Prediction of Adult Dyslipidemia Using Genetic and Childhood Clinical Risk Factors

The Cardiovascular Risk in Young Finns Study

Joel Nuotio, MD*; Niina Pitkänen, PhD*; Costan G. Magnussen, PhD; Marie-Jeanne Buscot, MSc; Mikko S. Venäläinen, MSc; Laura L. Elo, PhD; Eero Jokinen, MD, PhD; Tomi Laitinen, MD, PhD; Leena Taittonen, MD, PhD; Nina Hutri-Kähönen, MD, PhD; Leo-Pekka Lytytikäinen, MD; Terho Lehtimäki, MD, PhD; Jorma S. Viikari, MD, PhD; Markus Juonala, MD, PhD; Olli T. Raitakari, MD, PhD

Background—Dyslipidemia is a major modifiable risk factor for cardiovascular disease. We examined whether the addition of novel single-nucleotide polymorphisms for blood lipid levels enhances the prediction of adult dyslipidemia in comparison to childhood lipid measures.

Methods and Results—Two thousand four hundred and twenty-two participants of the Cardiovascular Risk in Young Finns Study who had participated in 2 surveys held during childhood (in 1980 when aged 3–18 years and in 1986) and at least once in a follow-up study in adulthood (2001, 2007, and 2011) were included. We examined whether inclusion of a lipid-specific weighted genetic risk score based on 58 single-nucleotide polymorphisms for low-density lipoprotein cholesterol, 71 single-nucleotide polymorphisms for high-density lipoprotein cholesterol, and 40 single-nucleotide polymorphisms for triglycerides improved the prediction of adult dyslipidemia compared with clinical childhood risk factors. Adjusting for age, sex, body mass index, physical activity, and smoking in childhood, childhood lipid levels, and weighted genetic risk scores were associated with an increased risk of adult dyslipidemia for all lipids. Risk assessment based on 2 childhood lipid measures and the lipidspecific weighted genetic risk scores improved the accuracy of predicting adult dyslipidemia compared with the approach using only childhood lipid measures for low-density lipoprotein cholesterol (area under the receiver-operating characteristic curve 0.806 versus 0.811; P=0.01) and triglycerides (area under the receiver-operating characteristic curve 0.740 versus area under the receiver-operating characteristic curve 0.758; P<0.01). The overall net reclassification improvement and integrated discrimination improvement were significant for all outcomes.

Conclusions—The inclusion of weighted genetic risk scores to lipid-screening programs in childhood could modestly improve the identification of those at highest risk of dyslipidemia in adulthood. (Circ Cardiovasc Genet. 2017;10:e001604. DOI: 10.1161/CIRCGENETICS.116.001604.)

Key Words: cholesterol ■ dyslipidemias ■ genetics ■ lipids ■ risk factor

Blood lipids are one of the major modifiable risk factors for cardiovascular disease (CVD). Lipid levels have the tendency to track from childhood to adulthood, and early identification of children and adolescents at increased risk could allow targeted prevention strategies earlier in the disease process with the potential to prevent or delay dyslipidemia-related complications. Previous child to adult studies have shown that the strongest predictors of dyslipidemia in adulthood are the corresponding lipid concentrations and body mass index (BMI) in childhood. Current pediatric guidelines for primary prevention of CVD recommend obtaining fasting lipid profiles from children aged ≥2 years who have a history of premature CVD or dyslipidemia in their family or have any other risk factors (obesity, hypertension, or diabetes mellitus). Universal screening of lipid levels is suggested among all children aged 9 to 11 years and adolescents and young adults aged 17 to 21 years.

See Editorial by Belsky
See Clinical Perspective

In recent genome-wide association studies, several single-nucleotide polymorphisms (SNPs) have been identified that are associated with blood lipid levels. The most recent study, which used a multistage design in 188,557 subjects, identified 157 SNPs significantly associated with blood lipid levels, including 62 novel SNPs. We earlier showed, using 95 SNPs associated with lipid levels, that the lipid-specific genetic risk scores did not significantly enhance the prediction of adult dyslipidemia compared with clinical lipid measurements, except for triglycerides. However, in earlier analyses, the number of SNPs was

Received August 29, 2016; accepted April 25, 2017.
*Dr Nuotio and Pitkänen contributed equally to this work.
Correspondence to Joel Nuotio, MD, Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Kiinamyllynkatu 10, 20520 Turku, Finland. E-mail jovanu@utu.fi
© 2017 American Heart Association, Inc.

Circ Cardiovasc Genet is available at http://circgenetics.ahajournals.org
DOI: 10.1161/CIRCGENETICS.116.001604
low compared with the number of SNPs currently known to be associated with blood lipid levels, and important nonlipid risk factors such as BMI were not taken into account. Whether the prediction of adult dyslipidemia could be improved by inclusion of novel SNPs remains unknown.

Therefore, in this study, we aimed to examine whether adding information on novel lipid-associated SNPs improves early identification of children and adolescents at increased risk for dyslipidemia in adulthood using data from the prospective follow-up of participants since childhood in the Cardiovascular Risk in Young Finns Study.13

Methods

Study Participants

The study population included 2422 participants (54% women, mean age 10.6 years in 1980) of the Young Finns Study who had been seen in clinical examinations in childhood and adolescence in 1980 (age 3–18 years), in 1986, or both in 1980 and 1986, and at least once in adulthood follow-up in 2001 (age 24–39 years), 2007 (age 30–45 years), or 2011 (age 34–49 years). A majority of the participants (75% for low-density lipoprotein cholesterol [LDL-C] and 76% for high-density lipoprotein cholesterol [HDL-C] and triglycerides) had 2 childhood lipid measurements from the years 1980 and 1986, 24% of the participants had one measurement from the year 1980, and 1% of the participants had 1 lipid measurement from the year 1986. Details of the study design and methods have been provided previously.14 Written informed consent was obtained from all participants or their parents, and the study was approved by local ethics committees in agreement with the Declaration of Helsinki.

