University of Helsinki
Dissertationes Universitatis Helsingiensis
318/2024

# NOVEL GENETIC AND BEHAVIORAL RISK FACTORS FOR HYPERCHOLESTEROLEMIA AND ATHEROSCLEROTIC CARDIOVASCULAR DISEASE

Nella Junna

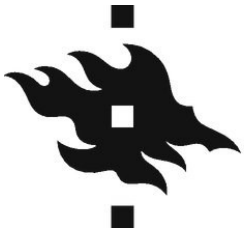
Institute for Molecular Medicine Finland (FIMM),
Helsinki Institute of Life Science (HiLIFE),
University of Helsinki, Finland

Doctoral Program in Population Health
University of Helsinki Doctoral School
University of Helsinki, Finland

DOCTORAL DISSERTATION

To be presented for public discussion with the permission of the Faculty of Medicine of the University of Helsinki, in Biomedicum Lecture Hall 2,
on the 3rd of December 2024, at 13 o'clock.

Helsinki 2024



HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

FIMM
Institute for Molecular Medicine Finland
Nordic EMBL Partnership for Molecular Medicine
HiLIFE UNIT

Publisher: Helsingin yliopisto
Series: Dissertationes Universitatis Helsingiensis 318/2024

ISBN 978-952-84-0426-2 (print)
ISBN 978-952-84-0425-5 (online)
ISSN 2954-2898 (print)
ISSN 2954-2952 (online)

PunaMusta, Joensuu 2024

The Faculty of Medicine uses the Ouriginal system (plagiarism recognition) to examine all doctoral dissertations.

**Supervisor**

Docent Elisabeth Widén, MD, PhD
Institute for Molecular Medicine Finland (FIMM)
HiLIFE, University of Helsinki, Finland

**Reviewers**

Professor Katriina Aalto-Setälä, MD, PhD
Faculty of Medicine and Health Technology
Tampere University
Tampere, Finland

Associate professor Matti Jauhiainen, PhD
Minerva Foundation Institute for Medical Research
Helsinki, Finland

**Opponent**

Professor Johannes Kettunen, PhD
Center for Life Course Health Research
Faculty of Medicine, and Biocenter Oulu
University of Oulu
Oulu, Finland

# ABSTRACT

Atherosclerotic cardiovascular disease (ASCVD), including coronary artery disease (CAD), is the most common cause of death worldwide. One of the most important risk factors is hypercholesterolemia, especially elevated low-density lipoprotein cholesterol (LDL-C). Identifying risk factors for hypercholesterolemia and ASCVD is important, to identify individuals at high risk to prevent detrimental outcomes of ASCVD.

The aims of this thesis were to identify novel risk factors for hypercholesterolemia and ASCVD and estimate whether returning personal genetic risk to study participants would motivate them to make positive health modifications and thus lower their risk for ASCVD. All separate studies composing this doctoral thesis used the GeneRISK study cohort (N=7 342), including baseline health check-up and e-questionnaire, and follow-up study after 1.5 years on average. All study participants received their personal risk for developing ASCVD, including traditional risk factors and polygenic risk.

High-impact mutations in genes associated with familial hypercholesterolemia (FH) can cause severely elevated LDL-C levels and significantly increase the risk for ASCVD. In Study I, the aim was to characterize the genetic architecture of severe hypercholesterolemia in Finland by whole-exome sequencing 351 cases with severe hypercholesterolemia (LDL-C ≥ 5 mmol/l) and further assess the impact of both novel and known hypercholesterolemia-causing mutations on ASCVD risk using a large population cohort. We identified two novel Finnish-enriched mutations in *LDLR* and one in *LDLRAP1* and additionally discovered that the impact on ASCVD risk varies by gene and mutation. ASCVD risk was also associated with polygenic clustering of variants associated with risk for hypercholesterolemia or CAD.

Behavioral factors, such as diet, affect LDL-C levels, but the impact of personal attitudes has not been as widely studied. In Study II, the purpose was to examine the effects of behavioral factors and attitudes on hypercholesterolemia. We discovered that cases with hypercholesterolemia (N=701, LDL-C ≥ 4.6 mmol/l) also exhibited more behavioral risk factors (e.g., less healthy diets) than their control counterparts with normal cholesterol levels (N=3 128, LDL-C < 3.4 mmol/l). Cases had significant sex differences in their behavioral factors as well. A poor attitude to normal

cholesterol was more common among males and was associated with a less healthy diet. In addition, we discovered that hypercholesterolemia was poorly treated at baseline despite most of the cases being previously diagnosed with hypercholesterolemia, and the same trend continued at follow-up.

There is evidence that adding polygenic risk to the already commonly used ASCVD risk models improves the risk estimate and helps to identify individuals at high risk. In Study III, we returned ASCVD risk, including polygenic risk, to study participants and explored how they perceived their personal risk and how it affected their health behavior. The attitudes toward personal risk were primarily positive, and particularly persons at high risk for ASCVD (N=683) were motivated to make positive lifestyle modifications, such as losing weight.

This doctoral thesis demonstrates that the effects of both known and novel hypercholesterolemia-causing mutations and polygenic clustering on ASCVD risk are variable. Additionally, individuals with hypercholesterolemia had less healthy lifestyles. Although it is important to identify high-risk persons, most cases with hypercholesterolemia did not receive recommendations to prevent ASCVD from the healthcare. Nevertheless, study participants with high risk for ASCVD found the risk information motivational for improving their health behavior.

# TIIVISTELMÄ

Ateroskleroottinen sydäntauti, sisältäen myös sepelvaltimotaudin, on maailmanlaajuisesti yleisin kuolinsyy. Hyperkolesterolemia, erityisesti kohonnut LDL kolesteroli (LDL-C), on yksi tärkeimmistä sepelvaltimotaudin riskitekijöistä. Hyperkolesterolemian ja sepelvaltimotaudin riskitekijöiden tunnistaminen on tärkeää, että korkeariskiset henkilöt voidaan tunnistaa ja sepelvaltimotaudin vakavat seuraukset ennaltaehkäistä.

Tämän väitöskirjan tavoitteena oli tunnistaa uusia hyperkolesterolemian ja sepelvaltimotaudin riskitekijöitä ja arvioida, vaikuttaisiko geneettisen riskin tiedostaminen tutkittavien henkilöiden motivaatioon tehdä positiivisia terveyskäyttäytymisen muutoksia ja täten alentaa heidän sydäntautiriskiään. Kaikissa tämän väitöskirjan osatöissä hyödynnettiin GeneRISK tutkimuskohorttia (N=7 342), jossa tutkittavat osallistuivat terveystarkastukseen ja täyttivät sähköisen kyselylomakkeen. Keskimäärin 1,5 vuoden jälkeen he osallistuivat seurantatutkimukseen. Kaikille tutkittaville palautettiin heidän yksilöllinen riskinsä sairastua sydäntautiin, riskiarvio sisälsi sekä perinteisen riskiarvion että polygeenisen riskisumman.

Korkeavaikutteiset mutaatiot familiaaliseen hyperkolesterolemiaan (FH) liittyvissä geeneissä voivat aiheuttaa merkittävästi kohonneita LDL-C-tasoja ja nostaa huomattavasti riskiä sairastua sydäntautiin. Ensimmäisessä osatyössä (I) tavoite oli kuvata vakavan hyperkolesterolemian geneettistä taustaa Suomessa. Eksomisekvensoimme 351 henkilöä, joilla oli vakava hyperkolesterolemia (LDL-C ≥ 5 mmol/l). Lisäksi arvioimme sekä uusien että tunnettujen hyperkolesterolemiaa aiheuttavien mutaatioiden vaikutusta sydäntautiriskiin käyttämällä aineistona suurta populaatiokohorttia. Löysimme kaksi uutta Suomeen rikastunutta mutaatiota LDL-reseptori (*LDLR*)-geenissä sekä yhden mutaation *LDLRAP1*-geenissä. Lisäksi havaitsimme, että mutaatioiden vaikutus sydäntautiriskiin vaihtelee geeneittäin sekä mutaatioittain. Myös hyperkolesterolemiaan tai sydäntautiriskiin liittyvien varianttien polygeenisellä kertymisellä oli merkittävä vaikutus sydäntautiriskiin.

Käyttäytymiseen liittyvät tekijät, kuten ruokavalio, tunnetusti vaikuttavat LDL-C-tasoon, mutta asenteiden vaikutusta ei ole tutkittu yhtä laajasti. Toisessa osatyössä (II) tarkoitus oli tutkia käyttäytymiseen liittyvien tekijöiden ja asenteiden vaikutusta

hyperkolesterolemiaan. Hyperkolesterolemiatapauksilla (N=701, LDL-C ≥ 4.6 mmol/l) oli useampia riskitekijöitä (esim. epäterveellinen ruokavalio) kuin normaalin LDL-C:n (N=3 128, LDL-C < 3.4 mmol/l) omaavilla verrokeilla. Tapauksilla oli myös merkittäviä sukupuolieroja riskitekijäprofiileissaan. Miehet olivat naisia useammin sitä mieltä, ettei kolesterolin pitäminen normaalilla tasolla ole tärkeää ja kyseinen asenne liittyi epäterveellisempään ruokavalioon. Lisäksi havaitsimme, että tutkimuksen alkaessa hyperkolesterolemia oli huonosti hoidettu, vaikka suurin osa tapauksista oli jo aiemmin tiennyt kohonneista kolesteroliarvoistaan. Seurantatutkimuksessa huono hoitotaso vaikutti jatkuvan.

Useissa tutkimuksissa on havaittu, että polygeenisen riskin lisääminen yleisesti käytettyihin sydäntaudin riskiarvioihin parantaa riskiarvion ennustetta, tunnistaen useampia henkilöitä, joilla on korkea riski sairastua. Kolmannessa osatyössä (III) toimme tutkittaville tietoon sydäntautiriskiarvion, joka sisälsi polygeenisen riskisumman. Kartoitimme, miten tutkittavat suhtautuivat yksilölliseen riskiarvioonsa, ja miten se vaikutti heidän terveyskäyttäytymiseensä. Asenteet yksilöllistä riskiarviota kohtaan olivat enimmäkseen positiivisia, ja erityisesti korkean riskin omaavat henkilöt (N=683) olivat motivoituneita tekemään positiivisia terveyskäyttäytymisen muutoksia, kuten pudottamaan painoaan.

Tämä väitöskirja osoittaa miten sekä tunnettujen että uusien hyperkolesterolemiaa aiheuttavien mutaatioiden ja polygeenisen kertymisen aiheuttama sydäntautiriskin suuruus vaihtelee. Lisäksi hyperkolesterolemiatapauksien terveyskäyttäytyminen oli epäterveellisempää. Vaikka korkeariskisten henkilöiden tunnistaminen on tärkeää, suurin osa tapauksista ei saanut suosituksia sydäntaudin ehkäisyyn terveydenhuollosta. Kuitenkin korkeariskiset tutkittavat kokivat sydäntautiriskitiedon motivoivana parantaa terveyskäyttäytymistään.

# CONTENTS

**ABSTRACT** ..... i

**TIIVISTELMÄ** ..... iii

**CONTENTS** ..... v

**LIST OF ORIGINAL PUBLICATIONS** ..... vii

**ABBREVIATIONS** ..... viii

**1 INTRODUCTION** ..... 1

**2 LITERATURE REVIEW** ..... 3

**2.1 Genetic predisposition to diseases** ..... 3

2.1.1 Sequencing of the human genome ..... 3

2.1.2 Genetic variation of the human genome ..... 4

2.1.3 Studying monogenic diseases ..... 5

2.1.4 Studying polygenic and multifactorial diseases ..... 10

2.1.5 Biobanks ..... 13

2.1.6 Complexities of mutational analysis ..... 14

**2.2 Atherosclerotic cardiovascular disease (ASCVD) and hypercholesterolemia** ..... 15

2.2.1 Epidemiology of cardiovascular disease ..... 15

2.2.2 Atherosclerosis ..... 18

2.2.3 Lipoproteins and cholesterol metabolism ..... 20

2.2.4 Monogenic risk factors for hypercholesterolemia and ASCVD ..... 25

2.2.5 Polygenic risk factors for hypercholesterolemia and ASCVD ..... 35

2.2.6 How to estimate personal ASCVD risk? ..... 37

2.2.7 Complex dynamics of risk factors for hypercholesterolemia and ASCVD ..... 38

**2.3 Behavioral factors and disease prevention** ..... 39

2.3.1 CVD prevention guidelines and lifestyle modifications ..... 39

2.3.2 Impact of attitudes and adherence on health behavior ..... 41

2.3.3 Returning personal disease risk ..... 43

**3 AIMS OF THE STUDY** ..... 45

**4 MATERIALS AND METHODS** ..... 46

**4.1 Study cohorts** ..... 46

4.1.1 The GeneRISK study ..... 46

4.1.2 The FinnGen study ..... 47

**4.2 Genetic analyses** ..... **49**
4.2.1 Whole-exome sequencing (WES) analysis ..... 49
4.2.2 Sanger sequencing ..... 50
4.2.3 Genotyping and imputation ..... 50
4.2.4 Polygenic risk scores (PRSs) ..... 50
**4.3 Phenotype analyses** ..... **51**
**4.4 Returning ASCVD risk information** ..... **52**
4.4.1 Estimation and definition of ASCVD risk ..... 52
4.4.2 Returning ASCVD risk via Kardiokompassi® ..... 53
4.4.3 Follow-up and definition of positive health behavior ..... 53
**4.5 Statistical analyses** ..... **54**
**5 RESULTS AND DISCUSSION** ..... **55**
**5.1 Novel genetic risk factors for severe hypercholesterolemia (I)** ..... **55**
5.1.1 Novel mutations causing severe hypercholesterolemia ..... 56
5.1.2 The Finnish spectrum of hypercholesterolemia-causing mutations broader than previously reported ..... 60
5.1.3 Utilizing a large population cohort to estimate the impact of genetic variants ..... 63
5.1.4 Individuals with high genetic risk for ASCVD are crucial to identify and treat ..... 68
**5.2 Impact of behavioral factors and attitudes on hypercholesterolemia and ASCVD prevention (II)** ..... **69**
5.2.1 Hypercholesterolemia cases have more behavioral risk factors ..... 69
5.2.2 Poor attitude to cholesterol more common in males ..... 71
5.2.3 ASCVD prevention leaves room for improvement ..... 73
5.2.4 The reasons behind the undertreatment of hypercholesterolemia remain unknown ..... 74
**5.3 Impact of returning polygenic risk for ASCVD on health behavior (III)** ..... **75**
5.3.1 Attitudes to polygenic risk mostly positive ..... 75
5.3.2 Having high risk for ASCVD motivates positive health behavior ..... 76
5.3.3 Guidelines for polygenic risk estimates still missing ..... 77
**5.4 Limitations** ..... **77**
**6 CONCLUSIONS AND FUTURE DIRECTIONS** ..... **78**
**ACKNOWLEDGEMENTS** ..... **81**
**REFERENCES** ..... **83**
**ORIGINAL PUBLICATIONS** ..... **106**

# LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications, which are referred to in the text by their Roman numerals:

I. **Junna, N.**, Ruotsalainen, S., Ripatti, P., Ripatti, S., & Widén, E. (2023). Novel Finnish-enriched variants causing severe hypercholesterolemia and their clinical impact on coronary artery disease. *Atherosclerosis*, 386, 117327.

II. **Junna, N.**, Ruotsalainen, S., Pöllänen, P., Partanen, J., Hotakainen, K., Ripatti, S. & Widén, E. (2024). Self-reported health behavior and attitudes towards the significance of cholesterol as risk factors for atherosclerotic disease among middle-aged persons with severe hypercholesterolemia. *Submitted*

III. Widén, E., **Junna, N.***, Ruotsalainen, S.*, Surakka, I., Mars, N., Ripatti, P., Partanen, J.J., Aro, J., Mustonen, P., Tuomi, T., Palotie, A., Salomaa, V., Kaprio, J., Partanen, J., Hotakainen, K., Pöllänen, P. & Ripatti, S. (2022). How communicating polygenic and clinical risk for atherosclerotic cardiovascular disease impacts health behavior: an observational follow-up study. *Circulation: Genomic and Precision Medicine*, 15(2), e003459. (* shared second authors)

Publication III was also used in Dr. Ruotsalainen's doctoral dissertation.

# ABBREVIATIONS

| | |
|---|---|
| ANGPTL3 | Angiopoietin-like 3 |
| APOB | Apolipoprotein B |
| APOC | Apolipoprotein C |
| APOE | Apolipoprotein E |
| ARH | Autosomal recessive hypercholesterolemia |
| ASCVD | Atherosclerotic cardiovascular disease |
| BAM | Binary Alignment Map |
| BMI | Body mass index |
| bp | Base pair |
| CAD | Coronary artery disease |
| CAD-PRS | Polygenic risk score for coronary artery disease |
| CHD | Coronary heart disease |
| CI | Confidence interval |
| CNV | Copy number variation |
| CVD | Cardiovascular disease |
| DNA | Deoxyribonucleic acid |
| EGF | Epidermal growth factor |
| ER | Endoplasmic reticulum |
| FH | Familial hypercholesterolemia |
| FRS | Framingham Risk Score |
| gnomAD | The Genome Aggregation Database |
| GWAS | Genome-wide association study |
| HDL-C | High-density lipoprotein cholesterol |
| HL | Hepatic lipase |
| HMGCR | 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase |
| HR | Hazard ratio |
| IDL | Intermediate-density lipoprotein |
| IGV | Integrative Genomics Viewer |
| IHD | Ischemic heart disease |
| kbp | Kilo base pair |
| LDL | Low-density lipoprotein |
| LDL-C | Low-density lipoprotein cholesterol |
| LDL-PRS | Polygenic score for low-density lipoprotein |
| LDLR | Low density lipoprotein receptor |
| LDLRAP1 | Low density lipoprotein receptor adaptor protein 1 |
| Lp(a) | Lipoprotein(a) |
| LPL | Lipoprotein lipase |
| LPS | Lipopolysaccharides |
| MACE | Major adverse cardiovascular events |
| MAF | Minor allele frequency |
| MI | Myocardial infarction |
| mmHg | Millimeters of mercury |
| mRNA | Messenger ribonucleic acid |

| | |
|---|---|
| NGS | Next generation sequencing |
| NO | Nitric oxide |
| NS | Non-significant |
| OR | Odds ratio |
| PA | Physical activity |
| PAD | Peripheral artery disease |
| PCSK9 | Proprotein convertase subtilisin/kexin type 9 |
| PGS | Polygenic score |
| PRS | Polygenic risk score |
| PTB | Phosphotyrosine-binding |
| QC | Quality control |
| RCT | Reverse cholesterol transport |
| RFDS | Recommended Finnish diet score |
| RR | Risk ratio |
| SD | Standard deviation |
| SNP | Single nucleotide polymorphism |
| SREBP2 | Sterol regulatory element binding protein 2 |
| TG | Triglycerides |
| VLDL | Very low-density lipoprotein |
| WES | Whole-exome sequencing |
| WGS | Whole-genome sequencing |
| WHO | World Health Organization |

# 1 INTRODUCTION

Cardiovascular disease (CVD) is the most common cause of death and disability worldwide, accounting for approximately 20 million deaths annually [1]. In accordance, CVD is the most common cause of death in Finland (~30% of deaths) [2]. Atherosclerotic CVD (ASCVD) is caused by atherosclerosis, which is the buildup of plaques inside arteries, causing narrowing of the arteries and risk for blood clots [3]. ASCVD includes coronary artery disease (CAD), cerebrovascular disease, and peripheral artery disease (PAD). ASCVD causes severe illness, such as myocardial infarction (MI), stroke, and sudden cardiac death. Risk factors for ASCVD can be divided into non-modifiable and modifiable factors. The first category includes sex, age, and genetic factors, whereas the second category includes hypertension, hyperlipidemia, diabetes, obesity, smoking, lack of exercise, and poor diet.

Hypercholesterolemia, elevated cholesterol levels in the blood, is a major risk factor for ASCVD. The causal relationship between low-density lipoproteins (usually measured by LDL-C, low-density lipoprotein cholesterol) and ASCVD has been proven through decades of studies with various viewpoints [4]. Moreover, the effect on ASCVD risk is cumulative, i.e., the longer the exposure to excess LDL-C and the greater the concentration, the higher the ASCVD risk. Therefore, early prevention, including medical treatment and lifestyle modifications, is crucial [5]. Hypercholesterolemia can be caused by monogenic mutations and polygenic predisposition, as well as behavioral factors, especially a diet rich in saturated fatty acids.

Hypercholesterolemia was already characterized in the 1920s [6]. Patients with familial hypercholesterolemia (FH) have a significantly increased risk for ASCVD. FH is categorized as a monogenic disease caused by dominant mutations, including large deletions and missense mutations. Mutations in genes *LDLR*, *APOB*, *PSCK9,* and *LDLRAP1* are associated with FH [7–9]. In Finland, seven mutations in *LDLR* have been identified and are estimated to cause approximately 80% of FH [10,11]. Although FH was originally considered a monogenic disease, accumulating evidence has inferred that the clinical FH phenotype can be caused by the clustering of polygenic, LDL-C increasing variants. In turn, genetic explanation cannot be identified for all patients with clinical FH [12–14]. Furthermore, the exact impact of FH mutations on the

phenotype, i.e., LDL-C level and ASCVD risk, is usually left uncovered. In Study I, we discovered novel hypercholesterolemia-causing mutations and examined their impact on ASCVD risk using a large population cohort.

In addition to genetic factors, behavioral factors, such as smoking, diet, and exercise, have been studied regarding their effects on ASCVD risk and hypercholesterolemia [1,15]. However, health-related attitudes have not been as widely studied; the main focus has been on adherence to medical treatment [16–18]. Moreover, several studies have demonstrated that patients with hypercholesterolemia are undertreated [19–21]. In Study II, we explored the association of behavioral factors and health-related attitudes with hypercholesterolemia. We also examined the ASCVD prevention measures (e.g., medical treatment) of hypercholesterolemia cases in healthcare.

Large GWASs and their results have allowed the development of polygenic risk scores (PRSs), which can estimate the genetic risk for ASCVD. The first CVD risk model, the Framingham risk score, was developed already in 1967 [22]. It has since been improved, and other risk models have been developed. CVD risk estimates are widely used in healthcare, aiding in identifying high-risk individuals and enabling targeted interventions. Regarding genetic risk, disease-causing monogenic mutations have long been sequenced and interpreted in healthcare. However, polygenic risk has not yet been integrated into ASCVD risk models. Furthermore, some studies have shown that incorporating PRSs into standard ASCVD risk models can identify more high-risk individuals who especially benefit from medical treatment [23,24]. Additionally, younger high-risk individuals could be identified, who remain unrecognized by current risk models [25]. Nonetheless, adding PRS to ASCVD risk models has raised concerns about how people would react to polygenic risk information. In Study III, the study participants received their ASCVD risk, including polygenic risk. We studied how the participants reacted to their risk, as well as how the risk information affected their health behavior.

# 2 LITERATURE REVIEW

## 2.1 Genetic predisposition to diseases

A wide range of genetic variants can affect the predisposition to diseases. Locus refers to the position on a chromosome. Alleles are one or more nucleotides residing at a specific locus. If alleles at a locus in both chromosomes are the same, the individual is homozygous for that locus and heterozygous if the alleles are different. If the disease is caused by a high-impact mutation in one gene, it is called a monogenic disease (i.e., Mendelian disease), whereas polygenic diseases are affected by multiple genes. In addition, millions of single nucleotide polymorphisms (SNPs) affect different diseases and traits. These polygenic variants can affect the predisposition, onset, and clinical phenotype of multifactorial diseases (i.e., diseases also affected by environmental factors). Traditionally, genetic research has focused on rare monogenic disorders to identify single genes and mutations underlying these syndromes. Later, the research focus has shifted to common diseases and variants.

However, this dichotomy of rare monogenic diseases and common multifactorial diseases is misleading. Genetic variants can modify the expression of other genes, the impact of variants is often variable, and environmental factors can also have an effect. Nonetheless, the genetic backgrounds of rare and common diseases have traditionally been studied using quite separate methods: rare variants with whole-exome sequencing (WES), for instance, and common diseases with genome-wide association studies (GWASs).

### 2.1.1 Sequencing of the human genome

Originally, studying the genetic background of diseases was quite laborious with, for instance, linkage analysis (searching for chromosomal regions that co-segregate with the phenotype). The sequencing of the human genome and the parallel technology development enabled the progress of more efficient analysis methods. The human genome was first sequenced in 2001 [26,27], covering approximately 94% of the human genome. Thus, a 2.91 billion base pair (bp) consensus sequence was formed,

leading the research of human genetics to a new level. The euchromatic sequence (i.e., actively transcribed sequence accounting for ~92% of the human genome) was completed in 2004, covering approximately 99% of the euchromatic genome and with an estimated 20 000 genes [28]. Later, in 2022, the human genome sequence was further improved by filling gaps, e.g., in the X chromosome and telomeres (the ends of chromosomes) [29].

### 2.1.2 Genetic variation of the human genome

One central purpose of human genetics has always been to identify mutations causing diseases. More samples were needed to decipher the amount of benign genetic variation in the population.

To investigate variation in the human genome, the International HapMap Project was launched in 2003 [30]. In its three phases [31–33], the HapMap Project provided information on human haplotypes in different geographical regions, recombination hotspots, and linkage disequilibrium structure of the human genome (i.e., groups of alleles inherited together). However, the sample size was modest, with ~300 samples in the 1st and 2nd phases and a little over 1000 in the last phase. In 2010, The 1000 Genomes Project was launched, with the aim to provide more detailed information on human genome variation [34]. In the end, in 2015, the 1000 Genomes Project database included 2 504 genomes or exomes from 26 populations [35–37]. Later, the numbers of publicly available whole-genome or whole-exome sequences have drastically increased. The purpose of the Exome Aggregation Consortium (ExAC) was to provide information on human genetic variation in an easily accessible web browser to enable, e.g., filtering of candidate variants based on frequency [38]. ExAC database included data of over 60 000 whole-exome sequenced individuals. With the Genome Aggregation Database (gnomAD), the number of publicly accessible exomes has increased to over 730 000 and genomes to over 76 000 [39–41].

As a result, large genetic databases have revealed that an average person carries 250-300 loss-of-function variants (i.e., frameshifting insertions or deletions, mutations causing premature STOP codons, or disrupting splice sites). In addition, an average person carries around 24-30 heterozygous variants in genes associated with

genetic diseases [34,36]. Some of these variants are putative recessive mutations, whereas others are not likely to cause illness.

Since Finnish people differ genetically from other European populations [42,43], it is important to have genetic data on Finns when studying the genetics of Finnish people. Finnish individuals were included already in the 1000 genomes study, but only with a modest sample size (N~100). In 2014, The Sequence Initiative Suomi (SISu) was launched, including over 10 000 individuals [44].

These publicly available databases have provided essential information on human genetic variation that can be utilized when studying possibly pathogenic variants.

