean kg/m ² (SD) ^a	27.4 (4.8)
BMI in women, mean kg/m ² (SD)	27.0 (5.1)
BMI in men, mean kg/m ² (SD)	27.9 (4.2)
Waist-to-hip ratio, mean (SD) ^b	0.87 (0.09)
Waist-to-hip ratio in women, mean (SD)	0.82 (0.07)
Waist-to-hip ratio in men, mean (SD)	0.94 (0.07)
Waist circumference, mean cm (SD) ^c	90 (13.5)
Waist circumference in women, mean cm (SD)	85 (12.5)
Waist circumference in men, mean cm (SD)	97 (11.4)
Hip circumference, mean cm (SD) ^d	103 (9.2)
Hip circumference in women, mean cm (SD)	103 (10.3)
Hip circumference in men, mean cm (SD)	104 (7.6)
Systolic blood pressure, mean mmHg (SD) ^e	138 (19)
Systolic blood pressure in women, mean mmHg (SD) ^e	135 (19)
Systolic blood pressure in men, mean mmHg (SD) ^e	141 (17)
Diastolic blood pressure, mean mmHg (SD) ^f	82 (10)
Diastolic blood pressure in women, mean mmHg (SD) ^f	81 (10)
Diastolic blood pressure in men, mean mmHg (SD) ^f	84 (10)

520

521

522

523

524

525

526

527

528

a Missing in 1,594 participants (0.4%).

b Missing in 883 participants (0.2%).

c Missing in 790 participants (0.2%).

d Missing in 838 participants (0.2%).

e Missing in 863 participants (0.2%).

f Missing in 850 participants (0.2%).

Exact numbers of participants included in each genetic analysis are in eTable 1.

Abbreviations: N, number of participants; BMI, body mass index; SD, standard deviation.

529 **Figure legends**

530
531
532 **Figure 1. Associations with compartmental fat mass of polygenic scores for higher WHR. Panel A**
533 shows associations with compartmental fat mass for the 202-variant general polygenic score for higher
534 WHR. Associations are reported in clinical or standardized units of continuous outcome per 1 SD increase
535 in $WHR_{BMI-adjusted}$ (corresponding to 0.056 ratio units of age-, sex- and BMI-residualized WHR in UK
536 Biobank) due to the polygenic score. The statistical significance threshold for analyses reported in this
537 panel was $P < .0016$. **Panel B** shows associations with compartmental fat mass for the waist- (orange) or
538 hip- (dark blue) specific polygenic scores for higher WHR. Associations were estimated in up to 18,330
539 European ancestry individuals from the UK Biobank,¹⁶ Fenland¹¹ and EPIC-Norfolk¹⁷ studies.
540 Associations are reported in clinical or standardized units of continuous outcome per 1 SD increase in
541 $WHR_{BMI-adjusted}$ (corresponding to 0.056 ratio units of age-, sex- and BMI-residualized WHR in UK
542 Biobank) due to the polygenic score used in a given analysis. The statistical significance threshold for
543 analyses reported in this panel was $P < .0016$. Abbreviations: N, number of participants; SD, standard
544 deviation; CI, confidence interval; WHR, waist-to-hip ratio; BMI, body mass index.