Study Variables

In childhood, venous blood samples were drawn after a 12-hour fast. Serum samples were stored at −20°C until thawed for the first time for analyses in 1980 and 1986. Total cholesterol concentrations were measured using a fully enzymatic CHOD-PAP method (Boehringer Mannheim, Germany) with OLLI 3000 and Kone CD analyzers (Kone, Co, Espoo, Finland). Serum HDL-C concentrations were measured from the supernatant after precipitation of very-low-density lipoprotein cholesterol and LDL-C with dextran sulfate and MgCl₂ (Pharmacia, Uppsala, Sweden). Serum triglycerides concentrations were determined by using a fully enzymatic method (Boehringer Mannheim). The concentration of LDL-C was estimated by using the Friedewald formula in participants with triglycerides levels <4.0 mmol/L.15

In adulthood, venous blood samples were drawn after an overnight fast, and serum was separated, aliquoted, and stored at −70°C until analysis. Triglycerides concentration was determined using the enzymatic glycerol kinase–glycerol phosphate oxidase method (Triglyceride Reagent; Beckman Coulter Biomedical). The same reagent was used for estimating cholesterol levels by the enzymatic cholesterol esterase–cholesterol oxidase method (Cholesterol Reagent; Beckman Coulter Biomedical, Ireland). Total cholesterol levels were measured by the enzymatic cholesterol esterase–cholesterol oxidase method (Cholesterol Reagent; Beckman Coulter Biomedical). The same reagent was used for estimating HDL-C levels after precipitation of LDL-C with dextran sulfate and MgCl₂.16 All the above assays were performed on an AU400 instrument (Olympus, Japan), and the same methods were used both in 2007 and 2011. Because of changes in reagents or methods in 2001 to 2007, triglycerides values were corrected by using correction factor equations.16 The analysis methods for total cholesterol and triglycerides have been accredited by the Finnish Accreditation Service according to standard ISO/IEC17025. Use of lipid-lowering medication in adulthood was determined from self-administrated questionnaires. Height and weight were measured, and BMI was calculated as weight (kg)/height (m)². The physical activity questionnaire consisted of the following variables: intensity of physical activity, frequency of moderate or vigorous activity, and hours spent on moderate or vigorous activity per week and was calculated as described earlier.17,18 Smoking was assessed by a questionnaire in subjects aged ≥12 years. In the baseline 1980 study when the participants were aged 3, 6, 9, 12, 15, and 18 years, questionnaires on smoking were collected from the older half of the participants, that is, those aged 12, 15, and 18 years. In year 1983, questionnaires on smoking were not collected from the 2 youngest age groups (aged from 6–9 years), and in 1986, questionnaires were not collected from the youngest age group of participants aged 9 years at the time. Nevertheless, the youngest age group who were 3 years old at baseline answered the questionnaire in the year 2001 follow-up when they were aged 24 years. Smoking was defined as positive if participants had smoked daily at some stage before or at age 24 years.

Genetic Analyses

For this study, we used 58 SNPs associated with LDL-C levels, 71 SNPs associated with HDL-C, and 40 SNPs associated with triglycerides levels identified in a large and recent genome-wide association studies.11 Weighted genetic risk scores were calculated for each of the 3 lipids as sums of genotyped risk alleles or imputed allele dosages carried by an individual each multiplied by the published effect estimates.11 Each wGRS was Z scored for analyses. Genotyping of the SNPs included in the lipid wGRSs was performed with the Illumina Human 670K Bead Chip, and imputation was performed with the 1000 Genomes Project. For LDL-C, APOE allele combinations of SNPs rs7412 and rs429358 (coded as 1=ε2/ε2, 2=ε2/ε3, 3=ε2/ε4, 4=ε3/ε3, 5=ε3/ε4, and 6=ε4/ε4), which have been shown to have a linear relationship with LDL-C levels, were added to the model including the genetic risk score.19 APOE genotyping was performed by using Taqman SNP Genotyping Assays (rs429358 assay C 3084793_20; rs7412 assay C 904973_10) and the ABI Prism 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA). No discrepancies emerged in the genotyping results of duplicate samples.20

Definition of High-Risk Groups in Adulthood

European cut points were used in adulthood to denote abnormal serum lipid values: LDL-C >3.0 mmol/L (≈115 mg/dL), HDL-C in men <1.0 mmol/L (≈40 mg/dL) and in women <1.2 mmol/L (≈45 mg/dL), and triglycerides >1.7 mmol/L (≈150 mg/dL).21 Lipid measurements in adulthood were from the 2011 follow-up (≈71% of the adulthood measurements for LDL-C and ≈72% for HDL-C and triglycerides), except in case of missing data from 2011, measurements from 2007 (≈17% for LDL-C and 16% for HDL-C and triglycerides) or 2001 (≈12% for all lipids) were used. In addition, participants who reported use of lipid-lowering medication in any of the adult follow-up studies were classified as having high-risk LDL-C. The use of lipid-lowering medication was only taken into account in the LDL-C analysis as the LDL-C is the main target of statin therapy.