## 2.1.3 Studying monogenic diseases

### 2.1.3.1 First-generation sequencing (Sanger sequencing)

First-generation DNA sequencing was developed by Sanger et al. in 1977 [45], and it is still considered the golden standard of DNA sequencing. However, it is far too slow and laborious for large sample sizes and long DNA sequences. Yet, it is still widely used, e.g., to determine known mutations of monogenic diseases.

Sanger sequencing is based on DNA synthesis by DNA polymerase (**Figure 1**, next page). A primer (a short, single-stranded DNA sequence) attaches to the template DNA. DNA polymerase adds nucleotides based on the template, extending the DNA sequence. When a fluorescently labeled nucleotide is added, the synthesis terminates. Then, the resulting DNA sequence is read. Initially, the sequence was deduced individually from an electrophoresis gel. Later, the sequencing was automated with DNA sequencers, including reading the sequence, which made the method faster and easier to use.

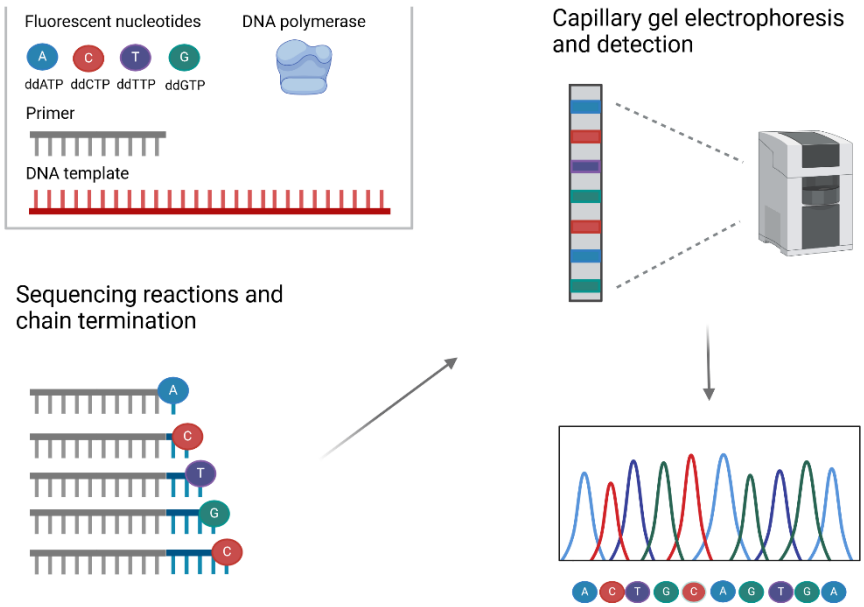



**Figure 1. Sanger sequencing.** Sanger sequencing, i.e., capillary gel electrophoresis, is based on DNA synthesis by DNA polymerase. Reagents include fluorescently labeled, chain-terminating dideoxy nucleotides (ddNTPs: ddATP, ddCTP, ddTTP, ddGTP), regular nucleotides (not shown), DNA polymerase, a primer (short, single-stranded DNA sequence), and template DNA. In the first step (chain-termination polymerase chain reaction), the primer attaches to the template DNA, and DNA polymerase extends the DNA from the primer, adding nucleotides. The synthesis terminates when a ddNTP is added. In the second step, the chain-terminated oligonucleotides (short DNA molecules) are separated by size with capillary gel electrophoresis. Then, the gel is read to infer the DNA sequence based on the fluorescent ddNTPs, producing a chromatogram. Figure created with Biorender.com.

#### 2.1.3.2 Next-generation sequencing (NGS)

Sanger sequencing is not adequate for sequencing whole genomes of large numbers of samples: costs are too high, and the duration is too lengthy. This was noticed especially when sequencing the human genome for the first time, which accelerated the development of new sequencing technologies called next-generation sequencing (NGS). NGS is also known as massive parallel sequencing, referring to the fact that billions of DNA fragments can be sequenced in parallel. However, the same principle as in Sanger sequencing is used in many newer technologies. On a general level, NGS protocols consist of DNA sample preparation, sequencing, and analysis of sequence data using bioinformatics tools. There are several different approaches and platforms, of which different Illumina platforms

(https://www.illumina.com/systems/sequencing-platforms.html) are the most common ones in use (**Figure 2**). NGS technologies can be divided into short-read and long-read technologies (read refers to the sequenced DNA fragment).

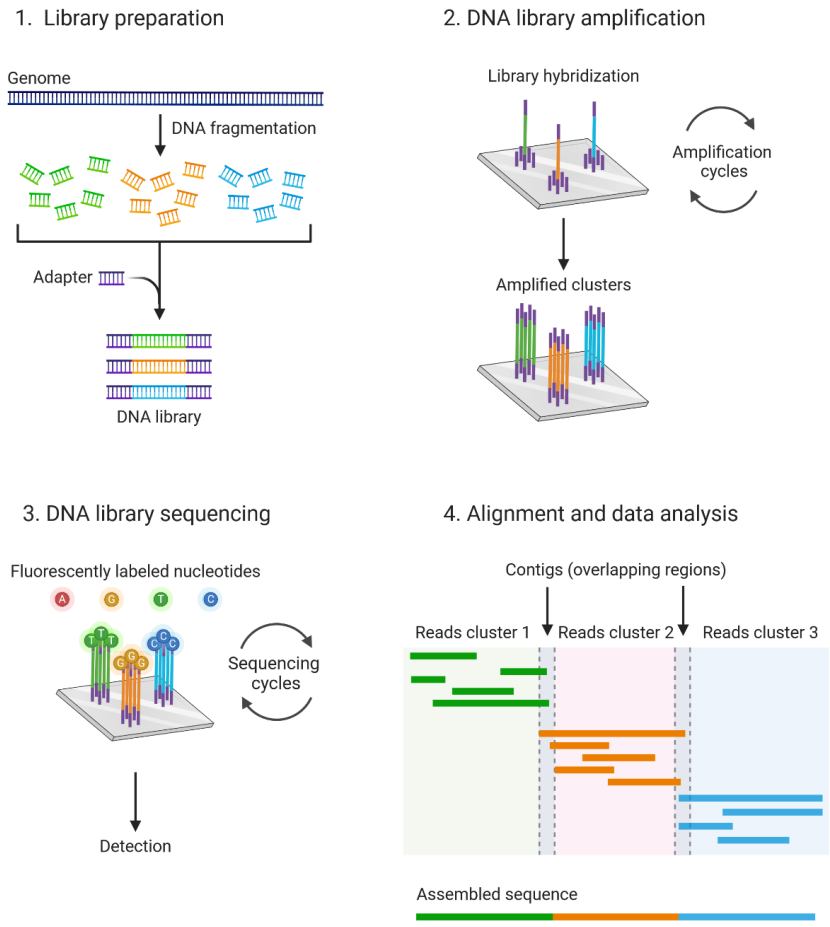



**Figure 2. Next-generation sequencing (NGS): sequencing by synthesis technique.** Sequencing by synthesis method is similar to Sanger sequencing and is utilized by Illumina platforms, for example. The first step, DNA library preparation: DNA is fragmented and amplified, creating a DNA library. Adapter sequences are added to the ends of the DNA fragments. The second step, DNA library amplification: DNA library, i.e., the DNA fragments containing adapters, ligate to the corresponding oligonucleotides (short DNA molecules) on a solid surface. The DNA library is then amplified by DNA polymerases. The third step, DNA library sequencing: DNA polymerases move along DNA strands, and fluorescent nucleotides are added and detected, similar to Sanger sequencing. The fluorescent nucleotides terminate DNA elongation and are cleaved off after detection, allowing DNA elongation to continue. Final step, alignment, and data analysis: the short contiguous sequences (contigs) are read and aligned, and the complete sequence is assembled using bioinformatics tools. Figure created with Biorender.com.

In short-read NGS (**Figure 2**), DNA is first fragmented and amplified to create a DNA library. The sequencing platforms mainly used today are based on sequencing by synthesis- technique, which is based on DNA polymerase, similar to Sanger sequencing. Briefly, the DNA library is ligated to a solid surface, and then DNA polymerases move along the DNA strands. When a new nucleotide is added, it is detected. For example, in Illumina platforms, DNA fragments are ligated to flow cells. DNA polymerase moves along the DNA strands and adds modified nucleotides containing blocking agents. These nucleotides terminate DNA elongation and are detected by fluorescence. Nucleotides are then cleaved off, and the synthesis continues. While the short-read technologies are most commonly used, their disadvantage is the short read length (usually less than 800 bp): larger structural variations are missed [46,47].

Long-read NGS was later developed to compensate for the deficiencies of short-read technologies. Although long-read techniques are better at identifying larger variations in the genome, their accuracy has not been as good as that of, e.g., Illumina's platforms. One difference compared with short-read technologies is that no DNA amplification is required, which is often the source of errors. DNA polymerase remains in place, and the DNA strand moves along it [46,47].

In NGS, errors often occur. The DNA polymerases used in sequencing can produce artifacts. Typically, there are more errors at the beginnings and ends of the reads. Moreover, different NGS technologies are prone to their specific errors. Generally, areas with the most difficulties in sequencing are rich in adenine (A) and thymine (T) nucleotides or guanine (G) and cytosine (C). Additionally, insertions and deletions are more challenging to detect correctly [46,47].

When interpreting massive parallel sequencing, quality control is a critical step. In the first step, low-quality reads are removed. Then, the reads are aligned to the reference genome. This part can be difficult, e.g., in regions with high numbers of repetitive sequences. The last step is variant calling, where false positives are removed [48]. After these initial steps, the actual interpretation of the variants can begin.

The variants are annotated using, for example, Annovar [49] or VEP [50] bioinformatics tools. Variant annotation refers to attaching functional information to the variants, for example, if the variant causes protein-coding changes. If the variant is previously associated with the disease in question, it can be queried from mutation

databases (e.g., ClinVar [51]). There are also several tools for predicting the pathogenicity of a variant, such as SIFT, PolyPhen-2, MutationTaster, and CADD [52–55]. Furthermore, if a variant is common in the population, it is unlikely to cause a rare disease. Thus, the previously mentioned human variation databases, such as gnomAD, are helpful tools for assessing the pathogenicity of a variant.

#### 2.1.3.3 Targeted NGS

In addition to whole-genome sequencing (WGS), only specific regions of the genome can be targeted for sequencing. In targeted NGS, only the region of interest in the genome is captured (e.g., with hybridization), enriched, and sequenced. However, a weakness of targeted NGS is that larger structural variations are missed [56,57].

In whole-exome sequencing (WES), only exons are sequenced. Only about 1% of the genome consists of exons; therefore, the cost of WES is a fraction of that of WGS. When investigating monogenic diseases, it is usually assumed that the mutation is located in the coding region of a gene, i.e., in an exon. In 2009, the first three genes associated with monogenic diseases were identified using WES [58–60]. Since then, the number of genes discovered that affect monogenic diseases has increased immensely.

In addition to WES, different disease-specific gene panels for targeted NGS exist and are mainly used in clinical settings. Gene panel sequencing is especially practical when several genes are associated with the genetic disease in question, compared to just one gene that could be sequenced using Sanger sequencing. Moreover, gene panel sequencing costs are lower than WES's [61].

### 2.1.4 Studying polygenic and multifactorial diseases

NGS technologies have successfully identified hundreds of genes associated with monogenic diseases. However, to discover disease-associated, common variants with modest impacts, large sample sizes are required to detect their small effects. Using NGS for this purpose would require considerable resources; thus, other methods have been developed.

#### 2.1.4.1 DNA microarrays and imputation

DNA microarrays, also called chips, are a relatively old technique: modern DNA arrays were developed in 1994 [62]. DNA microarrays are based on measuring the relative concentrations of different nucleotides, usually by fluorescence (**Figure 3**). DNA, which is typically first fragmented, hybridizes with short DNA molecules called oligonucleotides. Often, there are two different probes (labeled oligonucleotides) for two different alleles at the same position. Then, after hybridization, the concentrations of these two alleles are measured automatically [63]. Modern DNA microarrays can measure millions of SNPs parallelly from large sample sizes. They can also be used in SNP genotyping to produce genetic information much faster and cheaper than NGS. Although, in practice, they usually do not cover nearly as many variants as NGS.

The yield of SNP genotyping can be improved using imputation. Imputation utilizes a haplotype reference panel where genotypes for a vast number of SNPs are known. Genotypes for missing SNPs can be inferred from fewer SNPs using the reference information. SNP imputation is especially useful for GWAS because an increase in the number of known genotypes can increase the statistical power of the analysis [64].

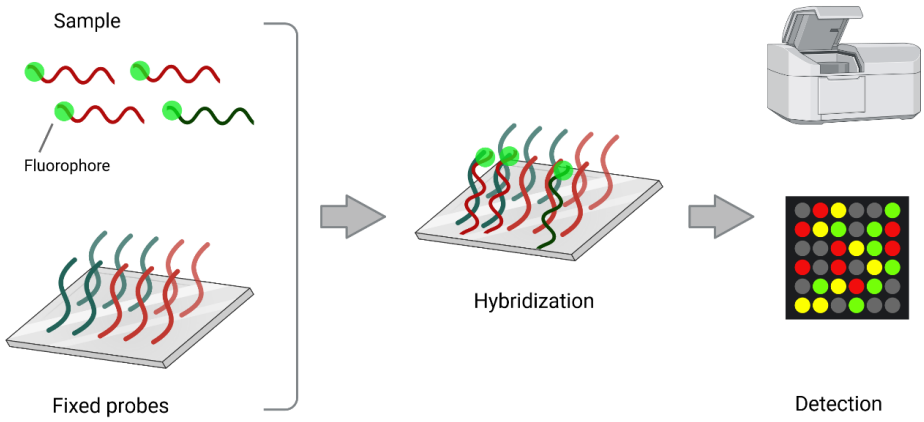



**Figure 3. DNA microarrays and genotyping.** Probes, i.e., single-stranded oligonucleotides, are fixed to a solid surface or a bead. Often, there are two different probes corresponding to two different alleles at the same position. Single-stranded sample DNA is labeled with a fluorophore (fluorescent chemical) and hybridized with complementary probes to form double-stranded DNA. The unattached samples are then washed away. Fluorescent labels are detected, and their intensity, i.e., concentration, is measured. Figure created with Birender.com.

#### 2.1.4.2 Genome-wide association study (GWAS)

The first GWAS was published in 2005 [65]. Since then, the number of GWASs and significant associations has rocketed.

"Common Disease, Common Variant"-hypothesis states that common diseases are affected by common variants, which have only a small impact on the trait in question [66,67]. Genome-wide association studies (GWASs) are used to study the associations between SNPs and diseases or traits. The associations are statistically calculated in a population, and the studied trait can be quantitative or binary. Thus far, the larger the sample size has been, the larger the number of statistically significant SNPs that have been discovered [68]. DNA microarrays are commonly used for SNP genotyping GWAS study samples, but other genetic data can also be used (e.g., WGS). As mentioned previously, not all SNPs need to be genotyped due to imputation: around 1 million SNPs can cover more than 80% of common SNPs found in European populations [69].

Typically, SNPs are tested individually, and the model is additive. Testing single SNPs causes multiple testing burden: the commonly used significance level for P value (0.05) must be corrected. Bonferroni correction is used with 1 million independent tests, corresponding to the estimated number of independent SNPs in the human

genome. The correction results in a significance level of $p < 5e\text{-}8$, which is commonly used as the genome-wide significance level [69,70].

Because of linkage disequilibrium (LD, non-random association of alleles in different loci), the SNP with the highest P value is not necessarily the causal SNP: only a third of the top hits are actually near the gene from where the association signal originates. The causal SNP or gene can be identified using fine-mapping (i.e., inferring the likely causal variant behind the GWAS signal) or performing functional studies [68]. Moreover, replication of the association in another study cohort is needed to verify the association. GWAS results can be erroneous if the cases and controls are not carefully defined, the model correctly adjusted (e.g., age and sex), and the population structure taken into account [69,70].

Besides causal SNPs and genes, GWASs can identify biological pathways and potential drug targets. GWAS summary statistics can also be used to estimate genetic correlations, SNP heritability, and polygenic scores (PGSs), as well as to research population history and structure and to perform Mendelian randomization studies [68,71].

The shortages of GWAS include rare variants whose carriers are sparse in population cohorts and thus unlikely to reach significance level in the results. Rare diseases are also challenging to research in a GWAS setting due to the lack of cases [71]. Additionally, most GWASs have been conducted in populations of European ancestry and might not replicate in other populations. Consequently, GWASs that include samples from other ancestries may produce novel genetic findings [72,73].

#### 2.1.4.3 Polygenic scores (PGSs)

The effect sizes of single SNPs calculated using GWAS are typically very small. However, when the effects of hundreds or thousands of SNPs are added together, the combined polygenic score (PGS) can add up to a significantly predictive sum. PGSs can be used to predict the risk of different diseases (often called polygenic risk scores, PRS) or to assess the overall polygenic component of a disease or trait [74].

Typically, the PGSs are a summation of SNPs, and their effect sizes (some positive and some negative) are used as weights in the calculation. Because of LD, several SNPs can be correlated with each other, which must be considered, or the results can be

distorted. Different algorithms use different approaches: some exclude extra SNPs in the same region, whereas others adjust the effect sizes of SNPs in LD. PGSs also have to be validated in another cohort than in which they were formed in [75,76]. In practice, PGSs can be used to identify high-risk individuals before the onset of the disease, thereby allowing the initiation of treatment or screening before harmful manifestations of the disease occur [74,75].

As mentioned, since GWASs have mostly been done with people of European origin and PGSs are based on GWAS results, they are not as accurate in non-European populations [77]. Also, common guidelines and standards for constructing and interpreting PGSs are missing [74,76].

### 2.1.5 Biobanks

Besides genetic variation databases, large biobank-based databases have emerged, helping even further discover genetic effects on different diseases and traits, both monogenic and multifactorial in origin. Biobanks are composed of biological samples that often include different types of genetic data (e.g., genotype and NGS). Phenotype data can include diverse questionnaires covering lifestyle factors, such as various measurements and laboratory values. Additionally, samples are often linked to health registries, thereby enabling longitudinal analyses of different diseases. One of the most utilized biobanks is the UK Biobank, which was initialized in 2006 and includes ~500 000 samples [78]. Other emerged biobanks include, for example, Estonian Biobank [79] and BioBank Japan Project [80].

In Finland, the FinnGen study was launched in 2017. It is a public-private partnership, including Finnish biobanks, universities, and pharmaceutical companies. The aim was to gather 500 000 samples, then genotype and link them to health registry data [81]. The FinnGen study enables, in particular, exploring the impact of Finnish-enriched variants that may not be included in cohorts of other populations.

### 2.1.6 Complexities of mutational analysis

Despite all existing databases and methods for investigating the genetic background of diseases, the interpretation of the results of genetic analyses can be difficult.

Although monogenic disorders refer to mutations in a single gene, the same phenotype can be caused by mutations in different genes, as is the case with FH. It applies to some other diseases as well, e.g., congenital deafness and hereditary spastic paraplegia. Moreover, the penetrance of monogenic mutations is often less than 100%. Penetrance refers to the proportion of individuals expressing the clinical phenotype of their genotype. For example, the penetrance of different mutations in the breast cancer gene *BRCA1* ranges from 24% to 71%. The mutation type can have an effect: truncating mutations can have greater penetrance and impact on disease severity than missense variants (e.g., cystic fibrosis) [82].

Furthermore, in the case of recessive diseases, if the two mutations are in different genes, the phenotype can differ significantly from that of mutations in the same gene (e.g., hereditary pancreatitis). The gene expression level of a gene can be affected by transcription factors and their variants. In diseases caused by repeat expansions, penetrance depends on the number of repeats (e.g., Huntington's disease). Moreover, there are cases in which a duplication of one gene can complement the loss-of-function of another gene, as in a rare case of spinal muscular atrophy. Additionally, some diseases exhibit sex-specific (e.g., hypertrophic cardiomyopathy caused by mutations in *MYBPC3* and *MYH7*) as well as age-related (e.g., breast cancer mutations) influences on penetrance or expression. Besides these factors, X-linked diseases often exhibit variable penetrance of mutations in female carriers [82].

Early population sequencing projects have discovered many presumably healthy individuals among cohorts carrying pathogenic disease mutations [82]. Chen et al. found 13 healthy adults with pathogenic variants in monogenic disease-associated genes (e.g., cystic fibrosis), attempting to identify resilient individuals. They speculated that the resilience could be explained by the effects of genetic modifiers [83].

## 2.2 Atherosclerotic cardiovascular disease (ASCVD) and hypercholesterolemia

Cardiovascular disease (CVD) is very common and the most common cause of death worldwide, including ASCVD, congenital heart diseases, cardiomyopathies, cardiac arrhythmias, and rheumatic heart disease. ASCVD includes coronary artery disease (CAD, also referred to as coronary heart disease, CHD), as well as cerebrovascular disease and peripheral artery disease (PAD) [84]. Ischemic heart disease (IHD), sometimes used congruently with CAD or CHD, refers to conditions resulting from reduced blood flow to the heart.

CAD is a multifactorial disease. Risk factors can be classified into non-modifiable and modifiable ones. Non-modifiable risk factors for CAD include age, sex, ethnicity, and family history of CAD. Modifiable risk factors include hypertension, hyperlipidemia, diabetes, obesity, smoking, poor diet, sedentary lifestyle, and stress [84]. CAD also has a strong genetic component. Families with prevalent early-onset CAD have been reported, besides hundreds of SNPs associated with CAD or myocardial infarction (MI) at a population level in GWASs. Heritability estimates calculated from twin studies are around 40-60% [85], and from GWASs of the same magnitude, 60-70% [86]. This thesis focuses on ASCVD, especially CAD, as well as hypercholesterolemia.

### 2.2.1 Epidemiology of cardiovascular disease

CVD accounts for approximately a third of all deaths worldwide [87]. As the most common CVD, CAD had an estimated worldwide prevalence of 315 million cases and 9.4 million deaths in 2022 [1,88]. The situation is similar in Finland, where CVD caused ~30% of all deaths in 2022. After CVD mortality was in a downward trend, there was a slight increase since 2021 by ~2% [89].

Males have a higher risk for CAD than females. For example, in the USA, the prevalence of CAD in adult males was 8% and 6% in females, respectively [90]. Likewise, in Europe (reported as IHD), the median age-standardized incidence in 2019 was 363:100 000 in males and 226:100 00 in females. CVD accounted for 45% of deaths in females and 39% in males, whereas IHD caused 17% of deaths in both sexes

[91]. The sex difference is partly explained by males having more CAD risk factors (e.g., smoking) than females. Moreover, female estrogen levels have a protective effect against atherosclerosis, yet the risk increases after menopause due to declining estrogen levels [92].

Globally, CVD mortality has been increasing for decades, which is largely explained by aging populations. For example, in Finland in 2020, the median ages of males and females who died of CAD were 79 years and 88 years [93]. In addition, certain CVD risk factors, such as hypertension and diabetes, tend to become more common with age. As populations age, CVD-related deaths are expected to continue to increase in the future [94].

Furthermore, socioeconomic factors (e.g., education, occupation, income) have been associated with CVD risk. More specifically, the prevalence of CVD risk factors, such as smoking, obesity, and diabetes, is associated with socioeconomic status (e.g., lower education and income) [95–98]. The association has also been directly observed with CVD prevalence and mortality [98–102]. Interestingly, the association with socioeconomic factors has been reported to be stronger in females [103].

### Modifiable risk factors for ASCVD

The most significant modifiable risk factors for ASCVD are smoking, hyperlipidemia, hypertension, and diabetes. Cigarette smokers have an increased risk for ASCVD, with a relative risk of 2.0 – 4.0, compared with never smokers. Additionally, even if only a single cigarette is smoked per day, the risk is still 40-50% of the risk of persons smoking 20 cigarettes per day. Molecules found in cigarette smoke promote atherosclerosis, and these atherosclerotic pathways appear to need only a little exposure to be enhanced [104].

The fact that LDL causes ASCVD has been confirmed by numerous studies using different approaches [4]. For instance, mutations in the LDL receptor gene (*LDLR*), which increase plasma LDL-C levels, also drastically increase ASCVD risk (discussed in more detail in chapter 2.2.4.1, page 25). The same is recognized in Mendelian randomization studies, in which variants affecting LDL-C levels also affect the risk for CAD [105,106]. Moreover, epidemiological studies have clearly shown an association between LDL-C and ASCVD risk [107,108], and medications that lower LDL-C decrease the risk for ASCVD as well [109–112]. Furthermore, the higher the cumulative effect of LDL-C, i.e., lifetime exposure to excess LDL-C, the higher the risk for ASCVD [5,113,114].

Hypertriglyceridemia and low HDL-C have also been associated with CAD risk, but their causality remains unclear. High LDL-C, low HDL-C, and high triglycerides often go together; consequently, their distinct effects are difficult to separate [106,115]. Additionally, elevated remnant cholesterol (plasma cholesterol without LDL-C and HDL-C) increases the risk for ASCVD by enhancing the progression of atherosclerosis. Remnant cholesterol is carried by lipoproteins that are rich in triglycerides: VLDL (very low-density lipoprotein) remnants and IDL (intermediate-density lipoprotein) particles, as well as chylomicron remnants [116,117]. Lipoprotein(a), Lp(a), is a cholesterol-containing LDL-like particle [118], and elevated Lp(a) levels have been associated with increased risk for CAD (up to OR 2.5) [119,120].

Hypertension is an independent risk factor for the development and progression of atherosclerosis. Hypertension also causes hypertensive heart disease (complications of prolonged hypertension), which is a risk factor for CAD [121]. Moreover, even with somewhat normal blood pressure levels, CVD-associated mortality increases as blood pressure increases. For example, among healthy 40-60-year-olds, a 20 mmHg (millimeters of mercury) increase in systolic blood pressure increased the risk for CAD-related death by 2-fold [122].

Diabetes is also a strong, independent risk factor, with about 2-fold excess risk for CVD: hazard ratio (HR) for CAD among patients with diabetes was 2.00, which was interestingly higher in females than males (2.6 vs. 1.9). Moreover, among patients with diabetes, the risk for coronary death was higher (HR 2.3) than that of non-fatal myocardial infarction (HR 1.8) [123].

Other factors increase ASCVD risk as well, such as unhealthy diet (e.g., excessive use of saturated fats or salt), physical inactivity, obesity, metabolic syndrome, and excessive alcohol consumption [1]. Diet can affect LDL-C levels directly, especially saturated fatty acids are associated with elevated LDL-C levels [15].

### 2.2.2 Atherosclerosis

Atherosclerosis refers to the buildup of plaques in the arteries. It can cause disease by narrowing the arteries, thereby reducing blood flow. The plaques can also rupture, leading to the formation of thrombi (blood clots). Atherosclerosis is a complex and slow process with many stages. LDL is a crucial part of the process, but atherosclerosis is viewed simultaneously as an inflammatory disease involving different immune cells and responses. These different immune-related pathways interact with other risk factors, thereby affecting plaque formation and progression in the arteries [3].