545
546
547 **Figure 2. Associations with cardio-metabolic risk factors and disease outcomes of waist- or hip-**
548 **specific polygenic scores for higher WHR. Panel A** shows associations with cardio-metabolic risk
549 factors for the waist- (orange) or hip- (dark blue) specific polygenic scores for higher WHR. Associations
550 are reported in clinical or standardized units of continuous outcome per 1 SD increase in $WHR_{BMI-adjusted}$
551 (corresponding to 0.056 ratio units of age-, sex- and BMI-residualized WHR in UK Biobank) due to the
552 polygenic score used in a given analysis. Data on blood pressure were from UK Biobank¹⁶; data on LDL-
553 C and triglycerides were from Global Lipids Genetics consortium²²; data on fasting insulin and fasting
554 glucose were from the Meta-analyses of Glucose and Insulin-related traits consortium^{20,21}. The statistical
555 significance threshold for analyses reported in this panel was $P < .0021$. **Panel B** shows associations with
556 type 2 diabetes and coronary artery disease risk for the waist- (orange) or hip- (dark blue) specific
557 polygenic scores for higher WHR. Associations are reported in odds ratio or absolute risk increase per 1
558 SD increase in $WHR_{BMI-adjusted}$ (corresponding to 0.056 ratio units of age-, sex- and BMI-residualized
559 WHR in UK Biobank) due to the polygenic score used in a given analysis. Associations with type 2
560 diabetes were estimated in 69,677 cases and 551,081 controls from the DIAGRAM consortium²³, EPIC-
561 InterAct¹⁸ and UK Biobank¹⁶. Associations with coronary artery disease were estimated in 85,358 cases
562 and 551,249 controls from UK Biobank¹⁶ and the CARDIoGRAMplusC4D consortium²⁴. The statistical
563 significance threshold for analyses reported in this panel was $P < .0063$. Abbreviations: N, number of
564 participants; SD, standard deviation; CI, confidence interval; LDL-C, low-density lipoprotein cholesterol;
565 WHR, waist-to-hip ratio; BMI, body mass index; OR, odds ratio; ARI, absolute risk increase; py,
566 participant-years of follow-up.

