


RESEARCH ARTICLE

# The Combined Effect of Common Genetic Risk Variants on Circulating Lipoproteins Is Evident in Childhood: A Longitudinal Analysis of the Cardiovascular Risk in Young Finns Study

Marie-jeanne Buscot<sup>1\*</sup>, Costan G. Magnussen<sup>1,2</sup>, Markus Juonala<sup>2,3,4</sup>, Niina Pitkänen<sup>2</sup>, Terho Lehtimäki<sup>5</sup>, Jorma S. A. Viikari<sup>2,3</sup>, Mika Kähönen<sup>6</sup>, Nina Hutri-Kähönen<sup>7</sup>, Nicholas J. Schork<sup>8</sup>, Olli T. Raitakari<sup>2,3</sup> , Russell J. Thomson<sup>1</sup> 

**1** Menzies Institute for Medical Research, University of Tasmania, Hobart, Australia, **2** Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland, **3** Department of Medicine, University of Turku and Division of Medicine, Turku University Hospital, Turku, Finland, **4** Murdoch Children Research Institute, Parkville, Australia, **5** Fimlab Ltd, University of Tampere and Tampere University Hospital, Tampere, Finland, **6** Department of Clinical Physiology, University of Tampere School of Medicine and Tampere University Hospital, Tampere, Finland, **7** Department of Pediatrics, University of Tampere School of Medicine and Tampere University Hospital, Finland, **8** Human Biology, The J. Craig Venter Institute, La Jolla, CA, United States of America

 These authors contributed equally to this work.

\* [m.buscot@utas.edu.au](mailto:m.buscot@utas.edu.au)



CrossMark  
click for updates

 OPEN ACCESS

**Citation:** Buscot M-j, Magnussen CG, Juonala M, Pitkänen N, Lehtimäki T, Viikari JSA, et al. (2016) The Combined Effect of Common Genetic Risk Variants on Circulating Lipoproteins Is Evident in Childhood: A Longitudinal Analysis of the Cardiovascular Risk in Young Finns Study. PLoS ONE 11(1): e0146081. doi:10.1371/journal.pone.0146081

**Editor:** David Fardo, University of Kentucky, UNITED STATES

**Received:** June 24, 2015

**Accepted:** December 11, 2015

**Published:** January 5, 2016

**Copyright:** © 2016 Buscot et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data analysed in the manuscript was obtained from the Cardiovascular Risk In Young Finns study after submission and approval of our study plan by the Young Finns Study coordinators. Due to ethical restrictions related to patient privacy, data is available to interested researchers upon request only. Data requests should be addressed to the project coordinator, Dr. Olli Raitakari (Email: [olli.raitakari@utu.fi](mailto:olli.raitakari@utu.fi)), Tel: +358 2 333 7220 Fax: +358 2 333 7270. Written requests should be sent to the following postal address: Research

## Abstract

Low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and triglycerides (TG) are modifiable risk factors for cardiovascular disease. Several genetic loci for predisposition to abnormal LDL-C, HDL-C and TG have been identified. However, it remains unclear whether these loci are consistently associated with serum lipid levels at each age or with unique developmental trajectories. Therefore, we assessed the association between genome wide association studies (GWAS) derived polygenic genetic risk scores and LDL-C, HDL-C, and triglyceride trajectories from childhood to adulthood using data available from the 27-year European ‘Cardiovascular Risk in Young Finns’ Study. For 2,442 participants, three weighted genetic risk scores (wGRSs) for HDL-C (38 SNPs), LDL-C (14 SNPs) and triglycerides (24 SNPs) were computed and tested for association with serum lipoprotein levels measured up to 8 times between 1980 and 2011. The categorical analyses revealed no clear divergence of blood lipid trajectories over time between wGRS categories, with participants in the lower wGRS quartiles tending to have average lipoprotein concentrations 30 to 45% lower than those in the upper-quartile wGRS beginning at age 3 years and continuing through to age 49 years (where the upper-quartile wGRS have 4–7 more risk alleles than the lower wGRS group). Continuous analyses, however, revealed a significant but moderate time-dependent genetic interaction for HDL-C levels, with the association between HDL-C and the continuous HDL-C risk score weakening slightly with age. Conversely, in males, the association between the continuous TG genetic risk score

Centre of Applied and Preventive Cardiovascular, Medicine University of Turku, Kiinamyllynkatu 10, FIN-20520 Turku, FINLAND.

**Funding:** The Young Finn study was financially supported by the Academy of Finland (grants 126925, 121584, 124282, 129378, 117797, 265877, and 41071), the Social Insurance Institution of Finland, the Turku University Foundation, Paavo Nurmi Foundation, Juho Vainio Foundation, Sigrid Juselius Foundation, Maud Kuistila Foundation, Research funds from the Kuopio, Turku and Tampere University Hospitals, the Finnish Foundation of Cardiovascular Research, the Finnish Medical Foundation, the Orion-Farmos Research Foundation, and the Finnish Cultural Foundation. CGM holds a National Health and Medical Research Council Early Career Fellowship (Public Health Fellowship, APP1037559). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the present manuscript. Fimlab Ltd provided support in the form of salaries for author TL, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

**Competing Interests:** Co-author Terho Lehtimäki is employed by Fimlab Ltd. There are no patents, products in development, or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

and triglycerides levels tended to be lower in childhood and become more pronounced after the age of 25 years. Although the influence of genetic factors on age-specific lipoprotein values and developmental trajectories is complex, our data show that wGRSs are highly predictive of HDL-C, LDL-C, and triglyceride levels at all ages.

## Introduction

Cardiovascular disease (CVD) is the leading cause of death and a major health burden worldwide [1]. Although modified by diet, obesity, lifestyle and other environmental factors, circulating lipoproteins represent a crucial partly heritable risk factor for atherosclerosis and CVD [2, 3]. Notably, elevated levels of total cholesterol and low-density lipoprotein cholesterol (LDL-C), have shown association with preclinical atherosclerosis in children and adolescents [4], significantly contributing to adult atherosclerosis. LDL-C, in particular, plays a major role in the initiation and progression of atherosclerotic lesions [5, 6]. The relevance of high-density lipoprotein cholesterol (HDL-C) and triglycerides (TG) to cardiovascular risk has been extensively debated in the last two decades. Although recent findings have questioned the causal relationship between genetically-defined HDL-C levels and atherosclerosis [7, 8], numerous prospective and case-control epidemiological studies have reported an inverse association between HDL-C levels and the risk of CVD [9–11]. Low HDL-C is therefore considered an independent risk factor for an increased risk of coronary artery disease, although pathways to its potential antiatherogenicity, are still not well understood [12]. High triglyceride (TG) levels are markers for several types of atherogenic lipoproteins involved in atherosclerosis. In contrast to LDL-C, the epidemiologic evidence-base for elevated TG as a CVD risk factor is generally less clear [13–15]. However, recent evidence suggests that genetically-defined plasma TG levels are associated with coronary artery disease risk, even after correcting for confounding effects due to LDL-C or HDL-C levels [16]. In light of these associations, managing dyslipidemia remains a continuing trend both in primary and secondary prevention and risk reduction of CVD worldwide [17].

While the greatest deviations from normal levels of lipoproteins are principally monogenic, the majority of adverse circulating lipid profiles arise as polygenic disorders with a substantial environmental component (e.g. diet, smoking, obesity) [18]. Although dyslipidemia is common, the complex interplay between various genetic and environmental risks that lead to onset and progression of the condition are still poorly understood. In the past 10 years, multiple genetic linkage analyses, candidate gene analyses, and large-scale genome-wide association studies (GWAS) have pinpointed a number of common genetic variants of candidate genes associated with inter-individual variation in plasma lipid levels [2, 19–24], indicating a genetic predisposition to dyslipidemia. Most lipid-associated SNPs (single nucleotide polymorphisms) are characterized by relatively small effect sizes, however some of the reported loci contain genes of clear biological and clinical importance, implicated in established mechanisms of lipoprotein metabolism [25]. Because most individual risk variants only explain a small fraction of those traits' heritability, the development of multilocus genetic risk scores that combine or accumulate the influence of validated susceptibility markers have proliferated in the hope of improving genetic CVD and other disease risk prediction [26–28].

Unfortunately, it is currently largely unknown whether reported lipid-associated risk alleles have any clinical relevance for a genetic predisposition to elevated adult or childhood lipid levels. With most GWAS leveraging cross-sectional samples from middle-aged adult populations,

the relative contribution of these genetic factors to the early stages and development of dyslipidemia between childhood and adulthood remains poorly understood. Moreover, whether identified risk alleles or GWAS-derived genetic risk scores predict lipid trait levels at all ages or whether they are associated with the intra-individual progression of serum lipids over time is unknown.

We examined the combined time-averaged and time-dependent effect of validated genetic risk alleles on lipoprotein trajectories across the life-course in a well-studied prospective population sample: the Cardiovascular Risk in Young Finns Study. Tikkanen et al. have reported population specific cross-sectional associations of 95 GWAS-derived individual SNPs with lipid levels in the Finnish cohort [29]. However, no study to date has investigated the association between lipid genetic risk scores and the development of lipid-trajectories across the life course at the individual level. Our primary aim was to quantify the contribution of a multigenic lipoprotein score to elevated levels of the LDL-C (14 risk SNPs) and triglycerides (24 risk SNPs) as well as depressed levels of the atheroprotective HDL-C (38 risk SNPs) from childhood to adulthood. We also investigated whether a multigenic predisposition to adult dyslipidemia might be modified by a lifestyle trajectory indicator such as the magnitude of BMI change from childhood to adulthood.

## Materials and Methods

### Study sample

The Cardiovascular Risk in Young Finn Study is an ongoing population-based prospective cohort that started in Finland in 1980. It was designed as a collaborative effort between the 5 medical schools in Finland to investigate cardiovascular risk factors and their determinants from childhood to adulthood [30]. A varying number of participants from the original cohort (consisting of 3596 children aged 3 to 18 years in 1980) were measured through middle adulthood (maximum age 49 years) in 2011 for numerous traits related to CVD development, and have as many as seven follow-up measurements (Tables A and B in [S1 File](#), Fig B in [S2 File](#)). Among the traits measured at multiple times, serum lipoproteins (plasma concentrations of LDL-C, HDL-C and TG) were collected at baseline and all seven follow-ups. For this study, analyses of the association between polygenic risk scores and circulating lipoprotein trajectories from 1980 to 2011 were pursued. These analyses included up to 2442 participants for whom genetic information was available for each of 76 risk SNPs identified in the literature. Participants reporting use of lipid lowering medication in 2001, 2007 and 2011 were excluded from the analyses ( $n = 7$ , 46 and 77 participants excluded respectively for the analysis of LDL-C, HDL-C and TG life course trajectories). Participants or their parents provided written informed consent, and the study was approved by local ethics committees (The Ethics Committee of the Hospital District of Southwest Finland) in agreement with the Declaration of Helsinki.

