Genome-wide association study of anti-Müllerian hormone levels in pre-menopausal women of late reproductive age and relationship with genetic determinants of reproductive lifespan

Katherine S. Ruth[1], Ana Luiza G. Soares[2,3], Maria-Carolina Borges[2,3], A. Heather Eliassen[4], Susan E. Hankinson[4,5], Michael E. Jones[6], Peter Kraft[7], Hazel B. Nichols[8], Dale P. Sandler[9], Minouk J. Schoemaker[6], Jack A. Taylor[10], Anne Zeleniuch-Jacquotte[11], Deborah A. Lawlor[2,3,12], Anthony J. Swerdlow[6,13], Anna Murray[1,*]

[*]To whom correspondence should be addressed at: Genetics of Complex Traits, University of Exeter Medical School, RILD Level 3, Royal Devon & Exeter Hospital, Barrack Road, Exeter, EX2 5DW, UK; Tel: +44 (0)1392 408257; Email: A.Murray@exeter.ac.uk

[1]Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, UK.

[2]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK.

[3]Population Health Science, Bristol Medical School, University of Bristol, Bristol, UK.

[4]Department of Epidemiology, Harvard T.H. Chan School of Public Health, and Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.

[5]Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA, USA.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

23 [6]Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK.

24 [7]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

25 [8]Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA.

26 [9]Epidemiology Branch, National Institute of Environmental Health Sciences, Research
27 Triangle Park, NC, USA.

28 [10]Epigenetic and Stem Cell Biology Laboratory, National Institute of Environmental Health
29 Sciences, NIH, Research Triangle Park, NC, USA.

30 [11]Department of Population Health and Perlmutter Cancer Center, New York University
31 School of Medicine, New York, NY, USA.

32 [12]National Institute for Health Research Bristol Biomedical Research Centre, University
33 Hospitals Bristol NHS Foundation Trust and University of Bristol, Bristol, UK.

34 [13]Division of Breast Cancer Research, The Institute of Cancer Research, London, UK.

35

36

37

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

38 **Abstract**

39 Anti-Müllerian hormone (AMH) is required for sexual differentiation in the fetus, and in
40 adult females AMH is produced by growing ovarian follicles. Consequently, AMH levels are
41 correlated with ovarian reserve, declining towards menopause when the oocyte pool is
42 exhausted. A previous genome-wide association study identified three genetic variants in and
43 around the *AMH* gene that explained 25% of variation in AMH levels in adolescent males but
44 did not identify any genetic associations reaching genome-wide significance in adolescent
45 females. To explore the role of genetic variation in determining AMH levels in women of late
46 reproductive age, we carried out a genome-wide meta-analysis in 3,344 pre-menopausal
47 women from five cohorts (median age 44–48 years at blood draw). A single genetic variant,
48 rs16991615, previously associated with age at menopause, reached genome-wide significance
49 at P=3.48×10^{-10}, with a per allele difference in age-adjusted inverse normal AMH of 0.26 SD
50 (95% CI [0.18,0.34]). We investigated whether genetic determinants of female reproductive
51 lifespan were more generally associated with pre-menopausal AMH levels. Genetically-
52 predicted age at menarche had no robust association but genetically-predicted age at
53 menopause was associated with lower AMH levels by 0.18 SD (95% CI [0.14,0.21]) in age-
54 adjusted inverse normal AMH per one-year earlier age at menopause. Our findings provide
55 genetic support for the well-established use of AMH as a marker of ovarian reserve.

56

57

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

58 **Introduction**

59 Anti-Müllerian hormone (AMH) is a member of the transforming growth factor-beta
60 superfamily that regulates the growth and development of ovarian follicles in females and is
61 required for sexual differentiation in the fetus, causing regression of the Müllerian ducts
62 during testis development (1). In males, AMH is required for testes development and function
63 and levels increase rapidly shortly after birth, peaking at 6 months of age, and then decline to
64 low levels in during puberty (2). In women, AMH is produced by the granulosa cells of
65 growing follicles and levels are correlated with the number of growing follicles and are used
66 as a clinical measure of ovarian reserve (3). AMH levels increase in women from birth until
67 their 20s, before declining gradually with age until levels are undetectable after menopause
68 when ovarian reserve is exhausted (1, 3–6). Since AMH levels are stable throughout the
69 menstrual cycle, they can be used as a measure of fertility in women of late reproductive age
70 and to predict response to fertility treatment (7).

71 AMH levels vary widely between women and genetic variation is thought to be important,
72 though few genetic studies have been conducted. Rare *AMH* mutations have been found with
73 functional effects on AMH signalling (8, 9), while polymorphisms in *AMH* or the gene
74 coding for its receptor, *AMHR2*, have been associated with response to ovarian stimulation,
75 infertility, follicle recruitment, primary ovarian insufficiency and polycystic ovary syndrome
76 in candidate gene studies (10).

77 A previous genome-wide association study (GWAS) in 1,360 adolescent males and 1,455
78 adolescent females from a single cohort identified three genetic variants in and around the
79 *AMH* gene that were independently associated with higher levels of AMH in adolescent
80 males (P=2×10^{-49} to P=3×10^{-8} for each variant when jointly included in the regression model)
81 (11). None of these variants showed strong evidence of statistical association in adolescent

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

82 females (P=8×10^{-4} to P=0.9 for each variant when jointly included in the regression model),
83 with considerably weaker effect estimates than in males. For all three variants there was
84 strong statistical evidence of a sex difference (P_{HET}=3×10^{-4} to P_{HET}=6×10^{-12}), with the three
85 cumulatively explaining 24.5% of the variation in AMH levels in males compared with 0.8%
86 in females. No cohorts were available for replication of this initial study and it is unknown
87 whether the weak or absent association in adolescent females persists into older ages, as
88 would be expected since differences in ovarian decline result in variation in AMH levels
89 between women.

90 We undertook a GWAS meta-analysis of 3,344 women from five cohorts – the Generations
91 Study, Sister Study, Nurses' Health Study, Nurses' Health Study II and Avon Longitudinal
92 Study of Parents and Children (ALSPAC) − to investigate genetic determinants of AMH
93 levels in pre-menopausal women of late reproductive age (median age at blood draw 44–48
94 years). We aimed to identify novel genetic variants associated with AMH levels and to
95 explore the effects of published genetic variants associated with AMH levels in previous
96 GWAS and candidate gene studies.

97

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

98 **Results**

99 *AMH is associated with a single significant signal in a known menopause locus*

100 In our genome-wide meta-analysis (Table 1), a single genetic variant in the *MCM8* gene at
101 20p12.3 reached genome-wide significance at $P<5\times10^{-8}$ (rs16991615, $P=3.48\times10^{-10}$) (Figure
102 1, Supplementary Figure 1). Within each of the five genotyped cohorts, we inverse-normally
103 transformed AMH (to ensure normality of the residuals in the association analysis) and tested
104 the association of over 11 million autosomal genetic variants imputed to HRC r1.1 2016 (12)
105 adjusted for age and genetic relatedness (13). We performed inverse variance weighted meta-
106 analyses of the genome-wide results from the five cohorts, filtering our results to include only
107 variants present in three or more of the five cohorts analysed, resulting in a total of 8.4
108 million variants in the final results dataset (summary statistics to be made available through
109 the EMBL-EBI GWAS catalogue at https://www.ebi.ac.uk/gwas/summary-statistics). A total
110 of 242 variants had $P<1\times10^{-5}$, resulting in 24 signals following distance-based clumping of
111 variants within 500kb, with the top ten signals presented in Table 2. The minor A allele of
112 rs16991615 increased age-adjusted inverse normal AMH by 0.26 SD per allele (95% CI
113 [0.18,0.34], $P=3.48\times10^{-10}$) (Table 2).

114

115 *Variants previously shown to be strongly associated with AMH levels in adolescent males*
116 *had weak effects in pre-menopausal adult women in our study*

117 For three genetic variants in and around the *AMH* gene that were previously found to be
118 independently associated with higher levels of AMH in adolescent males (11), we estimated
119 the effects in pre-menopausal women when the variants were jointly included in the
120 regression model (joint model), by carrying out approximate conditional analyses using the
121 software GCTA (14). To allow comparison between our results and the original study's

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

122 estimates, we generated effect estimates for age-adjusted inverse normal AMH in the
123 adolescent males and females from the original study sample (ALSPAC offspring), since
124 results from the original study were unadjusted and presented in natural log-transformed
125 AMH.