#### 2.2.2.1 Many stages of atherosclerosis

The progression of atherosclerosis has many stages (**Figure 4**). In the initial stage of atherosclerosis, cholesterol-containing lipoproteins (mostly LDL) accumulate in the arterial intima, forming a fatty streak. LDL particles can be modified in the arterial intima via oxidation, proteolysis, etc., to form atherogenic minimally modified LDL particles that are prone to be taken up by monocyte-macrophages. These processes stimulate immune responses. Monocytes differentiate into macrophages that take in lipids and then transform into foam cells, which are characteristic of atheromas, i.e., atherosclerotic plaques. In addition, plaques include T-cells, neutrophils, and mast cells.

As the plaque progresses, a necrotic core is formed. It consists of apoptotic and necrotic foam cells and smooth muscle cells. Lipid-rich material also keeps piling into the intima; even cholesterol crystals can be detected, and immune cells produce inflammatory molecules. Additionally, a fibrous cap is formed, and at this stage, the plaque is referred to as a fibroatheroma. There is also an angiogenic process ongoing in the plaque, which allows more immune cells to arrive, thereby creating more immune responses, but also hemorrhage inside the plaque. As a result, the necrotic core keeps growing, and the inflammation progressing. In the final stage, the plaque may calcify.

The plaque can rupture, meaning that the fibrotic core disrupts, revealing the core to the blood, which can cause thrombosis (formation of blood clots). It´s worth noting that not all plaques progress to the end stages, and progression can occur irregularly. Moreover, plaque rupture can be asymptomatic and heal itself [124,125].

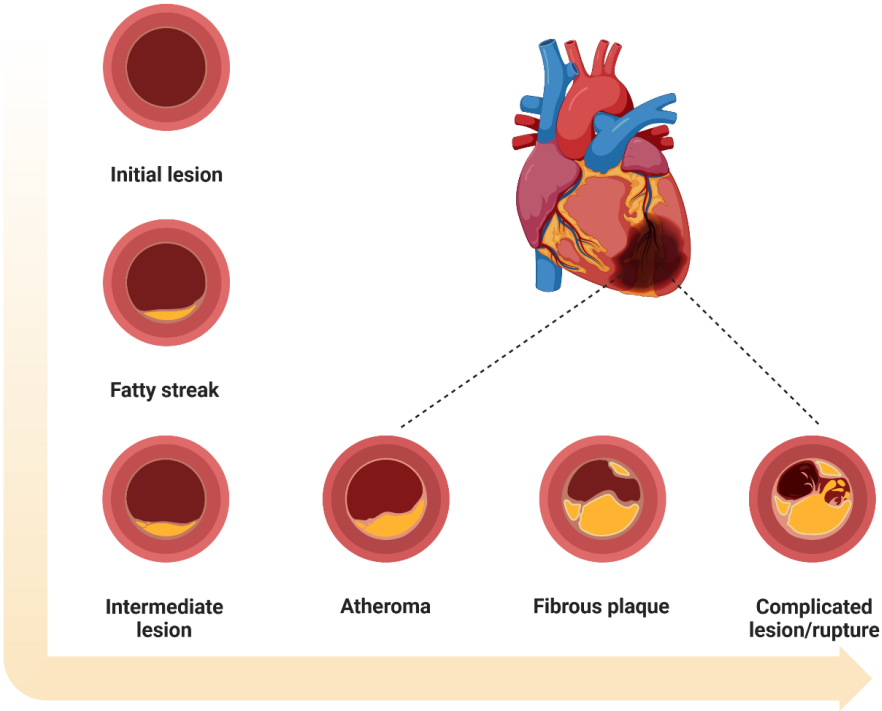



**Figure 4. Six stages of atherosclerosis.** First, an initial lesion forms. Second, cholesterol-containing lipoproteins accumulate and form a fatty streak. The third phase is called an intermediate lesion. Fourth, atheroma forms and its characteristics include foam cells. Fifth, a fibrous cap develops, which is a characteristic of a fibrous plaque. Finally, in a complicated lesion, rupture can occur. The last three stages can lead to ischemia, myocardial infarction and/or sudden cardiac death. Figure was created with Biorender.com.

#### 2.2.2.2 Symptoms and outcomes

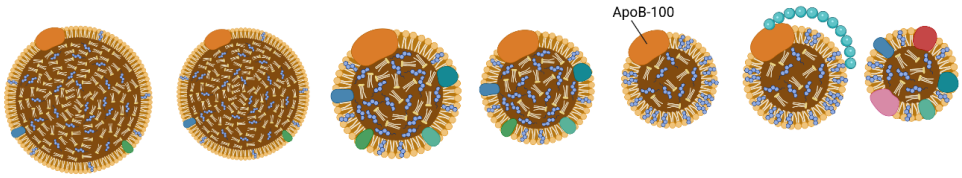
Atherosclerosis can occur in any part of the vasculature; the outcome depends on the location. As mentioned, atherosclerosis itself can cause reduced blood flow, ischemia, and endothelial dysfunction in the coronary arteries. The most common symptom is chest pain, which sometimes appears with physical exertion, and the pain may radiate to the upper limbs (usually to the left). This symptom is known as angina pectoris. Symptoms can also include difficulty breathing, nausea, and weakness. Acute disease is generally due to plaque rupture and thrombi. If this occurs in the coronary arteries, it usually causes acute coronary syndrome, in which blood flow to the heart is severely reduced or completely blocked. If the artery in question is situated in the brain, stroke can occur, and if the artery is located in the lower extremities, PAD can happen

[124,125]. The term MACE (major adverse cardiovascular events) is a composite outcome with varying definitions. Often, MACE includes non-fatal stroke, non-fatal MI, and cardiovascular death [126].

### 2.2.3 Lipoproteins and cholesterol metabolism

Hypercholesterolemia, the presence of excess cholesterol in the blood, is a significant risk factor for ASCVD, but cholesterol is also a vital component in cells. Cholesterol metabolism involves a vast number of enzymes and transporter proteins. This thesis focuses mainly on LDL.

#### 2.2.3.1 Lipoproteins




| Chylomicron | Chylomicron remnant | VLDL | VLDL remnant and IDL | LDL | Lp(a) | HDL |
|---|---|---|---|---|---|---|
| Size: 75-1200 nm | 30-80 nm | 30-80 nm | 25-35 nm | 18-25 nm | 30 nm | 5-12 nm |
| Major lipids: Triglycerides | Triglycerides, Cholesterol | Triglycerides | Triglycerides, Cholesterol | Cholesterol | Cholesterol | Cholesterol |
| Apolipoproteins: ApoB-48, ApoC, ApoE, ApoA-I, A-II, A-IV | ApoB-48, ApoE | ApoB-100, ApoE, ApoC | ApoB-100, ApoE, ApoC | ApoB-100 | ApoB-100, Apo (a) | ApoA-I, A-II, ApoC, ApoE |

**Figure 5. Lipoprotein classes.** Size in nanometers, main lipid content (triglycerides and/or cholesterol), and apolipoproteins are listed below the lipoprotein. Chylomicrons are the largest lipoproteins, and they are rich in triglycerides. They are secreted by the small intestine. Chylomicrons contain ApoB-48 and several other apolipoproteins. When LPL removes triglycerides from chylomicrons, smaller chylomicron remnants are formed. VLDL is produced in the liver and mainly contains triglycerides. Removal of triglycerides forms VLDL remnants and IDL. LDL can be formed from these particles. LDL contains primarily cholesterol and one ApoB-100 molecule. Lp(a) is an LDL particle containing apolipoprotein(a), attached to the ApoB-100. HDL contains cholesterol, phospholipids, and several different apolipoproteins. HDL is the only anti-atherogenic lipoprotein, while others promote atherogenesis. Apo: apolipoprotein, HDL: high-density lipoprotein, IDL: intermediate-density lipoprotein, LDL: low-density lipoprotein, Lp(a): lipoprotein(a), LPL: lipoprotein lipase, nm: nanometer, VLDL: very low-density lipoprotein. Figure created with Biorender.com.

Humans have several different lipoproteins in circulation (**Figure 5**), consisting of a hydrophilic layer of phospholipids, cholesterol, and different apolipoproteins, as well as a core of more hydrophobic lipids. Lipoproteins differ by their size, electric charge,

triglyceride, or cholesterol content, as well as by apolipoproteins on the surface of lipoprotein particles. They also have several subgroups of particles with different particle sizes. Lipoproteins include chylomicrons, chylomicron remnants, IDL (intermediate-density lipoprotein), VLDL (very low-density lipoprotein), and VLDL remnants, LDL (low-density lipoprotein), HDL (high-density lipoprotein) and lipoprotein(a), Lp(a).

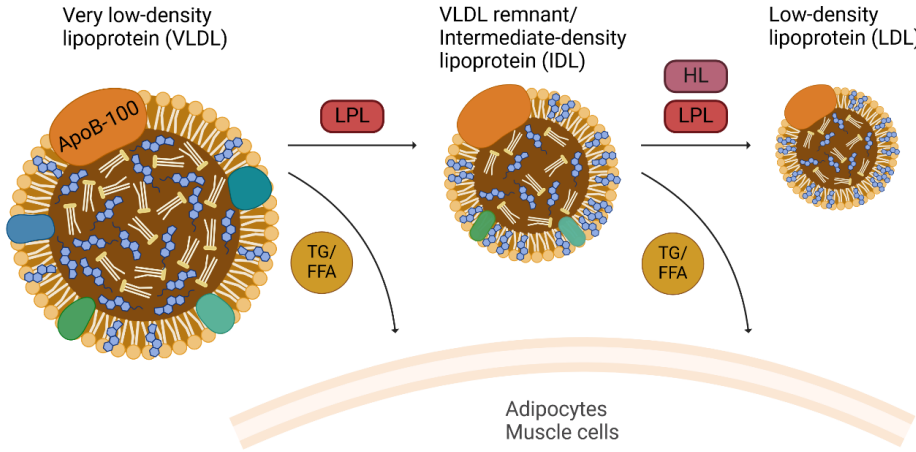



**Figure 6. Pathway of the formation of low-density lipoprotein (LDL).** Low-density lipoprotein (LDL) is formed in the circulation from VLDL particles secreted from the liver. First, VLDL particle triglycerides (TG) are hydrolyzed by lipoprotein lipase (LPL), and VLDL remnant/intermediate-density lipoprotein (IDL) is formed. Again, triglycerides are hydrolyzed via LPL and further via hepatic lipase (HL), and LDL is formed. Free fatty acids (FFA) after triglyceride hydrolysis by LPL are directed to adipose tissue and/or muscle cells. Figure created with Biorender.com.

The largest lipoproteins, chylomicrons, are responsible for transporting exogenous cholesterol and triglycerides (derived from food) from the gut to the liver and other tissues after modifying these particles in circulation mainly via lipoprotein lipase (LPL) function. However, most of the cholesterol transport (i.e., free cholesterol and cholesterol esters) is handled by LDL and HDL. LDL is formed in the circulation from liver-secreted VLDL particles, which are the target of LPL and hepatic lipase (HL) (**Figure 6**). LDL particle synthesis is tightly regulated in these LDL-HL-associated pathways. Each LDL particle contains one apolipoprotein B-100 (ApoB-100) in its surface layer. The primary function of LDL is to transport cholesterol to be used and converted to steroid hormones in adrenal cortical cells, and mainly back to the liver to be taken up via LDL receptor-mediated endocytosis. Under pathophysiological

conditions with elevated LDL-C levels, LDL accumulates in the subendothelial space of arteries and enhances atherosclerosis. On the other hand, HDL transports cholesterol from peripheral cells to the liver through reverse cholesterol transport (RCT), which is considered an anti-atherogenic function of HDL particles [118].

#### 2.2.3.2 Cholesterol biosynthesis

Cholesterol is an important component of cell membranes, while it is also needed to produce certain steroid hormones and bile acids. The primary source of cholesterol is biosynthesis in the liver, although some is absorbed from the intestine, derived from food. The synthesis of cholesterol is a complex process involving over 20 enzymes in the endogenous production of cholesterol. One of the most important steps is facilitated by the rate-limiting enzyme 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR), which controls the speed of the synthesis. The rate of cholesterol synthesis is regulated by many different factors, with the transcription factor sterol-regulatory element binding protein 2 (*SREBP2*) being the most important one. If there is too much cholesterol in the liver, the excess amount can be used and stored as cholesterol ester lipid droplets and to synthesize bile acid, which can be secreted to the gallbladder. The need for cholesterol regulates the expression of *LDLR* in the hepatocytes via a feedback mechanism [8].

#### 2.2.3.3 Formation of LDL and the LDL receptor pathway

LDL particles are generated in the circulation from VLDL particles that are secreted from the liver. VLDL particle triglycerides are hydrolyzed via LPL and further processed using HL, and the product from these reactions is LDL (**Figure 6**). The LDLR pathway is illustrated in **Figure 7.** LDL receptor (LDLR) resides on the cell membrane, and its primary function in hepatocytes is to remove LDL particles from circulation. LDLR binds to a specific domain of ApoB-100 of LDL particles, and the LDLR-LDL-complex moves to clathrin-coated pits that are then transferred to endosomes. In the endosome, LDLR and LDL are detached from each other due to a reduction in pH, after which the LDL particle is transported to a lysosome where ApoB-100 is hydrolyzed to amino acids and cholesterol esters cleaved to free cholesterol and fatty acids. LDLR

can be recycled back to the cell membrane multiple times or transferred to lysosome for degradation.

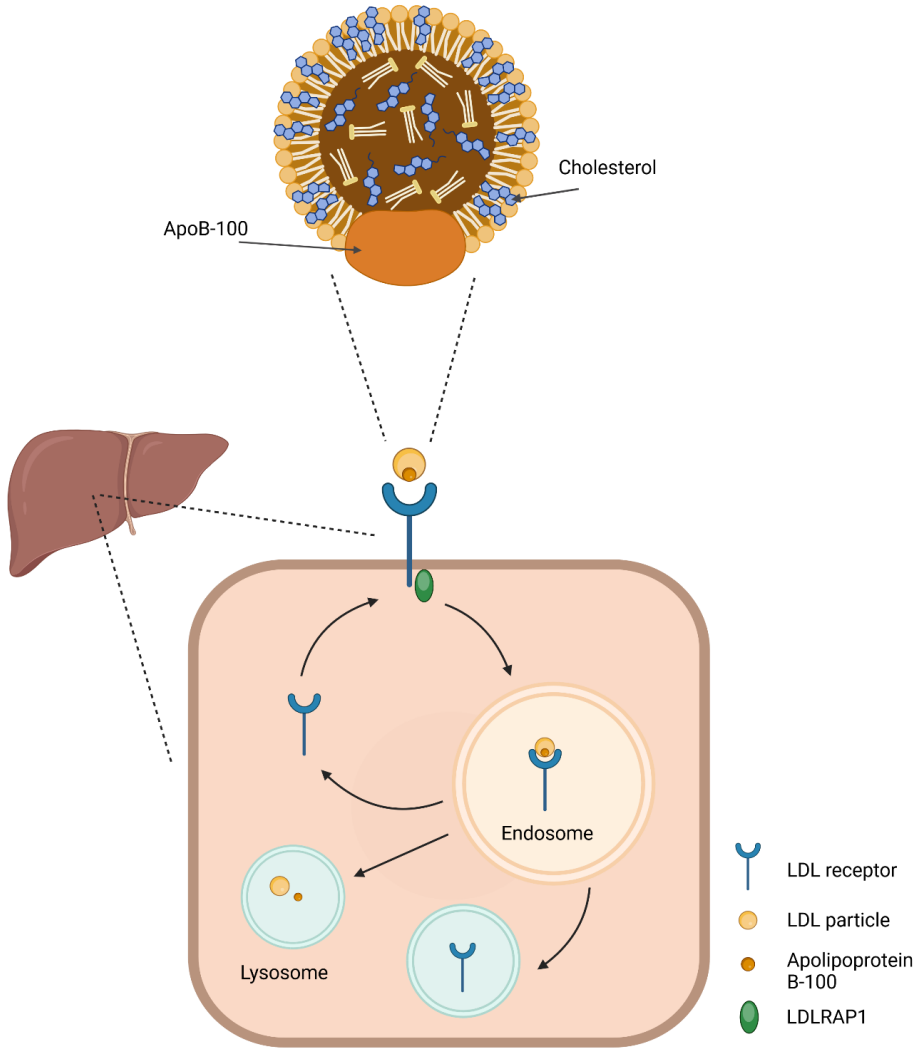



**Figure 7. Low-density lipoprotein (LDL) and the LDL receptor pathway.** On top, a low-density lipoprotein (LDL) particle containing high levels of cholesterol esters and a small amount of triglycerides in the core of the particle, cholesterol in its outer surface, as well as one apolipoprotein B-100 (ApoB-100). LDL particles are generated in the circulation via lipolytic enzymatic (LPL, HL) reactions from VLDL particles that are secreted from the liver. Below, the low-density lipoprotein receptor (LDLR) pathway in the liver: LDL receptor (LDLR) resides on the cell surface. LDLR binds to ApoB-100 on the surface of the LDL particle. Then, the LDL-LDLR complex is internalized in an endocytosis process, facilitated by low-density lipoprotein receptor adaptor protein 1 (LDLRAP1), and transported to an endosome, where LDL and LDLR are detached from each other. LDL (containing cholesterol esters, ApoB-100) is transported to a lysosome where lysosomal hydrolases degrade ApoB-100 protein to amino acids, and cholesterol esters are hydrolyzed to cholesterol and fatty acids. LDLR can be transported back to the cell surface or degraded in lysosomes. Figure created with Biorender.com

Transcription of the *LDLR* gene is regulated most importantly by *SREBP2* [127]. This transcription factor is located in the endoplasmic reticulum (ER) of hepatocytes as an inactive precursor form which, upon transport to the Golgi apparatus, is cleaved in tandem, allowing the N-terminal fragment to enter the nucleus and upregulate transcription of *LDLR* and other genes in the cholesterol pathway. Several other molecules are involved in the LDL metabolism. For example, PCSK9 (proprotein convertase subtilisin/kexin type 9) assists in the degradation of LDLR, and LDLRAP1 (low-density lipoprotein receptor adaptor protein 1) is involved in the endocytosis of LDLR [128].

#### 2.2.3.4 Lipid-lowering medications

The most prescribed group of lipid-lowering medications is statins. Statins inhibit HMGCR (the rate-limiting enzyme in cholesterol biosynthesis), leading to reduced cholesterol synthesis in the liver and subsequent upregulation of *LDLR* expression and transport of LDL receptors to the plasma membrane, thus increasing LDL uptake from the plasma. There are several statins: simvastatin, atorvastatin, and rosuvastatin, for instance. Among all statins, the mean decrease of LDL-C has been 35-55%, regardless of the LDL-C level before the beginning of the medication [129]. Moreover, a 1 mmol/l reduction in LDL-C, achieved by statins, corresponded to a 25% decrease in major CVD events (e.g., MI) per year of statin use [110]. Additionally, on average, the reduction of LDL-C has been 1.1 mmol/l, with a 2–6-year follow-up time. The incidence of major CAD events has decreased by approximately 20% per 1 mmol/l decrease in LDL-C [130]. In addition to reducing LDL-C levels, statins may have other effects on CAD risk, such as anti-inflammatory effects [3]. Regarding Lp(a), statins have been reported to increase Lp(a) levels by increasing *LPA* expression in the liver [131].

In addition to statins, other lipid-lowering medications exist, such as bile acid sequestrants and fibrates. Ezetimibe inhibits cholesterol absorption from the intestine and can be used in addition to statins [132]. Newer medications include PCSK9 inhibitors (evolocumab and alirocumab). PCSK9 promotes LDLR degradation; thus, inhibiting PCSK9 enhances LDL uptake. PCSK9 inhibitors can decrease LDL-C by up to 60% [133].

Recent studies have identified two relevant proteins, angiopoietin-like 3 protein (ANGPTL3) and apolipoprotein C-III (ApoC-III), as lipid-lowering drug targets. Evinacumab is a monoclonal antibody that inhibits ANGPTL3, whereas ANGPTL3 inhibits LPL, increasing lipid levels and CVD risk. In patients with homozygous FH already using lipid-lowering medication, evinacumab has further decreased their LDL-C levels by over 40%. Additionally, levels of triglycerides and Lp(a) have decreased significantly [134,135]. ApoC-III is a component of triglyceride-rich lipoproteins (e.g., VLDL), and its elevated plasma levels have been associated with increased CVD risk. Volanesorsen and olezarsen are antisense oligonucleotide-based drugs that target *APOC3* mRNA, lowering triglycerides by up to 77%, hereby decreasing CVD risk [136–138].

### 2.2.4 Monogenic risk factors for hypercholesterolemia and ASCVD

Although ASCVD is a multifactorial disease, ASCVD risk is particularly high among persons with monogenic mutations causing severe hypercholesterolemia.

Dyslipidemias are conditions in which the amount of lipids in the blood, for example, cholesterol, triglycerides, or Lp(a), is abnormal. The most common genetic dyslipidemia causing high LDL-C levels is familial hypercholesterolemia (FH). Other genetic dyslipidemias associated with elevated LDL-C levels are either very rare or also associated with abnormal levels of other lipids, such as triglycerides [139].

#### 2.2.4.1 Familial hypercholesterolemia (FH)

FH is a genetic condition caused by dominant mutations. Thus, patients with heterozygous FH carry one pathogenic allele, whereas patients with homozygous FH carry two pathogenic alleles. Because homozygous FH is very rare, most of this chapter refers to heterozygous FH.

**History of FH**

As early as the 1920s, there were reports of patients with symptoms including tendon xanthoma (depositions of cholesterol) and hypercholesterolemia. Symptoms often were also associated with CVD events, including sudden death. Moreover, the syndrome was suspected to be hereditary [6]. Later, the syndrome, called essential

familial hypercholesterolemia, was confirmed to be inherited as an autosomal dominant disease with incomplete penetrance [140]. Next, the cholesterol metabolic pathway, including LDL, was characterized, and the LDL receptor was identified. Defects in LDLR were identified as culprits behind the FH phenotype [141].

Intriguingly, studies tracing FH mutations in the pedigrees have discovered that there has not necessarily been increased mortality in the mutation carriers. For example, in a large Dutch pedigree, all-cause mortality was lower among FH mutation carriers than in the general population before 1915 [142]. Additionally, studies using LDLR knockout mice have shown that hypercholesterolemia may be protective against some bacterial infections and their toxic effects [143]. Furthermore, different lipoproteins (e.g., LDL and HDL) can bind to and thus inactivate bacterial lipopolysaccharides (LPS) on the membrane of Gram-negative bacteria, protecting against LPS-induced, harmful immune reactions [144]. Therefore, the FH phenotype could have been beneficial in human evolution and could thus be relatively common as a genetic disease.

### Diagnosing FH

There are several diagnostic scoring systems for clinical FH (**Table 1**), such as The Dutch Lipid Clinic Network (DLCN), Simon Broome, and MED-PED (Make early diagnosis - prevent early death) criteria that differ somewhat. All scoring systems include hypercholesterolemia (LDL-C > 4.9 mmol/l), and additional points come from tendon xanthomas or arcus cornealis (cholesterol depositions in the cornea), family history of premature CAD or hypercholesterolemia and a known genetic mutation of FH [7]. The scoring systems produce a diagnosis of definite, probable, possible, or unlikely FH.

**Table 1. Diagnostic scoring systems for clinical familial hypercholesterolemia**. The Dutch Lipid Clinic Network (DLCN)- criteria calculates points based on LDL-C levels, clinical history and physical examination of patients, family history of certain conditions, and carrying a pathogenic mutation in FH-associated genes. The points determine the diagnosis: definite, probable, possible, or unlikely FH. The Simon Broome- criteria is similar to DLCN but definite, and possible diagnosis requires LDL-C level above 4.9 mmol/l. The MED-PED (Make early diagnosis - prevent early death) criteria is the simplest, and it is based solely on LDL-C level. LDL-C cutoff values are based on age and whether the patient has a family member with diagnosed FH. Adapted from Nordestgaard et al., and McGowan et al. **[21,145]**, ASCVD: atherosclerotic cardiovascular disease, CAD: coronary artery disease, MI: myocardial infarction.

| Criteria | DLCN | Simon Broome | MED-PED |
|---|---|---|---|
| **LDL-C (mmol/l)** | >8.5 (8 points)<br>6.5-8.4 (5 p.)<br>5.0-6.4 (3 p.)<br>4.0-4.9 (1 p.) | >4.9 | >5.7-9.3 (depending on family history of FH and age of the patient) |
| **Clinical history, physical examination** | Premature ASCVD (1-3 p.)<br>Tendon xanthoma (6 p.)<br>Arcus cornealis (4 p.) | Tendon xanthoma | - |
| **Family history** | Premature CAD<br>Hypercholesterolemia<br>Tendon xanthoma, arcus cornealis<br>Child with hypercholesterolemia (1-2 p.) | Tendon xanthoma<br>Premature MI<br>LDL > 7.5 mmol/l | Diagnosed FH |
| **Genetics** | Mutation in *LDLR*, *APOB*, or *PCSK9* (8 p.) | Mutation in *LDLR*, *APOB*, or *PCSK9* | - |
| **Diagnosis** | Definite: >8 points<br>Probable: 6-8 p.<br>Possible: 3-5 p.<br>Unlikely: 0-2 p. | Definite: LDL-C + xanthoma or mutation<br>Possible: LDL-C + family history of premature MI or hypercholesterolemia | Based on LDL-C |

### Prevalence and ASCVD risk

Prevalence estimates of FH vary between countries, not the least because of the lack of FH registries in most countries: about 90% of the countries worldwide do not have a prevalence estimate for FH [146]. Moreover, different countries use different scoring systems that affect diagnosis counts. In any case, distinct studies and meta-analyses estimate that the mean prevalence of heterozygous FH is about 1:250 – 1:330 [19,146–148]. The prevalence of the homozygous FH patients is about 1: 1 000 000. However, most individuals with FH do not have a diagnosis. For instance, comparing the prevalence estimate to the number of diagnosed cases, <1% of estimated FH cases had a diagnosis (e.g., pure hypercholesterolemia) in most countries [20,21]. Moreover, not all mutation carriers fulfill the diagnosis criteria. For instance, the carriers of known Danish FH mutations would have received a result of definite or probable FH in only 0

– 23% of occasions, depending on the diagnosis criterium [149]. Carriers might not have high enough LDL-C levels due to a younger age [148] or reduced penetrance of the mutation [19].

As hypercholesterolemia is the main finding of FH, the risk for ASCVD among FH patients is significantly elevated. Indeed, in cases with ASCVD, the prevalence of heterozygous FH is 3.2 – 6.7%, i.e., significantly increased [146,148]. Moreover, the OR (odds ratio) for CAD in FH mutation carriers has been estimated as 1.7 – 3.3 and 3.0 – 3.7 for premature CAD, with the OR being the highest among *LDLR* mutation carriers [19,149,150]. Furthermore, it has been shown that CAD risk is increased among FH mutation carriers independently of measured LDL-C levels. FH mutation carriers are likely to have higher LDL-C levels already from childhood, leading to a higher lifetime cumulative LDL-C exposure [150]. CAD risk is also increased compared to individuals with hypercholesterolemia caused by a polygenic background [151]. Thus, it is important to identify and treat persons carrying FH mutations.