### Measures

**Serum lipoproteins.** All serum lipid determinations were performed in duplicate on fasting samples using standard methods in the same laboratory. HDL-C was determined enzymatically after precipitation of very low-density lipoprotein and LDL-C with dextran sulfate 500 000 (Olympus System Reagent, Olympus Diagnostica, Hamburg, Germany) in a clinical chemistry analyzer (AU400, Olympus Optical, Mishima, Japan) [31]. The concentration of LDL-C was determined indirectly by the Friedewald formula, so those participants with triglycerides  $>4.0$  mmol/L ( $n = 32$ ) were not included in the LDL-C evaluation [32]. Specific details on the lipid determination methods used in earlier—[33] and later follow-up studies [34] have been

published previously. To adjust for changes in kits and determination methods across study years, lipoprotein levels from early follow-ups were corrected to those measured in the most recent follow-up using correction factor equations, which were determined with linear regression analysis utilizing standardized principal component adjustment [34, 35]. For each lipid, the specific calibrating equations and the dates at which analyzers and reagent suppliers changed along study years are presented in the Appendix 1 of [33] and in [35]. Note: We focused on HDL-C profiles in the present study because measures of other HDL-C fractions were not available at all timepoints in the YF study sample.

Adverse levels of lipoproteins (i.e. abnormally high LDL-C/triglycerides, and low HDL-C), which identify participants at increased risk of developing atherosclerotic CVD (normal risk vs. high risk), were defined using the NCEP adolescent and childhood cut points (for participants under 19 years) [36] and NCEP adult-treatment panel guidelines (for participants 19 years old and older) [37]. The NCEP has not defined desirable and undesirable triglycerides levels for children and adolescents, so high risk triglycerides levels are classified using cutoffs suggested in AAP and AHA pediatric guidelines [38]. The cut points used to define 'normal' and 'high risk' lipid levels are shown in [S3 Table](#).

**Genetic measures and genetic risk scores.** In the 2001 follow-up, a subset of original participants (1,123 males, 1,319 females) were successfully genotyped using a custom Illumina BeadChip containing 670,000 SNPs and CNV probes, for a final list of 546,677 SNPs that passed QC and allele frequency filters. The exact custom content of the probes, as well as initial clustering, filtering, and data exclusion are described by Smith et al. [10]. Genotype imputation was performed using MACH[39] with the HapMap haplotypes as a reference panel (phase II, release 22 CEU, NCBI build 36, dbSNP 126). In the present analyses, we used 38 HDL-C, 24 TG and 14 LDL-C associated SNPs identified by genome wide meta-analysis conducted by the Global Lipids Genetics Consortium (GLGC) on 46 lipid GWAS carried out in over 100,000 European individuals of Caucasian descent [23]. Three composite genetic risk scores (wGRSs) (LDL-C, HDL-C and TG wGRS) were constructed as the arithmetic sums of these 38 HDL-C lowering, 24 TG-raising and 14 LDL-C raising alleles respectively, weighted by their effect sizes (in  $\text{mg}\cdot\text{dl}^{-1}$ ) as established from a published large-scale meta-analysis [23, 28] ([S2 Table](#)). To avoid redundancy and overlap of genetic information, in each lipid wGRS, we chose to include only the SNPs with which it showed the strongest independent associations among the 3 lipid traits in the meta-analysis [23]. The variant rs9411489 was not included in the LDL-C wGRS because it was missing on the chip and not available in the HapMap 2 reference panel. For comparability of metrics and to estimate the wGRSs ability to discriminate between extreme lipoprotein phenotypes, participants were categorized into 'high' and 'low' genetic risk groups categories defined as the cohort-specific lower (25<sup>th</sup> percentile) and upper (75<sup>th</sup> percentile) quartile of each composite risk score variables (HDL-C wGRS, TG wGRS and LDL-C wGRS). In each case, the other 50% of participants, lying in the interquartile range, were classified as 'medium' genetic risk ([Table 1](#)). Histograms of each lipid's wGRS are presented in Fig A in [S2 File](#). This approach is commonly preferred to case-control dichotomy when investigating the association between genetic factors and disorders implicating quantitative traits continuously distributed over the population (such as dyslipidemia), because it increases the statistical power of testing the variants for association [40].

## Statistical analyses

**Association between longitudinal lipoprotein profiles and composite genetic risk scores.** The principal outcome was the association between the categorical polygenic risk score status (High vs. Low wGRSs) and longitudinal trends in HDL-C, LDL-C and triglycerides

**Table 1. Average lipid concentrations in childhood, young adulthood and middle adulthood, across 1980–2011 (in mmol/L), and genetic risk factors considered in the longitudinal lipoprotein profile analyses.**

	Males	Females
<b>HDL Analysis</b>	(N = 1064**)	(N = 1244**)
<b>Average HDL-C*</b>	1.37 (0.36) (N† = 9043)	1.51 (0.32) (N† = 10540)
<b>3–15 years</b>	1.58 (0.34) (N† = 2649)	1.57 (0.30) (N† = 3078)
<b>18–30 years</b>	1.28 (0.30) (N† = 3374)	1.52 (0.33) (N† = 3937)
<b>33–49 years</b>	1.19 (0.29)(N† = 3020)	1.42 (0.31) (N† = 3525)
<b>Genetic risk:</b>		
Average HDL wGRS	32.46 (3.36)	32.62 (3.41)
High score (wGRS >34.84)	N = 253	N = 324
Mid score (30.1<wGRS≤34.8)	N = 541	N = 613
Low score (wGRS ≤30.1)	N = 270	N = 307
<b>LDL Analysis</b>	(N = 1121**)	(N = 1314**)
<b>Average LDL-C*</b>	3.22 (0.86) (N† = 9530)	3.17 (0.81) (N† = 10834)
<b>3–15 years</b>	3.12 (0.83) (N† = 2781)	3.32 (0.84) (N† = 3255)
<b>18–30 years</b>	3.07 (0.85) (N† = 3563)	3.06 (0.80) (N† = 3856)
<b>33–49 years</b>	3.41 (0.85) (N† = 3186)	3.07 (0.74) (N† = 3723)
<b>Genetic risk:</b>		
Average LDL wGRS	42.1 (6.6*)	41.9 (6.9*)
High score (wGRS >46.1)	N = 278 (25%)	N = 332 (25%)
Mid score (37.5<wGRS ≤46.2)	N = 553 (50%)	N = 665 (50%)
Low score (wGRS ≤37.5)	N = 290 (25%)	N = 317 (25%)
<b>Triglycerides Analysis</b>	(N = 1121*)	(N = 1314*)
<b>Average Triglycerides*</b>	1.17 (0.96) (N† = 9513)	1.00 (0.56) (N† = 11148)
<b>3–15 years</b>	0.73 (0.32) (N† = 2776)	0.79 (0.34) (N† = 3257)
<b>18–30 years</b>	1.17 (0.69) (N† = 3557)	1.06 (0.53) (N† = 4163)
<b>33–49 years</b>	1.56 (1.10) (N† = 3180)	1.15 (0.89) (N† = 3728)
<b>Genetic risk:</b>		
Average TGwGRS	32.71 (15.81)	131.91 (15.72)
High score (wGRS >142.37)	N = 280 (25%)	N = 334 (25%)
Mid score (121.61<wGRS ≤142.37)	N = 280 (25%)	N = 660(50%)
Low score (wGRS ≤121.61)	N = 275 (25%)	N = 322 (25%)

\*Data are sex-specific averages (SD) for lipoprotein concentration and for continuous genetic risks cores (wGRSs) collected for the entire study sample between 1980 and 2011 (the average age of male participants was 24.2 (11.8) years and the average age of female participants was 24.2 (11.8) years over the study period, which was not significantly different). We also present average (SD) lipoprotein levels stratified by age group (i.e. childhood (3–15 years), young adulthood (18–27 years), and middle adulthood (30–49 years)). The grouping of wGRSs into categories was based on whole-cohort 25<sup>th</sup> and 75<sup>th</sup> percentiles (See [methods](#))

Abbreviations: HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.

\*\*Indicates the number of participants included in the longitudinal lipoprotein profile analyses.

† Indicates the number of available measurements for the calculation of each average lipoprotein concentrations.

doi:10.1371/journal.pone.0146081.t001

levels from 1980 through 2011. To determine whether sex and genetic risk group membership modifies average lipoprotein level or the growth parameters of the participants' lipoprotein trajectories over time, we used individual growth curve analysis (IGC), an advanced multilevel mixed effect regression technique that allows to model simultaneously inter-individual differences in intra-individual systematic changes over time (i.e. repeated individual measurements) [41–44]. An IGC model comprises 3 main components: (A) the functional form of the

response variable, which partitions and quantifies the variance across people and time, (B) the fixed effects (i.e. group-level predictors of change), and (C) the stochastic part of the model, which includes the random effects (i.e. individual effects on growth parameters), and the residual error covariance structure. While there is, in the literature, a few variants in the specification and the procedure of IGC model building (sometimes also referred to as ‘growth curve analysis’ (GCA)[43]), we followed the modeling strategy suggested by Singer and Willett [41], with a few adaptations (see [S1 Appendix](#) for a step by step protocol of our modeling approach, and explicit parametrization of the IGC submodels).

Prior IGC analysis, individual empirical growth plots and generalized additive mixed models (GAMMs) were used to explore the functional form (shape) lipid profiles across the life course in the YF cohort [45] and inform the modeling procedure. For each lipid, the IGC analyses then consisted in testing several submodels as follow: (1) an unconditional mean (UM) model (i.e. null model), examining any difference in average lipid levels between individuals, (2) a linear unconditional growth (UG) model (with no group-level predictors), used as a reference to explore the functional shape of the lipid growth overtime (3) two or more higher-order polynomial UG models to test if the lipid rate of change was accelerated or decelerated as subjects aged (i.e. curvilinear age-related change), (4) models for slope(s) variability, to test for random trajectory parameters between participants, (5) a set of models to assess the within-subject error structure of the best UG model to test if incorporating (a) an autoregressive structure with serial correlation, and (b) heterogeneity of the residual error will improve model fit, and finally (5) a conditional growth (CG) model, where wGRSs, sex and their cross-product are introduced as subject-level predictors of each growth parameter variability (i.e. intercept, linear -, quadratic- and cubic (and quartic) growth). CG models allows assessing average wGRSs effect effect at baseline and whether or not there was an age-dependent effect of wGRS score on each lipid’s trajectory parameters (i.e  $wGRS * age$ ,  $wGRS * (age)^2$ , . . .). CG models also examine whether individual variability in lipid intercept and slopes estimates can be accounted for by the interaction of wGRSs and sex. Throughout the IGC analyses, when comparing increasingly complex submodels, the improvement in model fit is assessed by likelihood ratio test (LR-test) or using Akaike’s and Bayesian Information Criterion (AIC and BIC). The significance of each estimated model growth parameter in the final CG model is assessed with t-statistics (i.e. defined as the ratio of parameter estimate and SE) ([S1 Appendix](#)). A flowchart of the IGC modeling approach is outlined in [S1 Fig](#).

Prior to introducing sex and wGRSs as time-independent predictors of individual variability in lipid trajectories, we considered the need to minimize for confounding by (1) birth cohort and (2) period effect. We tested whether “year of birth” and calendar “year” at examination respectively modified the age-related trajectory of lipids across the life course. This was done by adding the variables (1) birth year (“yob”, categorical variable with 6 levels) and (2) year at follow-up (“year”, centered around baseline (1980)), and their interactions with trajectory parameters to each lipid’s sex-specific UG model ([S1 Appendix](#)). Birth cohort does not appear to significantly modify the lipid profiles across the life course in this study sample (i.e. later birth cohorts do not show significantly different lipid trajectories compared to earlier birth cohorts). However, as we found significant linear yearly secular trends for each lipid (see [Results](#)), we adjusted for “year” in all subsequent steps of the IGC analyses.