126 The effect estimates from the joint model for the three published genetic variants were
127 directionally concordant across adolescent males and females (ALSPAC offspring cohort)
128 and the pre-menopausal women in the current study (from five cohorts including ALSPAC
129 mothers), but had about one-fifth of the effect on the level of AMH compared with the effect
130 in adolescent males (P_{HET}<0.001) (Figure 2). The weak or null effect sizes for rs4807216,
131 rs8112524 and rs2385821 were similar in adolescent and pre-menopausal females (P_{HET}>0.05
132 for all). Genetic variant rs2385821 had the strongest effects in females of the three variants
133 from the previous publication (11), but did not reach genome-wide significance in pre-
134 menopausal females (for joint model, per allele difference in age-adjusted inverse normal
135 AMH of 0.27 SD (95% CI [0.13,0.41]), P=4.0×10^{-5}) (Supplementary Table 1).

136

137 *Genetic variants for early menopause are associated with reduced levels of AMH*

138 Since the only genetic variant to reach genome-wide significance in our study (rs16991615)
139 has previously been reported as associated with menopause timing (15–17), we investigated
140 the association of AMH levels with all 56 genetic variants associated with menopause timing
141 (17). The effect estimates of these genetic variants on AMH level were positively correlated
142 with the published effects on age at menopause (r=0.83) and there were consistent directions
143 of effect for 50/56 variants (P=1×10^{-9} for binomial sign test; χ^2_{56}=194.39, P=4×10^{-17} for
144 global chi-squared test of association) (Figure 3, Supplementary Table 2). There were no

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

145 obvious outliers among the 56 menopause timing variants and generally variants with large

146 effects on age at menopause also had large effects on AMH levels.

147 Two-sample Mendelian randomization analysis by inverse variance weighted (IVW) and

148 Egger estimation supported a causal relationship between genetically-predicted age at

149 menopause and pre-menopausal AMH level (Figure 3, Supplementary Table 2). For a one-

150 year increase in genetically-predicted age at menopause, age-adjusted inverse normal AMH

151 was increased by 0.18 SD (95% CI [0.14,0.21]) with no horizontal pleiotropy detected by

152 Egger analysis (in Egger analysis, 0.20 SD (95% CI [0.13,0.27]) age-adjusted inverse normal

153 AMH per a one-year increase in genetically-predicted age at menopause one-year increase in

154 genetically-predicted age at menopause, *P-intercept*=0.49). This relationship remained

155 similar even when rs16991615 was excluded from the analysis (Supplementary Table 3).

156

157 *Genetic variants for age at menarche are not associated with AMH levels*

158 We investigated the effect of genetic variants associated with age at menarche on AMH

159 levels (18), since menarche marks the start of cyclic selection of ovarian follicles from the

160 growing follicle pool. We identified 327 of 389 published independent signals (18) in our

161 meta-analysis results. For the 327 variants, there was little correlation between the published

162 effect on age at menarche and the effect on AMH levels (r=-0.05) and the directions of

163 effects were not consistent, with 158/327 (48%) in the same direction (P=0.58 for binomial

164 sign test; χ^2_{327}=328.62, P=0.46 for global chi-squared test of association) (Figure 3,

165 Supplementary Table 4). Two-sample Mendelian randomization analysis by IVW and Egger

166 estimation found no causal relationship between age at menarche and AMH level (difference

167 in mean age-adjusted inverse normal AMH per one-year increase in genetically-predicted age

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

168 at menopause was for IVW, -0.05 SD (95% CI [-0.12,0.02]) and for Egger, -0.03 SD (95% CI

169 [-0.22,0.15], *P-intercept*=0.87) (Figure 3, Supplementary Table 3).

170

171 *Genetic variant for follicle-stimulating hormone levels is not associated with AMH levels*

172 Since levels of follicle-stimulating hormone (FSH) and luteinising hormone (LH) rise around

173 menopause, we tested the association of a genetic variant at the *FSHB* locus that affects levels

174 of these hormones (19) with AMH levels in pre-menopausal women. The *FSHB* promoter

175 polymorphism (rs10835638; -211G>T) was not associated with AMH levels (per allele

176 difference in age-adjusted inverse normal AMH of 0.01 SD (95% CI [-0.05,0.07]), P=0.79).

177

178 *Sensitivity analysis*

179 Results for genetic variants with $P<5\times10^{-5}$ in the main analysis were well-correlated whether

180 we adjusted for age or not (r=0.99) (Supplementary Figure 2), used our favoured inverse

181 normal transformation or a natural log transformation (as in reference (11)) (r=1.00)

182 (Supplementary Figure 3), and when we excluded women whose AMH level was imputed as

183 it was below the lower limit of detection (r=0.99) (Supplementary Figure 4).

184

185

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

186 **Discussion**

187 Our results indicate that variation in AMH levels in pre-menopausal women is contributed to
188 by the underlying biology of ovarian reserve, as shown by the correlation between
189 genetically-predicted age at menopause and AMH levels, supporting the use of AMH as a
190 means of measuring ovarian reserve. The only signal passing genome-wide significance in
191 our analyses was rs16991615 in *MCM8*, a published menopause timing variant (15–17), with
192 the same allele associated with earlier menopause and lower AMH levels. Genome-wide
193 analyses of menopause timing, a proxy measure for ovarian reserve, have identified 56
194 genetic variants and highlighted the importance of DNA damage response pathways during
195 follicle formation *in utero* and for follicle maintenance during a woman's lifetime (17).
196 Additionally, 389 genetic variants have been identified for menarche timing, the age at which
197 cyclic selection of ovarian follicles from the pool of growing, AMH-producing follicles starts
198 (18). Therefore, it is plausible that genetic determinants of menarche and menopause timing
199 could affect ovarian reserve and influence AMH levels in pre-menopausal women, many
200 years prior to menopause. For genetic variants associated with age at menopause, the
201 published effect estimates were positively correlated with the effects on AMH levels and
202 there was evidence of a causal relationship between genetically-predicted earlier menopause
203 and lower pre-menopausal AMH levels, which remained even when rs16991615 was
204 excluded. We did not find any association between genetically-predicted age at menarche and
205 pre-menopausal AMH levels. We interpret these results as suggesting that AMH levels in
206 pre-menopausal women are determined by declining ovarian reserve as a result of
207 reproductive ageing but not menarche timing, and that women with lower AMH are nearer to
208 the end of their reproductive lifespan.

209 Variant rs16991615 has previously been found to be associated with differences in
210 menopause timing by 0.9 years per allele and is a missense variant in exon 9 of *MCM8*

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

211 (E341K), which is required for homologous recombination (20). Other mutations in *MCM8*
212 causing reduced double strand break repair have been found in women with premature
213 ovarian failure (21) and follicle development is arrested at an early stage in *MCM8* knockout
214 mice (22). Pathway analysis has shown that the menopause timing variants identified from
215 genome-wide analyses are enriched for genes involved in DNA damage response, including
216 double-strand break repair during meiosis, suggesting that the genetic determinants of age at
217 menopause act during ovarian follicle formation or maintenance, potentially affecting ovarian
218 reserve from before birth until menopause (18). Therefore, it seems likely that rs16991615
219 affects AMH levels through differences in ovarian reserve.

220 Although the three published GWAS signals in and around the *AMH* gene (11) did not reach
221 genome-wide significance when jointly included in the regression analysis, they did show
222 directional consistency and were nominally associated in the pre-menopausal women in our
223 study. The previous GWAS of adolescent females included fewer samples (n=1,455) than our
224 analysis (n=3,344), hence we were better powered to detect the effects of these variants in
225 females. The three published variants for AMH had smaller effects in pre-menopausal
226 women compared with adolescent males but were consistent with the associations seen
227 previously in adolescent females. The strongest signal in the GWAS of AMH levels in
228 adolescents, rs4807216, was not associated with age at menopause in the most recent
229 genome-wide meta-analysis (per allele difference in age-adjusted inverse normal AMH of
230 0.05 SD (95% CI [-0.05, 0.15]), P=0.37) (17). Differences in genetic regulation of AMH
231 levels in males and females are plausible given AMH's different function in men and women.
232 In males, AMH is required for regression of the Müllerian ducts during testis development in
233 the fetus, and is involved in testicular development and function (2). In females, AMH is
234 produced by granulosa cells of primary, pre-antral and small antral follicles, inhibiting both
235 the further recruitment of primordial follicles from the follicle pool and also FSH-dependent

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

236 selection of follicles for growth during the menstrual cycle (3, 7, 23, 24). AMH expression
237 starts *in utero* at 36 week's gestation, peaking at around 25 years before declining to
238 undetectable levels at menopause (3, 5, 7). In our analysis, a polymorphism in the promoter
239 of *FSHB* (-211G>T) that affects FSH levels (25–27) had no effect on AMH levels in pre-
240 menopausal women, supporting the absence of direct negative feedback of FSH on AMH.