### Treatment of FH

As with hypercholesterolemia in general, the most important treatment of FH is statin medication, while other options exist for most severe cases, such as homozygous FH. For instance, novel ANGPTL3 inhibitors can decrease LDL-C by an additional 49% in homozygous FH patients already using other lipid-lowering medication [152]. ASCVD mortality in FH patients has decreased over time, possibly due to the introduction of statins [153,154]. Moreover, when FH patients are treated with statins, their risk for MI can be lowered by up to 76%, decreasing their risk to the same level as an average person [155]. Despite the importance of lipid-lowering medication among FH patients, about half of them remain untreated [19–21]. Additionally, a large proportion of the patients do not reach the target level of LDL-C [21,156], commonly at least < 2.6 mmol/l, which is achieved most efficiently with more effective medication [157]. Specifically, PCSK9 inhibitors, the most potent medication, are not available in all countries, and reimbursement practices vary [20]. Furthermore, especially women are treated less-efficiently [20,153].

#### 2.2.4.2 Genetics of familial hypercholesterolemia

Mutations in at least four genes (*LDLR*, *APOB*, *PCSK9*, and *LDLRAP1*) can cause FH. *LDLR* mutations are the most common cause of FH, and *LDLR* harbors the highest number of different FH mutations. However, the proportions of distinct mutations and genes vary by population. Phenotype, i.e., the severity of the disease, varies considerably by the specific mutations and mutation types. Although, the effect of a specific mutation is difficult to predict, e.g., because of variable penetrance [9,158]. Moreover, not all patients with clinical FH carry any known FH mutations. For instance, Saadatagah et al. reported finding causative mutations for only 26% of patients with clinical FH [12]. Consequently, not all FH mutations are probably known, but it is also possible that not all hypercholesterolemia-associated genes are known either [159]. Nonetheless, there appears to be a poor overlap between clinical FH and FH mutation carriers.

##### Low density lipoprotein receptor (*LDLR*)

Low density lipoprotein receptor (*LDLR*) was the first gene identified to be associated with FH [141,160]. Mutations in *LDLR* are the most common cause of FH with the highest number of identified pathogenic mutations: there are over 2200 variants in the UCL *LDLR* variant database, of which ~1500 are evaluated as likely pathogenic [161]. Pathogenic mutations include point mutations as well as small and large insertions and deletions. Loss-of-function mutations, e.g., large deletions, typically cause a more severe phenotype [162].

*LDLR* is a 45 kbp (kilo base pair) gene in chromosome 19 with 18 exons. The protein consists of 860 amino acids and has five domains [8,163,164]. The LDL receptor is synthesized in the endoplasmic reticulum of hepatocytes, modified in the Golgi apparatus, and then transported to the cell membrane [128]. On the membrane, LDLR binds LDL particle from plasma and facilitates its uptake to endosomes via endocytosis process. Then, LDLR either goes to lysosomes for degradation or is recycled back to the cell membrane (**Figure 7**, page 23).

Pathogenic mutations can occur in any of the *LDLR* exons. FH mutations can be divided into five groups (**Figure 8**) depending on the impact of the mutation. Class 1 mutations are null mutations where no functional LDLR protein is produced. Class 2 mutations are caused by a defect in transporting the receptor to the cell membrane. In class 3 mutations, the defect is the failure to bind the LDL particle properly. In class 4, the LDL receptor is able to bind the LDL particle, but the mutation affects the internalization of the LDLR-LDL complex. In class 5 mutations, the LDL receptor is not correctly recycled further. Class 1 mutations cause the most severe effect, and classes 3-5 are likely to cause a milder phenotype [163,165].

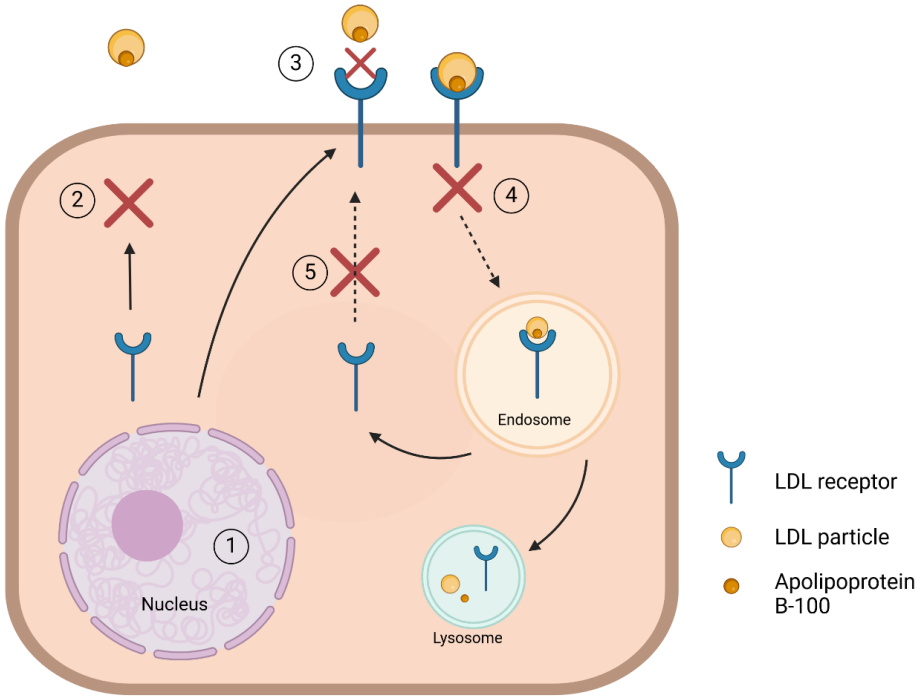



**Figure 8. Different types of FH-causing *LDLR* mutations.** Mutations in *LDLR* can be divided into five classes: Class 1 includes null mutations, where no functional protein is produced. In class 2, the LDL receptor is produced, but it is not transported to the cell membrane. Class 3 mutations inhibit LDLR from binding with the LDL particle. In class 4 mutations, LDLR binds with LDL, but there is a defect in the internalization of the LDLR-LDL complex. In class 5 mutations, the recycling of LDLR is impaired. Figure created with Biorender.com.

### Apolipoprotein B (*APOB*)

Mutations in the *APOB* gene can cause hypercholesterolemia, also called familial defective apolipoprotein B-100. Most FH cases caused by *APOB* mutations carry a mutation residing in a hotspot region of the gene around amino acid 3527.

*APOB* is a 43 kbp gene in chromosome 2 with 29 exons. Two protein products are produced from *APOB*: the smaller ApoB-48, which is synthesized in the small intestine (part of chylomicron production), and ApoB-100, which is synthesized in the liver. ApoB-100 is an essential component of lipid particles LDL, VLDL, IDL, and Lp(a). In the lipid particles, ApoB-100 can bind to LDLR. The binding site of LDLR is in the $\beta_2$-domain of ApoB-100, which is not included in the smaller ApoB-48-protein [8,166].

Mutations in *APOB* can cause symptoms similar to those of defects in *LDLR*, including increased ASCVD risk [166]. The most common FH-causing *APOB* mutation in Europe is Arg3527Gln (earlier Arg3500Gln). The amino acid 3527 resides in the $\beta_2$-domain, and the substitution disrupts the binding to the LDLR [167]. All carriers of this European mutation share ancestry, and the mutation is roughly 6000-7000 years old [168]. Furthermore, other hypercholesterolemia-mutations of the same amino acid exist, for example, Arg3527Trp in the Chinese population [169]. Conversely, truncating mutations in *APOB* can cause hypocholesterolemia [166].

### Proprotein convertase subtilisin/kexin type 9 (*PCSK9*)

The third gene found to be associated with dominant FH was *PCSK9* [170], although hypercholesterolemia-causing *PCSK9* mutations have been reported to be much rarer than mutations in *LDLR* or *APOB*.

*PCSK9* resides in chromosome 1, containing 3 637 bp and 12 exons [171]. The synthesis of PCSK9 takes place in the liver. The precursor polypeptide is 692 amino acids. The prodomain is cleaved off, and then the mature, functional protein is complete. PCSK9 binds to the LDLR-LDL complex and aids the degradation of LDLR, thus decreasing the amount of LDL receptors and increasing LDL-C in the plasma. The binding site of the LDLR-LDL complex is located in LDLR in the first EGF domain and in the catalytic domain of PCSK9 [172].

Several functional studies indicate that the mechanism of hypercholesterolemia-causing mutations in *PCSK9* is gain-of-function [173]. Moreover, hypercholesterolemia caused by these mutations appears to respond poorly to statin treatment, at least in some cases [172]. While hypercholesterolemia is caused by gain-of-function

mutations, loss-of-function mutations of *PCSK9* can cause decreased LDL-C. Some of the gain-of-function mutations are relatively common in the general population, and they seem to affect the processing of PCSK9. This finding obviously made PCSK9 a compelling drug candidate [172], leading to the development of PCSK9 inhibitors.

### Low density lipoprotein receptor adaptor protein 1 (*LDLRAP1*)

Mutations in *LDLRAP1* can cause autosomal recessive hypercholesterolemia (ARH) [174]. *LDLRAP1* is a 25 kbp gene in chromosome 1, including 9 exons. The protein product consists of 308 amino acids. LDLRAP1 is a chaperone protein that assists the internalization of LDLR-LDL complex in hepatocytes. The binding occurs between the PTB domain of LDLRAP1 and the cytoplasmic tail of LDLR [175].

Originally, recessive loss-of-function mutations in *LDLRAP1* were associated with hypercholesterolemia phenotype, including xanthomas and premature CAD. The first families carrying these mutations were identified in Sardinia and Lebanon [174]. Later, *in silico* analyses have suggested that missense mutations could also affect the LDLRAP1 protein, possibly affecting the binding of LDLRAP1 to the LDLR-LDL complex [176].

### Homozygous FH

The prevalence of homozygous FH is very low: the traditional estimate has been 1: 1 000 000, which was based on the previous prevalence estimate of heterozygous FH. Since the prevalence estimates have changed, the new estimates for homozygous FH range from 1: 170 000 – 630 000 [177,178].

The patient can be homozygous for the same mutation (usually in *LDLR*), or compound heterozygous for two different mutations, or the mutations can be in different FH-associated genes. LDLR function in homozygous FH patients can be entirely absent. Therefore, symptoms of homozygous FH are similar to heterozygous cases but more severe: LDL-C levels are often above 10 mmol/l, tendon xanthomas are more likely, and ASCVD can occur a lot earlier and be more severe than in heterozygous cases [177,178]. However, depending on the exact mutations and their impact on LDL metabolism, the severity of the disease varies.

Because LDLR is usually not functional, standard lipid-lowering treatments may not work. Besides different lipid-lowering medications, e.g., statins, ezetimibe, and PCSK9 inhibitors, lipoprotein apheresis can be used, which removes LDL directly from the plasma. These treatment options have significantly improved the life expectancy

of homozygous FH patients [177,178]. If the treatments do not decrease LDL-C levels enough, new lipid-lowering therapies, such as ANGPTL3-inhibiting medication, should be considered [179,180].

#### 2.2.4.3 Familial hypercholesterolemia in Finland

The spectrum of FH mutations in Finland is different from other countries. Seven Finnish-enriched mutations in *LDLR* (**Table 2**) have been reported, explaining about 80% of the Finnish FH cases.

**Table 2. Finnish-enriched mutations causing familial hypercholesterolemia.** The most common seven Finnish FH mutations in *LDLR* (FH-Fin 11 and FH-Fin 12 are very rare). Position refers to the genomic position in chromosome 19. *Prevalence estimates from Lahtinen et al. [11], NF: carriers of these mutations were not found in the study. CRCh38: Genome Reference Consortium Human Build 38. Modified from the original publication (I) under the terms of the Creative Commons Attribution License (CC BY 4.0).

| Name | Position (build GRCh38) | Mutation | Frequency in Finland* |
|---|---|---|---|
| FH-Helsinki | 19:11123595-1132129 | g.11234271_11242805del8535 | 0.00043 |
| FH-Pogosta | 19:11116937 | c.1784G>A (p.Arg595Gln) | 0.00036 |
| FH-North Karelia | 19:11107499-1107505 | c.925_931delCCCATCA (p.Pro309Lysfs) | 0.00028 |
| FH-Turku | 19:11129654 | c.2531G>A (p.Gly844Asp) | NF |
| FH-Pori | 19:11113293 | c.1202T>A (p.Leu401His) | NF |
| FH-Fin 11/Cincinnati-4 | 19:11116888 | c.1735G>A (p.Asp579Asn) | NF |
| FH-Fin 12/Avellino | 19:11110767 | c.1056C>G (p.Cys352Trp) | NF |

The first FH mutation found in Finland was named FH-Helsinki: it is a large deletion, where exons 16 and 17 and a part of exon 18 are missing from the final receptor protein [181,182]. Later, the FH-Helsinki mutation was estimated to cause about 30% of the Finnish FH cases [10]. Next, a 7 bp deletion, FH-North Karelia (Pro309Lysfs) was identified. It causes a frameshift and a truncated protein. Interestingly, the prevalence of this mutation has significant regional variation, and it is most common in North Karelia of Eastern Finland [183,184].

Additionally, 24 mutations in *LDLR* have been reported, but most of them are rare and family-specific [10]. Besides the two deletions mentioned above, three missense mutations are enriched in Finland: FH-Pogosta (Arg595Gln) is the second most common cause of FH [10,11]. FH-Turku (Gly844Asp) and FH-Pori (Leu401His) are relatively common in the Southwestern part of Finland but account for only about 5%

of Finnish FH cases [10,185]. As mentioned, these five mutations, in addition to rarer FH-Fin 11 and FH-Fin 12, are thought to cause about 80% of FH cases, while the rest of the known FH mutations cover only a minor proportion [10,11].

The European *APOB* mutation, mentioned before, has not been found in the Finnish population despite many efforts [186,187]. Moreover, no pathogenic mutations in *PCSK9* or *LDLRAP1* have been published [11].

A study by Lahtinen et al. estimated the prevalence of FH in Finland to be about 1:600, based on the prevalence of the five Finnish mutations [11]. According to this estimate, FH is much rarer in Finland than in many other countries, i.e., 1:250 – 1:330.

#### 2.2.4.4 Other monogenic risk factors for coronary artery disease

Besides dyslipidemias, there have been reports of families with a high prevalence of early-onset CAD. An extended family was reported with prevalent early-onset MI. After whole-exome sequencing, two mutations were identified in two different genes: *GUCY1A3* and *CCT7.* These genes have been associated with nitric oxide (NO) signaling that affects thrombocytes and, thus, possibly the risk for MI [188]. In another family, there was also an excess prevalence of early-onset CAD. In this case, a mutation in *PDE5A* was co-segregating with the disease. Interestingly, *PDE5A* is also associated with NO signaling [189]. However, monogenic mutations beyond hypercholesterolemia-causing ones appear to be very rarely behind early-onset ASCVD.

### 2.2.5 Polygenic risk factors for hypercholesterolemia and ASCVD

At the population level, most of the genetic burden of hypercholesterolemia and ASCVD is polygenic, i.e., common variants with small effect sizes, and the risk is modest compared to the monogenic causes. The effect of a variant can be decreasing (or protective), neutral, or increasing.

#### 2.2.5.1 Polygenic risk factors for hypercholesterolemia

Numerous lipid-GWASs (including LDL-C) have been performed, and hundreds of loci have been discovered. In a recent GWAS, total number of causal variants was ~250 [190]. Top loci in GWASs on LDL-C measurement typically include genes known to be involved in lipid metabolism, such as *LDLR*, *APOB*, and *APOE*. Several LDL-PGSs have been calculated and published, produced with different methods and SNPs. LDL-C variance explained by the PGSs has been up to ~20% [190,191]. Moreover, persons with LDL-PGS in the top 5% have approximately 0.78 mmol/l (30 mg/dl) higher LDL-C levels [75]. Consequently, severe hypercholesterolemia can be caused by the clustering of LDL-C increasing SNPs [13].

**Monogenic and polygenic hypercholesterolemia**

In a study of the adult population with severe hypercholesterolemia, 1.5% carried a monogenic FH mutation, and 15.6% had LDL-PGS in the top 10%, i.e., most cases could not be genetically explained [12]. In turn, in samples of patients with clinical FH, polygenic clustering has explained up to ~50% of cases without monogenic mutations [14,192]. Despite the variable genetic background, several studies have shown that persons carrying monogenic FH mutations have a higher ASCVD risk than those with polygenic hypercholesterolemia with corresponding LDL-C levels. For example, HR for ASCVD in carriers of monogenic mutations was 1.93, and 1.26 in persons with LDL-PGS in the top 5% of the population distribution [151]. Compared to polygenic hypercholesterolemia, the severity of monogenic FH has been thought to result from higher cumulative lifetime exposure to high cholesterol levels [192]. However, the effects of polygenic variants might come from pathways beyond LDLR.

#### 2.2.5.2 Polygenic risk factors for ASCVD

As with lipids, many GWASs on different CVDs have been performed. For instance, in a recent GWAS, over 200 significant loci were associated with CAD [193]. An integral part of polygenic risk factors for CAD relates to lipid metabolism, accounting for at least 20% of loci [194]. Other GWAS loci are associated with, e.g., vascular regulation, blood pressure, inflammation, and the function of smooth muscle cells, although the function of many loci is unknown [195]. Many different PRSs for different CVDs have been constructed. For instance, the top 5% of CAD-PRS has been associated with approximately three times greater odds for CAD (1.4-1.6-fold risk per SD) [75].

Interestingly, since there are sex differences in the prevalence and manifestations of CAD, there are some findings in the sex chromosomes, which are usually excluded from GWASs. Of male Y chromosome haplotypes, one (haplogroup I1) out of seven has been associated with CAD [196]. Moreover, a recent GWAS discovered nine novel, significant loci from the X chromosome [197].

##### The combined effect of monogenic mutations and polygenic variants

When carriers of monogenic FH mutations have an increased ASCVD risk, CAD-associated SNPs, in turn, include variants not involved in lipid metabolism. Thus, polygenic SNPs can significantly increase ASCVD risk in carriers of FH mutations. While FH mutation carriers had an average 3.2-fold risk for CAD, mutation carriers with CAD-PRS in the lowest quintile had only a 1.3-fold CAD risk. However, mutation carriers with CAD-PRS in the highest quintile had a striking 12.6-fold CAD risk [194]. Also, LDL-PGS was associated with ASCVD risk among FH mutation carriers. The highest ASCVD risk was among mutation carriers, who also had LDL-PGS in the top 20%: HR 3.06 (compared to 1.96 among mutation carriers without a high LDL-PGS) [14].

## 2.2.6 How to estimate personal ASCVD risk?

### 2.2.6.1 ASCVD risk algorithms

There are several validated algorithms to estimate personal ASCVD risk, usually in the following 10 years. The first one was the Framingham Risk Score (FRS), developed already in 1967 (but since improved) [198,199]. Additional algorithms include, for example, SCORE [200] and QRISK [201,202]. SCORE is commonly used in many European countries, and QRISK is used in the UK. ASCVD risk algorithms, in general, usually include age, sex, systolic blood pressure, and smoking. Other commonly used variables include different cholesterol measurements and diabetes diagnosis. Moreover, family history of CAD is included in the QRISK, for instance. In different risk models, the weights of the variables are usually calculated slightly differently, and the estimated endpoints also vary a little (any ASCVD or a specific outcome). It is worth noting that the scores are estimated based on populations of European ancestry and may not be as accurate in other populations [22]. Furthermore, it has been suspected that traditional risk models underestimate ASCVD risk in females, and thus, females at high risk are not necessarily recognized in the healthcare system. Additionally, some ASCVD risk factors are female-specific, such as gestational diabetes [203].

#### Does polygenic risk information improve ASCVD risk models?

Even though the risk of developing ASCVD has an evident genetic background, commonly used risk models do not include any genetic information except the family history of ASCVD in some models. Still, family history does not purely represent genetic background because families can share behavioral risk factors (e.g., dietary habits) [22,25]. Using genetic risk information in risk models is appealing because it does not change as lifestyle factors do, thus enabling the early identification of high-risk persons. After the development of PGSs, studies have been made to investigate the value of adding PGSs to ASCVD risk estimates. However, the added value of PGSs has been modest, ~4% reclassification improvement [204].

Again, with CAD-PRS, it is possible to identify persons who, in addition to having higher genetic ASCVD risk, would benefit more from statin therapy. In persons with high CAD-PRS, CAD risk decreased up to 46% when using statins, compared to 26% among all the others. The amount of LDL-C change does not explain the difference and

possibly implies subclinical atherosclerosis [23,24]. Additionally, since traditional risk models have usually been calibrated with middle-aged people and age is a key variable, younger persons tend to get low risk estimates. Because traditional and genetic risk factors are independent of each other, adding CAD-PRS to risk models allows the identification of younger individuals at high risk [25].

### 2.2.7 Complex dynamics of risk factors for hypercholesterolemia and ASCVD

As described above, the monogenic mutations causing FH can significantly elevate ASCVD risk, making the identification of mutation carriers and their early treatment essential. However, accumulating evidence shows that polygenic variants may influence the penetrance and impact of these mutations. Consequently, it is difficult to separate the phenotype of monogenic FH from polygenic hypercholesterolemia. Moreover, the precise effects and ASCVD risks of different FH mutations vary considerably.

Regarding CAD, lipids and factors affecting lipids are only a part of the complexity of the disease. Other factors, such as inflammation and blood pressure, are at play, and they are again affected by different genetic and environmental factors. As a result, the individual CAD risk is a complicated interplay of various genetic and lifestyle factors that is difficult to dissect.

In addition to genetic factors affecting the phenotype, lifestyle factors also have an impact. Among FH patients, many traditional risk factors for ASCVD, such as smoking, are independent risk factors for the development of ASCVD [205,206]. FH mutation carriers with less healthy lifestyles did indeed have a higher risk for CAD [207]. Since FH primarily causes hypercholesterolemia, other risk factors can increase the risk for CAD through mechanisms beyond lipid metabolism [208].

Khera et al. studied the connection between polygenic risk for CAD and several independent lifestyle factors (smoking, obesity, exercise, and diet). For example, persons who had high CAD-PRS and healthy lifestyles had 1.90 HR for CAD, compared to 3.50 in persons having unfavorable lifestyles. Thus, even with elevated genetic risk, lifestyle modification can lower ASCVD risk significantly [209].

## 2.3 Behavioral factors and disease prevention

Although behavioral factors can significantly affect hypercholesterolemia and ASCVD risk, lifestyle modifications are often difficult to accomplish. ASCVD risk is commonly estimated and returned to the patient in the hope of changing health behavior or initiating medical treatment according to the recommendations.

### 2.3.1 CVD prevention guidelines and lifestyle modifications

#### Guidelines for the prevention of CVD

Current guidelines for the prevention of CVD, by The American Heart Association (AHA) and The European Society for Cardiology (ESC), make recommendations for the prevention of CVD based on personal risk of developing CVD in the following 10 years. Overall, the emphasis is on lifestyle factors. Both guidelines recommend similar diets (although ESC specifically mentions a Mediterranean style diet), e.g., rich in vegetables, fruits, and whole grains but low in red or processed meat. ESC additionally recommends replacing saturated fatty acids with unsaturated ones, reducing salt intake as well as alcohol use and eating fish weekly. Exercise recommendations include at least 150 minutes of moderate-intensity or 75 minutes of vigorous-intensity physical activity (PA) per week. Additionally, smokers are encouraged to quit, and obese persons to lose weight. Also, blood pressure should be maintained at – or lowered to - a normal level (optimally systolic blood pressure ~120-130 mmHg and diastolic ~80 mmHg). Since diabetes is a strong risk factor for CVD, patients with diabetes should have their disease in control, either via lifestyle modifications or medication. Regarding LDL-C, lipid-lowering medication is generally recommended for persons with a high risk for CVD [210,211]. According to The Task Force for the management of dyslipidaemias of The European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS), the LDL-C target should always be at least less than 3 mmol/l. However, the higher the CVD risk is, the lower the LDL-C target. For persons with high CVD risk, LDL-C goal is < 1.8 mmol/l. For instance, persons with LDL-C ≥ 5 mmol/l are regarded as having high risk. The low target LDL-C level can be difficult to achieve without lipid-lowering medication. Moreover, lipid-lowering

medication is usually initiated or enhanced after a CVD event, such as MI (secondary prevention) [212].

### Impact of lifestyle modifications on ASCVD risk

Indeed, lowering LDL-C is an effective way to reduce one's ASCVD risk: it has been reported that total cholesterol reduction by 0.6 mmol/l would translate into a 20-50% reduction in CAD risk [213]. This level of reduction could be achieved with dietary modifications: In a hospital environment, LDL-C can be lowered by diet by 0.62 mmol/l (~10-15%) [214]. Correspondingly, in real-life situations, a mean LDL-C reduction of 13% has been achieved [215].

As included in the ESC guidelines, strong evidence supports the Mediterranean diet in lowering CVD risk: HR for CAD has been around 0.7 in persons on the Mediterranean diet [216,217]. Also, the DASH diet (Dietary Approaches to Stop Hypertension), designed for lowering blood pressure, reduces CAD risk by up to 20% (RR ~0.80). The effect on CAD risk is probably mediated by improving lipid values and decreasing diabetes risk, besides the impact on blood pressure [218,219]. Additionally, vegetarian diets are associated with lower LDL-C levels [220]. As many dietary studies have investigated the effect of baseline diets, the effect of changing one's diet has produced more conflicting results. For instance, improving diet quality within a 12-year follow-up period reduced CVD mortality by up to 15% (HR 0.85-0.96) [219]. However, improving dietary quality does not always yield statistically significant results [221].

Regarding exercise, again, many studies indicate that individuals with high physical activity have lower CVD risk and mortality; thus, increasing PA would decrease CVD risk [222–224]. However, there is a lack of evidence of increasing participants' PA and lowering CAD risk. The reduction in CVD risk possibly comes from a more extended time period that can be observed in studies [225]. The effects of aerobic exercise have been more widely studied than resistance training. Many studies indicate that aerobic exercise, or aerobic training combined with resistance training, is more effective in improving CVD risk profiles. Regarding lipids, aerobic exercise explicitly increases HDL-C, thus attenuating coronary plaque progression. However, resistance training may be more feasible for people with reduced mobility [226,227].

Overweight and obesity, considered a global epidemic, have been associated with CVD risk for long. Some of the risk comes from obesity-associated factors, e.g., hypertension, diabetes, fatty liver, and inflammation, but behavioral factors such as

unhealthy diet and lack of exercise as well. However, obesity itself has an independent effect on CVD risk; some studies estimate it to be around half of the whole association [228].