Statistical Analysis

The mean age- and sex-specific Z scores of childhood lipid levels and BMI measured in 1980 and 1986 were used to represent childhood risk factor levels. If either measurement was missing, a Z score of a single measurement was used. For physical activity during childhood, an age- and sex-specific Z score was calculated. The normality of the wGRS distributions was tested using the Shapiro–Wilk test of normality separately in individuals with dyslipidemia and normal lipid levels. Association of lipid-specific wGRS with the risk of adult dyslipidemia was analyzed using logistic regression. Multivariable logistic regression models, with and without lipid-specific wGRS, were constructed for all outcomes separately including the following childhood risk factors: age, sex, BMI, smoking status, and physical activity index. The association of lipid-specific wGRSs with childhood lipid levels was evaluated using linear regression analysis adjusted for age and sex.

The additional value of wGRS in prediction of adult dyslipidemia was examined using the R packages PredictABEL,22 Hmisc, and pROC23 to estimate fit, calibration, and the differences in predictive abilities of the models. The discrimination performance of each model was estimated by calculating the area under the receiver-operating characteristic curve (AUC).24 The differences between the AUCs were tested using the method used by DeLong et al.25 The J
statistic of Youden" was used to determine the optimal cutoff value for sensitivity and specificity in each model. The improvement of prediction models was assessed using the continuous net reclassification improvement (NRI) and integrated discrimination improvement, and model calibration was tested using the Hosmer–Lemeshow goodness-of-fit test. Additional predictive models were built using the gradient boosting algorithm implemented in the R package gbm. In these analyses, the Bernoulli loss function was used, and the study cohort was first divided to a training data set (72% of the study population) and to a validation data set (28% of the study population). Gradient boosting is an ensemble learning technique that builds a final strong model over many middle-step weak models, and its superior performance has been demonstrated by earlier studies. To avoid model overfitting, cross-validation was used. The discrimination performance of the model was evaluated by the AUC in both the training and the validation data. Statistical significance was inferred at a 2-tailed P value <0.05. The statistical analysis was performed with R version 3.2.2.

Results
Baseline characteristics of the study participants are shown in Table 1. Lipid-lowering medication was used by 89 participants (3.7% of the study population) in adulthood. In adults, an abnormal lipid profile for LDL-C was observed in 1390 participants (57.5%), abnormal HDL-C in 585 participants (24.1%), and abnormal triglycerides in 502 participants (20.7%).

The wGRS for HDL-C ranged from −3.42 to 3.13 (mean −0.36) in individuals with low HDL-C and −3.01 to 3.10 (mean 0.12) in individuals with high HDL-C and was normally distributed in both groups. The wGRS for LDL-C ranged from −3.02 to 3.50 (mean 0.19) in individuals with high LDL-C and −3.17 to 4.20 (mean −0.27) in individuals with low LDL-C and was normally distributed in both groups with high LDL-C, but not in individuals with low LDL-C (P=0.015). The wGRS for triglycerides ranged from −2.34 to 3.86 (mean −0.38) in individuals with high triglycerides and −3.13 to 3.34 (mean −0.10) in individuals with low triglycerides. The distribution of the wGRS for triglycerides deviated from normal in both groups (P<0.05).

To examine the prevalence of elevated LDL-C levels in adulthood according to childhood LDL-C levels and wGRS, participants were divided into 4 groups using the 80th percentile of childhood LDL-C levels and median value of wGRS for LDL-C as cutoff (Figure). In adulthood, LDL-C values from year 2011 were primarily used, except in case of missing data from 2011, values from 2007 or from 2001 were used. Of those who had low wGRS and normal childhood LDL-C levels, 45% had adult LDL-C >3 mmol/L or

| Table 1. Baseline (1980) Characteristics of 2422 Participants in the Cardiovascular Risk in Young Finns Study |
|-----------------|-----------------|-----------------|
| Age, y          | Women (n=1316)  | Men (n=1108)    |
| LDL cholesterol, mmol/L | 3.49 (0.8) n=1306 | 3.36 (0.8) n=1101 |
| HDL cholesterol, mmol/L | 1.56 (0.3) n=1306 | 1.56 (0.3) n=1101 |
| Triglycerides, mmol/L | 0.69 (0.3) n=1308 | 0.64 (0.3) n=1101 |
| BMI, kg/m²      | 17.8 (3.0) n=1311 | 18.0 (3.2) n=1099 |

Values are presented as mean (SD) except for smoking prevalence as proportion. Smoking prevalence indicates the portion of individuals who have reported being daily smokers at any stage between ages 12 and 24 y. Physical activity was assessed by a self-report questionnaire. Subjects answered the questions themselves, with their parents’ assistance as necessary. The physical activity questionnaire consisted of the following variables: intensity of physical activity, frequency of moderate or vigorous activity, and hours spent on moderate or vigorous activity per week. BMI indicates body mass index; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

Figure. Prevalence of elevated low-density lipoprotein (LDL) cholesterol levels in adulthood according to childhood LDL cholesterol levels and weighted genetic risk score (wGRS). LDL cholesterol >3 mmol/L or reported use of lipid-lowering medication in adulthood. LDL cholesterol level <80th percentile point in childhood. wGRS under median value. wGRS over median value. LDL-cholesterol ≥80th percentile point in childhood.
had reported use of lipid-lowering medication in any of the adult follow-up studies compared with 92% in those who had both high wGRS and high childhood LDL-C levels. In separate analyses, the lipid-specific wGRSs were significantly associated with the corresponding childhood lipid measurements. The age- and sex-adjusted regression coefficients were \( \beta = 0.29 \) for LDL-C, \( \beta = 0.31 \) for HDL-C, and \( \beta = 0.21 \) for triglycerides (\( P < 0.001 \)).