241 AMH levels vary widely between women (6), reflecting factors such as variation in ovarian
242 reserve, age and ethnicity (13). We controlled for age and ethnicity by adjusting for age and
243 restricting our analyses to genetically European individuals. Adjustment for age will remove
244 a source of variation in AMH level, the effect of the decrease in the primordial follicle pool
245 with age, highlighting the effect of genetic variants responsible for variation in the initial size
246 of the primordial follicle pool or that either accelerate or protect against loss of ovarian
247 reserve with age.

248 We would have been unable to detect low frequency variants or those with a smaller effect
249 size since we were only powered (>80%) to detect a variant with a MAF of 5% and an effect
250 of 0.36 SD or greater in the sample size analysed. We were unable to evaluate the effect of
251 time from study participation to menopause on our results, to investigate whether the
252 association of the menopause variants was modified by proximity to menopause, since we did
253 not have sufficient follow-up data. Future analysis should consider stratifying by participants'
254 time to menopause. However, such analyses would require large numbers of women who had
255 pre-menopausal AMH measurements (AMH levels are generally undetectable post-
256 menopause), recorded age at menopause and varying times of follow-up since age at
257 menopause. We are not aware of any such study currently but with continued follow-up of
258 the women included in this study such analyses may be possible.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

259 This study confirms genetically that AMH levels are a marker for ovarian decline and
260 reproductive ageing in pre-menopausal women. In addition to its use as a marker of fertility,
261 there is evidence that AMH is a biomarker of breast cancer risk in pre-menopausal women. In
262 a recent study, odds of pre- and postmenopausal breast cancer were 60% higher in women in
263 the highest quartile of AMH level compared with the lowest, even after adjusting for
264 potential confounders such as age (28). Our study suggests that these effects could be
265 mediated through preserved ovarian reserve, or a correlate of ovarian reserve, as a result of
266 delayed reproductive ageing, supported by findings from a large scale genomic analysis that
267 showed a causal effect of later menopause on increased risk of breast cancer by 6% per year
268 (17) and strong epidemiological evidence that later age at menopause increases risk of breast
269 cancer (29). Our study provides genetic evidence that underlying biological factors
270 responsible for reproductive ageing contribute to AMH levels in pre-menopausal women and
271 are likely to be the main driver for the observed associations of AMH with health outcomes.

272

273

274

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

275 **Materials and Methods**

276 *Studies included*

277 The central analysis team at University of Exeter Medical School coordinated data collection
278 from the five studies. We included 3,344 women who had pre-menopausal AMH levels
279 measured, who were participants in the Generations Study (30), the Sister Study (31), the
280 Nurses' Health Study, the Nurses' Health Study II (32) and ALSPAC (33–35) (Table 1)
281 (Supplementary Information). For the Generations Study, the Sister Study and the Nurses'
282 Health Studies, genotype and phenotype data were provided to the central analysis team for
283 quality control, cleaning and analysis. For the ALSPAC study, quality control and genotype-
284 phenotype analyses were undertaken in house and summary descriptive, GWAS and
285 sensitivity analyses statistics were provided to the central analysis team for meta-analysis.

286

287 *Genetic data*

288 In the Generations Study, Sisters and the Nurses' Health Studies, samples were genotyped on
289 the OncoArray array (Table 1). For the Nurses' Health Studies, a further 225 samples were
290 genotyped on an Illumina array. For each cohort and array type, data were cleaned using a
291 standard quality control process in PLINK v1.9 (www.cog-genomics.org/plink/1.9/) (36).
292 SNPs were removed if they were poorly genotyped (missing in >5% samples) or were not in
293 Hardy-Weinberg equilibrium ($P<1\times10^{-6}$). Samples were removed if they were poorly
294 genotyped (missing >5% SNPs), were a sex mismatch or were outliers in terms of
295 heterozygosity. Within each cohort, samples that were related to each other as 3rd degree
296 relatives or closer were identified and the sample with the greater proportion of missing SNPs
297 was removed. Principal component analysis was carried out in FlashPCA in order to identify
298 and remove genetically non-European samples from the analysis. SNPs with MAF>1%

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

299 aligned to the correct strand in HRC v1.1 were used for imputation. Genotypes for
300 chromosomes 1–22 were phased using SHAPEIT and imputed to HRC r1.1 2016 using the
301 University of Michigan Imputation Server (https://imputationserver.sph.umich.edu/) (12, 37,
302 38).

303 ALSPAC mothers were genotyped using the Illumina Human660W-Quad array at Centre
304 National de Génotypage (CNG) and genotypes were called with Illumina GenomeStudio.
305 Quality control was performed in PLINK v1.07 (39) by removing poorly genotyped SNPs
306 (missing in >5% samples), not in Hardy-Weinberg equilibrium (P<1×10-6), or that had
307 MAF<1%. Samples were removed if they were poorly genotyped (missing >5% SNPs), had
308 indeterminate X chromosome heterozygosity or extreme autosomal heterozygosity. Samples
309 showing evidence of population stratification were identified by multidimensional scaling of
310 genome-wide identity by state (IBS) pairwise distances using the four HapMap populations
311 as a reference, and then excluded (IBS>0.125). Haplotypes were estimated using SHAPEIT
312 (v2.r644) (40) and phased haplotypes were then imputed to HRC panel (12) using IMPUTE
313 V3.

314

315 *AMH phenotype*

316 For the Generations Study, Sister Study and the Nurses' Health Studies, AMH levels were
317 measured in blood samples taken from pre-menopausal women before breast cancer
318 incidence by the individual studies as part of a collaborative, prospective study of AMH and
319 breast cancer risk (28). Serum AMH levels were measured using an ultrasensitive ELISA
320 (Ansh Labs, Webster, TX) (Sister Study) or a picoAMH enzyme-linked immunoabsorbent
321 assay (Ansh Labs, Webster, TX) (Generations Study and the Nurses' Health Studies, and
322 samples below limit of detection of ELISA in Sister Study) (28). In ALSPAC, blood samples

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

323 were taken following a standardized protocol in women who attended a series of clinic
324 assessments starting about 18 years after the index pregnancy and fasted (overnight or a
325 minimum of 6 hours for those assessed in the afternoon) serum AMH levels were measured
326 using the Beckman Coulter AMH Gen II ELISA assay (34, 35).

327 For samples with AMH below the lower limit of detection, levels were imputed: for the
328 Generations Study, the value was the midpoint between zero and the lower limit of detection
329 (0.00821 pmol/L); for the Sister Study, missing values were imputed as 0.0015 ng/mL to be
330 consistent with the previous analysis (28); for ALSPAC, measured AMH values <0.01 ng/mL
331 were imputed to be 0.01 ng/mL. A small number of women (n=24) in the Nurses' Health
332 Studies with AMH below the lower limit of detection (2.038 pg/mL) were excluded from the
333 analyses. For all studies, measured values of AMH were converted to pmol/L using 1
334 pg/mL=0.00714 pmol/L and 1 ng/mL= 7.14 pmol/L. AMH was transformed by inverse
335 normal transformation, in which the rank of the AMH value was converted to the z-score for
336 the corresponding quantile of a standard normal distribution. This was done in order to
337 approximate a normal distribution to ensure normality of the residuals in the association
338 analysis and to rank the measured AMH values from each cohort to a consistent scale,
339 avoiding the issue of consistently higher/lower AMH measurements due to the assay used.
340 AMH levels for each study are summarised in Table 1.