HDL-C and LDL-C distributions were reasonably close to normal ([S2 Appendix](#)). Because triglyceride levels showed very long tails skewed to the right, we applied the Box-Cox procedure to determine the optimal transformation to remediate deviations from the assumptions of the linear regression model [46]. As the best transformation ( $\lambda = -0.2$ ) was close to the logarithmic case, Natural logarithm was used to transform triglycerides levels prior all analyses. Ages of participants at each measurement were treated as continuous covariates, that we centered

around youngest age at baseline (3 years old) to avoid collinearity problems with higher-order polynomial age terms and their interactions in the multilevel mixed models [47]. As an index of fit of the different linear mixed effect models (i.e. final CG models), and to estimate how much the genetic predictors contribute to the variation of the lipoprotein profiles outcome, we used the novel conditional  $R^2$  and marginal  $R^2$  for linear mixed models developed by Nakagawa & Schielzeth, 2013 [48] and adapted by Johnson, 2014 [49] to accommodate for random slopes. For a given mixed effect model, the marginal  $R^2$  describes the proportion of variance explained by the fixed effects alone, while the conditional  $R^2$ , describes the proportion of variance explained by both the fixed and subject- level random factors. In the case of IGC models, which are typically hierarchical mixed models, these two novel coefficients of determination are superior to the pseudo- $R^2$  often reported for linear mixed model (i.e. squared correlation between the fitted and observed values) which ignores the variance components at multiple levels of the random factors by choosing to calculate  $R^2$  relative to only the residual variance. All analyses were performed in R 3.0.1 9 [50] using the nlme 3.1.102 [51] and mgcv [52] packages.

To complement the categorical analyses and make inferences at the population levels, we used a similar age- and sex- adjusted mixed modeling growth curve analysis approach to examine the association of the continuous wGRSs and the longitudinal trends of HDL-C, LDL-C and triglycerides. As per above, main effects (age-averaged) as well as age-dependent effects of continuous lipoprotein risk scores on lipoprotein trajectory parameters were assessed. Z-scores were calculated for the wGRS prior to these continuous analyses, so that for LDL-C, HDL-C and, TG the estimated effects (i.e., the regression parameters:  $s$ ) indicate the change in mmol/L lipoprotein level per 1-sd change in wGRSs. For triglycerides, the regression coefficient  $s$  were exponentiated for ease of interpretation, so that  $\exp(s)$  correspond to changes in the ratio of the expected triglyceride level per 1-sd change in wGRSs.

## Secondary analyses

**Age- and sex stratified linear regression analysis.** To ascertain the age at which the polygenic effect on the lipoproteins is first detectable and examine the strength of association between continuous genetic risk scores wGRSs and lipoprotein levels across age groups, we used sex-specific age-stratified linear regression models adjusted by study year. Trends in the associations between age-groups were assessed using LOESS curves. Additionally, we present a table summarizing sex-specific associations between categorical genetic risk scores and lipid levels at ages 3 years, 15 years, 24 years and 45 or 46 years (S4 Table). The reported mean effect sizes are in mmol/L for the number of risk allele differences between high and low wGRSs for each lipid.

**Indication of polygenic gene-lifestyle interaction on adult lipoprotein?** Because lifestyle factors relating to weight status (i.e. dietary, exercise, and sedentary habits) are known to strongly correlate with blood lipids, we asked whether the polygenic effect of risk loci on adult lipoproteins might be modified by an individuals' BMI trajectory from childhood to adulthood. For this, we test for an interaction between continuous wGRS and change in standardized BMI between childhood and adulthood,  $\Delta BMI_i$ , in sex- and age-adjusted linear regression models of adult lipoproteins ( $n = 2100$  for adult LDL-C model,  $n = 2062$  for adult triglycerides, and  $n = 2034$  for adult HDL-C model). For each participant  $i$ , this measure was calculated as:

$$\Delta BMI = z.score\ BMI_{(adult)_i} - average(z.score\ BMI_{(childhood)_i}) \quad (1)$$

with  $z.score\ BMI_{(adult)}$  defined as the sex-specific BMI z-scores observed at the latest follow-up attended in adulthood for adults 30 years old or older in 2001, 2007 or 2011 (i.e participants younger than 30 years old in 2001 were excluded); and  $average(z.score\ BMI_{(childhood)})$  defined

as the participant's average of sex-and age-specific BMI z-scores measured at multiple occurrences in childhood (ages at follow-ups ranging 3 to 18). The significance of the interaction term (wGRS \*  $\Delta$  BMI) as a predictor of either adult lipoprotein levels was assessed by a likelihood ratio test.

## Results

Participants' characteristics for lipoprotein levels and genetic risk scores considered in the longitudinal analyses are shown in [Table 1](#). The difference in the number of risk alleles between subjects in the High wGRS and low wGRS group ranged between 4 and 7 alleles in average alleles for LDL-C and HDL-C, respectively ([S4 Table](#)). When stratified by life stages (i.e. childhood (ages 3 to 15), young adulthood (ages 18 to 27) and middle adulthood (ages 30 to 49), average concentrations of each of the three lipoproteins were mostly consistent with pediatric and adult healthy cholesterol and fasting triglycerides levels (NECP guidelines 2010), and the standard deviations were homogeneous over time. For each lipid, a histogram of the continuous wGRSs distribution also showing the quartile stratification into 'low', 'mid' and 'high' genetic risk is shown in Fig A in [S2 File](#). Additional descriptive statistics showed that the wGRSs were not strongly linearly correlated with the lipoproteins overall when ignoring the effect of participants age (Pearson's  $r = 0.21$  for LDL-C, 0.19 for HDL-C and 0.18 for triglycerides).

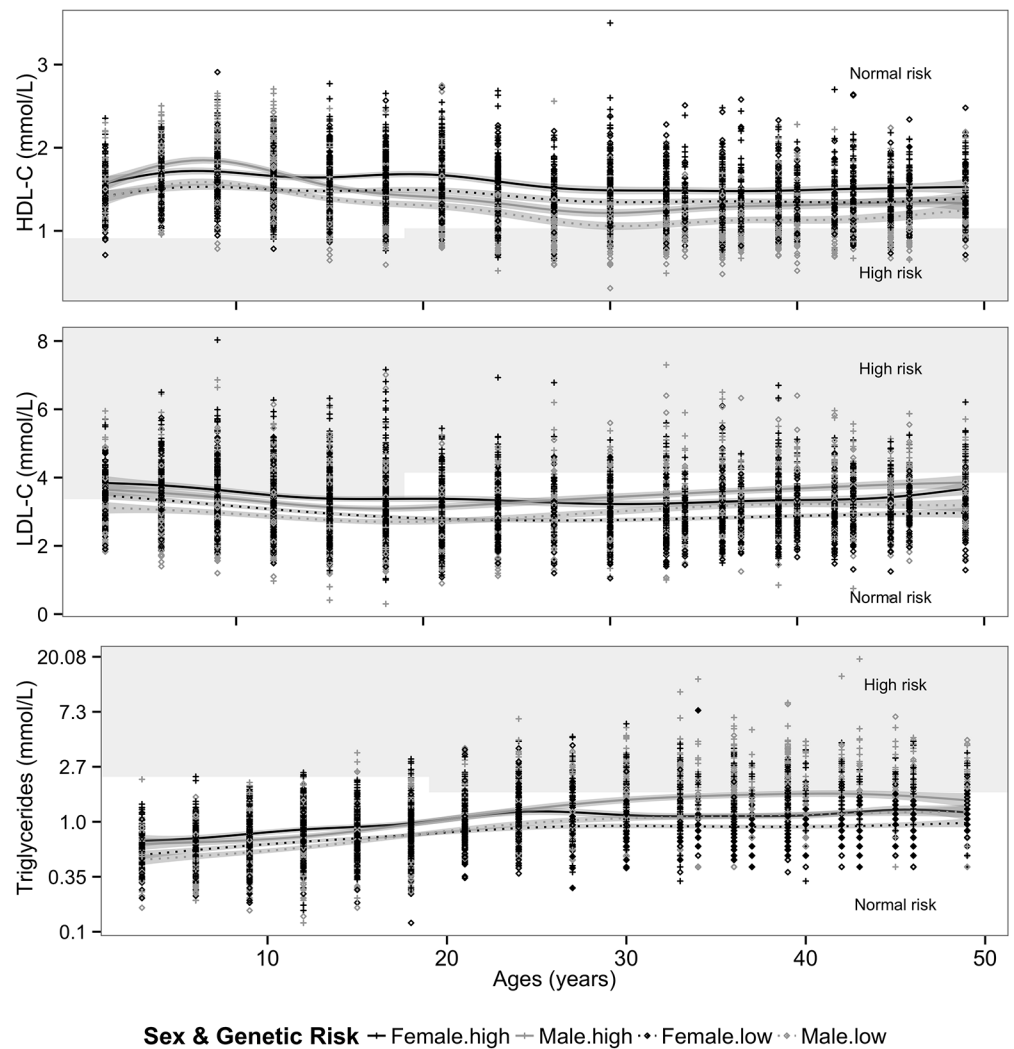
### Longitudinal lipoprotein profiles

We found significant decreasing secular trends for LDL-C and TG between 1980 and 2011 ( $\beta_{\text{year}_{\text{LDL-C}}} = -0.09$  (se = 0.007) and  $\beta_{\text{year}_{\text{TG}}} = -0.003$  (se = 0.0008) respectively, p-values < 0.05), but the decreasing trend was not significantly different between males and females. HDL-C, conversely, showed a modest yearly increase ( $\beta_{\text{year}_{\text{HDL-C}}} = 0.005$  (se = 0.001), p-value < 0.05) for both sexes. However, calendar year at examination did not appear to modify the average age-related trajectories of either lipid in the cohort (i.e. all 3-way year\*sex\*age-terms interactions were non-significant, p-values > 0.05).

The sigmoidal function of age developed by Wineinger et al. 2013 [[53](#)] did not fit the lipoprotein profiles in our study sample ( $\text{lipoprotein}_{(t)} \sim \text{intercept} + \sin(\pi/2 * ((2 * \text{age}_{(t)} - \text{max}(\text{age})/\text{min}(\text{age})) * \text{sex}))$ ), and the best non-linear fits were achieved by using a 4<sup>th</sup> polynomial age term for HDL-C profile and 3<sup>rd</sup> degree polynomial age term LDL-C and triglycerides profiles (i.e. models yielding the lowest AIC and BIC values). Final IGC models assess the effect of sex, wGRSs and their interaction as predictors of the individual variability in lipid growth parameters over the life course. Model selection for the optimal random effects structure revealed that a continuous first order autoregressive correlation structure was needed in each model for the error term, implying that the within-subject correlation between lipid measures drops exponentially with increasing temporal separation. [Fig 1](#) shows the predicted average sex-specific lipid trajectories for the participants in the upper and lower wGRSs quartiles with corresponding 95% confidence intervals, determined from the estimated parameters and SEs of the final IGC models (i.e. prototypical growth curves). Sex-specific prototypical growth curves show that trends in lipoproteins from age 3 to 49 were different for males and females in the cohort. When stratified by sex, participants from 'high' or 'low' genetic risk groups (upper and lower wGRSs quartiles) showed differences in average levels of each lipoprotein from age 3, but displayed average profiles of globally similar shapes over time ([Fig 1](#)).