### 2.3.2 Impact of attitudes and adherence on health behavior

In practice, lifestyle modifications can be challenging to carry out. WHO has defined adherence as "the extent to which a person's behaviour – taking medication, following a diet, and/or executing lifestyle changes, corresponds with agreed recommendations from a health care provider" [229].

#### Adherence to medical treatment

Among patients with cardiometabolic diseases, only 33% were completely adherent to statin medication [230]. Correspondingly, high adherence among hypercholesterolemia patients resulted in 25% lower CVD risk than those with low adherence [231]. Moreover, adherence to statin medication has been lower among females, which could be due to patients as well as physicians thinking that CVD risk in females is very low [232].

Reasons behind nonadherence have included, for example, suspicion towards the treatment, side effects, and lack of help and support [231]. Conversely, the effectiveness of medication, lack of side effects, a positive attitude to a healthy lifestyle, trusting the physician, being involved in the decision-making, and physician listening to the concerns of the patients are associated positively with adherence [18,230].

Among hypertensive patients, lack of motivation and knowledge were common in general [17], whereas hypertensive patients believing that treating hypertension would reduce their risk for MI were more likely to monitor their blood pressure at home [16].

#### Adherence to lifestyle modifications

Adherence related to weight-loss interventions has also been studied. Adherence has been reported to be a more significant factor than the actual diet [233], while the adherence rate has been only around 50-60% [234,235]. One factor in improving adherence has been different kinds of social support, such as group meetings or personal coaching [233,235]. Indicators of nonadherence include poor weight loss at the beginning of the intervention, which can then impair motivation. Even though

lifestyle modifications can improve disease risks, despite the actual weight loss [236]. Moreover, better adherence has been associated with healthier baseline lifestyle factors and measurements [234].

### Knowledge and health literacy

Furthermore, knowledge and health literacy have been associated with adherence to medication and lifestyle modifications. The American Heart Association defines health literacy as "the degree to which individuals are able to access and process basic health information and services and thereby participate in health-related decisions." Hence, increasing knowledge does not help if the patients' health literacy is poor. Health literacy is associated with lower education level and older age [237].

To summarize, adherence to medical treatment and lifestyle modifications is crucial to affect disease risk. Attitudes and motivation are important underlying factors of adherence; if a person believes that they can modify their disease risk, they are more likely to do so. Side effects of the medication also influence adherence. Many studies also show the impact of healthcare, that patients feel they can trust their physicians and are heard. If the persons who are more likely to be nonadherent would be recognized in the healthcare system, more support could be directed to them.

### 2.3.3 Returning personal disease risk

There is a long history of intervention studies of lifestyle modifications. Many are based on increasing knowledge and offering support in healthcare or support groups. There are positive results in smoking cessation, increasing PA, and improving diets, for example [238]. Estimating and returning personal risk for different diseases, such as CAD, is a standard procedure in healthcare. The goal regarding high-risk individuals is to motivate them to make appropriate lifestyle modifications or initiate medical treatment to lower their risk. However, returning personalized disease risk does not often affect behavior more than information on the risk on a general level. Although, how the risk is communicated might have an effect [239].

After developing PGSs for common diseases and understanding their predictive value, an interest in using them in healthcare has risen. Moreover, monogenic mutations have a long history of being tested in healthcare. Nonetheless, initial concerns have been how the individuals would react to their genetic risk information, whether carrying monogenic high-impact variants or polygenic risk. Would information about having low risk encourage careless behavior, or having high risk cause exaggerated concern? But most importantly, would returning genetic risk encourage positive health behavior?

#### Receiving genetic risk information

The central focus of studying the returning of genetic risk information has been on how this information is received. Regarding monogenic mutations (e.g., FH and familial cancers), patients who have received results of being mutation carriers have mostly had positive attitudes. In some cases, there have been concerns, which have faded during the following months [240,241]. Additionally, only a small minority regret being screened for mutations: for instance, only 2% of patients tested for FH mutations [240].

Regarding polygenic risk information, the responses have been largely positive as well. No anxiety or depression has been reported [242,243], despite some high-risk individuals reporting concern about their personal risk [244]. Thus, revealing genetic risk information to people does not appear to cause major psychological harm.

### Impact on health behavior

So far, studying the impact of genetic risk information on lifestyle modification has been sparse. Regarding FH mutations, confirming clinical diagnosis with a genetic test did not impact lifestyle modifications or adherence to lipid-lowering medication. However, mutation carriers relied more on medication (compared to lifestyle factors) than persons with just a clinical diagnosis [245]. Many studies have also explored the impact of mutation testing related to familial cancers. Mutation carriers attended screenings, such as mammograms and endoscopies, more often (up to 100%). However, non-carriers did not reduce their recommended screening frequencies [241,246].

A growing number of studies have been exploring returning PGSs to patients. PGSs have included risk for CVD or common cancers, for instance. Some studies have returned PGSs to all study participants, and some have randomized returning of the genetic risk. Many of the studies have not reported differences in lifestyle modifications, whether the person has a high genetic risk or not. However, high-risk persons have not worsened their lifestyles, i.e., fatalism regarding genetic risk does not appear to be true [243,244,247]. On the other hand, persons with high CVD risk, with high PGS or not, have received lipid-lowering medication, and thus, their LDL-C levels have decreased [242]. Interestingly, one study reported that persons with high genetic risk searched for more information on CVD, particularly related genetic factors [247]. The impact of this behavior remains elusive.

What has not been investigated is how the risk information is returned and how that affects the person's response to the risk. Thus far, there have been no guidelines on how the healthcare provider should present and interpret the risk information and educate the patient about their personal risk. It is a timely problem since there is already a substantial supply of different kinds of commercial genetic testing, including PGSs.

# 3 AIMS OF THE STUDY

Hypercholesterolemia is a major risk factor for ASCVD but a treatable one. Individuals at high risk for ASCVD are important to identify and treat to prevent premature illness and death. Carriers of monogenic hypercholesterolemia-causing mutations are at especially high risk. However, not all mutations are known, and thus their carriers not identified. Besides genetic factors, behavioral factors affect hypercholesterolemia as well, but their impact is not well characterized. While it is crucial to identify high-risk individuals, PRSs may improve ASCVD risk estimates. Yet, it is not well known how returning genetic risk scores would affect individuals.

The general aim of this thesis was to identify and characterize novel risk factors for hypercholesterolemia and ASCVD and to estimate if returning personal ASCVD risk, including genetic risk, would encourage persons to make healthy lifestyle modifications. The primary study cohort was the GeneRISK study, which includes genetic data, measurements, and questionnaires, consisting of 7 342 middle-aged persons from Southern Finland.

The specific aims were:

1. To characterize the genetic background of severe hypercholesterolemia in Finland by whole-exome sequencing cases with severe hypercholesterolemia, and further, to estimate the impact of both known and possible novel hypercholesterolemia-causing mutations on ASCVD risk (Study I).
2. To study the effect of behavioral factors on hypercholesterolemia, and to evaluate the ASCVD prevention measures received by hypercholesterolemia cases in healthcare (Study II).
3. To explore if returning personal ASCVD risk, including PRS, would promote positive health behavior and how study participants perceive their genetic risk (Study III).

# 4 MATERIALS AND METHODS

## 4.1 Study cohorts

All the studies composing this thesis used the GeneRISK study cohort. The FinnGen study was utilized in Study I to estimate the impact of genetic variants identified in the GeneRISK cohort.

### 4.1.1 The GeneRISK study

The GeneRISK study (**Figure 9,** www.generisk.fi) is a prospective observational study. It consists of 7 342 study participants (4 691 females and 2 651 males), aged 45-65 years at baseline, and recruited from Southern Finland in 2015-2017. Persons with previous ASCVD were excluded; pregnancy and being under guardianship were also exclusion criteria.

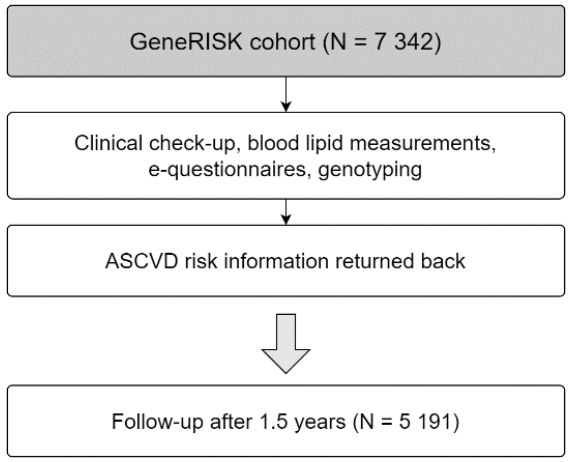



**Figure 9. Study design of the GeneRISK study.** At baseline, 7 342 individuals were recruited. The study participants filled out an e-questionnaire and underwent a clinical health check-up, including blood lipid measurements. All samples were also genotyped. The risk for developing atherosclerotic cardiovascular disease (ASCVD) in the following 10 years was estimated based both on traditional risk factors and a polygenic risk score. The personal ASCVD risk was returned to the study participants. After 1.5 years, on average, study participants were invited to a follow-up study, including an e-questionnaire and a clinical check-up. The follow-up e-questionnaire was filled out by 5 191 study participants.

At baseline, all the study participants underwent a health check-up, including weight, height, and blood pressure measurements, as well as lab measurements of blood lipids and fasting glucose. LDL-C value was calculated using the Friedewald formula [248]. Additionally, the study participants filled out an e-questionnaire including questions about, e.g., family history of different diseases, behavioral factors, attitudes, and socioeconomic factors. All the study participants were genotyped and received a 10-year risk estimate for ASCVD (described in chapter 4.4.1, page 52). After 1.5 years on average, the participants answered another e-questionnaire, which included the same questions as at baseline and additional questions regarding received ASCVD risk and lifestyle modifications after the baseline visit. Participants were also invited to attend a follow-up check-up, which included the same measurements as at the baseline visit. The GeneRISK study has also been linked to national health registries, including hospital visits and drug prescriptions, enabling longitudinal analyses.

Study I included study participants with LDL-C ≥ 5.0 mmol/l at baseline (N=366), Study II encompassed cases with LDL-C ≥ 4.6 mmol/l and controls with LDL-C < 3.4 mmol/l (701 cases and 3 128 controls, persons using lipid-lowering medication were excluded), and Study III was based on the whole cohort, but concentrating on individuals with high risk for ASCVD (1 760 high-risk and 5 582 average- or low-risk individuals). Characteristics of cases in the studies are presented in **Table 3** (next page), and variables used in the analyses are described in chapter 4.3 (page 51).

### 4.1.2 The FinnGen study

The FinnGen study ([81], www.finngen.fi) is a large public-private partnership that includes nine Finnish biobanks and aims to collect samples from 500 000 Finnish individuals. FinnGen includes samples from hospital biobanks as well as epidemiological and disease-based cohorts. The FinnGen study is linked with longitudinal registry data, including, e.g., national hospital-based registries, cause-of-death register, and drug reimbursement data from KELA (the Finnish social security institute). FinnGen samples have also been genotyped and imputed.

In Study I, the eighth FinnGen data release (R8) was used, including 356 082 individuals (198 484 females, 157 523 males).

**Table 3. Characterization of cases in the three studies composing this thesis.** All studies utilized GeneRISK study participants: Study I was based on whole-exome sequenced persons with LDL-C ≥ 5.0 mmol/l. Study II utilized cases with LDL-C ≥ 4.6 mmol/l and controls with LDL-C < 3.4 mmol/l. Persons using lipid-lowering medication at baseline were excluded, as well as persons carrying hypercholesterolemia-causing mutations, identified in Study I. Study III included all study participants but focused on persons with high risk for atherosclerotic cardiovascular disease (ASCVD). High risk refers to a 10-year ASCVD risk exceeding 10%. High risk also includes certain pre-existing conditions, such as diabetes, that cause high risk. Mean (standard deviation) is reported for continuous variables and N (%) for categorical variables.

| | **Study I: LDL-C ≥ 5.0 mmol/l** | **Study II: LDL-C ≥ 4.6 mmol/l** | **Study III: High risk for ASCVD** |
|---|---|---|---|
| **N** | 366 | 701 | 1 760 |
| **Males, N (%)** | 144 (39.3%) | 278 (39.7%) | 1 067 (60.6%) |
| **Age (years)** | 56.5 (5.3) | 56.4 (5.5) | 58.9 (5.2) |
| **BMI (kg/m$^2$)** | 27.9 (4.3) | 27.9 (4.3) | 29.3 (5.1) |
| **Total cholesterol (mmol/l)** | 7.7 (0.8) | 7.3 (0.7) | 5.7 (1.3) |
| **HDL-C (mmol/l)** | 1.52 (0.41) | 1.54 (0.44) | 1.37 (0.47) |
| **LDL-C (mmol/l)** | 5.5 (0.6) | 5.1 (0.6) | 3.6 (1.1) |
| **Triglycerides (mmol/l)** | 1.6 (0.6) | 1.5 (0.6) | 1.7 (1.4) |
| **Glucose (mmol/l)** | 5.8 (0.8) | 5.8 (0.8) | 6.3 (1.5) |
| **Systolic blood pressure (mmHg)** | 131.3 (17.2) | 130.5 (16.5) | 137.7 (17.1) |
| **Diastolic blood pressure (mmHg)** | 87.5 (10.6) | 86.7 (10.1) | 89.7 (10.3) |
| **Use of lipid-lowering medication, N (%)** | 15 (4.1%) | 0 (0.0%) | 386 (16.3%) |
| **High risk for ASCVD, N (%)** | 186 (50.8%) | 282 (40.2%) | 1 760 (100.0%) |
| **Previously diagnosed hypercholesterolemia, N (%)** | 313 (85.5%) | 581 (82.9%) | 369 (21.0%) |
| **Family history of early-onset MI, N (%)** | 90 (24.5%) | 162 (23.1%) | 593 (33.7%) |
| **Family history of hypercholesterolemia, N (%)** | 258 (70.5%) | 479 (68.3%) | 1 007 (57.2%) |
| **LDL-PRS** | 0.57 (0.85) | 0.51 (0.86) | 0.05 (1.0) |
| **CAD-PRS** | -0.10 (0.93) | -0.07 (0.95) | 0.12 (1.0) |
| **Daily smokers, N (%)** | 51 (13.9%) | 93 (13.3%) | 366 (20.8%) |
| **RFDS food score** | 7.7 (3.2) | 7.8 (3.2) | 7.8 (3.2) |
| **Exercise at free time** | 2.8 (1.0) | 2.8 (1.0) | 2.7 (1.0) |
| **Normal cholesterol not important** | 47 (12.8%) | 89 (12.7%) | 164 (9.3%) |
| **Self-reported health** | 7.7 (1.4) | 7.7 (1.3) | 7.5 (1.4) |
| **Marriage or cohabitation, N (%)** | 253 (69.1%) | 507 (72.3%) | 1 339 (76.1%) |
| **Higher education, N (%)** | 93 (25.4%) | 190 (27.1%) | 403 (22.9%) |

## 4.2 Genetic analyses

In Study I, WES, Sanger sequencing, and genotype data were used in the analyses. Studies II and III utilized PRSs calculated from imputed genotype data.

### 4.2.1 Whole-exome sequencing (WES) analysis

In Study I, 351 cases with severe hypercholesterolemia were whole-exome sequenced at the Technology Center at the Institute for Molecular Medicine Finland (FIMM), including target capturing with NimbleGen SeqCap EZ MedExome and sequencing with Illumina HiSeq sequencer. Alignment was to the human reference genome hg19, and variant calling was performed with the variant calling pipeline at FIMM, with a sequence coverage of at least 20 times.

#### Variant filtering

The WES analysis focused on genes associated with FH: *LDLR, APOB, PCSK9*, and *LDLRAP1*. I annotated the exomes with ANNOVAR ([49], http://annovar.openbioinformatics.org/en/latest/, downloaded 24.10.2019), including gnomAD frequencies and eight *in silico* prediction algorithms (PolyPhen-2, SIFT, Mutation Taster, FATHMM, Radial SVM, LRT, MutationAssessor and LR). In the initial filtering of the variants, I included exonic regions with a predicted change in the protein-coding sequence (i.e., missense, nonsense, or frameshifting deletion or insertion variants) and at least 20 times coverage. Then, I included only variants with minor allele frequency (MAF) of less than 0.002 (frequencies from gnomAD, either in all populations or in Finns). I checked whether these variants were reported as pathogenic in the ClinVar database. Next, I focused on variants with at least 50% of the *in silico* programs giving a damaging or probably damaging prediction.

#### Analysis of copy number variants (CNVs)

Since WES does not cover large insertions or deletions, I run two CNV analysis programs, trying to detect possible large insertions or deletions (Conifer (version 0.2.2) [249], Xhmm (version 1.1) [250]). I also visually examined aligned and sorted BAM files with IGV (Integrative Genomics Viewer, https://software.broadinstitute.org/software/igv/download).

### 4.2.2 Sanger sequencing

Sanger sequencing was performed to verify WES results in Study I.

First, part of the exome-sequenced samples (N=206) was verified for seven Finnish FH mutations in a clinical laboratory by routine clinical diagnostic genotyping (Yhdistyneet Medix Laboratoriot, YML). This analysis also captures the large FH-Helsinki deletion.

Second, since the amino acid position 574 in *LDLR* is multiallelic, all GeneRISK study participants with any variants in this position were Sanger sequenced to find possible carriers of the Arg574Leu variant, in addition to the index case identified with WES. Capillary Sanger sequencing was performed at the Technology Center at FIMM; sequencing reactions were run on an Applied Biosystems 96-capillary 3730xl DNA Analyzer using oligonucleotide primers CAGCACGTGACCTCTCCTTA and CAAAGACGGCCAAGGAGAAG.

### 4.2.3 Genotyping and imputation

All the GeneRISK samples were genotyped using the HumanCoreExome BeadChip (Illumina Inc, San Diego, CA, USA). Genotypes were called using GenomeStudio and zCall at FIMM, they were phased using SHAPEIT2, and imputed using IMPUTE2 and a combined reference panel of 1000 Genomes Phase I integrated haplotypes and 1 943 Finnish genomes [251].

FinnGen samples were genotyped with Illumina (Illumina) and Affymetrix arrays (Thermo Fisher Scientific). A detailed description of the FinnGen genotyping pipeline can be found in Kurki et al. [81].

### 4.2.4 Polygenic risk scores (PRSs)

PRS for LDL-C (LDL-PRS, Study II) in GeneRISK was calculated using LDPred [252] and weights derived from GWAS of 324 708 samples from the UK Biobank [253]. PRS for CAD (CAD-PRS, Studies II, and III) was based on CardiogramPlusC4D meta-analysis summary statistics and calculated with the sum of the risk alleles weighted by the logarithms of the odds ratio risk estimates [254].

CAD-PRS and LDL-PRS (Study I) in the FinnGen study were calculated using CS-PRS [255] and PLINK v.2.0 [256], and used weights were from GWAS by Surakka et al. [257] for LDL-PRS and Cardiogram and CARDIoGRAMplusC4D Consortium for CAD-PRS [197].

## 4.3 Phenotype analyses

In Study II, we utilized data from the GeneRISK e-questionnaires, which included questions about behavioral variables, socioeconomic factors, and attitudes. A detailed description of the questions is presented in the original publication II.

Previous hypercholesterolemia referred to the question: "Have you ever been diagnosed with high or elevated blood cholesterol?". Family histories of different diseases refer to father, mother, brothers, or sisters being diagnosed with the disease. Early-onset MI refers to "myocardial infarction or heart attack before age 60 years" and hypercholesterolemia to "hypercholesterolemia (elevated blood lipids)".

Exercise at free time was communicated as: "How often do you exercise and strain yourself physically during free time on average?". There were five options, from very little exercise to competitive training. I used the variable as a numeric, continuous variable (1 to 5). Self-reported health was based on: "Evaluate your state of health today by choosing a number that best reflects your current state of health (0-10). Number zero describes the worst possible state of health, and number ten the best possible one." I treated it as a numeric, continuous variable. The question about attitude was asked: "How important are the following health-related subjects to you? Keeping your cholesterol level within normal limits." The five options included: "Not at all important," "A little important," "Moderately important," "Quite important," and "Very important." I pooled together two options, not at all and a little important, to form the variable "Normal cholesterol not important".

Food scores were calculated based on the Recommended Finnish Diet Score, RFDS [258], derived from the answers regarding eating habits. RFDS is based on Finnish eating recommendations, including, e.g., consumption of fruits and vegetables and red meat. Because the GeneRISK questionnaire did not cover all the foodstuffs included in the RFDS, some minor modifications were made (detailed description in original publication II); thus, the possible range of RFDS was 0-17 (from least to most healthy).

## 4.4 Returning ASCVD risk information

In Study III, the GeneRISK study participants (N=7 342) received their 10-year risk of developing ASCVD after their baseline health check-up. The risk information, as well as the results of their laboratory results, was displayed via the interactive web tool Kardiokompassi® (**Figure 10**).

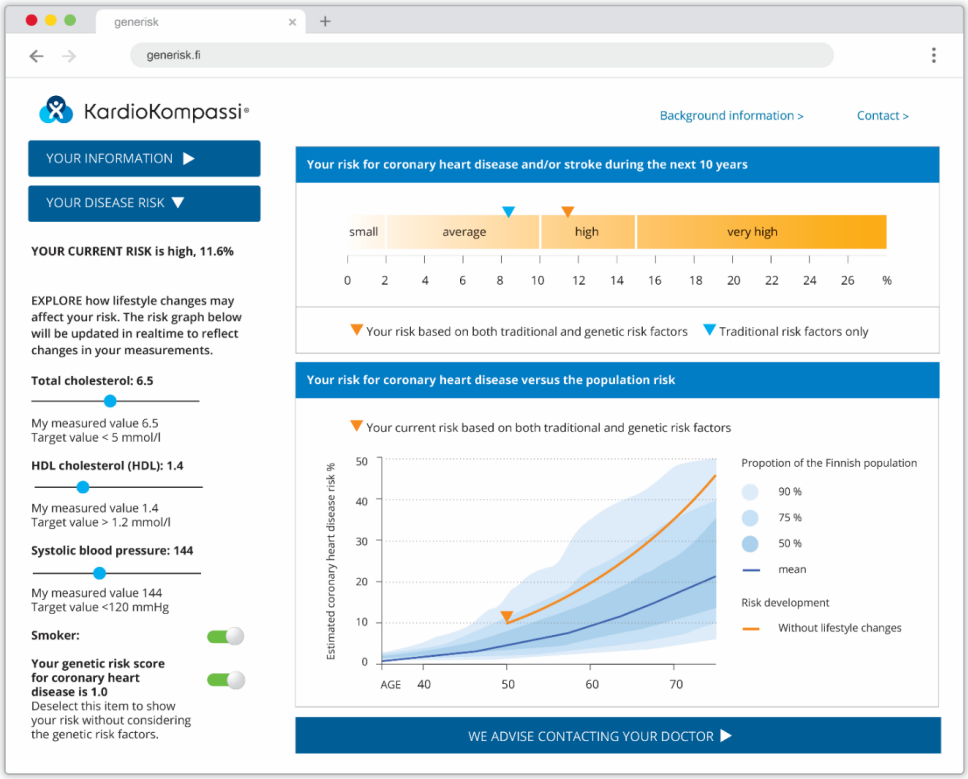



**Figure 10. Example view of Kardiokompassi®.** Kardiokompassi is a web-based tool to communicate personal risk for atherosclerotic cardiovascular disease (ASCVD) to the study participants. In this example, the current risk for ASCVD is high (11.6%), and the genetic risk score is 1.0, i.e., risk-increasing. From the original publication (III) under the terms of the Creative Commons Attribution License (CC BY).

### 4.4.1 Estimation and definition of ASCVD risk

In the GeneRISK study, two risk models were constructed. One included only traditional risk factors: systolic blood pressure, sex, age, total and HDL cholesterol, smoking status, use of antihypertensive medication, as well as family history of early-onset CAD. The second model additionally included CAD-PRS, and the risk was a

composite of traditional and genetic risk factors. According to the Finnish guidelines [259], high risk was defined as at least a 10% risk of developing ASCVD in the next 10 years. Average risk was defined as 2-10% and low risk less than 2%. Also, persons with certain conditions that increase ASCVD risk, such as diabetes, were assigned to the high-risk group, regardless of their other risk factors.

### 4.4.2 Returning ASCVD risk via Kardiokompassi®

The study participants were able to view their ASCVD risk as well as their laboratory results using the Kardiokompassi web application (**Figure 10**). Via Kardiokompassi®, ASCVD risk can be viewed with or without polygenic risk. The impact of different risk factors can be demonstrated, for example, how smoking cessation would impact the future risk, thus visualizing the effects of lifestyle change. Individuals at high risk for ASCVD were instructed to see a physician. However, they were instructed to contact their own physician, and no specific intervention instructions were given.

Besides receiving their ASCVD risk and laboratory results, all study participants were given the possibility to sign up for electronic health coaching modules offered by The Finnish Medical Society Duodecim, for supporting several lifestyle modifications, including weight control, healthy dietary habits, exercising, smoking cessation, stress reduction, and healthy sleeping.

### 4.4.3 Follow-up and definition of positive health behavior

After 1.5 years, on average, all the study participants were asked to fill out a follow-up e-questionnaire and attend a follow-up health check-up. The follow-up e-questionnaire included additional questions about the participants' attitudes toward receiving their risk information, as well as their health behavior, e.g., if they had tried changing their lifestyles or visited a physician. Altogether, 5 191 persons filled out the follow-up e-questionnaire, and 4 921 attended the follow-up visit.

We defined positive health behavior as signing up for health coaching, losing weight, smoking cessation, or visiting a physician. These behaviors were mainly assessed using e-questionnaires, although weight change was also examined utilizing the follow-up health check-up. Weight loss was defined as both reporting losing weight and weighing at least 0.1 kg less at the follow-up visit compared to the baseline.

## 4.5 Statistical analyses

In Study I, I analyzed the relationship between hypercholesterolemia-associated variants and ASCVD as well as the use of lipid-lowering medication by multivariate logistic regression. I adjusted the model with age and sex.

In Study II, I tested continuous variables with Student's t-test (except for skewed variables, where I used the Mann-Whitney-U test) and categorical variables with $\chi^2$ test.

In Study III, the comparison of health behavior at follow-up, high-risk individuals compared to low- or average-risk individuals, was performed with $\chi^2$ test. Multivariate logistic regression analysis was performed to evaluate factors associated with perceiving risk information as concerning, and the model was adjusted with the recruiting clinical center (i.e., Carea, Veripalvelu, or Mehiläinen).

# 5 RESULTS AND DISCUSSION

Both genetic and behavioral factors affect hypercholesterolemia and ASCVD risk. Genetic factors are unchangeable, but behavioral factors are not. First, novel genetic factors of severe hypercholesterolemia were identified, and their impact on ASCVD risk was studied. Secondly, the effect of behavioral factors on hypercholesterolemia, as well as ASCVD prevention measures in healthcare, were investigated. Lastly, personal ASCVD risk was returned to the study participants, and the effect of the risk information on the motivation to make lifestyle modifications was assessed.