341

342 *Genome-wide analysis*

343 Genome-wide linear regression analysis was carried separately for each of the five cohorts
344 for autosomal genetic variants with imputation quality>0.4 assuming an additive model. Age
345 at time the blood sample was taken was included as a covariate since age is known to be
346 strongly associated with AMH levels and was negatively correlated with AMH level in

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

347 exploratory analysis (median age in each study is summarised in Table 1). For the
348 Generations Study, Sisters and Nurses' Health Study, analysis was carried out using
349 GEMMA 0.94.1 (41), which calculates a genetic relationship matrix to account for cryptic
350 relatedness and population stratification between the samples. The genetic relationship matrix
351 was created from a pruned list of uncorrelated SNPs created in PLINK 1.9 (www.cog-
352 genomics.org/plink/1.9/) (36) using --indep-pairwise, excluding regions of long range linkage
353 disequilibrium, based on variants with MAF>0.01, excluding variants with r^2>0.5 (window
354 size of 1000, calculated in steps of 50). The analysis included approximately 11.8 million
355 genetic variants for the Generations Study, 12.9 million for the Sister Study, 12.3 million for
356 Nurses' Health Study OncoArray and 11.1 million for Nurses' Health Study Illumina. For
357 ALSPAC, the analysis was carried out in SNPTESTv2.5 (42) adjusting for the top ten
358 principal components of ancestry which resulted in approximately 14.7 million SNPs.

359 Standard error weighted meta-analysis of the individual GWAS results was carried out in
360 METAL (version 2011-03-25) (43) with genomic control applied to account for inflation due
361 to any remaining population stratification. Genetic variants included in the meta-analysis had
362 imputation>0.4 and minor allele count>5 (calculated from allele frequencies), resulting in a
363 total of 11.2 million autosomal SNPs that were analysed. Approximately 8.4 million variants
364 were present in three or more of the five datasets analysed and were included in our final
365 results. We identified independent signals as being suggestive of genome-wide association if
366 they had $P<1\times10^{-5}$ and were more than 500kb from another signal; from these, we identified
367 signals reaching genome-wide significance at $P<5\times10^{-8}$.

368 Manhattan and quantile–quantile plots for the genome-wide association results were created
369 using the package *qqman* (44) in R (The R Foundation for Statistical Computing).
370 LocusZoom v1.4 (45) was used to plot the association statistics with age-adjusted inverse
371 normal AMH for variants within 500kb of the top variant, showing linkage disequilibrium.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

372 Linkage disequilibrium was calculated in PLINK v1.9 (www.cog-genomics.org/plink/1.9/)
373 (36) from best guess genotypes for 1000 Genomes Phase 3/HRC imputed variants in
374 ~340,000 unrelated Europeans from the UK Biobank study (46).

375

376 *Generation of age-adjusted inverse normal effect estimates in ALSPAC offspring cohort*

377 For three published genetic variants that were associated with AMH levels in adolescent
378 males in the ALSPAC offspring cohort (11), we generated effect estimates for age-adjusted
379 inverse normal transformed AMH in the original study sample, since the original published
380 estimates were unadjusted and presented in natural log-transformed AMH. Analyses were
381 carried in SNPTEST v2.5 (42) adjusting for the top ten principal components of ancestry and
382 age and excluding the most extreme 1% of measured AMH values, resulting in 1,312 males
383 and 1,421 females from the ALSPAC offspring cohort for analysis. Other methods were as
384 described previously (11).

385

386 *Estimation of joint effects of variants in and around the* AMH *gene*

387 We used GCTA (version 1.25.0), using the command --cojo-joint (14), to carry out an
388 approximate conditional analysis to estimate the joint effects of three genetic variants in and
389 around the *AMH* gene (11). Linkage disequilibrium between the variants was estimated using
390 a random sample of 8,569 white British individuals from the UK Biobank May 2015 interim
391 release of imputed genetic data (47).

392

393 *Comparison of effects for published traits with AMH results*

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

394 We compared the published effect estimates for 56 genetic variants associated with age at
395 menopause (17) and 389 genetic variants associated with age at menarche (18) with their
396 effects on AMH level in our analysis, by carrying out a Binomial sign test of directional
397 consistency and calculating Pearson correlation coefficients. To explore whether age at
398 menopause or age at menarche causes differences in AMH levels, we used the genetic
399 variants associated with menopause and menarche timing as instruments for age at
400 menopause and age at menarche in two-sample Mendelian Randomization analyses. We used
401 the Stata package *mrrobust* (48) to carry out inverse variance weighted (IVW) and Egger
402 (which takes account of horizontal pleiotropy (49)) analyses. Analyses were carried out in
403 Stata MP 13.0 and Stata SE 14.2 (StataCorp, TX, USA).

404

405 *Sensitivity analysis*

406 The genome-wide analysis was repeated without adjustment for age, using a natural log
407 transformation (to be consistent with the previously published GWAS (11)), and also
408 excluding women whose AMH level was imputed as it was below the lower limit of
409 detection. We compared effect sizes in the main analysis with estimates from these alternate
410 analyses for genetic variants with $P<5\times10^{-5}$ in the main analysis.

411

412

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

413 **Funding**

414 This work was supported by Breast Cancer Now [to Generations Study]; British Heart
415 Foundation [SP/07/008/24066 to ALSPAC/DAL]; Gillings Family Foundation [to KSR];
416 National Institute of Health Research [NF-SI-0611-10196 to DAL]; Roche Diagnostics [to
417 ALSPAC]; The Institute of Cancer Research [to Generations Study]; UK Medical Research
418 Council [102215/2/13/2, G1001357 to ALSPAC/DAL; MC_UU_00011/6 to MCB, ALGS
419 and DAL; MR/P014054/1 to MCB]; UK National Health Service [to NIHR Biomedical
420 Research Centre Institute of Cancer Research]; University of Bristol [to ALSPAC and the
421 MRC Integrative Epidemiology Unit]; US National Cancer Institute [CA186107, CA176726,
422 CA49449, CA67262, CA178949 to the Nurses' Health Study and the Nurses' Health Study
423 II; CA178949 to AZ-J]; Wellcome Trust [102215/2/13/2, WT088806 and WT092830/Z/10 to
424 ALSPAC/DAL]. This work has also received support from NIHR Biomedical Research
425 Centre at the University Hospitals Bristol NHS Foundation Trust and the University of
426 Bristol [to DAL].

427 The work presented here is that of the authors. The views expressed in this publication are
428 those of the author(s) and not necessarily those of the UK Medical Research Council,
429 Wellcome Trust, British Heart Foundation, the National Institute for Health Research or the
430 UK National Health Service or Department of Health.

431

432 **Acknowledgements**

433 We thank Breast Cancer Now and the Institute of Cancer Research for support and funding of
434 the Generations Study, and the study participants, study staff, and the doctors, nurses and
435 other health care providers and health information sources who have contributed to the study.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

436 The Institute of Cancer Research acknowledges National Health Service funding to the NIHR
437 Biomedical Research Centre.

438 We would like to thank the participants and staff of the Nurses' Health Study and the Nurses'
439 Health Study II.

440 This research has been conducted using the UK Biobank Resource. We would like to thank
441 Dr R.N. Beaumont and Dr S.E. Jones (Genetics of Complex Traits, University of Exeter
442 Medical School) for assistance with calculating linkage disequilibrium in UK Biobank. The
443 authors would like to acknowledge the use of the University of Exeter High-Performance
444 Computing (HPC) facility in carrying out this work.

445 We are extremely grateful to all the families who took part in the ALSPAC study, the
446 midwives for their help in recruiting them, and the whole ALSPAC team, which includes
447 interviewers, computer and laboratory technicians, clerical workers, research scientists,
448 volunteers, managers, receptionists and nurses.

449

450

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

451 **Conflicts of Interest**

452 DAL has received support from Roche Diagnostics and Medtronic for research unrelated to
453 that presented here.

454 AM and KSR received support from Randox Laboratories Ltd for research unrelated to that
455 presented here.

456 No other conflicts of interest are declared.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

457 **References**

458 1. La Marca, A., Broekmans, F. J., Volpe, A., Fauser, B. C. and MacKlon, N. S. (2009) Anti-
459 Mullerian hormone (AMH): What do we still need to know? *Hum. Reprod.*, **24**, 2264–
460 2275.

461 2. Matuszczak, E., Hermanowicz, A., Komarowska, M. and Debek, W. (2013) Serum AMH
462 in Physiology and Pathology of Male Gonads. *Int. J. Endocrinol.*, **2013**, 128907.