### Association of Childhood Lipid Levels and wGRSs With Adult Dyslipidemias

Childhood lipid levels and lipid-specific wGRSs were significantly associated with abnormal lipid levels in adulthood (\( P < 0.001 \); Table 2). The odds ratios (95% confidence interval) for childhood lipid levels were 3.82 (3.27–4.45) for LDL-C, 0.28 (0.24–0.33) for HDL-C, and 1.85 (1.63–2.10) for triglycerides. The odds ratios (95% confidence interval) for lipid-specific wGRSs were 1.25 (1.12–1.39) for LDL-C, 0.80 (0.72–0.90) for HDL-C, and 1.50 (1.33–1.68) for triglycerides. In addition, \( APOE \) genotypes were independently associated (odds ratio 1.20 [1.06–1.36]; \( P < 0.001 \)) with elevated LDL-C levels in adulthood.

To determine whether there was linear interaction between wGRS and age, an age*wGRS interaction term was included in the multivariable models for each lipid, but statistically significant interactions were not observed (\( P \) always >0.05), and the interaction term was omitted from the final model. The odds ratios (95% confidence interval) for the best childhood prediction model (age, sex, childhood lipid level, and wGRS) in predicting abnormal levels in adulthood for different age groups (3, 6, 9, 12, 15, and 18) are shown in Figure I in the Data Supplement.

### Model Discrimination and Reclassification

Specificity, sensitivity, and AUCs for each of the models are shown in Table 3. Adding wGRS to the childhood risk factor model significantly improved the AUC for LDL-C (\( P = 0.01 \)) and triglycerides (\( P < 0.001 \)). For LDL-C, the AUC increased from 0.806 to 0.811, and for triglycerides from 0.740 to 0.758. Improvement for HDL-C model (from 0.771–0.775) did not reach statistical significance (\( P = 0.09 \)). In comparison with the model without the wGRS, the number of false-positive adult dyslipidemia cases was increased from 194 to 214, and the number of false negatives reduced from 466 to 433 when the LDL-C wGRS was added into the model, and the threshold corresponding to the best sum of sensitivity and specificity was used.

When the gradient boosting approach was used as an additional method to study the predictive value of SNPs, we similarly found that inclusion of wGRSs improved significantly the prediction of adult dyslipidemia for LDL-C in the training data (AUC=0.811 versus AUC=0.817; \( P = 0.01 \)), but the small improvement observed in the validation data was not significant (AUC=0.815 versus AUC=0.817; \( P = 0.49 \)). For HDL-C, significant improvements were observed in the training data (AUC=0.787 versus AUC=0.793; \( P = 0.01 \)) and in the validation data (AUC=0.750 versus AUC=0.760; \( P = 0.02 \)). Similarly, for triglycerides, significant improvements for the prediction of adult dyslipidemia were observed in both the training data (AUC=0.776 versus AUC=0.799; \( P = 0.001 \)) and the validation data (AUC=0.702 versus AUC=0.726; \( P = 0.03 \)). Furthermore, the model including wGRSs for both triglycerides and LDL-C significantly enhanced the prediction of adult type IIa dyslipidemia (combined outcome including both triglycerides >1.7 mmol/L and LDL-C >3 mmol/L) compared with clinical risk factors (model including both triglycerides and LDL-C levels as child) in the training data (AUC=0.798 versus AUC=0.813; \( P = 0.002 \)) and in the validation data (AUC=0.749 versus AUC=0.774; \( P = 0.01 \)). The results remained essentially similar when use of lipid-lowering medication was used as a criterion of adult dyslipidemia for triglycerides and HDL-C, but for LDL-C, the improvement in the training data became nonsignificant (\( P = 0.27 \)) when lipid-lowering medication was not used as a criterion of adult dyslipidemia.

The net percentage of individuals with abnormal lipid levels correctly classified upward (event NRI) was 4.4% for LDL-C, 15% for HDL-C, and 9.8% for triglycerides. Furthermore, the net percentage of individuals without abnormal lipid levels correctly classified downward (non-event NRI) was 17.2% for LDL-C, 6.5% for HDL-C, and 17.6% for triglycerides. These changes resulted in the overall statistically significant improvement in continuous NRI of 0.22 (\( P < 0.001 \)) for LDL-C, 0.22 (\( P < 0.001 \)) for HDL-C, and 0.29 (\( P < 0.001 \)) for triglycerides. Furthermore, the integrated discrimination improvement was 0.011 (\( P < 0.001 \)) for LDL-C, 0.007 (\( P < 0.001 \)) for HDL-C, and 0.022 (\( P < 0.001 \)) for triglycerides, indicating that the difference in average predicted risks between the individuals with and without the outcome increased significantly when the wGRS was included in the prediction model (Table 4).

---

**Table 2. Association of Childhood Lipid Levels* and Lipid-Specific wGRS With Abnormal Lipid Levels in Adulthood**

<table>
<thead>
<tr>
<th>Adult Outcome</th>
<th>Predictor</th>
<th>n</th>
<th>OR</th>
<th>95% CI</th>
<th>( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal LDL cholesterol</td>
<td>Childhood LDL cholesterol levels</td>
<td>2348</td>
<td>3.82</td>
<td>3.27–4.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>wGRS for LDL cholesterol</td>
<td>2348</td>
<td>1.25</td>
<td>1.12–1.39</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Abnormal HDL cholesterol</td>
<td>Childhood HDL cholesterol levels</td>
<td>2360</td>
<td>0.28</td>
<td>0.24–0.33</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>wGRS for HDL cholesterol</td>
<td>2360</td>
<td>0.80</td>
<td>0.72–0.90</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Abnormal triglycerides</td>
<td>Childhood triglyceride levels</td>
<td>2360</td>
<td>1.85</td>
<td>1.63–2.10</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>wGRS for triglycerides</td>
<td>2360</td>
<td>1.50</td>
<td>1.33–1.68</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Multivariable logistic regression models adjusted for age, sex, childhood BMI Z score, childhood physical activity Z score, and smoking status were used. LDL-C model additionally adjusted for \( APOE \) genotypes. LDL cholesterol >3 mmol/L or reported use of lipid-lowering medication in adulthood was defined as abnormal. For HDL cholesterol, adult values <1.0 mmol/L were defined as abnormal. For triglycerides, adult values >1.7 mmol/L were defined as abnormal. BMI indicates body mass index; CI, confidence interval; HDL, high-density lipoprotein; LDL, low-density lipoprotein; OR, odds ratio per unit increase in wGRS; and wGRS, weighted genetic risk score.

