SUPPLEMENTAL METHODS

Overview of consortium

The SUNLIGHT consortium (Study of Underlying Genetic Determinants of Vitamin D and Highly Related Traits) was formed in 2008. It represents a collaboration of 14 cohorts from the United Kingdom, United States, Canada, Netherlands, Sweden, and Finland. Three of the participating cohorts (Framingham Heart Study, Cardiovascular Health Study, and the Rotterdam Study) are also members of the CHARGE consortium (Cohorts for Heart and Aging Research in Genetic Epidemiology).[1]

Individual cohort descriptions

The **Framingham Heart Study** was initiated in 1948 and the Original cohort includes 5,209 participants from Framingham, Massachusetts, US, who have attended exams every other year to investigate cardiovascular disease and related risk factors. The Offspring cohort, recruited in 1971, is comprised of 5,124 children of the Original cohort and the children's spouses.[2] Participants from the Offspring cohort have undergone examinations roughly every four years. The Third Generation cohort includes 4,095 children of the Offspring cohort who attended their first exam in 2002.[3] The current study sample consists of 1,815 individuals from the Offspring Generation who attended the 5th examination (1991-1995) and 3,841 from the Third Generation who attended the 1st examination (2001-2005).

The **Twins UK** registry in St. Thomas' Hospital, King's College London recruited a total sample of 11,000 identical and non-identical, mostly female Caucasian, twins from across the UK through national media campaigns. Their age ranges between 16 and 85 years. Over 7,000 twins have attended detailed clinical examinations with a wide range of phenotypes over last 18 years. All participants were recruited without presence or interest in any particular disease or trait. A total of 5,065 healthy female Caucasian twins had measurement of 25-OH D and were included in the analysis.

The **Rotterdam Study** is a prospective population-based cohort study of chronic disabling conditions in Dutch elderly individuals aged 55 years and over. The study was initiated in 1990 and participants have undergone examinations roughly every 4 years. The current study sample consists of 1,237 individuals from the baseline examination for whom both 25-OH D levels and genotype data were available.

The **1958 British Birth cohort study (1958BC)** was started in 1958 as the Perinatal Mortality Study, where for one week in March all births (~17,000) in England, Wales and Scotland were surveyed. Participants have been subsequently followed up with DNA and blood sample data collection at ages 44 to 46 years (2002-2004).[4] Of those contacted, 78% responded, of whom 94% consented to the collection of blood samples and 90% to DNA analysis. A total of 3,837 1958BC participants were included in the present study and 2,715 in the replication.

The **Amish Family Osteoporosis Study (AFOS)** recruited Old Order Amish individuals from Lancaster, Pennsylvania, from 1997-2008. The cohort includes 1,520 individuals in

multigenerational families, all of whom can be connected in a single pedigree. Participants were at least 20 years old and generally healthy. Exclusion criteria included malignancy, liver and kidney disease (serum creatinine > 1.3 mg/dl), Paget disease of bone, hypercalcemia, untreated thyroid disease and severe hypertension. The current study includes 330 participants who also had GWAS performed.

The **Cardiovascular Health Study (CHS)** is a prospective cohort study of clinical and subclinical cardiovascular disease among older adults.[5] In 1989-1990, CHS enrolled 5,201 ambulatory men and women ages 65 and older from Medicare eligibility lists in Forsyth County, North Carolina; Sacramento County, California; Washington County, Maryland; and Pittsburgh, Pennsylvania. Exclusion criteria included the use of a wheelchair in the home, institutionalization, the need for a proxy respondent to provide informed consent, plans to move from the area within 3 years, or current treatment for cancer. For the purposes of these analyses we excluded CHS participants who had cardiovascular disease at the time of the 1992-1993 CHS exam, defined by a history of any one of the following conditions: coronary heart disease, heart failure, stroke, transient ischemic attack, claudication, atrial fibrillation, pacemaker, or implantable cardioverter defibrillator. The current study included 1,786 CHS participants.

The **Northern Finland Birth Cohort 1966 (NFBC)** includes 12,058 live born individuals, of European descent, with expected dates of birth during 1966 in the provinces of Oulu and Lapland (coverage 96% of eligible birth cohort). A wide range range of phenotypic, lifestyle, and demographic data were collected via health clinics and questionnaires at various ages, including

blood samples for DNA and 25-OH D analyses at age 31 years.[6] Both were available for 4,534 participants.