463 3. Visser, J. A., Schipper, I., Laven, J. S. E. and Themmen, A. P. N. (2012) Anti-Müllerian
464 hormone: an ovarian reserve marker in primary ovarian insufficiency. *Nat. Rev.*
465 *Endocrinol.*, **8**, 331–341.

466 4. Hagen, C. P., Aksglaede, L., Sørensen, K., Mouritsen, A., Andersson, A. M., Petersen, J.
467 H., Main, K. M. and Juul, A. (2012) Individual serum levels of anti-Mllerian hormone in
468 healthy girls persist through childhood and adolescence: A longitudinal cohort study.
469 *Hum. Reprod.*, **27**, 861–866.

470 5. Kelsey, T. W., Anderson, R. A., Wright, P., Nelson, S. M. and Wallace, W. H. B. (2012)
471 Data-driven assessment of the human ovarian reserve. *Mol. Hum. Reprod.*, **18**, 79–87.

472 6. de Kat, A. C., van der Schouw, Y. T., Eijkemans, M. J. C., Herber-Gast, G. C., Visser, J.
473 A., Verschuren, W. M. M. and Broekmans, F. J. M. (2016) Back to the basics of ovarian
474 aging: A population-based study on longitudinal anti-Müllerian hormone decline. *BMC*
475 *Med.*, **14**, 151.

476 7. Meczekalski, B., Czyzyk, A., Kunicki, M., Podfigurna-Stopa, A., Plociennik, L., Jakiel, G.,
477 Maciejewska-Jeske, M. and Lukaszuk, K. (2016) Fertility in women of late reproductive
478 age: the role of serum anti-Müllerian hormone (AMH) levels in its assessment. *J.*
479 *Endocrinol. Invest.*, **39**, 1259–1265.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

480 8. Alvaro Mercadal, B., Imbert, R., Demeestere, I., Gervy, C., De Leener, A., Englert, Y.,
481 Costagliola, S. and Delbaere, A. (2015) AMH mutations with reduced in vitro
482 bioactivity are related to premature ovarian insufficiency. *Hum. Reprod.*, **30**, 1196–
483 1202.

484 9. Gorsic, L. K., Kosova, G., Werstein, B., Sisk, R., Legro, R. S., Hayes, M. G., Teixeira, J.
485 M., Dunaif, A. and Urbanek, M. (2017) Pathogenic anti-Müllerian hormone variants in
486 polycystic ovary syndrome. *J. Clin. Endocrinol. Metab.*, **102**, 2862–2872.

487 10. Pabalan, N., Montagna, E., Singian, E., Tabangay, L., Jarjanazi, H., Barbosa, C. P. and
488 Bianco, B. (2016) Associations of Polymorphisms in Anti-Müllerian Hormone (AMH
489 Ile49Ser) and its Type II Receptor (AMHRII -482 A>G) on Reproductive Outcomes and
490 Polycystic Ovary Syndrome: a Systematic Review and Meta-Analysis. *Cell. Physiol.*
491 *Biochem.*, **39**, 2249–2261.

492 11. Perry, J. R. B., McMahon, G., Day, F. R., Ring, S. M., Nelson, S. M. and Lawlor, D. A.
493 (2016) Genome-wide association study identifies common and low-frequency variants at
494 the AMH gene locus that strongly predict serum AMH levels in males. *Hum. Mol.*
495 *Genet.*, **25**, 382–388.

496 12. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., Kang,
497 H. M., Fuchsberger, C., Danecek, P., Sharp, K., *et al.* (2016) A reference panel of
498 64,976 haplotypes for genotype imputation. *Nat. Genet.*, **48**, 1279–1283.

499 13. La Marca, A., Grisendi, V. and Griesinger, G. (2013) How much does AMH really vary
500 in normal women? *Int. J. Endocrinol.*, **2013**.

501 14. Yang, J., Lee, S. H., Goddard, M. E. and Visscher, P. M. (2011) GCTA: A Tool for
502 Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.*, **88**, 76–82.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

503 15. Stolk, L., Zhai, G., van Meurs, J. B. J., Verbiest, M. M. P. J., Visser, J. A., Estrada, K.,
504 Rivadeneira, F., Williams, F. M., Cherkas, L., Deloukas, P., *et al.* (2009) Loci at
505 chromosomes 13, 19 and 20 influence age at natural menopause. *Nat. Genet.*, **41**, 645–
506 647.

507 16. Stolk, L., Perry, J. R. B., Chasman, D. I., He, C., Mangino, M., Sulem, P., Barbalic, M.,
508 Broer, L., Byrne, E. M., Ernst, F., *et al.* (2012) Meta-analyses identify 13 loci associated
509 with age at menopause and highlight DNA repair and immune pathways. *Nat. Genet.*,
510 **44**, 260–8.

511 17. Day, F. R., Ruth, K. S., Thompson, D. J., Lunetta, K. L., Pervjakova, N., Chasman, D. I.,
512 Stolk, L., Finucane, H. K., Sulem, P., Bulik-Sullivan, B., *et al.* (2015) Large-scale
513 genomic analyses link reproductive aging to hypothalamic signaling, breast cancer
514 susceptibility and BRCA1-mediated DNA repair. *Nat. Genet.*, **47**, 1294–1303.

515 18. Day, F. R., Thompson, D. J., Helgason, H., Chasman, D. I., Finucane, H., Sulem, P.,
516 Ruth, K. S., Whalen, S., Sarkar, A. K., Albrecht, E., *et al.* (2017) Genomic analyses
517 identify hundreds of variants associated with age at menarche and support a role for
518 puberty timing in cancer risk. *Nat. Genet.*, **49**, 834–841.

519 19. Ruth, K. S., Campbell, P. J., Chew, S., Lim, E. M., Hadlow, N., Stuckey, B. G. A.,
520 Brown, S. J., Feenstra, B., Joseph, J., Surdulescu, G. L., *et al.* (2016) Genome-wide
521 association study with 1000 genomes imputation identifies signals for nine sex
522 hormone-related phenotypes. *Eur. J. Hum. Genet.*, **24**, 284–290.

523 20. Nishimura, K., Ishiai, M., Horikawa, K., Fukagawa, T., Takata, M., Takisawa, H. and
524 Kanemaki, M. T. (2012) Mcm8 and Mcm9 Form a Complex that Functions in
525 Homologous Recombination Repair Induced by DNA Interstrand Crosslinks. *Mol. Cell*,
526 **47**, 511–522.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

527 21. AlAsiri, S., Basit, S., Wood-Trageser, M. A., Yatsenko, S. A., Jeffries, E. P., Surti, U.,
528 Ketterer, D. M., Afzal, S., Ramzan, K., Faiyaz-Ul Haque, M., *et al.* (2015) Exome
529 sequencing reveals MCM8 mutation underlies ovarian failure and chromosomal
530 instability. *J Clin Invest*, **125**, 258–262.

531 22. Lutzmann, M., Grey, C., Traver, S., Ganier, O., Maya-Mendoza, A., Ranisavljevic, N.,
532 Bernex, F., Nishiyama, A., Montel, N., Gavois, E., *et al.* (2012) MCM8- and MCM9-
533 Deficient Mice Reveal Gametogenesis Defects and Genome Instability Due to Impaired
534 Homologous Recombination. *Mol. Cell*, **47**, 523–534.

535 23. Visser, J. A. and Themmen, A. P. N. (2014) Role of anti-Mullerian hormone and bone
536 morphogenetic proteins in the regulation of FSH sensitivity. *Mol Cell Endocrinol*, **382**,
537 460–465.

538 24. Kevenaar, M. E., Themmen, A. P. N., Laven, J. S. E., Sonntag, B., Fong, S. L.,
539 Uitterlinden, A. G., de Jong, F. H., Pols, H. A. P., Simoni, M. and Visser, J. A. (2007)
540 Anti-Müllerian hormone and anti-Müllerian hormone type II receptor polymorphisms
541 are associated with follicular phase estradiol levels in normo-ovulatory women. *Hum.*
542 *Reprod.*, **22**, 1547–54.

543 25. Grigorova, M., Punab, M., Ausmees, K. and Laan, M. (2008) FSHB promoter
544 polymorphism within evolutionary conserved element is associated with serum FSH
545 level in men. *Hum. Reprod.*, **23**, 2160–2166.