Participants 10 years and younger already had 'high-risk' or close to 'high-risk' average levels of LDL-C (especially for participants in the high genetic risk group, and females in the lowest LDL-C wGRS quartile) as defined by the NCEP pediatric and adolescent cut points [[36](#)]





**Fig 1. Scatterplot of serum lipoprotein longitudinal profiles of participants according to their sex and wGRSs status (High and Low wGRS\*) (N = 2435, N = 2308 and N = 2435 for LDL-C-, HDL-C- and triglyceride profiles respectively).** Solid and dotted lines represent estimated sex-specific average age-related lipid trajectories for participants in High and Low genetic risk score, respectively (i.e. prototypical growth curves); grey bands around the growth curves represent approximated 95% prediction CI. Overlaid with the prototypical lipid trajectories are the age-specific cut points for lipoprotein status (normal vs. high risk) as defined by the NCEP adolescent and childhood classification [36] and NCEP adult-treatment panel guidelines [37]. The cut points are represented in grey/white blocks, used to identify those at significantly increased risk of developing atherosclerotic CVD in adulthood. \* Mid wGRS risk group are not presented on the figure for the purpose of readability.

doi:10.1371/journal.pone.0146081.g001

(Fig 1). The decrease in HDL-C levels in the cohort was noticeable above the age of 30 years and was more pronounced among males, who in early life had higher average HDL-C levels compared with females, independent of their wGRS status. Similarly, triglyceride profiles in the cohort show a sex-dependent divergence over time, with males tending to exhibit higher average triglyceride levels compared to females from their early to mid 20's both in the high and low wGRS groups. Fig 1 also suggests the divergence by sex in adulthood is more pronounced in the 'high risk' TG wGRS participants, suggesting that males in this group have exacerbated average triglyceride levels from age 25 onwards.

The output of the wGRS group- and sex-adjusted individual growth curve models corroborate the observations that the genetic effect is already present in childhood for each of the three lipids, with significant time independent effects of polygenic genetic risk score categories on baseline lipoprotein levels (Table 2,  $\beta_{\text{high-low}}$  and  $\beta_{\text{high-mid}}$ ). Indeed, participants in the low and mid LDL-C wGRS group have LDL-C concentrations lowered by 46%, and 23% respectively (p-values <0.0001) as compared to participants in the high LDL-C wGRS. The analysis of HDL-C profiles revealed that participants in the low and mid HDL-C wGRS had average baseline HDL-C levels increased by ~ 16.8% and ~8% respectively as compared to participants in the high risk score group (p-values <0.0005) (Table 2). Similarly, participants in the low and mid TG wGRS had average baseline triglyceride levels lowered by ~ 20% (i.e. log (0.8)) and ~10% (i.e. log (0.9)) respectively as compared to participants in the high risk score group (p-values <0.0005).

Analyses of interactions terms revealed no time-dependent polygenic effects of the 14 risk SNPs on circulating LDL-C, implying that there is no worsening effect of LDL-C levels over time among those belonging to the high genetic risk group compared with those in the low or mid genetic risk group (all linear, quadratic and cubic  $\beta_{\text{wGRS} \times \text{age}}$  interactions p-values >0.3). On average at baseline, males had 7% higher LDL-C levels and 15% lower HDL-C levels compared with females (p-values <0.004). Sex appears to be the variable that drives the longitudinal trajectory of LDL-C levels in this cohort ( $\beta_{\text{sexMale} \times \text{f}(\text{age})} = 0.019$ , se = 0.0007, p-values <0.001), rather than wGRS group membership (linear, quadratic, and cubic rate of change not significant, Table 2).

For HDL-C however, we found a significant linear and quadratic age-dependent interactions between participants belonging to the high and low wGRS group ( $\beta_{\text{wGRS} \times \text{f}(\text{age})} = 0.022$  and  $\beta_{\text{wGRS} \times \text{f}(\text{age-square})} = -2.6 \times 10^{-04}$ ) (Table 2), suggesting that genetic group membership is a modifier of the HDL-C trajectory. The positive effect of high wGRS on the linear age term implies that, for children in the high wGRS category, the effect of the combined variants leads to an initial increase slightly faster (by 2.2%) per year as compared to children in the low wGRS group (i.e. slightly steeper linear increase). Similarly, the negative effect on the quadratic rate of change also indicates that for participants in the high wGRS group, the positive genetic effect on HDL-C levels will decelerate in time slightly slower (by 0.026%) than it does in the lowGRS group at around adolescence. This small difference in HDL-C trajectories between high and low wGRS groups is not easily distinguished in Fig 1, as the modifying effect of the variants on the trajectory parameters is relatively mild. These age-dependent interactions were not modified by sex (all 3-way interactions were not significant p>0.05), so that the effect of wGRS categories on HDL-C trajectory parameters were not significantly different in males and females.

For triglycerides, the linear, quadratic, and cubic age-dependent interactions between participants belonging to the high and low wGRS group are not significant in females (Table 2), however, for males, genetic risk group membership modified the linear change if triglycerides (3-way interactions between linear age-dependent change rate, sex, and TG wGRS group membership) were significant both for low vs. high genetic risk group ( $\exp \beta_{\text{male} \times \text{wGRS} \times \text{f}(\text{age})} = 0.96$ , se = 0.003, p-value = 0.0009), and for mid vs. high genetic risk group ( $\exp \beta_{\text{male} \times \text{wGRS} \times \text{f}(\text{age})} = 0.99$ , se = 0.002, p-value = 0.01). That is, males belonging to the high genetic risk group will tend to have a linear increase in TG levels by 4% (i.e. log(0.96)) and by 1% (i.e. log(0.99) for participants in the mid genetic risk group). As these effects are also moderate, it does not result in strong divergences in the prototypical triglyceride trajectories (Fig 1).

When using the wGRSs values as continuous predictors, the final IGC models for the age-related trajectories of LDL-C, HDL-C and triglycerides profiles achieved a conditional R<sup>2</sup> of 0.64, 0.68 and 0.61 respectively, with the fixed predictors (age, sex, and continuous wGRS) jointly accounting for 12%, 27% and 19% of deviance respectively in each model. Consistent

**Table 2. Time-averaged and time-dependent effects of the categorical combined genetic risk scores (HDL-C-, TG- and LDL-C wGRSs) on lipoprotein concentrations (mmol/L) from childhood through adulthood Regression coefficients  $\beta$ s are in % change against the reference group (i.e. High genetic risk group).**

Lipoprotein	Main wGRS effects <sup>a</sup>		Time-dependent wGRS effects <sup>b</sup>		Goodness of fit
	$\beta$ (se)	p-val	$\beta$ (se)	p-val	
<b>HDL-C†</b>					
<i>High vs. Low</i>	-0.17 (0.01)	0.0001*	L: 0.022 (0.0015)	0.005*	Marginal R <sup>2</sup> : 0.21
			Q: $-2.6 \times 10^{-4}$ ( $1.1 \times 10^{-3}$ )	0.02*	Conditional R <sup>2</sup> : 0.72
			C: $2.5 \times 10^{-6}$ ( $3.4 \times 10^{-6}$ )	0.47	
			4 <sup>th</sup> : $4.0 \times 10^{-7}$ ( $2.4 \times 10^{-7}$ )	0.09	
<i>High vs. Mid</i>	-0.08 (0.012)	$1.0 \times 10^{-4}$ *	L: 0.0014 (0.0009)	0.11	
			Q: $-1.0 \times 10^{-4}$ ( $9.0 \times 10^{-5}$ )	0.28	
			C: $2.5 \times 10^{-6}$ ( $5.1 \times 10^{-6}$ )	0.40	
			4 <sup>th</sup> : $1.0 \times 10^{-7}$ ( $2.1 \times 10^{-7}$ )	0.67	
<b>LDL-C††</b>					
<i>High vs. Low</i>	-0.46 (0.04)	$1.0 \times 10^{-4}$ *	L: $-7.0 \times 10^{-3}$ (0.002)	0.72	Marginal R <sup>2</sup> : 0.11
			Q: $2.8 \times 10^{-5}$ ( $9.5 \times 10^{-7}$ )	0.76	Conditional R <sup>2</sup> : 0.71
			C: $3.0 \times 10^{-6}$ ( $6.6 \times 10^{-7}$ )	0.71	
<i>High vs. Mid</i>	-0.23 (0.03)	$1.0 \times 10^{-4}$ *	L: $6.1 \times 10^{-2}$ ( $1.9 \times 10^{-2}$ )	0.52	
			Q: $6.3 \times 10^{-5}$ ( $6.1 \times 10^{-6}$ )	0.45	
			C: $1.3 \times 10^{-9}$ ( $1.0 \times 10^{-9}$ )	0.94	
<b>Triglycerides†††</b>					
<i>High vs. Low</i>	0.80 (0.025)	0.0001*	L: 0.995 (0.001)	0.11	Marginal R <sup>2</sup> : 0.18
			Q: 0.99 ( $6.8 \times 10^{-5}$ )	0.08	Conditional R <sup>2</sup> : 0.68
			C: 1.00 ( $4.9 \times 10^{-7}$ )	0.37	
<i>High vs. Mid</i>	0.90 (0.022)	$1.0 \times 10^{-4}$ *	L: 0.999 (0.001)	0.15	
			Q: $-0.997$ ( $5.8 \times 10^{-5}$ )	0.09	

Abbreviations: HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; *High/Mid/ Low*, categorical genetic risk score groups (for either lipoprotein trait); L, Q, C and 4<sup>th</sup>, Linear/quadratic/cubic/ and quartic rate of change (in either lipoprotein concentration as a function of age).

† / †† / ††† wGRS effects refer to the combined effect of the 38-, 14- and 24 SNPs associated respectively to HDL-C, LDL-C and fasting triglycerides levels (see [methods](#)).

\* Indicates that the estimated regression parameter is significant at the 0.05 significance level.