*Mean of age- and sex-standardized measures from 1980 and 1986 was used.

---

**Model Discrimination and Reclassification**

Specificity, sensitivity, and AUCs for each of the models are shown in Table 3. Adding wGRS to the childhood risk factor model significantly improved the AUC for LDL-C (\( P = 0.01 \)) and triglycerides (\( P < 0.001 \)). For LDL-C, the AUC increased from 0.806 to 0.811, and for triglycerides from 0.740 to 0.758. Improvement for HDL-C model (from 0.771–0.775) did not reach statistical significance (\( P = 0.09 \)). In comparison with the model without the wGRS, the number of false-positive adult dyslipidemia cases was increased from 194 to 214, and the number of false negatives reduced from 466 to 433 when the LDL-C wGRS was added into the model, and the threshold corresponding to the best sum of sensitivity and specificity was used.

When the gradient boosting approach was used as an additional method to study the predictive value of SNPs, we similarly found that inclusion of wGRSs improved significantly the prediction of adult dyslipidemia for LDL-C in the training data (AUC=0.811 versus AUC=0.817; \( P = 0.01 \)), but the small improvement observed in the validation data was not significant (AUC=0.815 versus AUC=0.817; \( P = 0.49 \)). For HDL-C, significant improvements were observed in the training data (AUC=0.787 versus AUC=0.793; \( P = 0.01 \)) and in the validation data (AUC=0.750 versus AUC=0.760; \( P = 0.02 \)). Similarly, for triglycerides, significant improvements for the prediction of adult dyslipidemia were observed in both the training data (AUC=0.776 versus AUC=0.799; \( P = 0.001 \)) and the validation data (AUC=0.702 versus AUC=0.726; \( P = 0.03 \)). Furthermore, the model including wGRSs for both triglycerides and LDL-C significantly enhanced the prediction of adult type IIb dyslipidemia (combined outcome including both triglycerides >1.7 mmol/L and LDL-C >3 mmol/L) compared with clinical risk factors (model including both triglycerides and LDL-C levels as child) in the training data (AUC=0.798 versus AUC=0.813; \( P = 0.002 \)) and in the validation data (AUC=0.749 versus AUC=0.774; \( P = 0.01 \)). The results remained essentially similar when use of lipid-lowering medication was used as a criterion of adult dyslipidemia for triglycerides and HDL-C, but for LDL-C, the improvement in the training data became nonsignificant (\( P = 0.27 \)) when lipid-lowering medication was not used as a criterion of adult dyslipidemia.
In the present study, we observed that the lipid-specific genetic risk scores improve significantly the risk discrimination of adult dyslipidemia compared with clinical childhood lipid measures. In addition, we recently demonstrated that genome-wide association study–derived genetic risk scores predict lipoprotein levels at all ages, 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, with no major differences in the genetic effect observed between child and adolescent age groups from 3 to 18 years.

Similarly, we showed for other cardiovascular risk factors, such as adult hypertension and type 2 diabetes mellitus, that genetic information provides incremental predictive information in addition to nongenetic childhood risk factors. In the future, novel genetic risk factors may help to identify children with high risk for adult dyslipidemia early in life. Our results are in line with this hypothesis by demonstrating that genetic information provides incremental information compared with clinical risk factors in identifying children and adolescents who are at risk of developing dyslipidemia in adulthood. As atherosclerosis has its roots in childhood, early identification and control of dyslipidemia throughout the life span would reduce the risk for later clinical CVD. Results from the Special Turku Coronary Risk Factor Intervention Project indicate that lower lipid levels in childhood and adolescence can be achieved by repeated dietary counseling without pharmacological interventions.

## Discussion

Whether recently discovered novel genetic risk scores can improve the accuracy of pediatric CVD risk prediction in childhood and adolescence is a clinically relevant question that can only be addressed using data from longitudinal studies extending throughout the life course from childhood to adulthood. In this study, we observed that the prediction of adult dyslipidemia can be improved significantly by incorporating lipid-specific genetic risk scores compared with prediction models consisting of only the corresponding lipid level in childhood, age, sex, and childhood BMI, smoking status, and physical activity during childhood. We used the mean of 2 childhood lipid measures as we have earlier showed that information obtained with 2 measures is sufficient, and current pediatric guidelines also recommend universal lipid screening with 2 lipid measures. Genetic risk factors differ from traditional childhood risk factors as their measurement is unequivocal, they remain stable throughout life, and are unconfounded by other lifestyle-related risk factors. In an earlier report, we have showed using 95 loci associated with lipid levels that the lipid-specific genetic risk scores did not significantly enhance the prediction of adult dyslipidemia compared with clinical lipid measurements, except for triglycerides. In the present study, using genetic risk scores based on 157 SNPs including 62 novel SNPs, we observed that the lipid-specific genetic risk scores improve significantly the risk discrimination of adult dyslipidemia compared with clinical childhood lipid measures. In addition, we recently demonstrated that genome-wide association study–derived genetic risk scores predict lipoprotein levels at all ages, 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, with no major differences in the genetic effect observed between child and adolescent age groups from 3 to 18 years.