546 26. Schuring, A. N., Busch, A. S., Bogdanova, N., Gromoll, J. and Tuttelmann, F. (2013)
547 Effects of the FSH-beta-subunit promoter polymorphism -211G->T on the
548 hypothalamic-pituitary-ovarian axis in normally cycling women indicate a gender-
549 specific regulation of gonadotropin secretion. *J Clin Endocrinol Metab*, **98**, E82-6.

550 27. La Marca, A., Papaleo, E., Alviggi, C., Ruvolo, G., De Placido, G., Candiani, M.,

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

551 Cittadini, E., De Michele, F., Moriondo, V., Catellani, V., *et al.* (2013) The combination
552 of genetic variants of the FSHB and FSHR genes affects serum FSH in women of
553 reproductive age. *Hum. Reprod.*, **28**, 1369–74.

554 28. Ge, W., Clendenen, T. V., Afanasyeva, Y., Koenig, K. L., Agnoli, C., Brinton, L. A.,
555 Dorgan, J. F., Eliassen, A. H., Falk, R. T., Hallmans, G., *et al.* (2018) Circulating Anti
556 Müllerian Hormone and Breast Cancer Risk: A Study in Ten Prospective Cohorts. *Int. J.*
557 *Cancer*, 10.1002/ijc.31249.

558 29. Collaborative Group on Hormonal Factors in Breast Cancer (2012) Menarche,
559 menopause, and breast cancer risk: individual participant meta-analysis, including 118
560 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol*, **13**,
561 1141–1151.

562 30. Swerdlow, A. J., Jones, M. E., Schoemaker, M. J., Hemming, J., Thomas, D., Williamson,
563 J. and Ashworth, A. (2011) The Breakthrough Generations Study: design of a long-term
564 UK cohort study to investigate breast cancer aetiology. *Br. J. Cancer*, **105**, 911–917.

565 31. Nichols, H. B., Baird, D. D., Stanczyk, F. Z., Steiner, A. Z., Troester, M. A., Whitworth,
566 K. W. and Sandler, D. P. (2015) Anti-Müllerian hormone concentrations in
567 premenopausal women and breast cancer risk. *Cancer Prev. Res.*, **8**, 528–534.

568 32. Eliassen, A. H., Zeleniuch-Jacquotte, A., Rosner, B. and Hankinson, S. E. (2016) Plasma
569 Anti-Mullerian Hormone Concentrations and Risk of Breast Cancer among
570 Premenopausal Women in the Nurses' Health Studies. *Cancer Epidemiol. Biomarkers*
571 *Prev.*, **25**, 854–860.

572 33. Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L.,
573 Ness, A., Ring, S. and Smith, G. D. (2013) Cohort profile: The 'Children of the 90s'-
574 The index offspring of the avon longitudinal study of parents and children. *Int. J.*

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

575 *Epidemiol.*, **42**, 111–127.

576 34. Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey smith, G.,
577 Henderson, J., Macleod, J., Molloy, L., Ness, A., *et al.* (2013) Cohort profile: The Avon
578 Longitudinal Study of Parents And Children: ALSPAC mothers cohort. *Int. J.*
579 *Epidemiol.*, **42**, 97–110.

580 35. Wang, Q., Ferreira, D. L. S., Nelson, S. M., Sattar, N., Ala-Korpela, M. and Lawlor, D.
581 A. (2018) Metabolic characterization of menopause: Cross-sectional and longitudinal
582 evidence. *BMC Med.*, **16**, 1–12.

583 36. Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M. and Lee, J. J.
584 (2015) Second-generation PLINK: Rising to the challenge of larger and richer datasets.
585 *Gigascience*, **4**.

586 37. Loh, P.-R., Danecek, P., Palamara, P. F., Fuchsberger, C., A Reshef, Y., K Finucane, H.,
587 Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G. R., *et al.* (2016) Reference-based
588 phasing using the Haplotype Reference Consortium panel. *Nat. Genet.*, **48**, 1443–1448.

589 38. Sayantan Das et al (2016) Next-generation genotype imputation service and methods.
590 *Nat. Genet.*, **48**, 1284–1287.

591 39. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller,
592 J., Sklar, P., de Bakker, P. I., Daly, M. J., *et al.* (2007) PLINK: a tool set for whole-
593 genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**,
594 559–575.

595 40. O'Connell, J., Gurdasani, D., Delaneau, O., Pirastu, N., Ulivi, S., Cocca, M., Traglia, M.,
596 Huang, J., Huffman, J. E., Rudan, I., *et al.* (2014) A General Approach for Haplotype
597 Phasing across the Full Spectrum of Relatedness. *PLoS Genet.*, **10**.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

598 41. Zhou, X. and Stephens, M. (2012) Genome-wide efficient mixed-model analysis for
599 association studies. *Nat. Genet.*, **44**, 821–824.

600 42. Marchini, J. and Howie, B. (2010) Genotype imputation for genome-wide association
601 studies. *Nat. Rev. Genet.*, **11**, 499–511.

602 43. Willer, C. J., Li, Y. and Abecasis, G. R. (2010) METAL: fast and efficient meta-analysis
603 of genomewide association scans. *Bioinformatics*, **26**, 2190–2191.

604 44. D. Turner, S. (2018) qqman: an R package for visualizing GWAS results using Q-Q and
605 manhattan plots. *J. Open Source Softw.*, **3**, 731.

606 45. Pruim, R. J., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., Gliedt, T. P.,
607 Boehnke, M., Abecasis, G. R., Willer, C. J. and Frishman, D. (2011) LocusZoom:
608 Regional visualization of genome-wide association scan results. In *Bioinformatics*.Vol.
609 27, pp. 2336–2337.

610 46. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliot, L. R., Sharp, K., Motyer, A.,
611 Vukcevic, D., Delaneau, O., Connell, J. O., *et al.* (2017) Genome-wide genetic data on
612 ~500,000 UK Biobank participants. *bioRxiv*, http://dx.doi.org/10.1101/166298.

613 47. UK Biobank (2015) Genotype imputation and genetic association studies of UK Biobank.

614 48. Spiller, W., Davies, N. M. and Palmer, T. M. (2017) Software Application Profile:
615 mrrobust - A Tool For Performing Two-Sample Summary Mendelian Randomization
616 Analyses. *bioRxiv*, dx.doi.org/10.1101/14212.

617 49. Bowden, J., Smith, G. D. and Burgess, S. (2015) Mendelian randomization with invalid
618 instruments: Effect estimation and bias detection through Egger regression. *Int. J.
619 Epidemiol.*, **44**, 512–525.

620

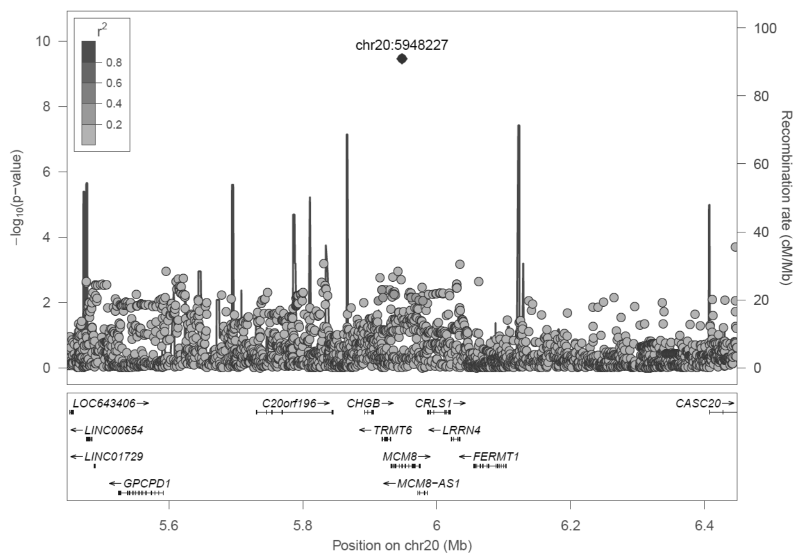
Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

621 **Legends to Figures**

622 *Figure 1. Association statistics with age-adjusted inverse normal AMH for variants within*

623 *500kb of rs16991615 (chr20:5948227) showing linkage disequilibrium with the top variant.*