<sup>a,b</sup> For ease of interpretation of the estimates of main and time-dependent effects of wGRSs, all age terms were centered around youngest childhood age at baseline (1980) in the cohort (3 years old) prior regression analysis. For triglycerides, regressions coefficients  $\beta$ s of main and time-dependent effects were exponentiated so they are presented in the original scale for ease of interpretation.

doi:10.1371/journal.pone.0146081.t002

with the categorical analyses, the time-averaged genetic effect of the combined genetic risk score on lipoprotein profiles was significant ( $p < 0.0001$ ) for all traits ([Table 3](#)). The polygenic effect size was stronger for LDL-C associated risk SNPs, with LDL-C wGRS increasing LDL-C levels by 18.2% per SD increase in score (as compared to a 6.8% increase in HDL-C levels per SD increase in wHDL-C GRS). Additionally, wHDL-C GRS was a significant predictor of the linear, quadratic and quartic changes in HDL-C levels over time ( $p$ -values  $< 0.01$ , [Table 3](#)), implying that the slope of the genetic risk score variable on HDL-C concentration changes as participants age from childhood through adulthood. These time-dependent interactions are best visualized by computing and plotting the marginal effect of the combined genetic risk score on HDL-C levels ([Fig 2A](#)). The downward trends of the slopes of the continuous risk score on HDL-C level with age for both males and females, suggests that the association

**Table 3. Time-averaged and time-dependent effects of the continuous combined genetic risk scores (HDL-C-, LDL-C—and TG wGRSs) on lipoprotein concentrations (in mmol/L) from childhood through adulthood.**

Lipoprotein	Main wGRS effects <sup>a</sup>		Time-dependent wGRS effects <sup>b</sup>		Goodness of fit
	$\beta$ (se)**	p-val	$\beta$ (se)**	p-val	
HDL-C †	0.0064(0.002)	1.0x10 <sup>-4</sup> *	L: -6.1 x10 <sup>-3</sup> (2.0 x10 <sup>-4</sup> )	0.003 *	Marginal R <sup>2</sup> : 0.27
			Q: 2.0 x10 <sup>-5</sup> (4.1 x10 <sup>-6</sup> )	0.014*	Conditional R <sup>2</sup> : 0.68
			C: 4.0 x10 <sup>-6</sup> (1.2 x10 <sup>-7</sup> )	0.07	
			4 <sup>th</sup> : -1.0 x10 <sup>-7</sup> (7.8 x10 <sup>-9</sup> )	0.05.	
LDL-C ††	0.182 (0.001)	0.0003*	L: 2.7x10 <sup>-4</sup> (1.2 x10 <sup>-3</sup> )	0.85	Marginal R <sup>2</sup> : 0.13
			Q: -9.5 x10 <sup>-5</sup> (1.5x10 <sup>-4</sup> )	0.44	Conditional R <sup>2</sup> : 0.64
			C: 3.0 x10 <sup>-6</sup> (2.6 x10 <sup>-6</sup> )	0.23	
Triglycerides †††	1.094 (0.008)	1.0x10 <sup>-4</sup> *	L: 1.0 (9.8 x10 <sup>-4</sup> )	0.47	Marginal R <sup>2</sup> : 0.20
			Q: 1.0(8.0 x10 <sup>-5</sup> )	0.31	Conditional R <sup>2</sup> : 0.65
			C: 0.99 (1.7 x10 <sup>-7</sup> )	0.20	

Abbreviations: HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; L, Q, C and 4<sup>th</sup>, Linear/quadratic/cubic/ and quartic rate of change (in either lipoprotein concentration as a function of age).

† / †† / ††† wGRS effects refer to the combined effect of the 38-, 14- and 24 SNPs associated respectively to HDL-C, LDL-C and fasting triglycerides levels (see [methods](#)).

\* Indicates that the estimated regression parameter is significant at the 0.05 significance level.

<sup>a,b</sup> For ease of interpretation of the estimates of main and time-dependent effects of wGRSs, all age terms were centered around youngest childhood age at baseline (1980) in the cohort (3 years old) prior regression analysis.

\*\*Regression coefficient  $\beta$ s are in mmol/L per 1-sd change in wGRS for HDL-C and LDL-C. For triglycerides, regressions coefficients  $\beta$ s of main and time-dependent effects were exponentiated so they are presented in the original scale for ease of interpretation.

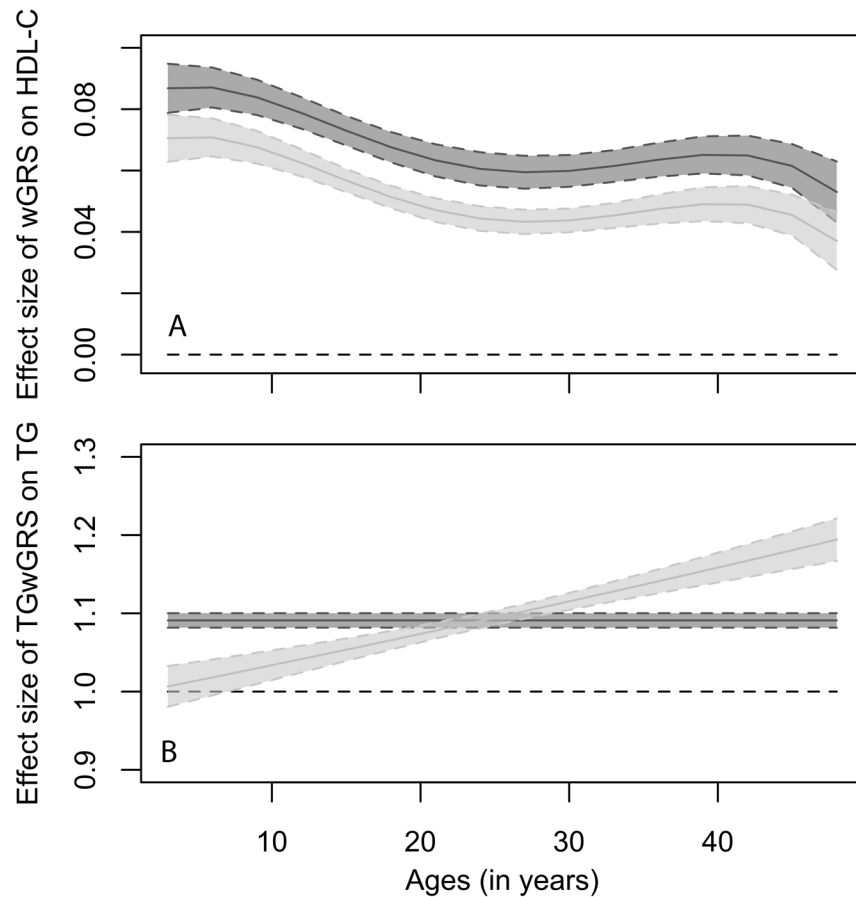
doi:10.1371/journal.pone.0146081.t003

between HDL-C levels and of the 38 HDL-C risk SNPs gets weaker as participants aged in this population, although the combined genetic effect of these loci on HDL-C was consistently stronger among females.

For triglycerides, continuous analyses revealed that while on average, TG wGRS did not modify the linear, quadratic or cubic change in triglyceride level, but when clustering participant by sex, TG wGRS effect on the linear change rate in triglyceride levels became significant for males (three way interaction  $\exp \beta_{wTGgrs24*sexMale*f(age)} = 1.0038$ , se = 0.0011, p-value = 0.001). This difference between sexes can also be visualised by plotting the marginal effect of the TG combined genetic risk score on triglyceride levels ([Fig 2B](#)). For females, a 1-sd increase in risk score will result in 9% higher serum triglyceride on average (regardless of their age;  $\log(1.1) = 0.09$ ). For males, the effect size of TG genetic risk score is age-dependent and lower in childhood than for females, it increases linearly to become larger in females from age 25 onwards, reaching 13% at age 49 years ( $\log(1.14) = 0.13$ ).

### Age-stratified linear regression

Cross sectional analyses confirmed that combined wGRSs were significantly associated with the three lipoproteins at all ages for both sexes ([Fig 3](#), all p-values <0.05). Confirming that the joint effect of the 14 risk SNPs on LDL-C levels is consistent across time, the fitting of a LOESS regression lines to the ages- and sex-specific regression parameters does not reveal any striking trend over time ([Fig 3](#), upper panel). We suggest that the variations in effect sizes between different age-groups were mostly attributable to differences in sample sizes and differences in the number of birth cohorts used in the regression analysis (only one cohort for the 3 year old age-group, against a mix of up to 5 cohorts for the 15 and 18 year old group). The age- and sex-



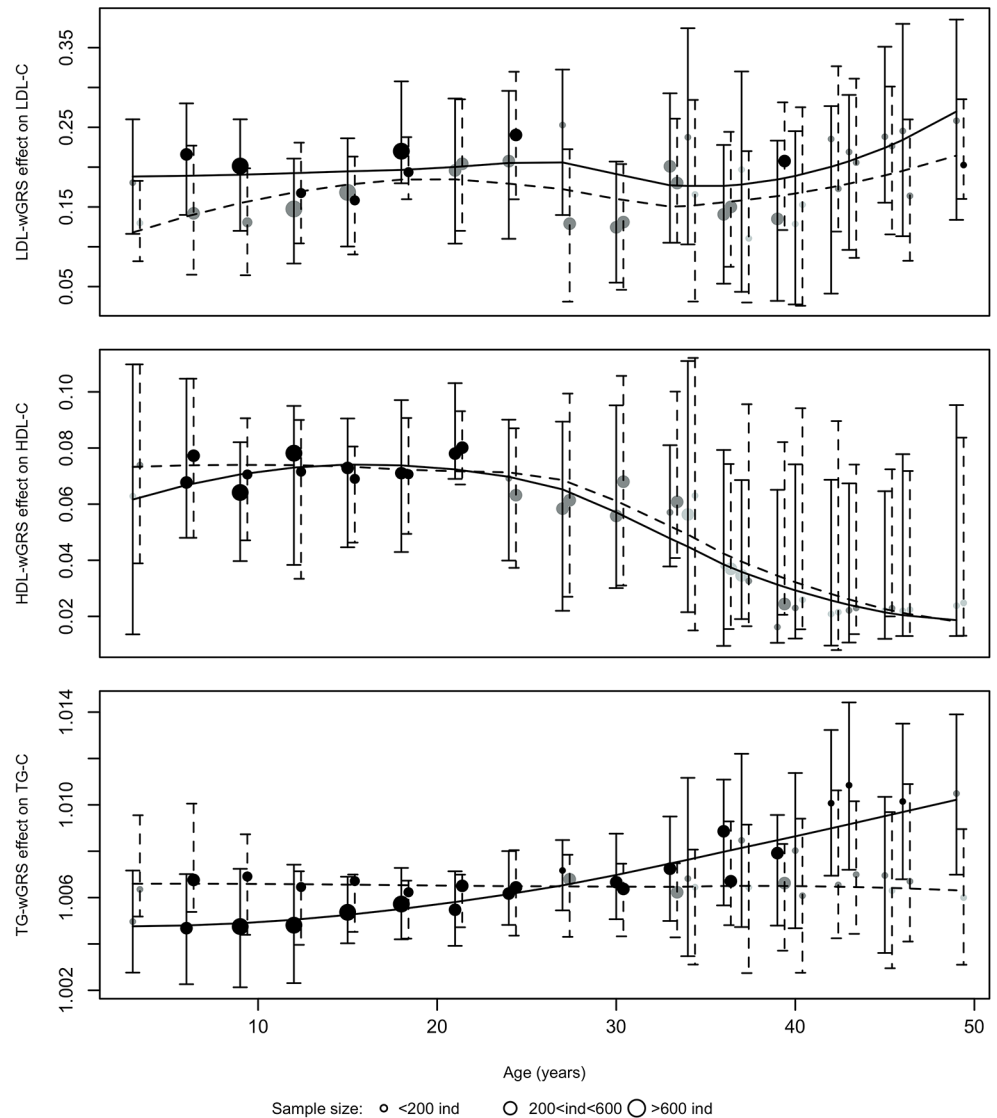
**Fig 2. Sex-specific marginal effect\* and 95% CI of (A) combined continuous HDL-C wGRS on HDL-C levels (effect size expressed in mmol/L lipoprotein level change per 1-sd change in wGRSs); and (B) combined continuous TG wGRS on fasting triglyceride levels (effect size expressed in odds ratio lipoprotein level change per 1-sd change in wGRSs).** Colour code: dark grey; females, light grey; males. (\*Plotted marginal effect includes the significant linear slope, quadratic and quartic rates of change (cubic trajectory parameter not significant in the final model); Horizontal black dashed line shows where the slopes are not significantly different from zero).

doi:10.1371/journal.pone.0146081.g002

stratified analysis of the association between HDL-C levels and HDL-C genetic risk however, confirmed that the effect size of HDL-C wGRS on HDL-C decreased almost by half on-average for age-groups >30 both in males and females. For triglycerides, the results of the age-stratified regression analysis are consistent with the results of the individual growth curve analyses, showing very stable effect estimates for females throughout age groups, and a clear upward trend in effect size estimates for males. Sex-specific mean effect sizes for categorical wGRS at chosen ages (i.e. 3, 15 24 and 45/46 years) are consistent with the results of the sex- and stratified cross sectional analyses of the continuous risk scores (S4 Table).