567 **References**

568
569

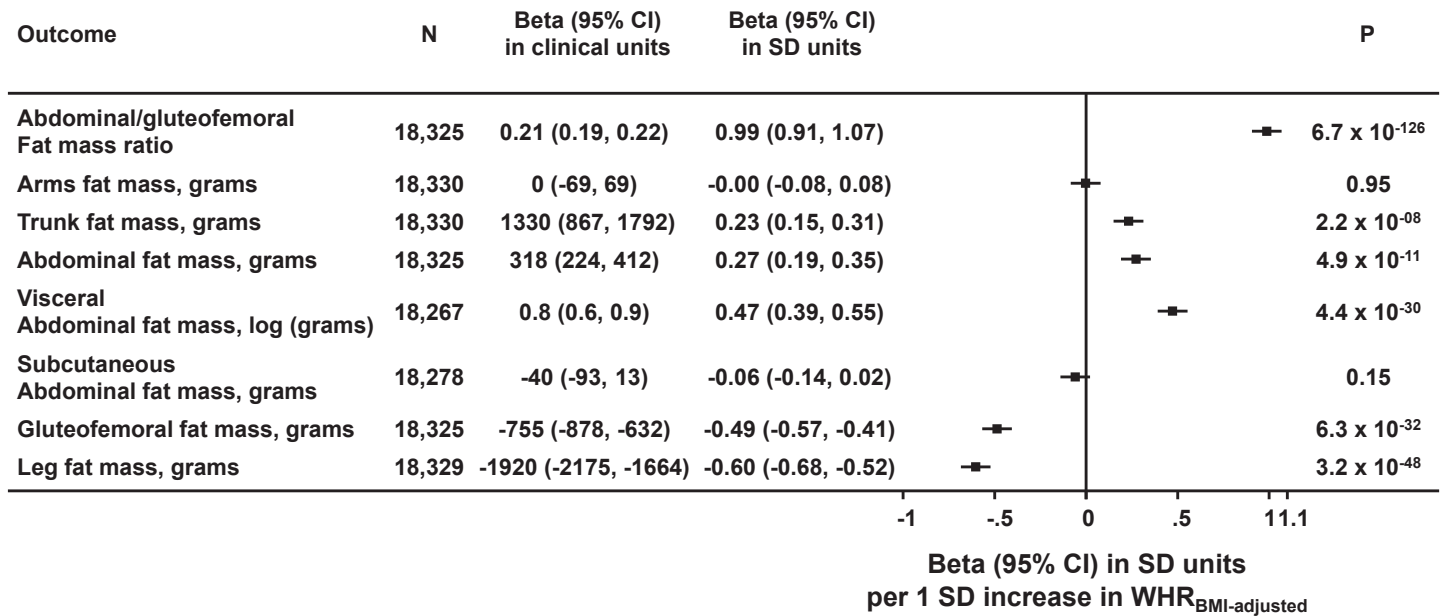
- 570 1. Vague J. The degree of masculine differentiation of obesities: a factor determining predisposition
571 to diabetes, atherosclerosis, gout, and uric calculous disease. *The American journal of clinical*
572 *nutrition*. Jan-Feb 1956;4(1):20-34.
- 573 2. Yusuf S, Hawken S, Ounpuu S, et al. Obesity and the risk of myocardial infarction in 27,000
574 participants from 52 countries: a case-control study. *Lancet*. Nov 5 2005;366(9497):1640-1649.
- 575 3. Biggs ML, Mukamal KJ, Luchsinger JA, et al. Association between adiposity in midlife and older
576 age and risk of diabetes in older adults. *JAMA : the journal of the American Medical Association*.
577 Jun 23 2010;303(24):2504-2512.
- 578 4. InterAct Consortium, Langenberg C, Sharp SJ, et al. Long-term risk of incident type 2 diabetes
579 and measures of overall and regional obesity: the EPIC-InterAct case-cohort study. *PLoS*
580 *medicine*. 2012;9(6):e1001230.
- 581 5. Stefan N, Häring HU, Hu FB, Schulze MB. Metabolically healthy obesity: epidemiology,
582 mechanisms, and clinical implications. *Lancet Diabetes Endocrinol*. 2013;1(2):10.
- 583 6. Garg A. Acquired and inherited lipodystrophies. *The New England journal of medicine*. Mar 18
584 2004;350(12):1220-1234.
- 585 7. Semple RK, Savage DB, Cochran EK, Gorden P, O'Rahilly S. Genetic syndromes of severe
586 insulin resistance. *Endocrine reviews*. Aug 2011;32(4):498-514.
- 587 8. Karpe F, Pinnick KE. Biology of upper-body and lower-body adipose tissue--link to whole-body
588 phenotypes. *Nature reviews. Endocrinology*. Feb 2015;11(2):90-100.
- 589 9. Ryden M, Andersson DP, Bergstrom IB, Arner P. Adipose tissue and metabolic alterations:
590 regional differences in fat cell size and number matter, but differently: a cross-sectional study. *The*
591 *Journal of clinical endocrinology and metabolism*. Oct 2014;99(10):E1870-1876.
- 592 10. Dahlman I, Ryden M, Brodin D, Grallert H, Strawbridge RJ, Arner P. Numerous Genes in Loci
593 Associated With Body Fat Distribution Are Linked to Adipose Function. *Diabetes*. Feb
594 2016;65(2):433-437.
- 595 11. Lotta LA, Gulati P, Day FR, et al. Integrative genomic analysis implicates limited peripheral
596 adipose storage capacity in the pathogenesis of human insulin resistance. *Nature genetics*. Jan
597 2017;49(1):17-26.
- 598 12. Scott RA, Fall T, Pasko D, et al. Common genetic variants highlight the role of insulin resistance
599 and body fat distribution in type 2 diabetes, independent of obesity. *Diabetes*. Dec
600 2014;63(12):4378-4387.
- 601 13. Yaghoobkar H, Scott RA, White CC, et al. Genetic evidence for a normal-weight "metabolically
602 obese" phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2
603 diabetes. *Diabetes*. Dec 2014;63(12):4369-4377.
- 604 14. Yaghoobkar H, Lotta LA, Tyrrell J, et al. Genetic Evidence for a Link Between Favorable
605 Adiposity and Lower Risk of Type 2 Diabetes, Hypertension, and Heart Disease. *Diabetes*. Aug
606 2016;65(8):2448-2460.
- 607 15. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin
608 biology to body fat distribution. *Nature*. Feb 12 2015;518(7538):187-196.
- 609 16. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the
610 causes of a wide range of complex diseases of middle and old age. *PLoS medicine*. Mar
611 2015;12(3):e1001779.
- 612 17. Day N, Oakes S, Luben R, et al. EPIC-Norfolk: study design and characteristics of the cohort.
613 European Prospective Investigation of Cancer. *British journal of cancer*. Jul 1999;80 Suppl 1:95-
614 103.
- 615 18. InterAct Consortium, Langenberg C, Sharp S, et al. Design and cohort description of the InterAct
616 Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type
617 2 diabetes in the EPIC Study. *Diabetologia*. Sep 2011;54(9):2272-2282.