## 5.1 Novel genetic risk factors for severe hypercholesterolemia (I)

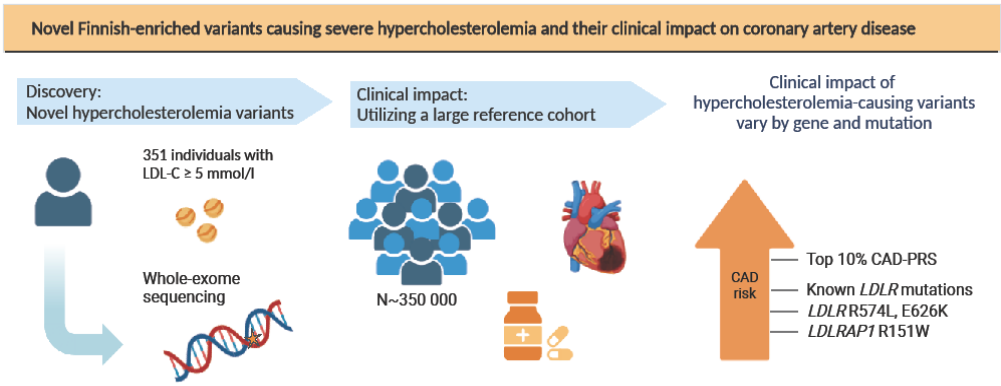



**Figure 11. Study I: the overall study design and main results.** Discovery: Novel hypercholesterolemia-associated variants were identified by whole-exome sequencing 351 GeneRISK study participants with severe hypercholesterolemia (LDL-C ≥ 5 mmol/l). Clinical impact (statin use and ASCVD) was examined by utilizing a large reference cohort. Result: Clinical impact of hypercholesterolemia-causing variants vary by gene and mutation. From the original publication (I) under the terms of the Creative Commons Attribution License (CC BY 4.0).

In Study I (**Figure 11**), whole-exome-sequencing was performed for 351 GeneRISK study participants with severe hypercholesterolemia (LDL-C ≥ 5 mmol/l). They had possible FH according to the Dutch Lipid Network Criteria [21], based solely on their LDL-C levels. The first aim was to identify carriers of known FH mutations, the second was to characterize novel mutations causing severe hypercholesterolemia, and the

third was to assess the clinical impact of both known and novel mutations using a large population cohort.

### 5.1.1 Novel mutations causing severe hypercholesterolemia

Among 351 whole-exome sequenced GeneRISK study participants, I identified three carriers of previously known FH mutations (**Table 2**, page 33): FH-Pogosta (*LDLR* Arg595Gln), FH-North Karelia (*LDLR* Pro309Lysfs) and the European *APOB* Arg3527Gln mutation. Since one of the Finnish-enriched FH mutations is a large deletion, not captured in WES, samples from 206 cases were also sent for clinical diagnostic testing. However, no carriers of, e.g., FH-Helsinki were identified. CNV analyses using *in silico* algorithms did not reveal any large insertions or deletions either.

The three carriers of known FH mutations had varying LDL-C levels: the carrier of FH-North Karelia had the highest LDL-C in the whole GeneRISK study, 12 mmol/l, and family history of early-onset MI as well as hypercholesterolemia. In turn, the carrier of FH-Pogosta had LDL-C 5.6 mmol/l – not particularly high in our WES subset –, and the carrier of the *APOB* Arg3527Gln had LDL-C 6.3 mmol/l. Both individuals had a family history of hypercholesterolemia but not MI. None of these three carriers were using lipid-lowering medication at baseline.

#### Novel variants associated with severe hypercholesterolemia

According to the filtering steps described in the Methods chapter (page 49) and **Figure 12**, I searched for novel hypercholesterolemia-associated variants, focusing on FH-associated genes: *LDLR*, *APOB*, *PCSK9*, and *LDLRAP1*. Filtering resulted in six variants (**Table 4**, page 58): two in *LDLR*, three in *APOB*, and one in *LDLRAP1*. I excluded two variants in *PSCK9* because *in silico* algorithms predicted loss-of-function, whereas gain-of-function mutations in *PCSK9* are associated with hypercholesterolemia [173]. I additionally searched for more carriers of the six variants from the genotype data of the whole GeneRISK study. Moreover, as mentioned in the Method chapter (page 50), *LDLR* amino acid position 574 was Sanger sequenced from all possible carriers.

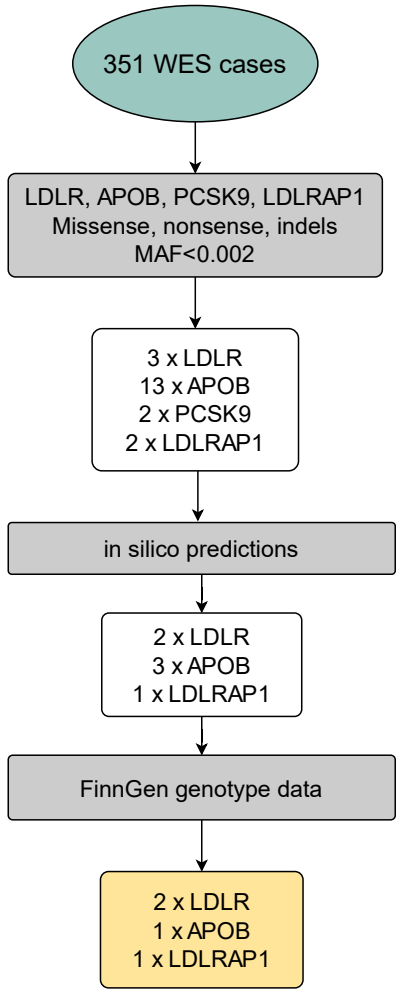



**Figure 12. Filtering of hypercholesterolemia-associated variants.** Variants, carried by 351 whole-exome sequenced (WES) GeneRISK study participants, were first filtered by gene (FH-associated genes *LDLR*, *APOB*, *PCSK9*, and *LDLRAP1*) and mutation type (missense, nonsense, insertions, or deletions) with minor allele frequency (MAF) < 0.002. This initial filtering resulted in three variants in *LDLR*, 13 in *APOB*, two in *PCSK9*, and two in *LDLRAP1*. Further filtering was based on *in silico* prediction algorithms: variants with at least 50% of the eight algorithms predicting damaging consequences were included, leaving two variants in *LDLR*, three in *APOB*, and one in *LDLRAP1.* These variants were searched from the FinnGen genotype data, and carriers of two *LDLR* variants, one *APOB* variant and one *LDLRAP1* variant were found.

**Table 4. Hypercholesterolemia-associated variants identified from the whole-exome sequenced GeneRISK study participants.** Hypercholesterolemia-associated variants in genes associated with FH: Nonsynonymous or indel variants, MAF (minor allele frequency) < 0.002, *in silico* predictions damaging at least four of the eight algorithms. "Index cases" refer to the whole-exome sequenced GeneRISK study participants. "Other carriers" refer to the number of additional variant carriers in the GeneRISK cohort identified from genotype data. Mean LDL-C is reported in mmol/l, and the number of LDL-C values used in the calculation is in parentheses. *Excluded index cases and persons using lipid-lowering medication at baseline. MAF was retrieved from the gnomAD database, NF: not found. Modified from the original publication (I) under the terms of the Creative Commons Attribution License (CC BY 4.0).

| Gene | Index cases (N) | Variant | MAF, all | MAF, FIN | Damaging *in silico* predictions | Other carriers (N) | Mean LDL-C* |
|---|---|---|---|---|---|---|---|
| *LDLR* | 1 | c.1721G>T (p.Arg574Leu) | 0.00002 | 0.0002 | 8/8 | 6 | 4.2 (N=3) |
| *LDLR* | 1 | c.1876G>A (p.Glu626Lys) | 0.0003 | 0.0002 | 4/8 | 10 | 3.6 (N=8) |
| *APOB* | 1 | c.9770A>G (p.Asn3257Ser) | 0.00005 | 0 | 4/8 | NF | - |
| *APOB* | 1 | c.9491C>T (p.Thr3164Met) | 0.0001 | 0 | 5/8 | 1 | 3.7 |
| *APOB* | 2 | c.2630C>T (p.Pro877Leu) | 0.0005 | 0.0004 | 5/8 | 8 | 3.2 (N=7) |
| *LDLRAP1* | 1 | c.451C>T (p.Arg151Trp) | 0.0003 | 0.0012 | 5/8 | 13 | 3.7 (N=11) |

To further estimate the clinical impact of the variants, we utilized the FinnGen cohort, which contains a larger number of individuals, as well as longitudinal health registry information. Of the six variants, FinnGen genotype data contained carriers of *LDLR* Arg574Leu and Glu626Lys, *LDLRAP1* Arg151Trp, as well as *APOB* Pro877Leu, but not the remaining two *APOB* variants. Thus, we were not able to assess their pathogenicity. Nonetheless, our analysis using the FinnGen data (described in more detail in chapter 5.1.3, page 63) indicated that *LDLR* Arg574Leu and Glu626Lys, as well as *LDLRAP1* Arg151Trp, are associated with hypercholesterolemia, demonstrated by increased odds for the use of lipid-lowering medication and ASCVD diagnosis of the carriers. Additionally, *LDLR* Arg574Leu has been previously reported in one patient with FH [260].

Based on the previous literature, the positions of both novel *LDLR* variants would affect the EGF-like domain, which is thought to affect the dissociation of LDLR-LDL complex in endosomes, and its defects could influence the recycling of the LDL receptor and LDL degradation [8,9]. In turn, *LDLRAP1* Arg151Trp might affect the PTB

domain, which is involved in the binding of LDLR [9]. However, to validate the impact of these mutations at a molecular level would require functional testing.

Traditional methods to study the effects of *LDLR* variants are laborious and usually require cells from the patients [261]. A novel, semi-automized platform has been developed, which enables the quantification of LDL uptake and lipid storage in a systematic way. This platform offers a more feasible way to study the effects of novel *LDLR* mutations [262].

#### *LDLRAP1*

Interestingly, one of the novel, hypercholesterolemia-causing mutations was *LDLRAP1* Arg151Trp, in a heterozygous state, although *LDLRAP1* mutations have previously been associated with recessive FH [7,174]. However, some studies show that heterozygous carriers of *LDLRAP1* mutations have higher LDL-C levels than controls [174,263,264], though other studies have not reported elevated LDL-C values [265]. Traditionally, FH mutation analyses have included only recessive *LDLRAP1* variants. Our results show that this strategy might not be correct, and heterozygous carriers should also be considered. However, to conclude that *LDLRAP1* variants could be pathogenic in a heterozygous state would require further functional testing, e.g., using cell models.

### Summary of the genetic findings

Only six of the 351 whole-exome sequenced GeneRISK study participants (1.7%) carried a probably hypercholesterolemia-causing mutation (previously known or novel mutation identified in Study I). Still, this is in line with similar population-based studies of severe hypercholesterolemia, where less than 2.5% of cases have been explained by monogenic etiology [12,19,266]. Besides the aforementioned mutations, 58 (16.5%) WES cases had LDL-PRS in the top decile (unpublished results), which is in accordance with some previous studies [14,267]. Consequently, we found a genetic explanation for 26.5% of the WES cases.

What could explain the severe hypercholesterolemia of the remaining individuals? Most of the cases had LDL-C around 5-6 mmol/l, whereas especially FH-causing frameshifting mutations, such as the Finnish FH-Helsinki and FH-North Karelia mutations, often cause LDL-C levels higher than that [182,183,268]. Their elevated LDL-C values could be due to interactive effects of variants with lower impacts and less healthy lifestyles. Variants with modest effect sizes are difficult to identify and

would require large sample sizes. On the other hand, there were six WES cases with LDL-C 7-10 mmol/l with no genetic explanation. It is possible that mutations in other genes, not yet associated with hypercholesterolemia, could cause their severe hypercholesterolemia [21,159].

### 5.1.2 The Finnish spectrum of hypercholesterolemia-causing mutations broader than previously reported

Previous studies have identified seven *LDLR* mutations (**Table 2,** page 33), which explain approximately 80% of FH cases in Finland [10,11]. Additionally, a few other *LDLR* mutations have been identified, but they have only been identified in single individuals or families and do not account for much of the whole Finnish FH prevalence [10]. Moreover, no mutations in *APOB*, *PCSK9*, or *LDLRAP1* have been identified [11,186,187]. Furthermore, previous studies have calculated the frequency of Finnish FH mutations to be around 1:600 [11]. That is much rarer than the average FH prevalence in many international studies, i.e., ~1:250-330 [19,146–148]. That raises the question of whether FH is truly less prevalent in Finland or whether there are unidentified mutations causing the disease.

#### Novel Finnish hypercholesterolemia-causing mutations

Our study adds two *LDLR* mutations and one in *LDLRAP1* to the Finnish mutation spectrum of hypercholesterolemia-causing mutations (**Table 4,** page 58). Moreover, as far as we know, we identified the European *APOB* mutation for the first time in Finland, though in only one carrier. Thus, the Finnish FH mutation spectrum appears to be broader than previous publications have estimated.

#### Prevalence of hypercholesterolemia-causing mutations in Finland

Besides previously mentioned novel variants, I searched for carriers of known Finnish FH mutations from the genotype data of the FinnGen study to estimate the variants' impact on ASCVD risk. Unfortunately, the FinnGen genotype chips did not cover FH-Helsinki and FH-North Karelia mutations. I found carriers of FH-Pogosta, FH-Pori, FH-Turku, and FH-Fin 11. Additionally, I identified 22 carriers of *LDLR* Asp579Tyr, which resides in the same position as FH-Fin 11 and has previously been associated with FH in Italy [269].

The most common FH mutation was FH-Pogosta, and its frequency was 0.0003. This mutation is also the only shared one with our study and Lahtinen et al. [11], where the frequency of FH-Pogosta was similar (0.00036). Adding up the frequencies of the novel mutations with the known FH mutations results in ~0.0015 frequency (1:667) without the *LDLRAP1* mutation and ~0.0043 (1:233) with it. Since the FinnGen data did not include the FH deletions, I have estimated their frequency from the Lahtinen et al. study: FH-Helsinki 0.00043 and FH-North Karelia 0.00028. Thus, the Finnish prevalence estimate adds up to 0.0022 without *LDLRAP1* and 0.0041 with it, corresponding to a prevalence estimate of 1:244-454 – closer to the international prevalence estimates of FH than previously reported. Regarding the *LDLRAP1* mutation, it is debatable whether to include its carriers as having FH.

Regarding Finnish-enriched FH mutations, with an estimated prevalence of 1:600, the whole GeneRISK study could have included 12 carriers, likely in the whole-exome sequenced subset, whereas I identified only two carriers. One reason behind this discrepancy could be that the GeneRISK study excluded persons with previous ASCVD, which could also exclude carriers of more severe FH mutations, such as FH-Helsinki, which we did not identify in this study.

### Geographical distribution of the mutations

Previous studies have shown that the prevalence of Finnish FH mutations varies by geographical region [10,11]. However, there were no apparent geographical differences in the birthplaces of the FinnGen carriers (unpublished results, **Figure 13**, next page)**.** It could be due to Finns gradually moving away from their birthplaces [270]. Likewise, the birthplaces of the carriers of the three novel mutations are scattered around Finland.

### Medical treatment of the mutation carriers

Based on the FinnGen data, there could be ~15 000 carriers of the three novel mutations in Finland, with elevated ASCVD risk, and only 37-53% of them using lipid-lowering medication (**Table 5**, page 64). Moreover, a significant proportion of carriers (42%) of known FH mutations were not using lipid-lowering medication either.

Study I provides new information on the Finnish FH mutation spectrum; the novel mutations could be added to diagnostic mutation testing panels to identify carriers previously not recognized. The next chapter shows that the carriers of these novel variants have an increased risk for ASCVD if left untreated.

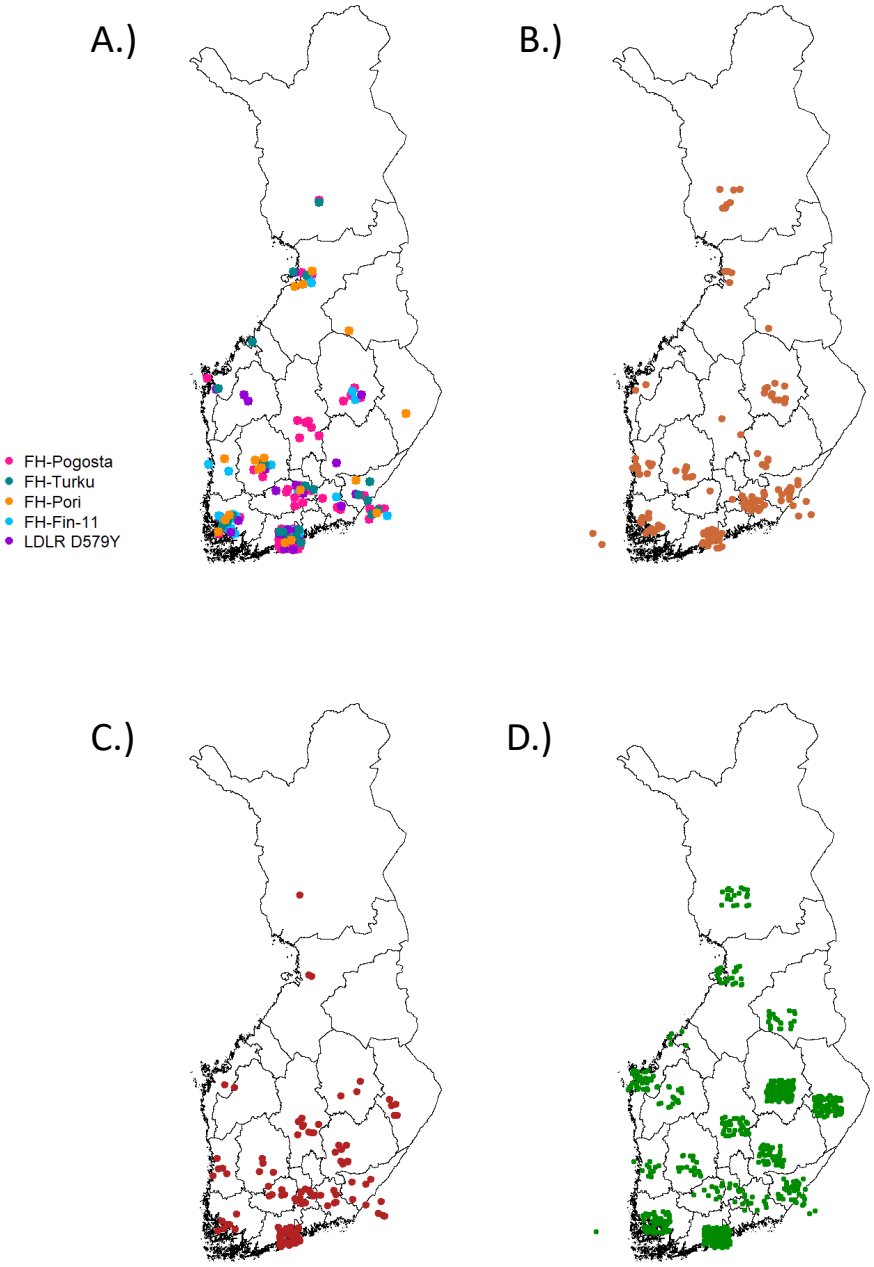



**Figure 13. Geographical distribution of the birthplaces of the FinnGen carriers of hypercholesterolemia-causing mutations.** Birthplaces of the FinnGen (release 8) carriers are in a province level, scatter added. Carriers of A.) five known FH mutations, B.) *LDLR* Glu626Lys, C.) *LDLR* Arg574Leu, D.) *LDLRAP1* Arg151Trp.

### 5.1.3 Utilizing a large population cohort to estimate the impact of genetic variants

After identifying novel hypercholesterolemia-associated variants, I wanted to explore their association with hypercholesterolemia and further analyze their impact on ASCVD risk on a continuous scale. I utilized the FinnGen study and its longitudinal health records and searched carriers of known FH mutations to compare the impact of known FH mutations with the novel variants.

#### How to study the impact of genetic variants on diseases?

One of the difficulties in interpreting novel genetic variants found from a typical WES study, for example, is deciphering their pathogenicity. Though there are tens of *in silico* prediction algorithms for this task, they are less than perfect. Besides the fact that the algorithms may produce conflicting results, they also do not consider the phenotype in question nor the effect on a continuous trait, such as LDL-C. Different functional studies, e.g., cell models, have been used, but they are often expensive and laborious and thus not practical for a vast number of variants. Likewise, in the case of FH, no information on ASCVD risk can be received.

We decided to utilize the FinnGen study with longitudinal health registry data and calculate the ASCVD risk associated with a given variant. Since the FinnGen data does not contain lipid measurements, I used prescriptions for lipid-lowering medication as a surrogate measure, as well as coronary revascularization and stroke as disease endpoints to study the impact on ASCVD risk. I defined the use of lipid-lowering medication as a minimum of three purchases (ATC code C10). The data was gathered from the prescription drug purchase register data (Kela), whereas the disease endpoints were retrieved from the National Hospital Discharge Register. Coronary revascularization included coronary angioplasty and coronary artery bypass grafting, and stroke endpoint excluded subarachnoidal hemorrhage.

#### Impact on ASCVD risk varies by gene and mutation

Only four out of the six variants identified in WES were included in the genotyping chips used in the FinnGen study: *LDLR* Arg574Leu, *LDLR* Glu626Lys, *APOB* Pro877Leu and *LDLRAP1* Arg151Trp (*LDLRAP1* Arg151Trp was also included in the imputed data, which I used in the later analyses because of higher number of variant carriers). Out of the seven Finnish-enriched, known FH mutations in *LDLR* (**Table 2,** page 33), I

found carriers of five mutations (**Table 5**), of which *LDLR* Asp579Tyr has the same allelic position as FH-Fin 11.

**Table 5. Carriers of hypercholesterolemia-associated variants in FinnGen.** FinnGen carriers of five previously known familial hypercholesterolemia (FH) mutations and four novel hypercholesterolemia-associated variants, identified in Study I. Lipid-lowering medication prescription purchases or coronary revascularization and stroke diagnosis among carriers. *LDLRAP1* Arg151Trp carriers were identified from imputed genotype data, others from genotype data. N (%) reported.

| Variant | N | Lipid-lowering medication | Coronary revascularization | Stroke |
|---|---|---|---|---|
| **FH-Pogosta** | 86 (44.2%) | 50 (58.1%) | 7 (8.1%) | 11 (12.8%) |
| **FH-Pori** | 18 (38.9%) | 5 (27.8%) | 1 (5.6%) | 1 (5.6%) |
| **FH-Turku** | 24 (33.3%) | 19 (79.2%) | 2 (8.3%) | 1 (4.2%) |
| ***LDLR* Asp579Tyr** | 22 (35.4%) | 11 (50.0%) | 3 (13.6%) | 1 (4.5%) |
| **FH-Fin 11** | 26 (34.6%) | 17 (65.4%) | 5 (19.2%) | 3 (11.5%) |
| ***LDLR* Arg574Leu** | 117 (35.0%) | 62 (53.0%) | 13 (11.1%) | 11 (9.4%) |
| ***LDLR* Glu626Lys** | 131 (43.5%) | 48 (36.6%) | 11 (8.4%) | 10 (7.6%) |
| ***APOB* Pro877Leu** | 218 (45.4%) | 67 (30.7%) | 7 (3.2%) | 15 (6.9%) |
| ***LDLRAP1* Arg151Trp** | 688 (42.2%) | 258 (37.5%) | 57 (8.3%) | 36 (5.2%) |

**Table 6. Odds ratios for different endpoints in carriers of hypercholesterolemia-associated variants in FinnGen.** Odds ratios [confidence intervals, CI] and P values of different endpoints of hypercholesterolemia-associated variants and persons with polygenic risk score (PRS) in the top 10% of the FinnGen cohort, carriers identified from the FinnGen study. Known familial hypercholesterolemia (FH) mutations include FH-Pogosta, FH-Pori, FH-Turku, *LDLR* Asp579Trp, and FH-Fin 11. Novel mutations in *LDLR* include *LDLR*Arg574Leu and Glu626Lys. Odds ratios (95% CI) were estimated using logistic regression and adjusted with age and sex. Persons using lipid-lowering medication were excluded from the calculation of coronary revascularization and stroke endpoints. Modified from the original publication (I) under the terms of the Creative Commons Attribution License (CC BY 4.0).

| Mutation | Lipid-lowering medication | *p* | Coronary revascularization | *p* | Stroke | *p* |
|---|---|---|---|---|---|---|
| **Known FH mutations (N=5)** | 4.7 [3.3-6.8] | **5.2e-17** | 2.7 [1.3-5.6] | **0.01** | 3.3 [1.7-6.4] | **0.0005** |
| **Novel mutations in *LDLR* (N=2)** | 2.1 [1.5-2.8] | **1.8e-6** | 2.0 [1.1-3.7] | **0.03** | 2.1 [1.2-3.8] | **0.01** |
| **All mutations in *LDLR* (N=7)** | 2.9 [2.3-3.7] | **1.1e-19** | 2.2 [1.4-3.7] | **0.001** | 2.5 [1.6-3.9] | **3.2e-5** |
| ***LDLRAP1* Arg151Trp** | 1.4 [1.1-1.6] | **0.0005** | 1.8 [1.2-2.7] | **0.008** | 0.8 [0.5-1.4] | 0.49 |
| **LDL-PRS top10%** | 1.9 [1.9-2.0] | **0.0** | 1.6 [1.5-1.7] | **8.5e-35** | 1.3 [1.2-1.4] | **1.1e-14** |
| **CAD-PRS top10%** | 1.9 [1.8-1.9] | **0.0** | 3.7 [3.5-3.9] | **0.0** | 1.4 [1.3-1.5] | **1.1e-27** |

First, we wanted to see if the novel variants were associated with the use of lipid-lowering medication (**Tables 5** and **6**). That was the case with known Finnish FH-mutations (OR 4.7, $p$ = 5e-17), as well as novel variants *LDLR* Arg574Leu and Glu626Lys (OR 2.1, $p$ = 2e-6), in addition to *LDLRAP1* Arg151Trp (OR 1.4, $p$ = 5e-4). However, among the *APOB* Pro877Leu carriers, only 30.7% used lipid-lowering medication, compared to 31.7% in the whole cohort. Thus, we did not consider this variant to be associated with hypercholesterolemia.