### Evidence for polygenic gene-lifestyle interaction on adult lipoprotein levels

Change in BMI z-score between adulthood and childhood ( $\Delta$  BMI) was highly predictive of each adult lipoprotein levels independently of the wGRSs in multivariable models adjusted for sex and age at baseline and follow-up (S1 Table). However, the LR tests revealed no evidence



**Fig 3. Age- and sex- stratified estimated effects of LDL-C wGRS (upper panel), HDL-C wGRS (middle panel) and TG wGRS (lower panel) on LDL-C, HDL-C and triglycerides levels respectively with color coded significance levels and studentized bootstrapped non-parametric 95% CI.** For each age, the continuous error bars correspond to males and the dashed error bars directly next to them correspond to female models. Effect sizes are in mmol/L change per 1-sd change in wGRS for LDL-C and HDL-C and in odds ratio lipoprotein level change per 1-sd change in wGRS for triglycerides. Point sizes of the beta estimates reflect sample size (number of participants included in each age- and sex-specific regression analysis) Parameter estimates significance: Lightgrey,  $0.05 < p\text{-val} < 0.01$ ; Darkgrey,  $1. \times 10^{-3} < p\text{-val} < 1. \times 10^{-6}$ ; Black,  $p\text{-val} \leq 1. \times 10^{-6}$ . Black lines: smooth trend curves fitted by LOESS (Locally weighted non-parametric regression) to help visualise trends in the cross-sectional associations.

doi:10.1371/journal.pone.0146081.g003

for gene-lifestyle interactions ( $\Delta$  BMI \*wGRSs) on the adult lipoprotein concentrations, as the combined effect of candidate genes on adult lipoprotein levels does not appear to be modulated by the trajectory of BMI from childhood in this cohort ( $\chi^2$ (df = 1) = 0.24 p = 0.62;  $\chi^2$ (df = 1) = 2.04, p-value = 0.15 and  $\chi^2$ (df = 1) = 1.17 p-value = 0.11 for HDL-C, LDL-C, and triglycerides respectively).

## Discussion

This is the first study to investigate the combined effect of dyslipidaemia-predisposing variants on longitudinal blood lipoprotein profiles from childhood to adulthood. Rather than considering individual-validated risk lipid-SNPs alone, usually characterized by weak to moderate effect sizes in large lipid GWASs, our approach used weighted genetic risk scores combining multiple loci identified in meta-analyses of large lipid GWAS. Collectively, our findings suggest that genetic factors influence age-specific lipoprotein values and developmental trajectories already from the age of 3 years.

The use of polygenic risk scores has become increasingly popular in recent years for the purpose of genetic prediction of a number of quantitative traits, with the increasing recognition that a substantial part of heritability comes from many susceptibility markers individually characterized by low predictive power [26]. For circulating lipids in particular, when combined, multiple common genetic variants with small effects on circulating lipid levels, were reported to be highly predictive of individual trait measures [28] and showed association with subclinical and clinical cardiovascular outcomes [27]. Despite these promising associations between SNPs (or genetic risk scores) and lipoproteins, susceptibility alleles are often identified from cross-sectional adult GWASs, and it remains unclear whether the within-individual level of genetic risk carried by these variants is stable through life or changes with age. This uncertainty greatly impedes the assessment of the abilities of individual SNPs (or polygenic GWAS-derived risk score) for prediction of quantitative traits collected across the lifecourse and their clinical usefulness in the primary prevention of dyslipidemia.

Independent of participants age-related lipid trends, we found significant decreasing secular linear trends in LDL-C and triglycerides across the 31 follow-up years, consistent with what has been reported in previous studies [54]. These trends may be partially due to a wider use of statins and/ or improvement in diet in the last decades.

Multilevel individual growth curve analysis revealed an expected difference in lipoprotein profiles from childhood to adulthood between males and females, and that while baseline HDL-C, LDL-C and triglycerides levels were not significantly different between sexes (age 3), the actual trajectories of lipid levels from childhood to adulthood were largely sex-driven. We observed a time-averaged association between lipid wGRSs and each of the three serum lipid levels, implying that composite genetic risk scores were robust predictors of average lipoprotein levels, as well as predictors of lipoprotein in childhood (from age 3 at baseline). The time-averaged effect of the 14 risk alleles on LDL-C levels was stronger than the effect of the 38 risk SNPs on HDL-C levels on average, suggesting that the genetic predisposition to high serum LDL-C is stronger compared to a predisposition to low HDL-C levels and high triglycerides levels. The LDL-C wGRSs did not modify the linear, quadratic or cubic trajectory of lipoprotein level, so that the combined effect of the 14 candidate SNPs is not only present in young childhood, but also consistent across a person's life course as she ages (from age 3 to 49 years). Categorical analyses revealed that the 38 risk SNPs modified the linear and quadratic component of HDL-C trajectory over time, although the effect sizes of the age-dependent terms were small, such that it does not translate to clearly divergent profiles between high- and low-risk participants. However, when plotting the estimated marginal effect of HDL-C wGRS over time (within-individual), for a given male participant the magnitude of the effect of the combined genetic risk score on his HDL-C levels will attenuate with age. This is also confirmed by the cross sectional age-stratified analyses, where childhood HDL-C levels were strongly related to genetics, but where the main effect of HDL-C wGRS diminished by half in age-groups over 30 years old. Although significant at all ages, the collective effect of the HDL-C risk score on HDL-C levels becomes weaker as participants age, with environmental, lifestyle and

behavioural factors such as diet, smoking, physical activity, potentially becoming more important determinants of adult circulating HDL-C.

For triglycerides, we found that the combined detrimental effect of the 24 risk SNPs on triglyceride levels appears consistent over time for females from childhood through adulthood. For males however, both the cross sectional regression and IGG analyses suggest that the association between TGwGRS and fasting triglycerides levels increases linearly with age, becoming stronger on average than in females from around 25 years of age. These findings suggest that in adulthood, males at risk may respond less efficiently than females to lifestyle interventions targeting the reduction of fasting triglyceride levels

For lipids, it is largely unknown how a conventional age-varying 'lifestyle-related' risk factor such as adiposity modifies the genetic risk of developing abnormal lipid over the life course. This is a particularly relevant topic as identifying people whose risk is amplified by a combination of genetic and behavioral factors might facilitate interventions to prevent or delay the onset of cardiometabolic diseases.  $\Delta$  BMI<sub>i</sub>, a proxy retrospective indicator of 'adiposity trajectory', was computed for each participant to summarize both the directionality and magnitude of change in their weight status-related lifestyle factors over their life course relative to the average change in the cohort. In our analyses, the change in BMI z-score from childhood to adulthood ( $\Delta$  BMI<sub>i</sub>) predicted adult lipoprotein levels in 2011. This observation is consistent with the fact that adult BMI is important for many adult metabolic factors including lipids and that associations between lipids and BMI generally strengthened with age [55]. However, the effect of BMI z-score change since childhood on adult concentration is not modified by the wGRS for any of the lipoproteins, signifying that the detrimental consequences of an above average change in BMI since childhood on circulating lipoprotein levels does not appear to be exacerbated in adults that are genetically predisposed to high-risk lipoprotein profiles. However, this analysis is preliminary, and although it considers the overall direction of change if BMI z-score between childhood and adulthood, it does not fully account for the age-varying nature of BMI as a confounder and how it might interact with wGRS at different age- or developmental stages (childhood, adolescence, young adulthood). A study looking at the importance of the age at which obesity developed in associations between adult lipids and weight status, showed that although obese adults had adverse levels of lipoproteins, these levels did not vary with childhood weight status or with the age at the onset of obesity [56]. It remains unknown however if specific age of onset obesity modifies the effect of a genetic predisposition to adverse lipoprotein levels, and whether primary prevention measures could be improved by specifically aiming at resolving obesity before a certain 'critical' age-window in individuals genetically at-risk to develop dyslipidemia, as has been shown for apparently healthy individuals [57].

The present work has a number of strengths and limitations. We studied a large, randomly selected cohort of men and women followed up at 8 occasions over the course of 30 years since early childhood. The extensive longitudinal lipoprotein phenotypic and genotypic data offered a rare opportunity for a refined analysis of the association between genetic risk and serum lipoprotein trajectories. The hierarchical mixed effect Individual Growth Curve modelling approach allowed us to comprehensively model between-individual changes in within individual outcome trajectories. The multivariable models of adult lipoprotein allowed estimating the ability of the polygenic GWAS-derived lipoprotein risk scores to predict adult lipoprotein concentration over other conventional childhood risk factors. Because the cohort is of European descent, our results are only generalizable to individuals with a similar ancestry. Loss to follow-up from the original cohort more often occurs for males, therefore the sex-specific time-averaged and time-dependent effect of multi-loci risk scores might be slightly biased. We also suggest that to validate causal inference, our findings should be tested for potential confounding effects of other variables such as additional adiposity indicators, which are known to correlate



highly with lipoproteins and vary over time. However, additional adiposity phenotypes have not been collected at each study wave in the Cardiovascular Risk in Young Finns Study.

## Conclusions

This study demonstrates the significance of GWAS-derived genetic risk scores as predictors of lipoprotein levels at all ages. Additionally, we report for the first time an age-dependent effect of the 38 HDL-C risk SNPs on HDL-C and the 25 TG risk SNPs on TG levels among males, suggesting that the genetically-determined effects on these lipoproteins tends to change as a person ages. Further studies are now needed to characterise how this polygenic effect translates in terms of disease status prediction from childhood to adulthood.

## Supporting Information

**S1 Appendix. Step by step IGC model-building procedure, explicit parametrization of the IGC submodels and references on IGC modelling.**

(DOCX)

**S2 Appendix. Kernel density plots and quantile-quantile plots of LDL-C, HDL-C, and triglyceride concentrations (in mmol/L).**

(DOCX)

**S1 Fig. Flowchart of the IGC analyses used for modelling the lipid trajectories over the life course in the cohort.**

(JPG)

**S1 File. Original cohort and design of the Young Finns Study from 1980 through 2011 (Table A) and number of non-missing lipid measurements per study year in the subsample of the cohort participants considered in the present analyses (Table B).** <sup>a, b</sup> Limitations in the sampling size in 1989 are due to physical examinations and blood tests collected only in one centre (Turku). In 1992, the limitation was due to economic constraints. \* Maximum of participants in the current study was 2442 (i.e. participants successfully genotyped in 2001).