Similarly, we showed for other cardiovascular risk factors, such as adult hypertension and type 2 diabetes mellitus, that genetic information provides incremental predictive information in addition to nongenetic childhood risk factors. In the future, novel genetic risk factors may help to identify children with high risk for adult dyslipidemia early in life. Our results are in line with this hypothesis by demonstrating that genetic information provides incremental information compared with clinical risk factors in identifying children and adolescents who are at risk of developing dyslipidemia in adulthood. As atherosclerosis has its roots in childhood, early identification and control of dyslipidemia throughout the life span would reduce the risk for later clinical CVD. Results from the Special Turku Coronary Risk Factor Intervention Project indicate that lower lipid levels in childhood and adolescence can be achieved by repeated dietary counseling without pharmacological interventions.

## Table 3. Discriminating Properties of the Pediatric Multivariable Prediction Models for Adult Dyslipidemia

<table>
<thead>
<tr>
<th>Adult Outcome</th>
<th>Without wGRS</th>
<th>With wGRS</th>
<th>P Value for Difference in AUCs*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal LDL cholesterol</td>
<td>0.796</td>
<td>0.667</td>
<td>0.806 (0.788–0.823)</td>
</tr>
<tr>
<td>Abnormal HDL cholesterol</td>
<td>0.702</td>
<td>0.718</td>
<td>0.773 (0.751–0.794)</td>
</tr>
<tr>
<td>Abnormal triglycerides</td>
<td>0.746</td>
<td>0.626</td>
<td>0.740 (0.715–0.764)</td>
</tr>
</tbody>
</table>

AUC indicates area under the receiver-operating curve; BMI, body mass index; HDL, high-density lipoprotein; LDL, low-density lipoprotein; and wGRS, weighted genetic risk score. Adjusted for age, sex, childhood BMI Z score, childhood physical activity Z score, and smoking status. LDL-C model additionally adjusted for APOE genotype. LDL cholesterol >3 mmol/L or reported use of lipid-lowering medication in adulthood was defined as abnormal. For HDL cholesterol, adult values <1.2 mmol/L (in women)/1.0 mmol/L (in men) were defined as abnormal. For triglycerides, adult values >1.7 mmol/L were defined as abnormal.

*Model with wGRS vs. model without wGRS.

The Hosmer–Lemeshow goodness-of-fit test P values were nonsignificant for all models, indicating that there was no evidence of poor fit for models (Table 4).

## Table 4. Improvement of Reclassification Properties of the Pediatric Adult Dyslipidemia Prediction Models Including wGRS Compared With Models Without wGRS

<table>
<thead>
<tr>
<th>Adult Outcome</th>
<th>Nonevent NRI (95% CI)</th>
<th>Event NRI (95% CI)</th>
<th>Overall NRI (95% CI)</th>
<th>IDI (95% CI)</th>
<th>H-L χ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal LDL cholesterol</td>
<td>0.17 (0.11–0.23)</td>
<td>0.044 (–0.009 to 0.096)</td>
<td>0.22 (0.13 to 0.30)</td>
<td>0.011 (0.007 to 0.015)</td>
<td>4.01</td>
</tr>
<tr>
<td></td>
<td>P&lt;0.001</td>
<td>P=0.1</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td>P=0.81</td>
</tr>
<tr>
<td>Abnormal HDL cholesterol</td>
<td>0.065 (0.019 to 0.11)</td>
<td>0.15 (0.07 to 0.23)</td>
<td>0.22 (0.12 to 0.31)</td>
<td>0.007 (0.003 to 0.01)</td>
<td>9.76</td>
</tr>
<tr>
<td></td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td>P=0.28</td>
</tr>
<tr>
<td>Abnormal triglycerides</td>
<td>0.176 (0.13 to 0.22)</td>
<td>0.098 (0.010 to 0.187)</td>
<td>0.29 (0.18 to 0.37)</td>
<td>0.022 (0.015 to 0.029)</td>
<td>3.36</td>
</tr>
<tr>
<td></td>
<td>P&lt;0.001</td>
<td>P=0.03</td>
<td>P&lt;0.001</td>
<td>P&lt;0.001</td>
<td>P=0.91</td>
</tr>
</tbody>
</table>

LDL cholesterol >3 mmol/L or reported use of lipid-lowering medication in adulthood was defined as abnormal. For HDL cholesterol, adult values <1.2 mmol/L (in women)/1.0 mmol/L (in men) were defined as abnormal. For triglycerides, adult values >1.7 mmol/L were defined as abnormal. Adjusted for age, sex, childhood BMI Z score, childhood physical activity Z score, and smoking status. LDL-C model additionally adjusted for APOE genotype. CI indicates confidence interval; HDL, high-density lipoprotein; H-L, Hosmer–Lemeshow goodness-of-fit test; IDI, integrated discrimination index; LDL, low-density lipoprotein; NRI, net reclassification index; and wGRS, weighted genetic risk score.
We used AUC, NRI, and integrated discrimination improvement to compare the risk assessment between of our risk prediction models. The AUC describes the overall performance of the model in discriminating individuals with and without the outcome, but it is relatively insensitive to change if risk factors with strong associations with the outcome are already included in the initial model. In our study, the increases in AUCs were statistically significant for LDL-C and triglycerides. For LDL-C, the genetic model identified 33 less false-negative cases compared with the model without genetic data. At the same time, the number of false-positive cases was increased by 20, when the genetic model was compared with the model consisting only of conventional risk factors. Nevertheless, the decrease in the number of false-negative cases was greater than the increase in the amount of false-positive cases.