- 618 19. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for
619 obesity biology. *Nature*. Feb 12 2015;518(7538):197-206.
- 620 20. Scott RA, Lagou V, Welch RP, et al. Large-scale association analyses identify new loci
621 influencing glyceic traits and provide insight into the underlying biological pathways. *Nature*
622 *genetics*. Sep 2012;44(9):991-1005.
- 623 21. Manning AK, Hivert MF, Scott RA, et al. A genome-wide approach accounting for body mass
624 index identifies genetic variants influencing fasting glyceic traits and insulin resistance. *Nature*
625 *genetics*. Jun 2012;44(6):659-669.
- 626 22. Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, et al. Discovery and refinement of
627 loci associated with lipid levels. *Nature genetics*. Nov 2013;45(11):1274-1283.
- 628 23. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights
629 into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics*. Sep
630 2012;44(9):981-990.
- 631 24. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide
632 association meta-analysis of coronary artery disease. *Nature genetics*. Oct 2015;47(10):1121-
633 1130.
- 634 25. Eastwood SV, Mathur R, Atkinson M, et al. Algorithms for the Capture and Adjudication of
635 Prevalent and Incident Diabetes in UK Biobank. *PLoS one*. 2016;11(9):e0162388.
- 636 26. Emdin CA, Khera AV, Natarajan P, et al. Genetic Association of Waist-to-Hip Ratio With
637 Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. *JAMA : the journal of the*
638 *American Medical Association*. Feb 14 2017;317(6):626-634.
- 639 27. Loh PR, Tucker G, Bulik-Sullivan BK, et al. Efficient Bayesian mixed-model analysis increases
640 association power in large cohorts. *Nature genetics*. Mar 2015;47(3):284-290.
- 641 28. Loh PR, Kichaev G, Gazal S, Schoech AP, Price AL. Mixed-model association for biobank-scale
642 datasets. *Nature genetics*. Jul 2018;50(7):906-908.
- 643 29. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide
644 association scans. *Bioinformatics*. Sep 1 2010;26(17):2190-2191.
- 645 30. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of
646 seven common diseases and 3,000 shared controls. *Nature*. Jun 7 2007;447(7145):661-678.
- 647 31. Aschard H, Vilhjalmsson BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable covariates can
648 bias effect estimates in genome-wide association studies. *American journal of human genetics*.
649 Feb 5 2015;96(2):329-339.
- 650 32. Stock JH WJ, Yogo M. A survey of weak instruments and weak identification in generalized
651 method of moments. *J Bus Econ Stat*. 2002;20(4):518-529.
- 652 33. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple
653 genetic variants using summarized data. *Genetic epidemiology*. Nov 2013;37(7):658-665.
- 654 34. Burgess S, Thompson DJ, Rees JMB, Day FR, Perry JR, Ong KK. Dissecting Causal Pathways
655 Using Mendelian Randomization with Summarized Genetic Data: Application to Age at
656 Menarche and Risk of Breast Cancer. *Genetics*. Oct 2017;207(2):481-487.
- 657 35. Dale CE, Fatemifar G, Palmer TM, et al. Causal Associations of Adiposity and Body Fat
658 Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A
659 Mendelian Randomization Analysis. *Circulation*. Jun 13 2017;135(24):2373-2388.
- 660 36. Danforth E, Jr. Failure of adipocyte differentiation causes type II diabetes mellitus? *Nature*
661 *genetics*. Sep 2000;26(1):13.
- 662 37. Virtue S, Vidal-Puig A. Adipose tissue expandability, lipotoxicity and the Metabolic Syndrome--
663 an allostatic perspective. *Biochimica et biophysica acta*. Mar 2010;1801(3):338-349.
- 664 38. Haslam DW, James WP. Obesity. *Lancet*. Oct 01 2005;366(9492):1197-1209.
- 665 39. Kan M, Auer PL, Wang GT, et al. Rare variant associations with waist-to-hip ratio in European-
666 American and African-American women from the NHLBI-Exome Sequencing Project. *European*
667 *journal of human genetics : EJHG*. Aug 2016;24(8):1181-1187.