A comparison of ASCVD prevalence by the use of lipid-lowering medication showed that medication use significantly decreased the risk for ASCVD (**Figure 14**, next page)**.** Hence, I excluded individuals using lipid-lowering medication before coronary revascularization or stroke from the analysis of ASCVD risk. Risk for both coronary revascularization and stroke was increased in the 94 carriers of known FH mutations (OR 2.7, $p$ = 0.01 and OR 3.3, $p$ = 0.0005) and 157 carriers of novel *LDLR* mutations (OR 2.0, $p$ = 0.03 and OR 2.1, $p$ = 0.01). Additionally, *LDLRAP1* Arg151Trp was associated with coronary revascularization (OR 1.8, $p$ = 0.008) in the 377 carriers.

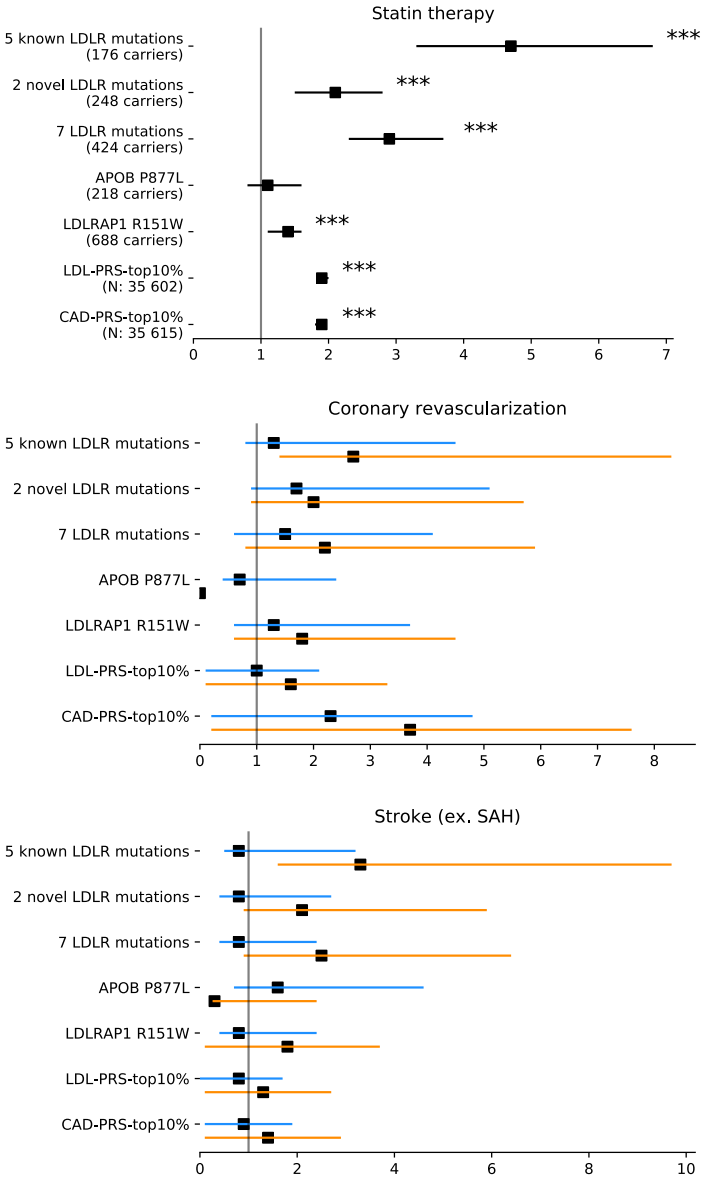



**Figure 14. Odds ratios for ASCVD-related endpoints in FinnGen.** Odds ratios of different endpoints of hypercholesterolemia-associated variants and persons with high polygenic risk were estimated with logistic regression and adjusted with age and sex. Known FH mutations include FH-Pogosta, FH-Pori, FH-Turku, *LDLR* Asp579Trp, and FH-Fin 11. Novel mutations in *LDLR* include *LDLR* Arg574Leu and Glu626Lys. P-values in top figure: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. Figure on top: use of lipid-lowering medication, middle: coronary revascularization, and at the bottom: stroke (excluding SAH, subarachnoidal hemorrhage). Horizontal lines represent 95% confidence intervals. Blue lines denote persons using lipid-lowering medication at the time of the disease event, and orange lines persons not using lipid-lowering medication (six months before the event). Top figure from the original publication (I) under the terms of the Creative Commons Attribution License (CC BY 4.0).

The approach of studying the impact of variants on ASCVD risk using a large population cohort proved to be a useful method. The limiting factor is that carriers of infrequent variants are difficult to find in population cohorts, especially in the numbers needed to reach statistical significance.

### Impact of polygenic variants on ASCVD risk

Further, I used previously calculated PRSs for LDL-C as well as CAD and included persons with PRSs in the top 10$^{th}$ percentile in the analyses, i.e., 35 602 individuals with high LDL-PRS and 35 615 with high CAD-PRS. Both groups had increased odds for lipid-lowering medication (OR 1.9).

Again, to calculate the risk for ASCVD, I excluded persons with previous use of lipid-lowering medication. Both groups had increased ASCVD risk: OR 1.6 ($p$ = 9e-35) for coronary revascularization and 1.3 ($p$ = 1e-14) for stroke in persons with high LDL-PRS (N=17 891), correspondingly OR 3.7 ($p$ = 0.0) and 1.4 ($p$ = 1e-27) in high CAD-PRS (**Table 6**, **Figure 14**). Hereby, the overall highest ASCVD risk was in individuals having high CAD-PRS, even when compared with carriers of FH mutations.

### 5.1.4 Individuals with high genetic risk for ASCVD are crucial to identify and treat

FH mutation carriers have a significant risk for early-onset ASCVD [20,150,266]. However, when treated with lipid-lowering medication, ASCVD risk is reduced to the level of a non-carrier [155]. Thus, it is crucial to identify the mutation carriers and treat them early to prevent detrimental ASCVD events. It is worth noting, however, that the two Finnish *LDLR* deletions were not included in our analyses since they could not be captured with genotype chips used in the FinnGen study. Via higher LDL-C levels, the ASCVD risk might be greater than the risk of missense mutations. Still, the highest ASCVD risk was among persons with high CAD-PRS. Regarding risk prevention, there are a lot more individuals in this group than there are FH mutation carriers (i.e., 10% with high CAD-PRS vs. <0.5% carrying FH mutations in the population). Nonetheless, it would be important to identify and treat both groups.

To summarize our genetic results, we could see a continuous scale. Different mutations in *LDLR* were associated with varying levels of ASCVD risk. However, polygenic clustering caused an equal or even higher risk. Traditionally, however, FH has been thought to be a distinct entity, severe, and separate from so called polygenic hypercholesterolemia. Our results demonstrate that even regarding FH mutations, the risk is variable, and the clustering of polygenic variants caused the highest ASCVD risk. If we would have had access to the lipid values of the mutation carriers, we could have similarly seen a continuous scale of the LDL-C levels of the carriers.

## 5.2 Impact of behavioral factors and attitudes on hypercholesterolemia and ASCVD prevention (II)

Several lifestyle factors, especially diet, affect LDL-C levels and ASCVD risk. Not as much is known about the effects of personal attitudes. For Study II, we selected cases with LDL-C ≥ 4.6 mmol/l (top 10$^{th}$ percentile) and controls with LDL-C < 3.4 mmol/l (mean LDL-C) from the GeneRISK study. After excluding individuals using lipid-lowering medication at baseline, as well as carriers of hypercholesterolemia-causing mutations (identified in Study I), we had 701 cases (423 females and 278 males) and 3 128 controls (2 181 females and 947 males). The aim of Study II was to assess the impact of behavioral factors and attitudes on hypercholesterolemia. Since the cases were undertreated at baseline, the further goal was to estimate ASCVD intervention measures in healthcare.

### 5.2.1 Hypercholesterolemia cases have more behavioral risk factors

At baseline, most of the cases (96.4%) were not using lipid-lowering medication despite 83% being previously diagnosed with hypercholesterolemia. Nonetheless, I excluded persons using lipid-lowering medication from the subsequent analyses. Because ASCVD risk and risk factors have evident differences between females and males [92,203], I conducted most analyses separately (**Table 7**, next page). In general, there was an apparent genetic background of hypercholesterolemia: hypercholesterolemia cases had higher LDL-PRSs and a family history of hypercholesterolemia and early-onset MI more frequently compared to controls. Still, cases had significant differences in behavioral factors that can affect their LDL-C levels as well as ASCVD risk. Interestingly, there were a lot of differences between female and male cases.

In the female-specific analysis, cases had more traditional ASCVD risk factors: they were older with more menopausal individuals, and they had more cardiometabolic risk factors. Regarding behavioral factors, female cases were more often smokers than controls (13.1% vs. 7.9%, $p$=0.0005), and they exercised less (2.8 vs. 3.0, $p$=2.9e-5). Also, their self-reported health was significantly lower (7.8 vs. 8.0, $p$=0.002). In the

male analysis, there were fewer differences between cases and controls. However, male cases had significantly less healthy diets (6.7 vs. 7.4, $p$=0.001) than male controls. Male cases also had less healthy diets than female cases (6.7 vs. 8.6, $p$=1.1e-13). Regarding socioeconomic factors, male cases were more often cohabiting or married than female cases (82.7% males vs. 65.5% females, $p$=9.2e-7).

**Table 7. Genetic factors, behavioral factors, and attitudes of cases with hypercholesterolemia, compared to controls, reported separately for females and males.** Cases were defined as LDL-C ≥ 4.6 mmol/l, controls LDL-C < 3.4 mmol/l. All persons using lipid-lowering medication at baseline or carrying hypercholesterolemia-associated mutations were excluded. Statistical tests: Student's t-test was used for continuous variables, except for variables BMI, glucose, triglycerides, and self-reported health, where the Mann-Whitney-U-test was used. $\chi^2$ test was used for categorical variables. Mean (standard deviation) is reported for continuous variables and N (%) for categorical variables. NS: non-significant, RFDS: Recommended Finnish Diet Score. Variables are described in more detail in chapter Methods 4.3, page 51. Modified from the manuscript Junna et al. (2024).

| | **Female cases** | **Female controls** | ***p*** | **Male cases** | **Male controls** | ***p*** |
|---|---|---|---|---|---|---|
| **N** | 423 | 2 181 | | 278 | 947 | |
| **Previously diagnosed hypercholesterolemia** | 352 (83.2%) | 580 (26.6%) | **1.2e-105** | 229 (82.4%) | 283 (29.9%) | **1.2e-55** |
| **Family history of early-onset MI** | 110 (26.0%) | 436 (20.0%) | **0.006** | 52 (18.6%) | 162 (17.1%) | NS |
| **Family history of hypercholesterolemia** | 309 (73.0%) | 1 265 (58.0%) | **9.5e-9** | 170 (61.2%) | 401 (42.3%) | **4.8e-8** |
| **LDL-PRS** | 0.52 (0.87) | -0.33 (1.03) | **7.1e-54** | 0.48 (0.85) | -0.43 (0.99) | **1.7e-39** |
| **Daily smokers** | 56 (13.1%) | 172 (7.9%) | **0.0005** | 37 (13.2%) | 114 (12.0%) | NS |
| **RFDS food score** | 8.6 (3.0) | 8.6 (3.2) | NS | 6.7 (3.1) | 7.4 (3.3) | **0.001** |
| **Exercise at free time** | 2.8 (0.9) | 3.0 (0.9) | **2.9e-5** | 2.9 (1.1) | 3.0 (1.0) | NS |
| **Self-reported health level** | 7.8 (1.2) | 8.0 (1.2) | **0.002** | 7.7 (1.3) | 7.8 (1.3) | NS |
| **Normal cholesterol not important** | 36 (8.5%) | 191 (8.8%) | NS | 53 (19.1%) | 138 (14.6%) | NS |
| **Marriage or cohabitation** | 277 (65.5%) | 1 543 (70.7%) | **0.04** | 230 (82.7%) | 741 (78.2%) | NS |
| **Higher education** | 117 (27.7%) | 731 (33.5%) | **0.02** | 73 (26.3%) | 282 (29.8%) | NS |
| **Menopausal** | 337 (79.7%) | 1 346 (61.7%) | **1.6e-12** | | | |

Despite female cases being older and having worse cardiometabolic profiles than female controls, only 31.4% had been estimated to be at high risk for ASCVD, compared to 53.6% of male cases. The shortcomings of commonly used ASCVD risk models in estimating female risk are known. For example, they miss many female-specific risk factors, such as gestational diabetes [92,203]. Additionally, risk models commonly estimate a 10-year risk, whereas female ASCVD risk starts to increase later in life [212]. Since traditional CVD risk estimation models give significant weight to the male sex, these females with several risk factors for ASCVD might not be adequately recognized in healthcare. Furthermore, if the individuals themselves do not know to be at risk, they might not be motivated to make positive lifestyle modifications.

### 5.2.2 Poor attitude to cholesterol more common in males

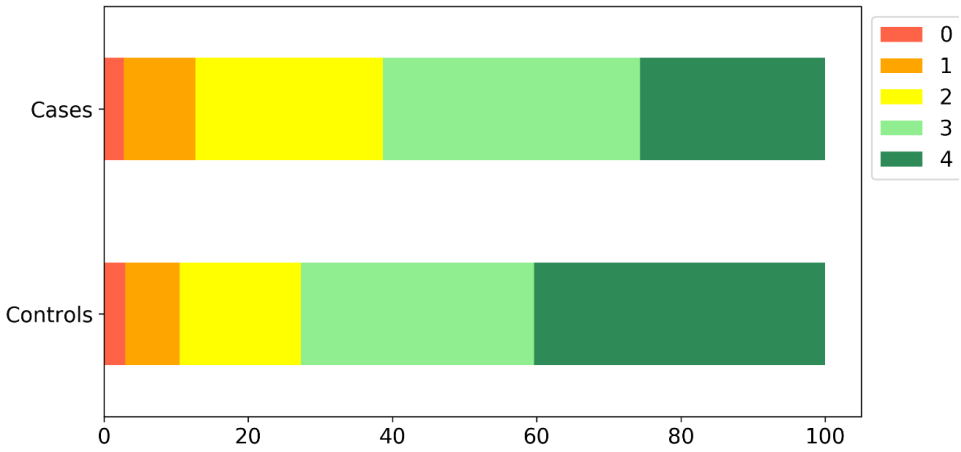



**Figure 15. Attitudes to normal cholesterol among hypercholesterolemia cases and controls.** Distribution of answers to the question: "How important are the following health-related subjects to you? Keeping your cholesterol level within normal limits.", cases (LDL-C ≥ 4.6 mmol/l, N=701) and controls (LDL < 3.4 mmol/l, N=3 128). Answer options: 0: "not at all important", 1: "a little important", 2: "moderately important", 3: "quite important", 4: "very important". The X-axis depicts percentages. From the manuscript Junna et al. (2024).

There is evidence that a positive attitude to the treatment of hypercholesterolemia is associated with LDL-C levels even in persons using lipid-lowering medication [18]. In the GeneRISK study, most study participants thought that keeping their cholesterol within normal limits is important (87.3% of cases, 89.5% of controls, **Figure 15**). Sex-specific analyses showed no significant difference in attitudes to normal cholesterol within the sexes. However, male cases had significantly more individuals than females

thinking that normal cholesterol is not important, e.g., 8.5% of female cases vs. 19.1% of male cases ($p$=6.6e-5). Moreover, a poor attitude (i.e., normal cholesterol not at all or only a little important) was more common in male cases than controls. However, the difference was not statistically significant (19.1% cases vs. 14.6% controls).

I performed a subgroup analysis of the hypercholesterolemia cases who had poor attitudes (36 females and 53 males) compared to cases with good attitudes (i.e., normal cholesterol quite or very important; 288 females and 142 males) (**Table 8**). Cases with poor attitudes had less healthy diets: 6.1 vs. 7.2 ($p$=0.05) in males and 7.4 vs. 8.8 ($p$=0.01) in females, correspondingly. Moreover, regarding cases with poor attitudes, males were more often smokers (18.9% vs. 9.9%, non-significant), and females exercised less (2.5 vs. 2.9, $p$=0.02). Our sample sizes were relatively small, which might have affected the significance of the results. Some studies imply that attitudes can impact health-related behavior [16,17]. Therefore, this group of individuals might be beneficial to identify in healthcare.

**Table 8. Comparison of cases with LDL-C ≥ 4.6 mmol/l, with a poor or good attitude to normal cholesterol.** Behavioral factors of hypercholesterolemia cases were divided by attitude to normal cholesterol (question "How important are the following health-related subjects to you? Keeping your cholesterol level within normal limits."). "No" refers to answers "not at all" or "a little" important. "Yes" refers to answers "quite" or "very" important. Statistical tests: Student's t-test was used for continuous variables, except for self-reported health, where the Mann-Whitney-U-test was used. $\chi_2$ test was used for categorical variables NS: non-significant.

| | Females | | | Males | | |
|---|---|---|---|---|---|---|
| **Normal cholesterol important** | **No** | **Yes** | ***p*** | **No** | **Yes** | ***p*** |
| **N (%)** | 36 (8.5%) | 288 (68.1%) | | 53 (19.1%) | 142 (51.2%) | |
| **Age** | 55.1 (4.4) | 57.5 (5.1) | **0.007** | 55.7 (5.4) | 55.6 (5.5) | NS |
| **Daily smokers** | 4 (11.1%) | 37 (12.8%) | NS | 10 (18.9%) | 14 (9.9%) | NS |
| **RFDS food score** | 7.4 (3.0) | 8.8 (3.0) | **0.01** | 6.1 (3.3) | 7.2 (3.1) | **0.05** |
| **Exercise at free time** | 2.5 (1.1) | 2.9 (0.9) | **0.02** | 2.9 (1.1) | 3.0 (1.0) | NS |
| **Self-reported health** | 7.7 (1.5) | 7.8 (1.2) | NS | 7.7 (1.4) | 7.7 (1.4) | 0.98 |

### 5.2.3 ASCVD prevention leaves room for improvement

#### ASCVD prevention measures before the study

As mentioned, 83% of the hypercholesterolemia cases had previously been diagnosed with hypercholesterolemia, although they were not using lipid-lowering medication at baseline. However, 73.5% (N=581) of them had received dietary instructions, and 46.8% recommendations to monitor their cholesterol levels. However, 19.3% reported not receiving any recommendations to lower their LDL-C levels. No tested variable (e.g., sex, attitude to cholesterol, ASCVD risk) was significantly associated with receiving or not receiving LDL-C-reducing recommendations.

#### ASCVD prevention measures after the baseline visit

After the baseline visit, all study participants with high ASCVD risk (40.2% of hypercholesterolemia cases) were recommended to visit their own physician. The study was continued by 484 (69.0%) cases and 2 197 (70.2%) controls. Unfortunately, only 26.7% of the hypercholesterolemia cases reported visiting their physician or a nurse. As a result of the visit, based on self-reported answers, 60.5% received recommendations for cholesterol or glucose monitoring, 15.5% dietary advice, and 34.1% prescription for lipid-lowering medication. However, 27.1% reported not receiving any of these recommendations to lower their ASCVD risk. Regarding sex differences, male cases received dietary advice more frequently (22.7% vs. 11.8%, non-significant), which could be explained by males having less healthy diets at baseline. As with baseline analysis, no tested variable could explain whether the person received recommendations.

#### Prescriptions for lipid-lowering medications

Given that not all study participants participated in the follow-up study, I utilized the National Health Registry, including all prescription medication purchases. According to the data, 36.7% of hypercholesterolemia cases not attending the follow-up received a prescription for lipid-lowering medication after the baseline visit, compared to 35.1% in cases attending the follow-up. Thus, we can assume that the cases not participating in the follow-up were as likely to see a physician as the other group. That also applies to attitude to cholesterol: persons with poor attitude were as likely to receive a prescription, i.e., visit a physician. Regarding prescriptions, cases with high

ASCVD risk were more likely to receive one (54.0% vs. 32.6%, *p*=4.9e-8), which is in accordance with national guidelines.

### 5.2.4 The reasons behind the undertreatment of hypercholesterolemia remain unknown

Study II confirms that middle-aged persons with hypercholesterolemia are undertreated in Finland. International guidelines [210,211], which Finnish ones lean on [259], emphasize lifestyle modifications in ASCVD prevention after evaluating the personal ASCVD risk. Our baseline analysis is in line with the guidelines; most had received some recommendations regarding their elevated LDL-C levels. Nonetheless, LDL-C ≥ 5 mmol/l produces high risk by itself, in which case lipid-lowering medication should be initiated. Still, most of our cases reaching this LDL-C level did not use lipid-lowering medication at baseline. The same trend continued at follow-up. Finnish guidelines state that if the LDL-C target (e.g., < 3 mmol/l for all adults, < 2.6 mmol/l if the ASCVD risk is average) is not reached within 3-6 months, more efficient measures should be started, generally meaning lipid-lowering medication. Yet, only ~35% of the hypercholesterolemia cases who had visited a physician received lipid-lowering prescriptions after baseline, although the target had clearly not been reached. Our analyses did not reveal the reason for the poor level of treatment. Perhaps the underlying reasons are so complex that we could not discover them in this study.

In the end, the point of recognizing risk factors and their impacts is to be able to modify the risk. Unfortunately, only ~25% of hypercholesterolemia cases went to see a physician despite all of them receiving information about their elevated LDL-C levels. Hypercholesterolemia cases with high risk – who were instructed to see a physician - were more likely to do so, but still only 36.5% followed the advice. Nonetheless, visiting a physician does not help if lipid-lowering medication is not prescribed as the guidelines recommend. However, since our analyses were based on self-reported data, it remains unclear how the visits went and if the cases declined medical treatment, for instance.

Besides the point of medication, behavioral factors were less healthy among hypercholesterolemia cases. Lifestyle modifications can decrease ASCVD risk, which we studied in more detail in Study III.

## 5.3 Impact of returning polygenic risk for ASCVD on health behavior (III)

Risk estimates for ASCVD, primarily based on non-genetic risk factors, have been reported to patients to educate them about their risk and motivate them to make healthy lifestyle modifications. Furthermore, adding PRS to traditional risk estimates can significantly improve their predictive value [25,271]. However, thus far, polygenic risk has not been added to risk estimates, nor are there guidelines for doing so. In Study III, we sought to investigate how persons would react to their personal ASCVD risk when polygenic risk is also included and whether this risk information would encourage positive health behavior, especially among individuals with high risk for ASCVD. At baseline, 7 342 individuals were recruited to the GeneRISK study, of which 5 191 individuals participated in the follow-up study, 1.5 years later, on average.

### 5.3.1 Attitudes to polygenic risk mostly positive

After the baseline visit, all study participants received their risk for ASCVD via the web tool Kardiokompassi®. Most of the participants (89.7%) viewed their results. The distribution of ASCVD risk (low, average, high, or high risk due to risk-increasing conditions) is shown in **Figure 16** (next page); average risk was the most common risk group (54.6%).

In the follow-up e-questionnaire, filled by 5 191 study participants, we asked how the participants perceived their risk results, and 88% thought their risk was easy to understand. Moreover, only 22% reported receiving unexpected risk results, and 29% were concerned. Expectedly, concern was associated with higher clinical ASCVD risk (OR 1.20, $p$=1.2e-3) and higher PRS (OR 1.54, $p$=3.3e-14), but with female sex (OR 0.50, $p$=1.4e-15), higher education (OR 1.98, $p$=1.5e-3), and age (OR 0.98, $p$=3.5e-2) as well. However, all participants received their polygenic risk, and we were not able to compare receiving only traditional risk to receiving also polygenic risk.

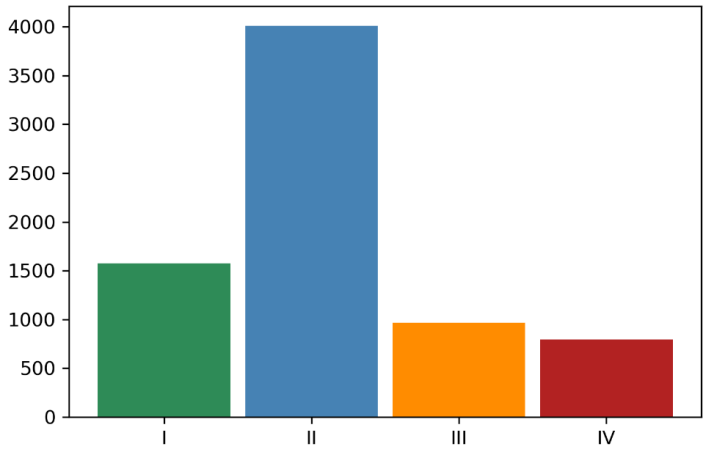



**Figure 16. Distribution of 10-year ASCVD risk categories of the GeneRISK study participants at baseline.** Risk of developing atherosclerotic cardiovascular disease (ASCVD) in the next 10 years; I: low risk, <2% (N=1 573); II: average risk, 2-10% (N=4 009); III: high risk (N=966), ≥10%; IV: high risk (N=794) due to risk-increasing condition (diabetes, angina pectoris, heart failure, fasting glucose ≥7 mmol/l or LDL-C ≥6 mmol/l).

Furthermore, 97% reported believing that genetic factors importantly influence their disease risk, but also 99% thought they can impact their disease risk through their lifestyle. Thus, genetic and behavioral factors were not considered to be mutually exclusive. Moreover, 89% felt that their personal risk information motivates them to take better care of their health.

### 5.3.2 Having high risk for ASCVD motivates positive health behavior

Next, we sought to explore whether the positive attitudes and motivation of the study participants would lead to making lifestyle modifications. Positive health behavior was defined as signing up for health coaching, losing weight, smoking cessation, or visiting a physician.

We were especially interested in the high-risk group (N=683), of which 42.6% made one or more of the abovementioned actions: 15.4% signed up for health coaching, 12.4% lost weight, 14.2% of smokers quit smoking, and 20.8% visited their physician. Study participants with ASCVD risk <10% (N=3 993) made significantly fewer behavioral changes (33.5%, $p$<0.001), e.g., visited a physician less frequently

(8.3%, $p$<0.001). Consequently, we speculate that having a high ASCVD risk had a motivating effect on making healthy lifestyle modifications.

We looked further into persons who reported visiting a physician. Among the high-risk group, persons who visited their physicians, LDL-C lowered by 0.5 mmol/l, on average, and systolic blood pressure by 7 mmHg, correspondingly. Results were similar among persons who lost weight.

### 5.3.3 Guidelines for polygenic risk estimates still missing

Our study shows the positive effects of returning polygenic risk to study participants. Generally, the risk was easy to understand and did not cause much concern. Moreover, the risk information was generally considered motivational to make lifestyle modifications – even if the risk was high. At present, writing this thesis, there are no official guidelines on how polygenic risk estimates should be constructed – or interpreted. However, commercial tests for polygenic risk for common diseases are available, and patients might approach their physicians with the results. Thus, it is important to form guidelines for interpreting PRSs and to evaluate their quality.