624 Note: Linkage disequilibrium shown is in unrelated Europeans in UK Biobank.



625

626

627

628

629

630

631

632

633

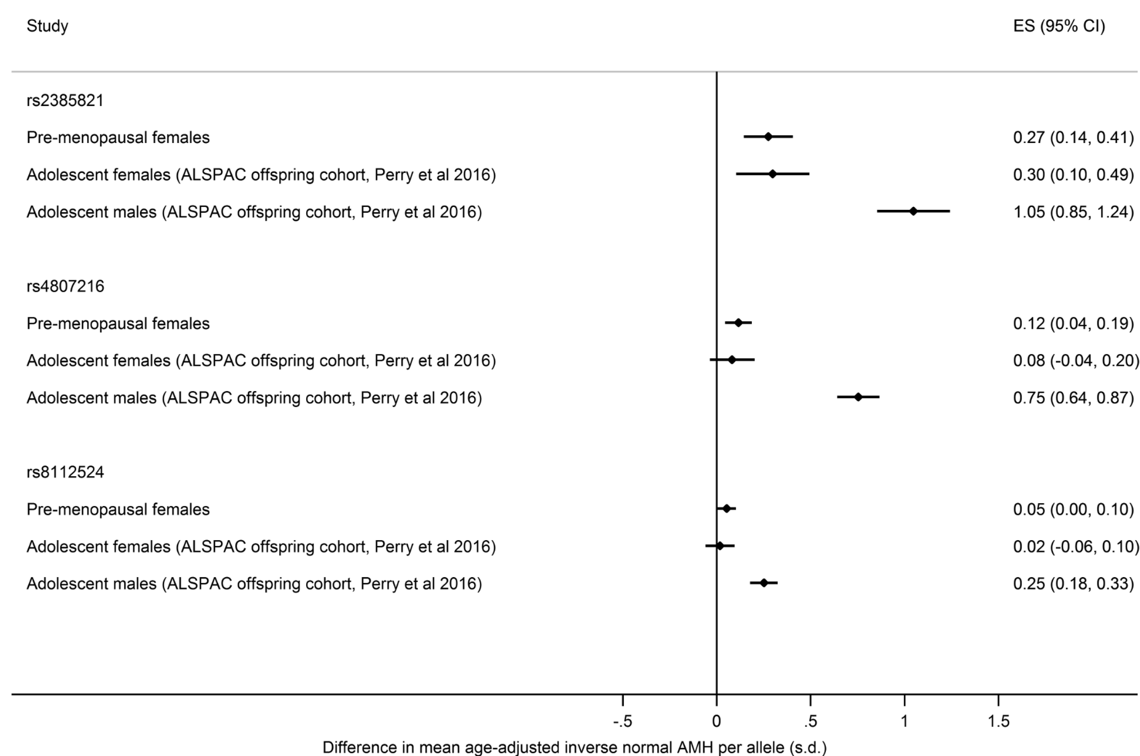
634

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

635 *Figure 2. Comparison of effect sizes of three genetic variants previously associated with*
636 *higher levels of AMH* (11) *when jointly included in the regression model: effect in adolescent*
637 *males and females from the ALSPAC offspring cohort (previous study, Perry et al 2016* (11)*)*
638 *and pre-menopausal women (current study).*

639 Note: Effect is difference in mean AMH per allele in standard deviations of age-adjusted
640 inverse normal AMH.

641 ES=effect size.



642

643

644

645

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

646 *Figure 3. Inverse variance weighted and Egger two-sample Mendelian randomization*

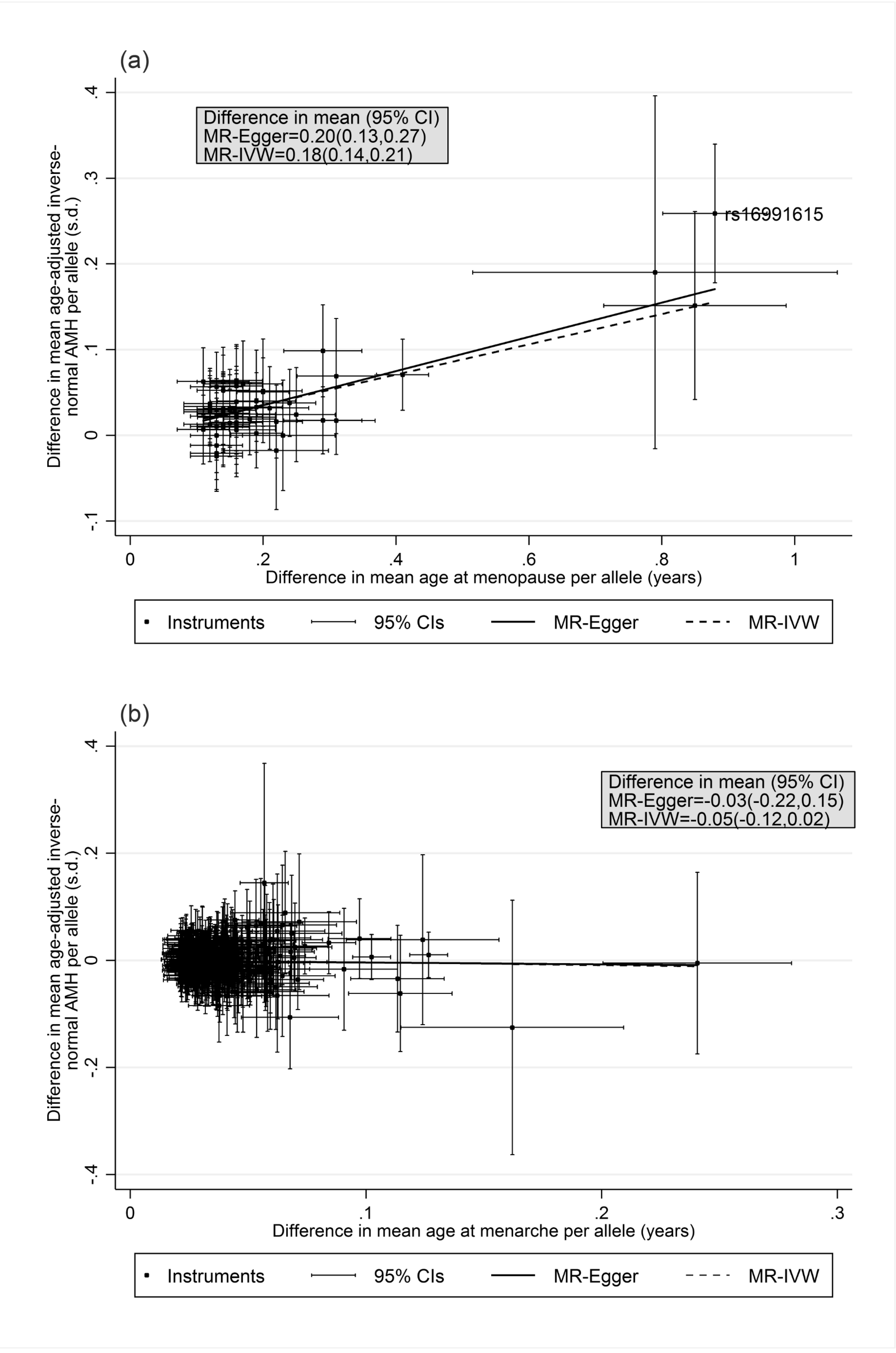
647 *analyses of the effect of genetically-predicted (a) age at menopause and (b) age at menarche*

648 *on age-adjusted inverse normal AMH levels in pre-menopausal women.*

649 Note: Difference in mean age-adjusted inverse-normal AMH per one-year increase in

650 genetically-predicted age at menarche or menopause.

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019



651

652

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

653 **Tables**

654 *Table 1. AMH levels and age of women in each of the five cohorts included in the genome-*

655 *wide analysis.*

		Median (interquartile range)	
Study	n	AMH (pmol/L)	Age (years) at blood draw
Generations Study	379	3.9 (0.8,11.7)	44 (40,48)
Sister Study	438	1.2 (0.1,6.0)	48 (45,51)
Nurses' Health Studies (Illumina)[1]	225	5.5 (1.5,12.4)	45 (42,48)
Nurses' Health Studies (OncoArray) [1]	417	6.7 (2.5,14.9)	44 (41,46)
Avon Longitudinal Study of Parents and Children	1,885	2.0 (0.4, 5.2)	46 (44,49)
Total	3,344		

656 [1]Data from the Nurses' Health Study and Nurses' Health Study II were combined and

657 genotyped on two different genome arrays, which were included as separate sub-samples in

658 this analysis.