The **study of Indiana women** was initiated to identify genes contributing to bone mineral density. Recruitment focused on families with two or more healthy sisters. Sisters were required to be within 10 years of each other in age. A sample of healthy brothers aged 18 to 61 years was also recruited, but not included in the present study. There was no exclusion based on the difference in age between the brothers. More recently, recruitment was expanded to include healthy men and women who met all study recruitment criteria but did not have a sibling eligible or desiring to participate. Exclusion criteria for both women and men included a history of chronic disease, taking medications known to affect bone mass or metabolism, unable to have BMD measured because of obesity (weight >136 kg) and abnormal blood biochemistry tests. Women who had irregular menses or a history of pregnancy or lactation within three months prior to enrollment were excluded; however, women taking oral contraceptives were not excluded. The final analytic sample included 567 women.

The **Health, Aging and Body Composition study (Health ABC)** is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in older adults. Health ABC enrolled well-functioning, community-dwelling black (n=1,281) and white (n=1,794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. Participants have undergone annual exams and semi-annual phone interviews. The

current study sample consists of 1,559 white participants who attended the second exam in 1998-1999 with available genotyping data.

The **Gothenburg Osteoporosis and Obesity Determinants (GOOD)** study was initiated to determine both environmental and genetic factors involved in the regulation of bone and fat mass.[7] Male study subjects were randomly identified in the greater Gothenburg area in Sweden using national population registers, contacted by telephone, and invited to participate. To be enrolled in the GOOD study, subjects had to be between 18 and 20 years of age. There were no other exclusion criteria, and 49% of the study candidates agreed to participate (n=921).

The **Canadian Multicentre Osteoporosis Study (CaMos)** is an ongoing, population-based prospective cohort study of 9,423 community-dwelling, randomly selected women (6,539) and men (2,884), aged 25 years and older at baseline (1995-1997), living within 50 km of nine urban regional centers.[8] CaMos objectives, methodology and sampling framework are described in detail elsewhere.[9] At an initial baseline visit, bone mineral density and vitamin D-related traits were measured in all available subjects and participants were interviewed by a trained interviewer to assess for vitamin D intake, osteoporosis, and fracture-related risk factors. A second intensive interview was conducted five and ten years after enrollment to reassess these risk factors. Participant retention was 67% at year 10. Blood samples for DNA and 25-OH D analyses were available for the present study from subjects enrolled at seven centres collected at year 10 (n=2,347).

The **Chingford study** is a well-established, prospective population-based study of middle-aged British women who have been evaluated intensively for osteoarthritis and osteoporosis since 1989. The cohort originally consisted of 1,003 women ages 45–64 years, selected through general practitioner records. There were no inclusion/exclusion criteria other than age. The cohort has been followed up biannually, and many clinical, anthropometric, psychosocial, radiologic, and metabolic variables have been recorded on at different points. The Chingford Study is currently approaching its twentieth year and has retained ~65% of the cohort, with a current age range of 60–82 years. All subjects still participating in the Chingford Study were re-examined for the present study in Chingford Hospital in North-East London. A total of 659 individuals were included in the current analyses.

The **Hertfordshire Cohort Study (HCS)** includes 3,000 men and women born in the county of Hertfordshire, UK, between 1931 and 1939. The selection procedure for these individuals was as follows: in brief, with the help of the National Health Service Central Registry at Southport, and Hertfordshire Family Health Service Association, men and women were identified who were born during 1931–39 in Hertfordshire, and still lived there during the period 1998–2003. The birthweight and weight at one year of age of each individual had been recorded in a ledger by a team of midwives and health visitors who had attended each birth in Hertfordshire in the 1930s and visited the child's home at intervals during the first year of life. This study included 1,204 HCS participants.

The **Aberdeen Prospective Osteoporosis Screening Study (APOSS)** is a population-based study of osteoporosis fracture risk. Between 1990 and 1994, 5,119 Caucasian women randomly

selected from the Community Health Index register attended for an initial visit. At this time, participants were aged 45-54 years. The women were invited to undergo further assessment between 1997 and 2000, when aged 56-60 years. A total of 3,883 women attended the second visit, and 3,113 donated blood samples for 25-OH D assessment, and DNA extraction and genotyping. A subset of 1,579 individuals was included in the present study.

Written informed consent was obtained from all subjects in the participating cohorts, and the study protocols were reviewed and approved by the local institutional review boards.