- 668 **40.** Chu AY, Deng X, Fisher VA, et al. Multiethnic genome-wide meta-analysis of ectopic fat depots
669 identifies loci associated with adipocyte development and differentiation. *Nature genetics*. Jan
670 2017;49(1):125-130.
671
672
673

Figure 1

A



B

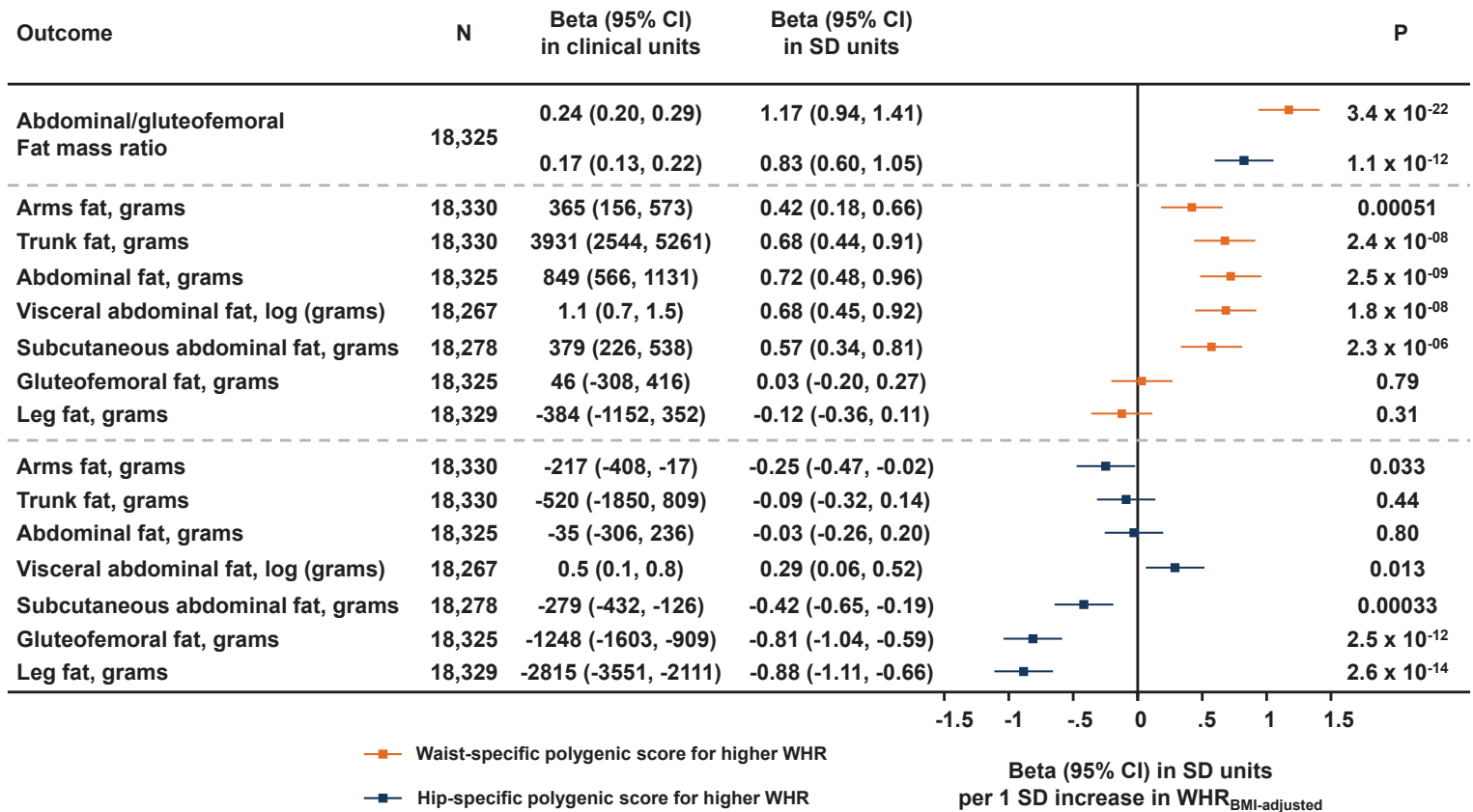
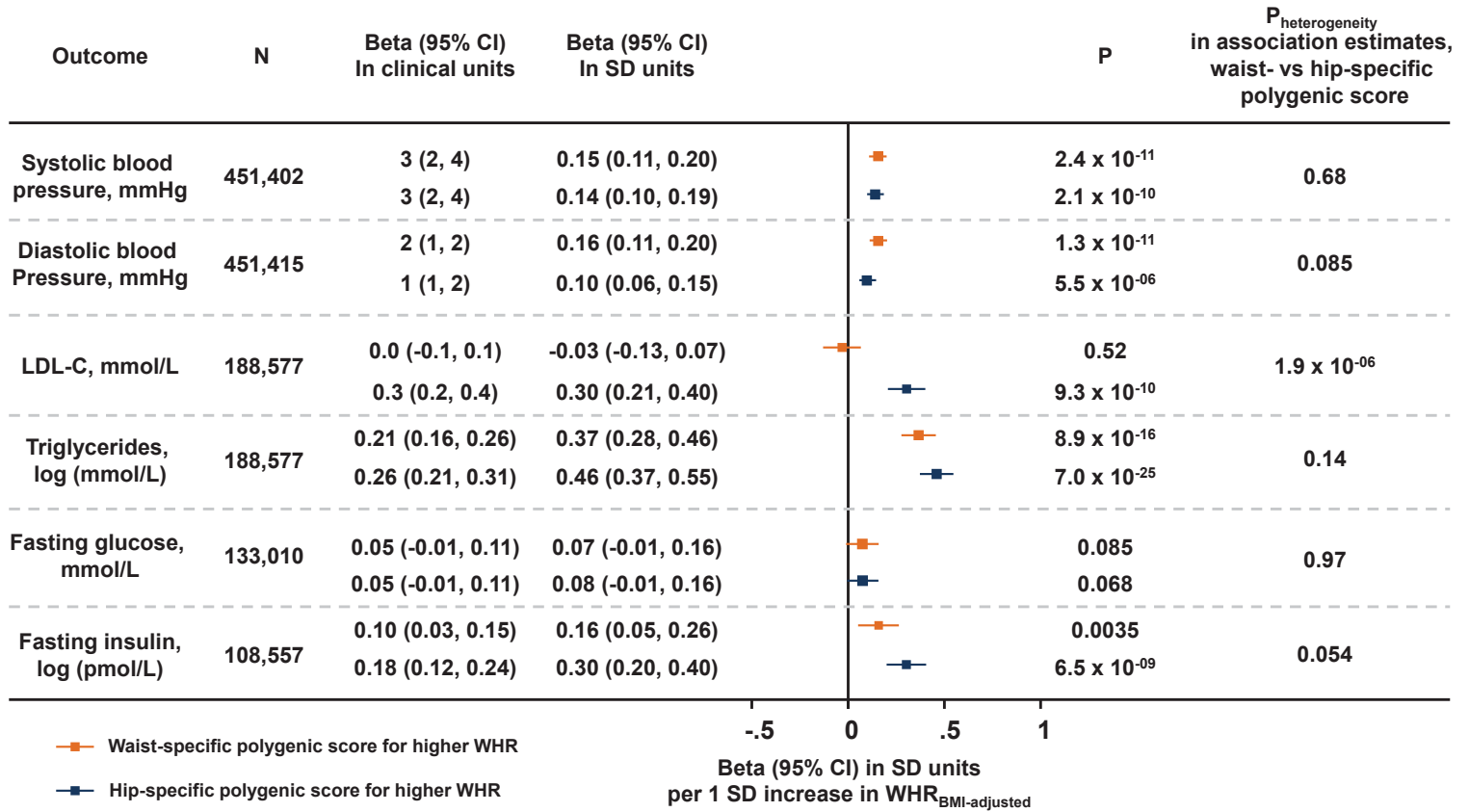


Figure 2

A



B