In additional analyses, we used the gradient boosting technique including internal validation in a subset of data that was not used when building the model. For triglycerides and HDL-C, statistically significant improvement in the AUC values was observed both in training and validation data when a prediction model of adult dyslipidemia that included wGRS was compared with a model consisting only of childhood clinical risk factors. Similar improvements were also observed in the prediction of adult type IIb dyslipidemia. For LDL-C, the improvement did not reach statistical significance in the validation data as it did in the training data. These results are consistent with the other analyses presented in this study, even though prediction of elevated adult LDL-C did not significantly improve in the validation data and the improvement in the prediction of low adult HDL-C reached significance both in the training data and in the validation data. Thus, in general, these results using different statistical methods are consistent with the other analyses presented in this study.

For LDL-C and triglycerides, the observed improvement in the overall NRI was mostly driven by the nonevent NRI indicating that the wGRS correctly decreased the risk estimates for nonevents. In contrast, for HDL-C, the changes were more dominant in event NRI. The integrated discrimination improvement was statistically significant for all models indicating that the difference in average predicted risks between the individuals with and without the outcome increased significantly when the wGRS was included in the models.

Our results also suggest that having low genetic risk for dyslipidemia and normal lipid phenotype in childhood does not strongly exclude adult dyslipidemia, as 45% of individuals with low genetic risk and normal LDL-C levels in childhood had elevated LDL-C in adulthood. At the same time, individuals who had elevated LDL-C levels in childhood and low genetic risk score had a higher risk for adult dyslipidemia (83%) than individuals with high genetic risk but low LDL-C levels (62%). Nevertheless, participants with both elevated LDL-C levels in childhood and high genetic risk had clearly highest risk for adult dyslipidemia, as 92% of these individuals had elevated LDL-C levels in adulthood.

The most recent genome-wide association studies performed by the Global Lipids Genetics Consortium identified 157 SNPs associated with lipid levels, including 62 novel SNPs that were not included in our previous analysis, which showed that the lipid-specific genetic risk scores based on 95 lipid trait associated SNPs only improved the prediction of adult dyslipidemia for triglycerides. By using approaches including literature review, pathway analysis, identification of coding variants, expression quantitative trait loci and overlapping genomic regulatory elements, the authors identified 70 potential candidate genes for the 62 identified novel association signals that provide new insights into lipid biology. In general, many of the lipid-associated loci contain genes that are of biological and clinical importance and are frequently associated with cardiovascular and metabolic traits, such as coronary artery disease, type 2 diabetes mellitus, blood pressure, waist–hip ratio, BMI, and known Mendelian lipid disorders. Some overlap exists in loci associated with different lipid traits. Of the newly identified HDL-C–associated SNPs, 3 were also associated with triglycerides and with LDL-C. One of the LDL-C SNPs was associated also with triglycerides and another one with HDL-C. Of the SNPs primarily associated with triglycerides, one was associated with LDL-C and 2 with HDL-C. These pleiotropic effects are not unexpected given the correlations between triglycerides, HDL-C, and LDL-C. For example, the list of candidate genes at novel loci includes RBM5 (RNA-binding motif protein 5; MIM 606884) involved in cell cycle arrest and apoptosis and CMTM6 (CKLF-like MARVEL; MIM 607889) with an unknown function, both of which also associate with coronary artery disease and a signal near VEGFA (vascular endothelial growth factor A; MIM 192240) locus associated with triglyceride levels, coronary artery disease, type 2 diabetes mellitus, and both systolic and diastolic blood pressure. It codes a growth factor involved in angiogenesis and endothelial cell growth.

In summary, we studied a large, randomly selected and carefully phenotyped cohort of young men and women prospectively followed for ≤31 years since early childhood. Extensive data were available on several possible childhood physical, environmental, and genetic determinants of dyslipidemia that could be comprehensively taken into account in multivariable models. Because our study cohort was racially homogeneous, the generalizability of our results is limited to whites. Other limitations include the loss of original participants during the long-term follow-up. Also, we lacked data of parental dyslipidemia and were not able to consider whether information on parental dyslipidemia may have impacted our model comparisons.

Conclusions
Childhood lipid levels and lipid-specific genetic risk scores were independently related to dyslipidemia in adulthood 21 to 31 years later. The inclusion of wGRSs to lipid-screening programs in childhood could modestly improve the identification of those at highest risk of dyslipidemia in adulthood.

Sources of Funding
The Young Finns Study has been financially supported by the Academy of Finland: grants 286284 (Dr Laitinen), 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi), 296801 (Dr Elo); the Social Insurance Institution of
Disclosures
None.

Appendix
From the Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland (J.N., N.P., C.G.M., M.J., O.T.R.); Menzies Institute for Medical Research, University of Tasmania, Hobart, Australia (C.G.M., M.-J.B.); Department of Biotechnology, University of Turku and Åbo Akademi University, Turku, Finland (M.S.V., L.L.E.); Children’s Hospital University, University Central Hospital Helsinki, Helsinki, Finland (E.J.); Department of Clinical Physiology and Nuclear Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland (T.L.); Department of Pediatrics, Vaasa Central Hospital, Vaasa, Finland (H.L.); Department of Pediatrics, University of Turku and Tampere University Hospital, Tampere, Finland (N.H.-K.); Department of Clinical Chemistry, Fimlab Laboratories and University of Tampere School of Medicine, Tampere, Finland (L.-P.L., T.L.); Department of Medicine, University of Turku, Turku, Finland (J.S.V., M.J.); and Division of Medicine (J.S.V., M.J.) and Department of Clinical Physiology and Nuclear Medicine (O.T.R.), Turku University Hospital, Turku, Finland.