(DOCX)

**S2 File. Histograms of composite genetic risk scores (wGRSs) (Fig A) and of the number of available lipid observations per participants (Fig B).** For each lipid, green colour in histogram A denotes the lower GRS quartile (i.e. the 25<sup>th</sup> percentile); red colour, the upper GRS quartile (i.e. the 75<sup>th</sup> percentile); and white colour, the interquartile range. (i.e. the 50% of the data lying between upper and lower quartile).

(DOCX)

**S1 Table. Estimation of interaction between polygenic lipoprotein risk scores and change in BMI z-score between adulthood (ages > 30 years) and childhood (ages up to 18 years) on adult lipoprotein concentration prediction.** *Abbreviations:* HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol, wGRS-HDL; *HDLwGRS3* continuous combined genetic risk score for HDL (38 SNPs); wGRS-HDL, *LDLwGRS14*, continuous combined genetic risk score for LDL-C (14 SNPs); wGRS-TG, *TGwGRS24*, continuous combined genetic risk score for triglycerides (24 SNPs); adult.age, age of participants at latest adult follow up (2001, 2007 or 2011),  $\Delta$  BMI, change in BMI z-score between childhood and adulthood as described in Methods section. \* For HDL-C and LDL-C, the effect of wGRSs is in mmol/L per 1-sd change in wGRS, and the effect of  $\Delta$  BMI is in mmol/L per 1-sd change in  $\Delta$  BMI. For triglycerides, the effect of wGRSs is in odds ratio per 1-sd change in wGRS, and the effect of  $\Delta$

BMI is in odds ratios per 1-sd change in  $\Delta$  BMI.  
(PDF)

**S2 Table. List of lipid-associated SNPs used to generate the genetic risk scores for HDL-C and LDL-C in the YF population (reported are the SNPs reference number, risk allele for the SNP, GWAS-derived effect size in  $\text{md.dl}^{-1}$ ).** The gene name listed under 'locus' is either plausible biological candidate gene in the locus or the nearest annotate gene to the SNP. 'chr' denote chromosome. 'Freq (%)' denotes the risk allele frequency in the YF population, Adapted from Teslovitch et al. (2010)[23] and Tikkanen et al. (2011) [29].  
(DOCX)

**S3 Table. Childhood and adult cutoffs used to define 'normal' and 'high-risk' serum lipid levels (in mmol/L).** Childhood cutoffs apply to participants 19 years and younger. Chosen cutoffs are based on NCEP adult-panel treatment guidelines, NECP pediatric and adolescent guidelines and AAP and AHA pediatric guidelines (see [methods](#)).  
(DOCX)

**S4 Table. Sex-specific associations between categorical genetic risk score and lipid levels at selected ages (3 years, 15 years, 24 years and 45 or 46 years).** Effect sizes  $\beta$  are in mmol/L for the number of risk allele differences between high and low wGRSs. Reported at each age are: Int (SE), the average lipid level in mmol/L (SE) in the high risk score group;  $\beta$  (SE), the difference in average lipid level in mmol/L for subjects in the low risk score group, the p-value for each cross-sectional association and the number of non-missing serum lipid observations considered in each regression (Nobs). (\*: For each sex, the number of risk alleles difference between high and low wGRS groups are calculated as: median (Number of alleles in the High wGRS)–median ((Number of alleles in the Low wGRS)).  
(DOCX)

## Acknowledgments

We thank the clinic and administrative staff for their contribution to data collection. Above all, we thank the participants of the Cardiovascular Risk in Young Finns Study.

## Author Contributions

Analyzed the data: MJB. Wrote the paper: MJB CGM RJT. Designed the study: MJB CGM RJT OTR. Collected the clinical data and biological samples: JSAV TL NP MJ MK NHK OTR. Processed samples and produced genetic scores: NP. Reviewed and approved the final manuscript: MJB CGM RJT OTR JSAV TL NP MK NHK NJS OTR.

## References

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* 2006;3(11):e442. Epub 2006/11/30. doi: [10.1371/journal.pmed.0030442](https://doi.org/10.1371/journal.pmed.0030442) PMID: [17132052](https://pubmed.ncbi.nlm.nih.gov/17132052/); PubMed Central PMCID: PMC1664601.
2. Kathiresan S. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med Genet.* 2007;8(Suppl 1):S17. PMID: [17903299](https://pubmed.ncbi.nlm.nih.gov/17903299/)
3. Choy PC, Siow YL, Mymin D, O K. Lipids and atherosclerosis. *Biochem Cell Biol.* 2004;82(1):212–24. Epub 2004/03/31. doi: [10.1139/o03-085](https://doi.org/10.1139/o03-085) PMID: [15052339](https://pubmed.ncbi.nlm.nih.gov/15052339/).
4. Berenson GS, Srinivasan SR, Bao W, Newman WP 3rd, Tracy RE, Wattigney WA. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The Bogalusa Heart Study. *N Engl J Med.* 1998;338(23):1650–6. Epub 1998/06/06. doi: [10.1056/nejm199806043382302](https://doi.org/10.1056/nejm199806043382302) PMID: [9614255](https://pubmed.ncbi.nlm.nih.gov/9614255/).

5. Badimon L, Vilahur G. LDL-cholesterol versus HDL-cholesterol in the atherosclerotic plaque: inflammatory resolution versus thrombotic chaos. *Ann N Y Acad Sci.* 2012;1254:18–32. Epub 2012/05/03. doi: [10.1111/j.1749-6632.2012.06480.x](https://doi.org/10.1111/j.1749-6632.2012.06480.x) PMID: [22548566](https://pubmed.ncbi.nlm.nih.gov/22548566/).
6. Graham I, Cooney MT, Bradley D, Dudina A, Reiner Z. Dyslipidemias in the prevention of cardiovascular disease: risks and causality. *Curr Cardiol Rep.* 2012;14(6):709–20. Epub 2012/09/12. doi: [10.1007/s11886-012-0313-7](https://doi.org/10.1007/s11886-012-0313-7) PMID: [22965836](https://pubmed.ncbi.nlm.nih.gov/22965836/).
7. van Capelleveen JC, Bochem AE, Motazacker MM, Hovingh GK, Kastelein JJ. Genetics of HDL-C: a causal link to atherosclerosis? *Curr Atheroscler Rep.* 2013;15(6):326. Epub 2013/04/18. doi: [10.1007/s11883-013-0326-8](https://doi.org/10.1007/s11883-013-0326-8) PMID: [23591671](https://pubmed.ncbi.nlm.nih.gov/23591671/).
8. Navab M, Reddy ST, Van Lenten BJ, Fogelman AM. HDL and cardiovascular disease: atherogenic and atheroprotective mechanisms. *Nat Rev Cardiol.* 2011;8(4):222–32. Epub 2011/02/10. doi: [10.1038/nrcardio.2010.222](https://doi.org/10.1038/nrcardio.2010.222) PMID: [21304474](https://pubmed.ncbi.nlm.nih.gov/21304474/).
9. Panel NE. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Final report. *Circulation.* 2002;(106):3143–421. PMID: [12485966](https://pubmed.ncbi.nlm.nih.gov/12485966/)
10. Gordon DH, Rifkind BM. High-density lipoprotein: the clinical implications of recent studies. *N Engl J Med* 1989;(321):1311–6. PMID: [2677733](https://pubmed.ncbi.nlm.nih.gov/2677733/)
11. Toth PP. High-density lipoprotein and cardiovascular risk. *Circulation.* 2004;109(15):1809–12. Epub 2004/04/21. doi: [10.1161/01.cir.0000126889.97626.b8](https://doi.org/10.1161/01.cir.0000126889.97626.b8) PMID: [15096460](https://pubmed.ncbi.nlm.nih.gov/15096460/).
12. Assmann G, Gotto AM Jr. HDL cholesterol and protective factors in atherosclerosis. *Circulation.* 2004;109(23 Suppl 1):Ii8–14. Epub 2004/06/17. doi: [10.1161/01.cir.0000131512.50667.46](https://doi.org/10.1161/01.cir.0000131512.50667.46) PMID: [15198960](https://pubmed.ncbi.nlm.nih.gov/15198960/).
13. Nordestgaard BG, Benn M, Schnohr P, Tybjaerg-Hansen A. Nonfasting triglycerides and risk of myocardial infarction, ischemic heart disease, and death in men and women. *JAMA.* 2007;298(3):299–308. Epub 2007/07/20. doi: [10.1001/jama.298.3.299](https://doi.org/10.1001/jama.298.3.299) PMID: [17635890](https://pubmed.ncbi.nlm.nih.gov/17635890/).
14. Miller M, Stone NJ, Ballantyne C, Bittner V, Criqui MH, Ginsberg HN, et al. Triglycerides and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation.* 2011;123(20):2292–333. Epub 2011/04/20. doi: [10.1161/CIR.0b013e3182160726](https://doi.org/10.1161/CIR.0b013e3182160726) PMID: [21502576](https://pubmed.ncbi.nlm.nih.gov/21502576/).
15. Talayero BG, Sacks FM. The role of triglycerides in atherosclerosis. *Curr Cardiol Rep.* 2011;13(6):544–52. Epub 2011/10/05. doi: [10.1007/s11886-011-0220-3](https://doi.org/10.1007/s11886-011-0220-3) PMID: [21968696](https://pubmed.ncbi.nlm.nih.gov/21968696/); PubMed Central PMCID: [PMC3234107](https://pubmed.ncbi.nlm.nih.gov/PMC3234107/).
16. Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet.* 2013;45(11):1345–52. Epub 2013/10/08. doi: [10.1038/ng.2795](https://doi.org/10.1038/ng.2795) PMID: [24097064](https://pubmed.ncbi.nlm.nih.gov/24097064/); PubMed Central PMCID: [PMC3904346](https://pubmed.ncbi.nlm.nih.gov/PMC3904346/).
17. Lardizabal JA, Deedwania PC. Benefits of statin therapy and compliance in high risk cardiovascular patients. *Vasc Health Risk Manag.* 2010;6:843–53. Epub 2010/10/20. doi: [10.2147/vhrm.s9474](https://doi.org/10.2147/vhrm.s9474) PMID: [20957130](https://pubmed.ncbi.nlm.nih.gov/20957130/); PubMed Central PMCID: [PMC2952453](https://pubmed.ncbi.nlm.nih.gov/PMC2952453/).
18. Jameson L, De Groot LJ. Genetic Basis of Lipid Disorders. *Endocrinology: Adult and Pediatric* 6 ed: Elsevier; 2010.
19. Kathiresan S. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet.* 2008;40:189–97. doi: [10.1038/ng.75](https://doi.org/10.1038/ng.75) PMID: [18193044](https://pubmed.ncbi.nlm.nih.gov/18193044/)
20. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet.* 2008;40(2):161–9. Epub 2008/01/15. doi: [10.1038/ng.76](https://doi.org/10.1038/ng.76) PMID: [18193043](https://pubmed.ncbi.nlm.nih.gov/18193043/).
21. Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet.* 2009;41(1):35–46. Epub 2008/12/09. doi: [10.1038/ng.271](https://doi.org/10.1038/ng.271) PMID: [19060910](https://pubmed.ncbi.nlm.nih.gov/19060910/); PubMed Central PMCID: [PMC2687077](https://pubmed.ncbi.nlm.nih.gov/PMC2687077/).
22. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, et al. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet.* 2009;41(1):47–55. Epub 2008/12/09. doi: [10.1038/ng.269](https://doi.org/10.1038/ng.269) PMID: [19060911](https://pubmed.ncbi.nlm.nih.gov/19060911/); PubMed Central PMCID: [PMC2687074](https://pubmed.ncbi.nlm.nih.gov/PMC2687074/).
23. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 2010;466(7307):707–13. Epub 2010/08/06. doi: [10.1038/nature09270](https://doi.org/10.1038/nature09270) PMID: [20686565](https://pubmed.ncbi.nlm.nih.gov/20686565/); PubMed Central PMCID: [PMC3039276](https://pubmed.ncbi.nlm.nih.gov/PMC3039276/).
24. Global Lipids Genetics C. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45(11):1274–83. doi: [10.1038/ng.2797](https://doi.org/10.1038/ng.2797) <http://www.nature.com/ng/journal/v45/n11/abs/ng.2797.html-supplementary-information>. PMID: [24097068](https://pubmed.ncbi.nlm.nih.gov/24097068/)