Genotyping and Quality Control Methods

In the Framingham Heart Study, Affymetrix 500K SNP arrays supplemented with the MIPS 50K array were used for genotyping. The quality control filters included call rate $\geq$ 97%, no excess Mendelian errors (< 1000) and average heterozygosity within 5 SD of mean (between 25.758% and 29.958%). Analysis was restricted to SNPs with minor allele frequency $\geq$ 1%, call rate $\geq$ 95% and HWE p $\geq 10^{-6}$. The genome wide association analyses were conducted on 425,593 genotyped SNPs in 5,656 Framingham participants with 25-OH D levels available. A subset of 378,163 SNPs passing more stringent filters including a minor allele frequency $\geq$ 0.01, SNP call rate $\geq$ 0.97, differential missingness p-value $\geq 10^{-9}$ and < 100 Mendelian errors were used for imputation based on the haplotypes of the HapMap CEU trios using the MACH software. A total of 2,003,437 imputed SNPs passed the quality control criterion (ratio of observed to expected variance $\geq$ 0.3) and were analyzed for association with 25-OH D levels.

Genotyping in the Twins UK study was performed using Illumina 317K and 610K SNP. Analyses were restricted to SNPs with call rate of genotype ≥ 98% with average heterozygosity within 3 SD of mean (< 0.3042438 and > 0.3267991). MAF was set as >1% and HWE P > 10^{-4}. Genome-wide association analyses were conducted on 545,026 SNPs for 610K and 313,575 SNPs for 317K in 5,180 Twins UK participants. IMPUTE software was used for imputation. A total of 2,548,976 SNPS passed the quality control criterion (posterior probability >98%) for imputation.

Rotterdam samples were genotyped using Illumina HumanHap550v3 Genotyping BeadChip at the Genetics Laboratory of the Department of Internal Medicine of Erasmus Medical Center, Rotterdam, Netherlands. The Beadstudio GenCall algorithm was used for genotype calling. The following exclusion criteria were applied: call rate <97.5%, gender mismatches with typed X-linked markers, excess autosomal heterozygosity >0.336, FDR >0.1%, duplicates and/or 1st or 2nd degree relatives using IBS probabilities >97% from PLINK, and ethnic outliers using IBS distances > 3SD from PLINK. Analysis was restricted to SNPs with call rate ≥98%, MAF ≥1%, P for Hardy-Weinberg equilibrium ≥10^{-6}. For imputation MACH software was used which resulted in a total of 2,543,887 imputed SNPs that were analysed for association with Vit D.

In the 1958BC, genome-wide data has been obtained through two sub-studies, both using the 1958BC members as a control population. First, 1,500 DNA samples were randomly selected as part of the Welcome Trust Case Control Consortium (WTCCC) and genotyped on the Affymetrix Genechip 500K. After genotyping, SNP exclusions were applied study-wide done as part of the WTCCC; the exclusions were HWE p-value of < 5.7 x 10^{-7}, study-wide SNP

missing data rate > 5% or study-wide SNPs missing data > 1% with minor allele frequency (MAF) < 5%, trend test for association between WTCCC control studies with 1 df p-value < 5.7e-7 or 2df p-value 5.7e-7 and MAF <0.01. A total of 383,325 SNPs remained after exclusions. Sample exclusions were based on the sample call rate < 97%, relatedness, external discordance, heterozygosity (outside of 22.5-30%) and non-European ancestry, leaving 1,480 subjects. Secondly, 2,592 DNA samples from the 1958BC were used as controls for a Type 1 diabetes case-control study.[10] Samples were genotyped through the JDRF/WT Diabetes and Inflammation Laboratory (DIL) using the Illumina Infinium 550K chip. SNP exclusions on the 2,592 DNA were based on the (HWE) p-value < 1e-7, SNP call rate < 95% and MAF < 1%, leaving 520, 413 SNPs taken into imputation. Sample exclusions were based on sample call rate < 97%, gender discrepancy, relatedness ($\hat{\pi}$ score > 0.2), heterozygosity (outside of 29-34%) and non-European ancestry, leaving a total of 2,530 subjects. For both datasets, subjects passing QC were used for imputation based on the HapMap CEU population using the software Impute.[11] A total of 2,413,400 and 2,510,365 imputed SNPs were available for the two 1958BC genotyping subcohorts. The number of participants with available data who were analyzed for association with 25(OH)D concentrations was 3,837. In addition, follow-up genotyping was performed for SNPs at 20q13 (rs6013897 and rs16982264) in another 2,715 participants from this cohort, using TaqMan (Applied Biosystems) at the Cambridge Institute for Medical Research, UK.