References
CLINICAL PERSPECTIVE

Serum lipoproteins are major modifiable risk factors for atherosclerotic cardiovascular disease. Recent genome-wide association studies have identified several single-nucleotide polymorphisms that are associated with lipid levels. The most recent study identified 157 single-nucleotide polymorphisms associated with lipid levels, including 62 novel single-nucleotide polymorphisms. We studied these genetic variants as lipid-specific weighted genetic risk scores and estimated their incremental value in the prediction of adult dyslipidemia when compared with approach using clinical childhood risk factors in the Cardiovascular Risk in Young Finns Study. The weighted genetic risk scores modestly enhanced the risk prediction of adult dyslipidemia over childhood lipid measurements, which suggests that weighted genetic risk scores might potentially provide useful incremental information in the risk assessment for atherosclerotic cardiovascular disease. In future, identifying children with high genetic risk for atherosclerotic cardiovascular disease could allow for targeting of more intense primary prevention to a high-risk population already in childhood. However, clinical use and cost effectiveness of weighted genetic risk scores in adult dyslipidemia prediction remains uncertain, and further studies with more complete single-nucleotide polymorphism panels and life course information are needed.
Prediction of Adult Dyslipidemia Using Genetic and Childhood Clinical Risk Factors: The Cardiovascular Risk in Young Finns Study
Joel Nuotio, Niina Pitkänen, Costan G. Magnussen, Marie-Jeanne Buscot, Mikko S. Venäläinen, Laura L. Elo, Eero Jokinen, Tomi Laitinen, Leena Taittonen, Nina Hutri-Kähönen, Leo-Pekka Lyytikäinen, Terho Lehtimäki, Jorma S. Viikari, Markus Juonala and Olli T. Raitakari

Circ Cardiovasc Genet. 2017;10:
doi: 10.1161/CIRCGENETICS.116.001604

Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2017 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/10/3/e001604

Data Supplement (unedited) at:
http://circgenetics.ahajournals.org/content/suppl/2017/06/15/CIRCGENETICS.116.001604.DC1
http://circgenetics.ahajournals.org/content/suppl/2017/11/30/CIRCGENETICS.116.001604.DC2

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Genetics can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Genetics is online at:
http://circgenetics.ahajournals.org//subscriptions/
**Supplemental Material**

**Additional Figure 1.** Odds ratios (95% CI) for the best childhood prediction model (age, sex, childhood lipid level, wGRS) for abnormal levels in adulthood* stratified by age.

*LDL-cholesterol >3 mmol/l or reported use of lipid lowering medication in adulthood was defined as abnormal. For HDL-cholesterol, adult values <1.2 (in women) /1.0 mmol/l (in men) were defined as abnormal. For triglycerides, adult values >1.7 mmol/l were defined as abnormal.
Predicción de dislipidemia en adultos mediante el uso de factores de riesgo clínicos genéticos y de la infancia

Estudio de riesgo cardiovascular en jóvenes finlandeses

El presente estudio estimó el valor cada vez mayor de las variantes genéticas como calificaciones de riesgo genético ponderado específico de lípidos para mejorar la predicción de la dislipidemia en adultos, en comparación con los factores de riesgo de la infancia en el estudio del riesgo cardiovascular de jóvenes finlandeses. Los resultados muestran que la calificación de riesgos genéticos ponderados mejora moderadamente la predicción del riesgo de la dislipidemia en adultos. No obstante, el uso clínico y la relación entre el costo y la efectividad siguen siendo inciertos, y se requieren más estudios.

ANTECEDENTES: La dislipidemia es un factor de riesgo modifiable principal para la enfermedad cardiovascular. Examinamos si la adición de polimorfismos de nucleótido simple nuevos para los niveles de lípidos en sangre mejora la predicción de la dislipidemia en adultos en comparación con las mediciones de lípidos en la infancia.

MÉTODOS Y RESULTADOS: Se incluyeron dos mil cuatrocientos veintidós participantes del Estudio de riesgo cardiovascular de jóvenes finlandeses que participaron en 2 encuestas realizadas durante la infancia (en 1980 cuando tenían 3-18 años y en 1986) y, al menos, en un estudio de seguimiento en la adultez (2001, 2007 y 2011). Examinamos si la inclusión de una calificación de riesgo genético ponderado específico de lípidos basada en 58 polimorfismos de nucleótido simple para el colesterol de la lipoproteína de baja densidad, 71 polimorfismos de nucleótido simple para el colesterol de la lipoproteína de alta densidad y 40 polimorfismos de nucleótido simple para los triglicéridos mejoraban la predicción de la dislipidemia en adultos en comparación con los factores de riesgo clínicos de la infancia. Los valores ajustados de la edad, el sexo, el índice de masa corporal, la actividad física y el tabaquismo en la infancia, los niveles de lípidos en la infancia y las calificaciones de riesgo genético ponderado se asociaron con un aumento en el riesgo de la dislipidemia en adultos para todos los lípidos. La evaluación de riesgo basada en 2 mediciones de lípidos en la infancia y la calificación de riesgo genético ponderado específico de lípidos mejoraron la precisión de la predicción de la dislipidemia en adultos en comparación con el enfoque que solo usa las mediciones de lípidos en la infancia para el colesterol de la lipoproteína de baja densidad (área bajo la curva de la característica operativa del receptor 0,806 contra 0,811; P = 0,01) y los triglicéridos (área bajo la curva de la característica operativa del receptor 0,740 contra el área bajo la curva de la característica operativa del receptor 0,758; P < 0,01). La mejoría neta general de la reclasificación y de la discriminación integrada fueron significativas para todos los resultados.

CONCLUSIONES: La inclusión de las calificaciones del riesgo genético ponderado para los programas de cribado de lípidos en la infancia podrían mejorar moderadamente la identificación de aquellos con mayor riesgo de dislipidemia en la adultez.

Correspondencia a: Joel Nuotio, MD, Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Kiantaayllynkatu 10, 20520 Turku, Finlandia. E-mail jovanu@utu.fi