659

660 *Table 2. Top ten signals from the genome-wide analysis of age-adjusted inverse normal AMH*

661 *in pre-menopausal women.*

SNPID	Chr	Pos	EA/OA/EAF	Beta (SE)	P-value	Dir.	Het. P	Imp. qual
rs16991615	20	5948227	A/G/0.068	0.26 (0.04)	3.5×10^{-10}	+++++	0.14	0.94
rs62236881	22	29450193	A/G/0.008	0.85 (0.16)	1.1×10^{-7}	+++??	0.77	0.84
rs186783371	5	88062223	T/A/0.011	0.64 (0.12)	1.4×10^{-7}	++++?	0.64	0.61
rs35829351	11	6120686	G/A/0.449	0.11 (0.02)	2.7×10^{-7}	++++-	0.02	0.89
rs10732995	1	175111574	T/C/0.966	0.29 (0.06)	3.9×10^{-7}	+++++	0.86	0.93
rs76673357	20	17831206	G/T/0.056	0.22 (0.04)	5.0×10^{-7}	+++++	0.47	0.89
rs7168070	15	93910220	T/C/0.329	0.11 (0.02)	5.5×10^{-7}	+++++	0.98	0.89

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

rs62237617	22	28761148	T/C/0.011	0.74 (0.15)	6.5×10^{-7}	+++??	0.89	0.82
rs141456816	17	10362307	C/T/0.013	0.50 (0.10)	6.7×10^{-7}	++++-	0.44	0.66
rs6743761	2	10393228	A/G/0.482	0.09 (0.02)	2.7×10^{-6}	++++-	0.11	0.99

662

663 Beta=difference in mean age-adjusted inverse normal AMH (SD) per allele.

664 Chr=chromosome; Dir. = positive (+) or negative (-) direction of effect in Avon Longitudinal
665 Study of Parents and Children, Sister Study, Generations Study, Nurses' Health Studies
666 (OncoArray), Nurses' Health Studies (Illumina) respectively, with "?" indicating that the
667 variant was absent; EA=effect allele; EAF=weighted average effect allele frequency across
668 the studies; Het. P=P-value from Cochrane's Q-test of heterogeneity of effects across the
669 studies; Imp. qual = mean imputation quality across the studies; Pos=position in
670 hg19/GRCh37; OA=other allele; SE=standard error.

671

672 **Abbreviations**

673 ALSPAC=Avon Longitudinal Study of Parents and Children

674 AMH=anti-Müllerian hormone

675 Chr=chromosome

676 CI=confidence interval

677 CNG=Centre National de Génotypage

678 EA=effect allele

679 EAF=weighted average effect allele frequency across the studies

680 ES=effect size

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

681 FSH=follicle-stimulating hormone

682 GWAS=genome-wide association study

683 Het. P=P-value from Cochrane's Q-test of heterogeneity of effects across the studies

684 IBS=identity by state

685 Imp. qual = mean imputation quality across the studies

686 IVW=inverse variance weighted

687 LH=luteinising hormone

688 OA=other allele

689 PCOS=polycystic ovary syndrome

690 Pos=position in hg19/GRCh37

691 SD=standard deviation

692 SE=standard error

693

694 **Supplementary Information**

695 *Details of studies included*

696 <u>ALSPAC</u>

697 The ALSPAC study (34–36) is a prospective population-based birth cohort study that
698 recruited 14,541 pregnant women resident in the South West of England with expected dates
699 of delivery from 1st April 1991 to 31st December 1992 (http://www.alspac.bris.ac.uk.). The
700 women and their offspring have been followed-up since that date and information presented
701 here is from a subgroup of the original mothers who were pre-menopausal at the time of

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

702 AMH blood sampling (34–36). Ethical approval for the study was obtained from the
703 ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Please note
704 that the study website contains details of all the data that is available through a fully
705 searchable data dictionary and variable search tool:
706 http://www.bristol.ac.uk/alspac/researchers/our-data.

707 Generations Study

708 The Generations Study (31) is a prospective population cohort study started in 2003 to
709 investigate the environmental, behavioural, hormonal and genetic causes of breast cancer
710 (31). The cohort includes over 110 000 women aged 16 and older at entry, recruited from the
711 general UK population through connections to the charity Breakthrough Breast Cancer (now
712 Breast Cancer Now) or who volunteered as a result of publicity, and female friends and
713 family members of participants. Follow-up questionnaires are mailed to participants about
714 every 3 years. The study received appropriate ethical approval from the South East MREC,
715 and informed consent was received from the participants. Detailed menstrual histories were
716 collected and blood samples were contributed by 92% of participants.

717 Nurses' Health Study and Nurses' Health Study II

718 In 1976, 121,701 female, registered nurses, ages 30 to 55 years, were enrolled in the Nurses'
719 Health Study (33). Biennially, participants complete mailed questionnaires on lifestyle, diet,
720 reproductive history, and disease diagnoses. In 1989–1990, 32,826 women ages 43 to 69
721 years (21% premenopausal) donated blood samples.

722 The Nurses' Health Study II was established in 1989, when 116,430 female registered nurses,
723 ages 25 to 42 years, completed and returned a questionnaire (33). The cohort has been
724 followed biennially following the methods of the NHS. Between 1996 and 1999, 23,393

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

725 premenopausal participants, who were cancer-free and between the ages of 32 and 54 years,

726 provided blood samples.

727 <u>Sister Study</u>

728 The Sister Study prospective cohort was designed to address genetic and environmental risk

729 factors for breast cancer. During 2003-2009, 50,884 U.S. and Puerto Rican women ages 35-

730 74 were recruited through a national multi-media campaign and network of recruitment

731 volunteers, breast cancer professionals and advocates. Eligible women had a sister who had

732 been diagnosed with breast cancer but did not have breast cancer themselves. This research

733 was approved by the Institutional Review Boards of the National Institute of Environmental

734 Health Sciences, NIH, and the Copernicus Group. All participants provided informed

735 consent. Data analysed in this study were from a subgroup of participants with a serum

736 sample who were premenopausal (32).

737

738 Supplementary Figure 1. (a) Manhattan and (b) QQ plot for GWAS of age-adjusted inverse

739 normal AMH in pre-menopausal women.

740 Supplementary Figure 2. Comparison of effect sizes in the main GWAS (SD of age-adjusted

741 inverse normal AMH) and the analysis not adjusted for age (SD of inverse normal AMH) for

742 genetic variants that were $P<5\times10^{-5}$ in the main GWAS.

743 Supplementary Figure 3. Comparison of effect sizes in the main GWAS (SD of age-adjusted

744 inverse normal AMH) and the natural log transformed analysis (SD of age-adjusted natural

745 log transformed AMH) for genetic variants that were $P<5\times10^{-5}$ in the main GWAS.

746 Supplementary Figure 4. Comparison of effect sizes in the main GWAS and the analysis

747 excluding women with AMH measured as below the lower limit of detection (effects in SD

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019

748 of age-adjusted inverse normal AMH for both) for genetic variants that were $P<5\times10^{-5}$ in the
749 main GWAS.

750

751 Supplementary Table 1. Comparison of effect sizes from univariate analyses and joint
752 analyses (approximate conditional analyses in GCTA) in pre-menopausal women and
753 adolescent males and females for three genetic variants associated with higher levels of AMH
754 in adolescent males (10).

755 Supplementary Table 2 Age-adjusted inverse normal AMH GWAS results and comparison of
756 effect directions for 56 published age of menopause variants (Day et al 2015).

757 Supplementary Table 3. Results of Mendelian Randomization analyses of the effect of
758 genetically-predicted age at menopause and age at menarche on age-adjusted inverse normal
759 AMH levels in pre-menopausal women.

760 Supplementary Table 4 Age-adjusted inverse normal AMH GWAS results and comparison of
761 effect directions for published age of menarche variants (Day et al 2017).

Downloaded from https://academic.oup.com/hmg/advance-article-abstract/doi/10.1093/hmg/ddz015/5289469 by Institute of Cancer Research user on 28 January 2019