In AFOS, two platforms were used for genotyping, Affymetrix 500K SNP arrays for 338 participants and 100K SNP arrays for 271 participants. The GeneChip Genotyping Analysis Software (GTYPE 4.0) was used to generate dynamic modeling algorithm– derived genotypes

that were reanalyzed with the BRLMM (Bayesian RLMM) genotype calling algorithm (confidence threshold of 0.33) to improve the proportion of heterozygote calls. As an initial quality-control measure, BRLMM-generated chip files with call rates <90% for both enzymes across all SNPs were excluded. The resulting median call rate across all of the remaining 419 case-control samples was 97.5% (97.6% for XbaI and 97.4% for HindIII). We further removed individual SNPs with genotype call rates <90%, monomorphic SNPs and SNPs with minor allele frequency <5%, and those deviating from Hardy-Weinberg equilibrium in control subjects ($P < 0.001$). The number of monomorphic and low-frequency SNPs (n=26,816) in the Amish was not appreciably different from that observed in more outbred Caucasians of the HapMap CEU sample. The concordance rate for 11 quality-control samples that were run twice on the Affymetrix GeneChip mapping panel was 97.5%. We also calculated a cross-platform concordance rate of 98% for 419 samples in which 61 SNPs were genotyped using the Affymetrix GeneChip Mapping 100K panel and an independent Illumina 1536-plex GoldenGate assay.

For CHS, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV BeadChip system. Genotypes were called using the Illumina BeadStudio software. Samples were excluded for genotypic sex mismatch, discordance with prior genotyping, or call rate <95%. Genotyping was successful in 96%. The following exclusions were then applied: call frequency <97%, Hardy Weinberg Equilibrium p-value $<10^{-5}$, duplicate error or Mendelian inconsistency, heterozygote frequency ≈ 0, or SNP not found in HapMap CEU. SNPs were further excluded from analysis if the ratio of the observed variance of the allele dosage to the expected variance of

the dosage was <0.01. Genotypes were imputed to approximately 2.5 million SNPs in HapMap, using BIMBAM software with Phase II CEU individuals as a reference panel. Imputation results are summarized as an "allele dosage" (a fractional value between 0 and 2), defined as the expected number of copies of the minor allele at that SNP.

For NFBC, genomic DNA was extracted from whole blood using standard methods. All DNA samples for the Illumina Infinium 370cnvDuo array were prepared for genotyping by the Broad Institute Biological Sample Repository (BSP). All individuals in the study were genotyped with call rates >95%. Individuals with discrepancy between their reported sex and the sex determined from the X chromosome were excluded from analysis. We used the identity-by-descent (IBD) analysis option of the software package PLINK to determine possible relatedness among our sample subjects and to identify sample duplications and sample contamination (the latter identified as individuals who seemed to be related to nearly everyone in the sample). If the sample duplication issue could not be resolved by external means, both samples were excluded. All apparently contaminated samples were excluded. We identified individuals related at the level of half-sibs or closer with the IBD analysis and excluded one subject from each pair (the subject with less complete genotyping). Subsequent to this overall exclusion, individuals may have been excluded from analysis of specific traits, as detailed above. SNPs were excluded from analysis if the call rate in the final sample was < 95%, if the *P* value from a test of Hardy-Weinberg Equilibrium (HWE) was <0.0001, or if the MAF was <1%.

In the Indiana women study, genotyping was performed on the Illumina Human610Quadv1_B BeadChips (Illumina, San Diego, CA, USA) by the Center for Inherited Disease Research

(CIDR) using the Illumina Infinium II assay protocol. This array contains 592,532 markers with a mean spacing of 5.8 kb. Allele cluster definitions for each SNP were determined using Illumina BeadStudio Genotyping Module version 3.2.32 and the combined intensity data from 100% of the study samples. The resulting cluster definition file was used on all study samples to determine genotype calls and quality scores. Genotype calls were made when a genotype yielded a quality metric (Gencall score) of 0.15 or higher. The final raw dataset released by CIDR to the investigators and to dbGaP contained called genotypes for 589,171 SNPs, with 8 samples not released due to inadequate quality of genotyping. Blind duplicate reproducibility was 99.99 % based on 36 paired samples. SNPs with a call rate of 98% or greater (n=581,255) were included and subjected to further quality control analyses. From these, SNPs were removed if: either the minor allele frequency was less than 0.01 in this dataset (n=32,948) or there was significant deviation ($p<0.00001$) from Hardy Weinberg equilibrium (n=1,998). The final dataset for analysis consisted of 547,971 SNPs that passed all quality control measures. Imputation was performed based on the haplotypes of the HapMap CEU trios using the MACH software.