25. Smith EN, Chen W, Kähönen M, Kettunen J, Lehtimäki T, Peltonen L, et al. Longitudinal Genome-Wide Association of Cardiovascular Disease Risk Factors in the Bogalusa Heart Study. *PLoS Genet.* 2010;6(9):e1001094. doi: [10.1371/journal.pgen.1001094](https://doi.org/10.1371/journal.pgen.1001094) PMID: [20838585](https://pubmed.ncbi.nlm.nih.gov/20838585/)
26. Chatterjee N, Wheeler B, Sampson J, Hartge P, Chanock SJ, Park JH. Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. *Nat Genet.* 2013;45(4):400–5, 5e1–3. Epub 2013/03/05. doi: [10.1038/ng.2579](https://doi.org/10.1038/ng.2579) PMID: [23455638](https://pubmed.ncbi.nlm.nih.gov/23455638/); PubMed Central PMCID: PMC3729116.
27. Isaacs A, Willems SM, Bos D, Dehghan A, Hofman A, Ikram MA, et al. Risk scores of common genetic variants for lipid levels influence atherosclerosis and incident coronary heart disease. *Arterioscler Thromb Vasc Biol.* 2013;33(9):2233–9. Epub 2013/06/15. doi: [10.1161/atvbaha.113.301236](https://doi.org/10.1161/atvbaha.113.301236) PMID: [23766260](https://pubmed.ncbi.nlm.nih.gov/23766260/).
28. Piccolo SR, Abo RP, Allen-Brady K, Camp NJ, Knight S, Anderson JL, et al. Evaluation of genetic risk scores for lipid levels using genome-wide markers in the Framingham Heart Study. *BMC Proc.* 2009;3 Suppl 7:S46. Epub 2009/12/19. PMID: [20018038](https://pubmed.ncbi.nlm.nih.gov/20018038/); PubMed Central PMCID: PMC2795945.
29. Tikkanen E, Tuovinen T, Widen E, Lehtimäki T, Viikari J, Kahonen M, et al. Association of known loci with lipid levels among children and prediction of dyslipidemia in adults. *Circ Cardiovasc Genet.* 2011;4(6):673–80. Epub 2011/10/11. doi: [10.1161/circgenetics.111.960369](https://doi.org/10.1161/circgenetics.111.960369) PMID: [21984478](https://pubmed.ncbi.nlm.nih.gov/21984478/).
30. Raitakari OT, Juonala M, Ronnema T, Keltikangas-Jarvinen L, Rasanen L, Pietikainen M, et al. Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol.* 2008;37(6):1220–6. Epub 2008/02/12. doi: [10.1093/ije/dym225](https://doi.org/10.1093/ije/dym225) PMID: [18263651](https://pubmed.ncbi.nlm.nih.gov/18263651/).
31. Kostner GM. Letter: Enzymatic determination of cholesterol in high-density lipoprotein fractions prepared by polyanion precipitation. *Clin Chem.* 1976;22(5):695. Epub 1976/05/01. PMID: [177229](https://pubmed.ncbi.nlm.nih.gov/177229/).
32. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the Concentration of Low-Density Lipoprotein Cholesterol in Plasma, Without Use of the Preparative Ultracentrifuge. *Clin Chem.* 1972;18(6):499–502. PMID: [4337382](https://pubmed.ncbi.nlm.nih.gov/4337382/)
33. Porkka KVK, Raitakari OT, Leino A, Laitinen S, Räsänen L, Rönnemaa T, et al. Trends in Serum Lipid Levels during 1980–1992 in Children and Young Adults: The Cardiovascular Risk in Young Finns Study. *Am J Epidemiol.* 1997;146(1):64–77. PMID: [9215224](https://pubmed.ncbi.nlm.nih.gov/9215224/)
34. Juonala M, Viikari JSA, Hutri-Kähönen N, Pietikäinen M, Jokinen E, Taittonen L, et al. The 21-year follow-up of the Cardiovascular Risk in Young Finns Study: risk factor levels, secular trends and east-west difference. *J Intern Med.* 2004;255(4):457–68. doi: [10.1111/j.1365-2796.2004.01308.x](https://doi.org/10.1111/j.1365-2796.2004.01308.x) PMID: [15049880](https://pubmed.ncbi.nlm.nih.gov/15049880/)
35. Raiko JR, Viikari JS, Ilmanen A, Hutri-Kahonen N, Taittonen L, Jokinen E, et al. Follow-ups of the Cardiovascular Risk in Young Finns Study in 2001 and 2007: levels and 6-year changes in risk factors. *J Intern Med.* 2010;267(4):370–84. Epub 2009/09/17. doi: [10.1111/j.1365-2796.2009.02148.x](https://doi.org/10.1111/j.1365-2796.2009.02148.x) PMID: [19754855](https://pubmed.ncbi.nlm.nih.gov/19754855/).
36. Adolescents NEPoBCLiCa. National Cholesterol Education Program (NCEP): Highlights of the Report of the Expert Panel on Blood Cholesterol Levels in Children and Adolescents. *Pediatrics.* 1992;89(3):495–501. PMID: [1741227](https://pubmed.ncbi.nlm.nih.gov/1741227/)
37. Program NCE. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Final Report. *Circulation.* 2002;106(25):3143.
38. Kavey RE, Daniels SR, Lauer RM, Atkins DL, Hayman LL, Taubert K. American Heart Association guidelines for primary prevention of atherosclerotic cardiovascular disease beginning in childhood. *Circulation.* 2003;107(11):1562–6. Epub 2003/03/26. PMID: [12654618](https://pubmed.ncbi.nlm.nih.gov/12654618/).
39. Li Y, Willer CJ, Ding J, Schee T P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol.* 2010;(34):816–34. doi: [10.1002/gepi.20533](https://doi.org/10.1002/gepi.20533) PMID: [21058334](https://pubmed.ncbi.nlm.nih.gov/21058334/)
40. Plomin R, Haworth CM, Davis OS. Common disorders are quantitative traits. *Nat Rev Genet.* 2009;10(12):872–8. Epub 2009/10/28. doi: [10.1038/nrg2670](https://doi.org/10.1038/nrg2670) PMID: [19859063](https://pubmed.ncbi.nlm.nih.gov/19859063/).
41. Singer JD, Willett JB. *Applied Longitudinal Data Analysis: Modeling change and event occurrence.* New York: Oxford University Press; 2003.
42. Mirman D, Dixon JA, Magnuson JS. Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *J Mem Lang.* 2008;59(4):475–94. Epub 2008/12/09. doi: [10.1016/j.jml.2007.11.006](https://doi.org/10.1016/j.jml.2007.11.006) PMID: [19060958](https://pubmed.ncbi.nlm.nih.gov/19060958/); PubMed Central PMCID: PMC2593828.
43. Mirman D. *Growth Curve Analysis and Visualization Using R.* Chapman and Hall / CRC.; 2014.
44. Curran PJ, Obeidat K, Losardo D. Twelve Frequently Asked Questions About Growth Curve Modeling. *Journal of cognition and development: official journal of the Cognitive Development Society.* 2010;11(2):121–36. doi: [10.1080/15248371003699969](https://doi.org/10.1080/15248371003699969) PMC3131138.

45. Wood S. Generalized additive models: an introduction with R: CRC press; 2006.
46. Box GEP, Cox DR. An analysis of transformations. *Journal of the Royal Statistical Society. 1964;Series B (26.):* 211–52.
47. Cheng J, Edwards LJ, Maldonado-Molina MM, Komro KA, Muller KE. Real longitudinal data analysis for real people: building a good enough mixed model. *Stat Med.* 2010;29(4):504–20. Epub 2009/12/17. doi: [10.1002/sim.3775](https://doi.org/10.1002/sim.3775) PMID: [20013937](https://pubmed.ncbi.nlm.nih.gov/20013937/); PubMed Central PMCID: PMC2811235.
48. Nakagawa S, Schielzeth H. A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods Ecol Evol.* 2013;4(2):133–42. doi: [10.1111/j.2041-210x.2012.00261.x](https://doi.org/10.1111/j.2041-210x.2012.00261.x)
49. Johnson PCD. Extension of Nakagawa & Schielzeth's R<sup>2</sup>GLMM to random slopes models. *Methods Ecol Evol.* 2014;5(9):944–6. doi: [10.1111/2041-210X.12225](https://doi.org/10.1111/2041-210X.12225) PMID: [25810896](https://pubmed.ncbi.nlm.nih.gov/25810896/)
50. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>. 2013.
51. Pinheiro J, Bates D, DebRoy S, Sarkar D. Linear and Nonlinear Mixed Effects Models. R package version 3.1–110. R Development Core Team (2013). 2013.
52. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society* 2011; 73(1):3–36.
53. Wineinger NE, Harper A, Libiger O, Srinivasan SR, Chen W, Schork NJ. Genomic risk models improve prediction of longitudinal lipid levels in children and young adults. *Frontiers in Genetics.* 2013;4. doi: [10.3389/fgene.2013.00086](https://doi.org/10.3389/fgene.2013.00086)
54. Bouillon K, Singh-Manoux A, Jokela M. Decline in low-density lipoprotein cholesterol concentration: lipid-lowering drugs, diet, or physical activity? Evidence from the Whitehall II study. *Heart.* 2011; (97):923–30. doi: [10.1136/hrt.2010.216309](https://doi.org/10.1136/hrt.2010.216309) PMID: [21487128](https://pubmed.ncbi.nlm.nih.gov/21487128/)
55. Pinto Pereira SM, Power C. Life course body mass index, birthweight and lipid levels in mid-adulthood: a nationwide birth cohort study 2013. 1215–24 p.
56. Freedman DS, Khan LK, Dietz WH, Srinivasan SR, Berenson GS. Relationship of childhood obesity to coronary heart disease risk factors in adulthood: the Bogalusa Heart Study. *Pediatrics.* 2001;108 (3):712–8. Epub 2001/09/05. PMID: [11533341](https://pubmed.ncbi.nlm.nih.gov/11533341/).
57. Juonala M, Magnussen CG, Berenson GS, Venn A, Burns TL, Sabin MA, et al. Childhood adiposity, adult adiposity, and cardiovascular risk factors. *N Engl J Med.* 2011;365(20):1876–85. Epub 2011/11/18. doi: [10.1056/NEJMoa1010112](https://doi.org/10.1056/NEJMoa1010112) PMID: [22087679](https://pubmed.ncbi.nlm.nih.gov/22087679/).