Genotyping in Health ABC was performed by the Center for Inherited Disease Research (CIDR) using the Illumina Human1M-Duo BeadChip system. Samples were excluded from the dataset for the reasons of sample failure, genotypic sex mismatch, and first-degree relative of an included individual based on genotype data. Genotyping was successful in 1,663 Caucasians. Analysis was restricted to SNPs with minor allele frequency $\geq$ 1%, call rate $\geq$97% and HWE $p \geq 10\text{-}6$. Genotypes were available on 914,263 high quality SNPs for imputation based on the

HapMap CEU (release 22, build 36) using the MACH software (version 1.0.16). A total of 2,543,888 imputed SNPs were analyzed for association with 25-OH D levels.

In the GOOD study, genotyping was performed at the Genetic Laboratory, Department of Internal Medicine, Erasmus Medical Center, Rotterdam, the Netherlands using Illumina HumanHap 610 arrays. Genotypes from 938 individuals passed the sample quality control criteria (exclusion criteria: sample call rate < 97.5%, gender discrepancy with genetic data from X-linked markers, excess autosomal heterozygosity > 0.33 ~ FDR < 0.1%, duplicates and/or first degree relatives identified using IBS probabilities [> 97%], ethnic outliers [3 SD away from the population mean] using multi-dimensional scaling analysis with four principal components). Genotypes were imputed for all polymorphic SNPs (521,160 with MAF $\geq$ 1%, SNP call rate $\geq$ 98% and HWE p value $\geq 10^{-6}$) using the MACH software, based upon phased autosomal chromosomes of the HapMap CEU Phase II panel (release 22, build 36), orientated on the positive strand. GWA analysis was conducted on 2,543,887 imputed SNPs in 921 GOOD participants with vitamin D levels available. No quality control criteria (ratio of observed to expected variance or MAF) were used on the imputed SNPs. The uncertainty in genotype prediction was accounted for by utilizing the dosage information from MACH.

In CaMOS, genotyping was performed using the Applied Biosystems (ABI) TaqMan protocol. SNPs rs12785878 and rs7944926 were genotyped and the genotype success rates were 99.74% and 99.91% for these SNPs respectively. For both SNPs, the minor allele frequencies closely matched HapMap samples (0.275 for both SNPs) and there were no significant deviations from HWE ($p = 0.67$).

In Chingford, Hertfordshire, and APOSS genotyping was performed using TaqMan (Applied Biosystems) at the Cambridge Institute for Medical Research, UK. All genotyping data were scored blind to sample status and double scored by a second operator to minimize error. Genotyping was assessed for Hardy-Weinberg equilibrium ($p>0.05$) and call rate was above 98%.

Statistical Power

At the threshold $\alpha= 5 \times 10^{-8}$, using a conservative discovery sample size of 14,000, our study had 80% power to detect SNPs explaining 0.28% of the total variance in 25-OH D levels, and 90% power to detect SNPs explaining 0.32% of the total variance.[12]

Supplemental Table 1A: Characteristics of GWAS cohorts ("discovery cohorts")

	Framingham	Twins UK	Rotterdam Study	1958BC	AFOS
Country	USA	UK	Netherlands	UK	USA
Genotyping platform	Affymetrix 500K and MIPS 50K	Illumina 317K or 610K	Illumina HumanHap 550K	Affymetrix 500K or Illumina 550K	Affymetrix 500K and 100K
N	5,656	5,065	1,237	3,837	330
Female (%)	2,998 (53%)	5,065 (100%)	730 (59%)	1957 (51%)	188 (57%)
Mean age (sd), years	47 (13.1)	49 (13.6)	66 (6.8)	45 (0.34)	49 (14.12)
Mean BMI(sd), kg/m^2	27 (5.4)	26 (4.7)	26.4 (3.7)	28 (4.9)	27.2 (4.2)
Winter exam (%)	1301 (23%)	1495 (30%)	297(24%)	729 (19%)	66 (20%)
25-OH D assay	RIA	RIA	RIA	ELISA	RIA
Mean 25-OH D (sd), nmol/L	78 (37.3)	71 (32.9)	66 (27.7)	58 (25.2)	57 (18.7)

Supplemental Table 1B: Characteristics of GWAS cohorts ("*in silico* replication" cohorts)

	Health ABC	CHS	GOOD	Indiana	NFBC
Country	USA	USA	Sweden	USA	Finland
Genotyping platform	Illumina 1M	Illumina 370K CNV chip	Illumina HumanHap 610	Illumina Human610-Quadv1	Illumina Infinium cnvDuo array
N	1,559	1,786	921	567	4,534
Female, %	733 (47%)	1248 (70%)	0 (0%)	567 (100%)	2312 (51%)
Mean age (sd), years	75 (2.9)	74 (4.7)	19 (0.6)	33 (7.2)	31 (0.36)
Mean BMI (sd), kg/m^2	26.4 (4.1)	26.9 (4.6)	22.4 (3.2)	26.2 (5.9)	24.1 (4.3)
Winter exam, %	452 (29%)	429 (24%)	120 (13%)	68 (12%)	272 (6%)
25-OH D assay	RIA	Mass Spec	RIA	Mass Spec	Mass Spec
Mean 25-OH D (sd), nmol/L	72 (27.6)	66 (25.6)	65 (22.0)	73 (32.7)	62 (23.9)

Supplemental Table 1C: Characteristics of non-GWAS cohorts ("*de novo* replication" cohorts)

	CaMos	Chingford	Hertfordshire	APOSS	Additional 1958BC sample
Country	Canada	UK	UK	UK	UK
Genotyping platform	TaqMan	TaqMan	TaqMan	TaqMan	TaqMan
N	2,347	659	1,204	1,579	2,715
Female, %	1619 (69%)	659 (100%)	638 (53%)	1579 (100%)	1358 (50%)
Mean age (sd), years	64 (17.5)	57.1 (6.1)	66 (2.6)	55 (2.2)	45 (0.4)
Mean BMI (sd), kg/m^2	27 (5.1)	25.7 (4.3)	27.3 (4.4)	26.6 (4.8)	27.2 (4.8)
Winter exam, %	516 (22%)	158 (24%)	470 (39%)	379 (24%)	624 (23%)
25-OH D assay	CLIA	Mass Spec	Mass Spec	HPLC	ELISA
Mean 25-OH D (sd), nmol/L	69 (25.2)	69 (30)	47 (23.83)	54 (21.9)	59 (26.2)

REFERENCES

1. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, Uitterlinden AG, Harris TB, Witteman JCM, Boerwinkle E, on Behalf of the CHARGE Consortium. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of Prospective Meta-Analyses of Genome-Wide Association Studies From 5 Cohorts. *Circ Cardiovasc Genet.* 2009;2:73-80.

2. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families: the Framingham Offspring Study. *Am J Epidemiol.* 1979;110:281-290.

3. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'Agostino RB, Sr., Fox CS, Larson MG, Murabito JM, O'Donnell CJ, Vasan RS, Wolf PA, Levy D. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am.J Epidemiol.* 2007;165:1328-1335.

4. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol.* 2006;35:34-41.

5. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol.* 1991;1:263-276.

6. Jarvelin MR, Sovio U, King V, Lauren L, Xu B, McCarthy MI, Hartikainen AL, Laitinen J, Zitting P, Rantakallio P, Elliott P. Early life factors and blood pressure at age 31 years in the 1966 northern Finland birth cohort. *Hypertension.* 2004;44:838-846.

7. Lorentzon M, Swanson C, Andersson N, Mellstrom D, Ohlsson C. Free testosterone is a positive, whereas free estradiol is a negative, predictor of cortical bone size in young Swedish men: the GOOD study. *J Bone Miner Res.* 2005;20:1334-1341.

8. Richards JB, Papaioannou A, Adachi JD, Joseph L, Whitson HE, Prior JC, Goltzman D. Effect of selective serotonin reuptake inhibitors on the risk of fracture. *Archives of internal medicine.* 2007;167:188-194.

9. Jackson SA, Tenenhouse A, Robertson L. Vertebral fracture definition from population-based data: Preliminary results from the Canadian Multicenter Osteoporosis Study (CaMos). *Osteoporosis International.* 2000;11:680-687.

10. Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, Maier LM, Smink LJ, Lam AC, Ovington NR, Stevens HE, Nutland S, Howson JM, Faham M, Moorhead M, Jones HB, Falkowski M, Hardenbol P, Willis TD, Todd JA. Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat Genet.* 2005;37:1243-1246.

11. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet.* 2007;39:906-913.

12. Gauderman WJ, Morrison JM. QUANTO 1.1: A computer program for power and sample size calculations for genetic epidemiology studies. Available at: http://hydra.usc.edu/gxe.