## 5.4 Limitations

This thesis has some limitations. The studies composing this thesis utilized mainly the GeneRISK cohort, consisting of middle-aged Finns without previously diagnosed ASCVD. Thus, the results may not be generalizable in other populations or age groups. Exclusion of patients with previous ASCVD might have excluded cases with FH, particularly carriers of large deletions, which can cause early-onset ASCVD. In Study I, WES and genotyping were used to discover mutations, leaving possible novel large insertions and deletions uncaptured. Moreover, no functional testing on the novel mutations was performed, and further studies are needed to investigate their functional effects. In Study II, the sample sizes were modest. A larger sample could bring more insights into the behavioral factors affecting hypercholesterolemia, which may have been missed due to the lack of statistical power. Further, Study II focused on self-reported variables prone to possible biases: participants might not be truthful in their answers or just not remember their previous behaviors accurately.

# 6 CONCLUSIONS AND FUTURE DIRECTIONS

The studies composing this thesis highlight the complexity and variable effect of genetic and behavioral factors on hypercholesterolemia and ASCVD risk. High-impact mutations, polygenic variants, and behavioral factors all have their impact.

In Study I, the genetic architecture of severe hypercholesterolemia in Finland turned out to be more complex than previously thought, including three novel mutations. I studied the impact of novel and known hypercholesterolemia-causing mutations on ASCVD risk by utilizing a large population cohort. I discovered that the risk varies by gene and mutation, and high CAD-PRS causes the highest risk. In Study II, I focused on behavioral factors and attitudes, which differed in hypercholesterolemia cases compared to controls and between the sexes. Additionally, ASCVD prevention measures in Finland leave room for improvement: most cases were not treated at baseline, and the same trend continued at follow-up. Study III explored how the participants perceived their personal ASCVD risk, which also contained polygenic risk. Reactions to risk information were mainly positive, and individuals with high risk were particularly motivated to make lifestyle modifications, thus lowering their ASCVD risk.

In Study I, I revealed new information on the genetic architecture of severe hypercholesterolemia in Finland. Despite the complexity of the severe hypercholesterolemia phenotype, carriers of high-impact mutations are still worth identifying because of the higher ASCVD risk that the mutations predispose to [150,151,266]. However, the odds for ASCVD in individuals with high LDL-PRS were also increased. Traditionally, monogenic FH has been separated from hypercholesterolemia with polygenic origin, but accumulating evidence from studies has shown that LDL-C levels vary between mutations and individual carriers in addition to polygenic clustering causing a similar phenotype [7,12,13,158,266]. Our results are in line with these previous findings and show that the effect on ASCVD risk varies between different genes, variants, and polygenic clustering.

Our method to estimate the effect of different mutations on ASCVD risk utilizing a large population cohort FinnGen proved to be very useful. This method could be used to assess other hypercholesterolemia-associated variants, and it could be utilized with

other phenotypes as well. Even with our method to estimate the impact of different variants, functional testing still has its place. Especially in the case of *LDLRAP1* Arg151Trp, further functional tests would be required to determine its pathogenicity conclusively, which we were not able to do in Study I. It would also have been interesting to see the association of different hypercholesterolemia mutations with LDL-C levels, where we probably would have seen a continuous effect.

In Study II, I focused on behavioral factors and attitudes. While behavioral factors have been studied to some extent, to my knowledge, attitudes related to health behavior have not been widely researched. My analyses showed that a poor attitude to cholesterol was associated with less healthy lifestyles, even though most study participants thought that keeping their cholesterol within normal limits was important. Since persons with poor health-related attitudes might be less likely to attend health-related studies, such as the GeneRISK, the proportion of people with poor attitudes in the general population might be higher.

In accordance with previous studies [92,203], female ASCVD risk appeared to be underestimated in our study, though we did not analyze the used risk model and the effect of sex specifically. Accurate risk estimates are influential since persons aware of their high risk are more likely to modify their health behavior (as seen in Study III).

Poor level of treatment in hypercholesterolemia cases, as well as ASCVD prevention measures, was evident, but our data did not reveal the reasons behind it. The emphasis in international and national prevention guidelines is on lifestyle modifications. However, the effect of lifestyle modifications on LDL-C comes mostly through diet and is limited to around 0.6 mmol/l decrease [213–215], which is not enough for patients with severe hypercholesterolemia. Yet, since our data was self-reported by the participants, we cannot be certain what happened at the clinical visit.

Study III was among the first to return polygenic risk information to the study participants and to study their reactions and health behavior. The study showed that participants perceived their personal ASCVD risk primarily positively, and most considered it motivational in improving their health behavior, especially high-risk persons. Thus, receiving personal risk for ASCVD, including polygenic risk, might be beneficial for health and preventing diseases. Some other studies have later emerged, returning polygenic risk to study participants. Similar to Study III, participants perceived their risk information as useful and not very concerning [79,272].

As more studies have emerged, adding PRSs to risk estimates has started to prove their worth. Polygenic risk information could, e.g., aid in recognizing persons who would benefit from early treatment [75,273]. However, we are far from applying PRSs in standard healthcare. Regardless of the low costs of genotyping, there is a continuous demand for Finnish public healthcare to cut their costs. Moreover, CVD is a global burden [84,274]. Bringing PRSs to clinics in low-income countries is not likely to be feasible, and risk models should first be recalibrated to the population and country in question [75]. Furthermore, guidelines for calculating, utilizing, and interpreting polygenic risk information in the healthcare system are still lacking when writing this thesis.

# ACKNOWLEDGEMENTS

This doctoral thesis was carried out at the Institute for Molecular Medicine Finland (FIMM), Helsinki Institute of Life Sciences (HiLIFE), University of Helsinki. I want to acknowledge the directors of FIMM during my time, Olli Kallioniemi, Jaakko Kaprio, Mark Daly, and Samuli Ripatti, for providing a high-quality research facility and a scientifically inspiring work environment.

This thesis was financially supported by the Doctoral Programme in Population Health of the University of Helsinki, for which I am thankful.

I want to thank Johannes Kettunen, for accepting to act as my opponent. I am grateful to Katriina Aalto-Setälä and Matti Jauhiainen for taking the time to review my thesis and for the constructive comments that improved this thesis. I thank Eveliina Salminen for accepting to serve as the faculty representative at my defense.

I want to express my gratitude to my thesis committee members, Annukka Marjamaa and Maija Wessman, for their interest in my work and for always being positive about my progress.

This thesis would not have been possible without my supervisor, Elisabeth Widén. Thank you for giving me support and feedback over the years while still allowing me independence throughout my thesis project and enabling me to participate in interesting projects beyond this thesis.

I want to thank Samuli Ripatti for agreeing to be the custos (once again). Samuli has also been a vital part of all the studies composing this thesis and always had compact but insightful comments.

My co-author, Sanni Ruotsalainen, was a crucial part of Study III and has been a significant statistical advisor for me during these years. Sanni also helped me tremendously with the FinnGen data and its many quirks.

During my time at FIMM, I have had the privilege to work with many talented people. I appreciate all the collaboration, assistance and learning opportunities: Johanna Aro, FIMM-IT team, FIMM Tech Center, Heidi Hautakangas, Sini Kerminen, Sari Kivikko, Jaakko Leinonen, Nina Mars, Hanna M. Ollila, Aarno Palotie (and the FinnGen team), Matti Pirinen, Pietari Ripatti, Susanna Rosas, Sanni Ruotsalainen, Ida Surakka and Taru Tukiainen. Thank you for sharing your knowledge and expertise – related to and beyond my thesis - and settling in at FIMM.

I also want to thank all the other people who I shared space, time, and thoughts with: Elmo, Felix, Hande, Jade, Jana, Joel, Joonas, Juha, Jukka, Juulia, Linda, Max, Meri, Om, Paula, Pyry, Rubina, Sannimari, Satu, Shabbeer, and so many others.

Last but not least, I want to thank my family for their support over the years. Especially Eija and Kim for the many lengthy conversations, and Jaska, Raksu, Mörkö, Rusina, and Rambo for stress relief and relaxation.

Nella Junna
October 2024
Helsinki

# REFERENCES


1. Vaduganathan M, Mensah GA, Turco JV, Fuster V, Roth GA. The Global Burden of Cardiovascular Diseases and Risk: A Compass for Future Health. J Am Coll Cardiol. 2022 Dec 20;80(25):2361–71.

2. Kuolleisuus kasvoi vuonna 2021 eniten muistisairauksiin ja koronavirustautiin edellisvuodesta - Tilastokeskus [Internet]. 2022 [cited 2024 Apr 4]. Available from: https://www.stat.fi/julkaisu/cktdxrx6o4sv90b62jy6t7qbg

3. Hansson GK. Inflammation, Atherosclerosis, and Coronary Artery Disease. New England Journal of Medicine. 2005 Apr 21;352(16):1685–95.

4. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. Eur Heart J. 2017 Aug 21;38(32):2459–72.

5. Ference BA, Graham I, Tokgozoglu L, Catapano AL. Impact of Lipids on Cardiovascular Health: JACC Health Promotion Series. J Am Coll Cardiol. 2018 Sep 4;72(10):1141–56.

6. Müller C. Xanthomata, Hypercholesterolemia, Angina Pectoris. Acta Medica Scandinavica. 1938;95(S89):75–84.

7. Berberich AJ, Hegele RA. The complex molecular genetics of familial hypercholesterolaemia. Nat Rev Cardiol. 2019 Jan;16(1):9–20.

8. Faiz F, Hooper AJ, van Bockxmeer FM. Molecular pathology of familial hypercholesterolemia, related dyslipidemias and therapies beyond the statins. Crit Rev Clin Lab Sci. 2012 Feb;49(1):1–17.

9. Soutar AK, Naoumova RP. Mechanisms of disease: genetic causes of familial hypercholesterolemia. Nat Clin Pract Cardiovasc Med. 2007 Apr;4(4):214–25.

10. Vuorio AF, Aalto-Setälä K, Koivisto UM, Turtola H, Nissen H, Kovanen PT, et al. Familial hypercholesterolaemia in Finland: common, rare and mild mutations of the LDL receptor and their clinical consequences. Finnish FH-group. Ann Med. 2001 Sep;33(6):410–21.

11. Lahtinen AM, Havulinna AS, Jula A, Salomaa V, Kontula K. Prevalence and clinical correlates of familial hypercholesterolemia founder mutations in the general population. Atherosclerosis. 2015 Jan;238(1):64–9.

12. Saadatagah S, Jose M, Dikilitas O, Alhalabi L, Miller AA, Fan X, et al. Genetic basis of hypercholesterolemia in adults. NPJ Genom Med. 2021 Apr 14;6(1):28.

13. Talmud PJ, Shah S, Whittall R, Futema M, Howard P, Cooper JA, et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study. Lancet. 2013;381(9874):1293–301.

14. Trinder M, Li X, DeCastro ML, Cermakova L, Sadananda S, Jackson LM, et al. Risk of Premature Atherosclerotic Disease in Patients With Monogenic Versus Polygenic Familial Hypercholesterolemia. J Am Coll Cardiol. 2019 Jul 30;74(4):512–22.

15. Mensink RP, Organization WH. Effects of saturated fatty acids on serum lipids and lipoproteins: a systematic review and regression analysis [Internet]. World Health Organization; 2016 [cited 2023 Jan 14]. 72 p. Available from: https://apps.who.int/iris/handle/10665/246104

16. Ayala C, Tong X, Neeley E, Lane R, Robb K, Loustalot F. Home blood pressure monitoring among adults-American Heart Association Cardiovascular Health Consumer Survey, 2012. J Clin Hypertens (Greenwich). 2017 Jun;19(6):584–91.

17. Jokisalo E, Kumpusalo E, Enlund H, Takala J. Patients' perceived problems with hypertension and attitudes towards medical treatment. J Hum Hypertens. 2001 Nov;15(11):755–61.

18. Kitagawa Y, Teramoto T, Daida H, APPROACH-J Study group. Adherence to preferable behavior for lipid control by high-risk dyslipidemic Japanese patients under pravastatin treatment: the APPROACH-J study. J Atheroscler Thromb. 2012;19(9):795–805.

19. Abul-Husn NS, Manickam K, Jones LK, Wright EA, Hartzel DN, Gonzaga-Jauregui C, et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. Science. 2016 Dec 23;354(6319):aaf7000.

20. EAS Familial Hypercholesterolaemia Studies Collaboration, Vallejo-Vaz AJ, De Marco M, Stevens CAT, Akram A, Freiberger T, et al. Overview of the current status of familial hypercholesterolaemia care in over 60 countries - The EAS Familial Hypercholesterolaemia Studies Collaboration (FHSC). Atherosclerosis. 2018 Oct;277:234–55.

21. Nordestgaard BG, Chapman MJ, Humphries SE, Ginsberg HN, Masana L, Descamps OS, et al. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. Eur Heart J. 2013 Dec;34(45):3478–3490a.

22. Bitton A, Gaziano T. The Framingham Heart Study's Impact on Global Risk Assessment. Progress in cardiovascular diseases. 2010 Aug;53(1):68.

23. Mega JL, Stitziel NO, Smith JG, Chasman DI, Caulfield M, Devlin JJ, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. Lancet. 2015 Jun 6;385(9984):2264–71.

24. Natarajan P, Young R, Stitziel NO, Padmanabhan S, Baber U, Mehran R, et al. Polygenic Risk Score Identifies Subgroup With Higher Burden of Atherosclerosis and Greater Relative Benefit From Statin Therapy in the Primary Prevention Setting. Circulation. 2017 May 30;135(22):2091–101.

25. Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, et al. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. Nat Med. 2020 Apr;26(4):549–57.

26. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860–921.

27. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science. 2001 Feb 16;291(5507):1304–51.

28. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 2004 Oct 21;431(7011):931–45.

29. Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, et al. The complete sequence of a human genome. Science. 2022 Apr;376(6588):44–53.

30. International HapMap Consortium. The International HapMap Project. Nature. 2003 Dec 18;426(6968):789–96.

31. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007 Oct;449(7164):851–61.

32. Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. Nature. 2010 Sep;467(7311):52–8.

33. Altshuler D, Donnelly P, The International HapMap Consortium. A haplotype map of the human genome. Nature. 2005 Oct;437(7063):1299–320.

34. 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. Nature. 2010 Oct 28;467(7319):1061–73.

35. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. Nature. 2015;526(7571):75–81.

36. A global reference for human genetic variation. Nature. 2015;526(7571):68–74.

37. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012 Nov 1;491(7422):56–65.

38. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016 Aug 18;536(7616):285–91.

39. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020 May;581(7809):434–43.

40. Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, et al. A genomic mutational constraint map using variation in 76,156 human genomes. Nature. 2024 Jan;625(7993):92–100.

41. About gnomAD | gnomAD [Internet]. [cited 2024 Jan 11]. Available from: https://gnomad.broadinstitute.org/about

42. Nelis M, Esko T, Mägi R, Zimprich F, Zimprich A, Toncheva D, et al. Genetic Structure of Europeans: A View from the North–East. PLOS ONE. 2009 May 8;4(5):e5472.

43. Uusimaa J, Kettunen J, Varilo T, Järvelä I, Kallijärvi J, Kääriäinen H, et al. The Finnish genetic heritage in 2022 – from diagnosis to translational research. Disease Models & Mechanisms. 2022 Oct 26;15(10):dmm049490.

44. Lim ET, Würtz P, Havulinna AS, Palta P, Tukiainen T, Rehnström K, et al. Distribution and Medical Impact of Loss-of-Function Variants in the Finnish Founder Population. PLOS Genetics. 2014 Jul 31;10(7):e1004494.

45. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463–7.

46. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. Hum Immunol. 2021 Nov;82(11):801–11.

47. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet. 2016 May 17;17(6):333–51.

48. Pfeifer SP. From next-generation resequencing reads to a high-quality variant data set. Heredity (Edinb). 2017 Feb;118(2):111–24.

49. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.

50. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biology. 2016 Jun 6;17(1):122.

51. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014 Jan;42(Database issue):D980-985.

52. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003 Jul 1;31(13):3812–4.

53. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010 Apr;7(4):248–9.

54. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014 Apr;11(4):361–2.

55. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Research. 2019 Jan 8;47(D1):D886–94.

56. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 2011 Sep 27;12(11):745–55.

57. Goh G, Choi M. Application of whole exome sequencing to identify disease-causing variants in inherited human diseases. Genomics Inform. 2012 Dec;10(4):214–9.

58. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature. 2009 Sep 10;461(7261):272–6.

59. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, et al. Exome sequencing identifies the cause of a mendelian disorder. Nat Genet. 2010 Jan;42(1):30–5.

60. Hoischen A, van Bon BWM, Gilissen C, Arts P, van Lier B, Steehouwer M, et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. Nat Genet. 2010 Jun;42(6):483–5.

61. Sikkema-Raddatz B, Johansson LF, de Boer EN, Almomani R, Boven LG, van den Berg MP, et al. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. Hum Mutat. 2013 Jul;34(7):1035–42.

62. Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. Proc Natl Acad Sci U S A. 1994 May 24;91(11):5022–6.

63. Bumgarner R. Overview of DNA microarrays: types, applications, and their future. Curr Protoc Mol Biol. 2013 Jan;Chapter 22:Unit 22.1.

64. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet. 2010 Jul;11(7):499–511.

65. Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, et al. Complement Factor H Variant Increases the Risk of Age-Related Macular Degeneration. Science. 2005 Apr 15;308(5720):419–21.

66. Reich DE, Lander ES. On the allelic spectrum of human disease. Trends Genet. 2001 Sep;17(9):502–10.

67. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. Rare Allele Hypotheses for Complex Diseases. Curr Opin Genet Dev. 2009 Jun;19(3):212–9.

68. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. Am J Hum Genet. 2017 Jul 6;101(1):5–22.

69. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. PLoS Comput Biol. 2012;8(12):e1002822.

70. Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA. 2008 Mar 19;299(11):1335–44.

71. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019 Aug;20(8):467–84.

72. Fitipaldi H, Franks PW. Ethnic, gender and other sociodemographic biases in genome-wide association studies for the most burdensome non-communicable diseases: 2005–2022. Human Molecular Genetics. 2023 Feb 1;32(3):520–32.

73. Popejoy AB, Fullerton SM. Genomics is failing on diversity. Nature. 2016 Oct;538(7624):161–4.

74. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet. 2018 Sep;19(9):581–90.

75. O'Sullivan JW, Raghavan S, Marquez-Luna C, Luzum JA, Damrauer SM, Ashley EA, et al. Polygenic Risk Scores for Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation. 2022 Aug 23;146(8):e93–118.

76. Klarin D, Natarajan P. Clinical utility of polygenic risk scores for coronary artery disease. Nat Rev Cardiol. 2022 May;19(5):291–301.

77. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019 Apr;51(4):584–91.

78. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Medicine. 2015 Mar 31;12(3):e1001779.

79. Leitsalu L, Haller T, Esko T, Tammesoo ML, Alavere H, Snieder H, et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. International Journal of Epidemiology. 2015 Aug 1;44(4):1137–47.

80. Nagai A, Hirata M, Kamatani Y, Muto K, Matsuda K, Kiyohara Y, et al. Overview of the BioBank Japan Project: Study design and profile. J Epidemiol. 2017 Feb 8;27(3 Suppl):S2–8.

81. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023 Jan;613(7944):508–18.

82. Cooper DN, Krawczak M, Polychronakos C, Tyler-Smith C, Kehrer-Sawatzki H. Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. Hum Genet. 2013 Oct;132(10):1077–130.

83. Chen R, Shi L, Hakenberg J, Naughton B, Sklar P, Zhang J, et al. Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. Nat Biotechnol. 2016 May;34(5):531–8.

84. Mendis S, Puska P, Norrving B, Organization WH, Federation WH, Organization WS. Global atlas on cardiovascular disease prevention and control [Internet]. World Health Organization; 2011 [cited 2023 Jan 21]. Available from: https://apps.who.int/iris/handle/10665/44701

85. Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Faire U. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. J Intern Med. 2002 Sep;252(3):247–54.

86. Won HH, Natarajan P, Dobbyn A, Jordan DM, Roussos P, Lage K, et al. Disproportionate Contributions of Select Genomic Compartments and Cell Types to Genetic Risk for Coronary Artery Disease. PLOS Genetics. 2015 Oct 28;11(10):e1005622.

87. The top 10 causes of death [Internet]. [cited 2024 Apr 15]. Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

88. Mensah GA, Fuster V, Murray CJL, Roth GA, null null, Mensah GA, et al. Global Burden of Cardiovascular Diseases and Risks, 1990-2022. Journal of the American College of Cardiology. 2023 Dec 19;82(25):2350–473.

89. Kuolleisuuden kasvun syynä ennen kaikkea koronavirustauti vuonna 2022 - Tilastokeskus [Internet]. 2023 [cited 2024 Apr 15]. Available from: https://www.stat.fi/julkaisu/cl8mlgiehwn8z0cvzmey6j7sr

90. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association. Circulation. 2022 Feb 22;145(8):e153–639.

91. Timmis A, Vardas P, Townsend N, Torbica A, Katus H, De Smedt D, et al. European Society of Cardiology: cardiovascular disease statistics 2021. Eur Heart J. 2022 Feb 22;43(8):716–99.

92. Gerdts E, Regitz-Zagrosek V. Sex differences in cardiometabolic disorders. Nat Med. 2019 Nov;25(11):1657–66.

93. Pajunen A. Tilastokeskus - 2. Kuolleisuus verenkiertoelinten sairauksiin väheni naisilla, miehillä ennallaan [Internet]. Tilastokeskus; [cited 2023 Jan 21]. Available from: https://www.stat.fi/til/ksyyt/2020/ksyyt_2020_2021-12-10_kat_002_fi.html

94. Wang Z, Du A, Liu H, Wang Z, Hu J. Systematic Analysis of the Global, Regional and National Burden of Cardiovascular Diseases from 1990 to 2017. J Epidemiol Glob Health. 2022 Mar;12(1):92–103.

95. Winkleby MA, Jatulis DE, Frank E, Fortmann SP. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. Am J Public Health. 1992 Jun;82(6):816–20.

96. Luoto R, Pekkanen J, Uutela A, Tuomilehto J. Cardiovascular risks and socioeconomic status: differences between men and women in Finland. J Epidemiol Community Health. 1994 Aug;48(4):348–54.

97. Yarnell J, Yu S, McCrum E, Arveiler D, Hass B, Dallongeville J, et al. Education, socioeconomic and lifestyle factors, and risk of coronary heart disease: the PRIME Study. Int J Epidemiol. 2005 Apr;34(2):268–75.

98. Havranek EP, Mujahid MS, Barr DA, Blair IV, Cohen MS, Cruz-Flores S, et al. Social Determinants of Risk and Outcomes for Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation. 2015 Sep 1;132(9):873–98.

99. Crombie IK, Kenicer MB, Smith WC, Tunstall-Pedoe HD. Unemployment, socioenvironmental factors, and coronary heart disease in Scotland. Br Heart J. 1989 Feb;61(2):172–7.

100. Leren P, Helgeland A, Hjermann I, Holme I. The Oslo study: CHD risk factors, socioeconomic influences, and intervention. Am Heart J. 1983 Nov;106(5 Pt 2):1200–6.

101. Mackenbach JP, Cavelaars AE, Kunst AE, Groenhof F. Socioeconomic inequalities in cardiovascular disease mortality; an international study. Eur Heart J. 2000 Jul;21(14):1141–51.

102. Schultz WM, Kelli HM, Lisko JC, Varghese T, Shen J, Sandesara P, et al. Socioeconomic Status and Cardiovascular Outcomes: Challenges and Interventions. Circulation. 2018 May 15;137(20):2166–78.

103. Backholer K, Peters SAE, Bots SH, Peeters A, Huxley RR, Woodward M. Sex differences in the relationship between socioeconomic status and cardiovascular disease: a systematic review and meta-analysis. J Epidemiol Community Health. 2017 Jun;71(6):550–7.

104. Hackshaw A, Morris JK, Boniface S, Tang JL, Milenković D. Low cigarette consumption and risk of coronary heart disease and stroke: meta-analysis of 141 cohort studies in 55 study reports. BMJ. 2018 Jan 24;360:j5855.

105. Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J Am Coll Cardiol. 2012 Dec 25;60(25):2631–9.

106. Holmes MV, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP, et al. Mendelian randomization of blood lipids for coronary heart disease. Eur Heart J. 2015 Mar 1;36(9):539–50.

107. The Emerging Risk Factors Collaboration*. Lipid-Related Markers and Cardiovascular Disease Prediction. JAMA. 2012 Jun 20;307(23):2499–506.

108. Prospective Studies Collaboration, Lewington S, Whitlock G, Clarke R, Sherliker P, Emberson J, et al. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. Lancet. 2007 Dec 1;370(9602):1829–39.

109. Cholesterol Treatment Trialists' (CTT) Collaboration, Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010 Nov 13;376(9753):1670–81.

110. Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. Lancet. 2016 Nov 19;388(10059):2532–61.

111. Cannon CP, Blazing MA, Giugliano RP, McCagg A, White JA, Theroux P, et al. Ezetimibe Added to Statin Therapy after Acute Coronary Syndromes. N Engl J Med. 2015 Jun 18;372(25):2387–97.

112. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, et al. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. N Engl J Med. 2017 May 4;376(18):1713–22.

113. Duncan MS, Vasan RS, Xanthakis V. Trajectories of Blood Lipid Concentrations Over the Adult Life Course and Risk of Cardiovascular Disease and All-Cause Mortality: Observations From the Framingham Study Over 35 Years. Journal of the American Heart Association. 2019 Jun 4;8(11):e011433.

114. Zhang Y, Pletcher MJ, Vittinghoff E, Clemons AM, Jacobs DR, Allen NB, et al. Association Between Cumulative Low-Density Lipoprotein Cholesterol Exposure During Young Adulthood and Middle Age and Risk of Cardiovascular Events. JAMA Cardiol. 2021 Dec 1;6(12):1406–13.

115. Reiner Ž. Hypertriglyceridaemia and risk of coronary artery disease. Nat Rev Cardiol. 2017 Jul;14(7):401–11.

116. Doi T, Langsted A, Nordestgaard BG. Elevated Remnant Cholesterol Reclassifies Risk of Ischemic Heart Disease and Myocardial Infarction. Journal of the American College of Cardiology. 2022 Jun 21;79(24):2383–97.

117. Wadström BN, Pedersen KM, Wulff AB, Nordestgaard BG. Elevated remnant cholesterol, plasma triglycerides, and cardiovascular and non-cardiovascular mortality. European Heart Journal. 2023 Apr 21;44(16):1432–45.

118. Feingold KR. Lipid and Lipoprotein Metabolism. Endocrinol Metab Clin North Am. 2022 Sep;51(3):437–58.

119. Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, Heath SC, et al. Genetic Variants Associated with Lp(a) Lipoprotein Level and Coronary Disease. New England Journal of Medicine. 2009 Dec 24;361(26):2518–28.

120. Saleheen D, Haycock PC, Zhao W, Rasheed A, Taleb A, Imran A, et al. Apolipoprotein(a) isoform size, lipoprotein(a) concentration, and coronary artery disease: a mendelian randomisation analysis. Lancet Diabetes Endocrinol. 2017 Jul;5(7):524–33.

121. Alexander RW. Hypertension and the Pathogenesis of Atherosclerosis. Hypertension. 1995 Feb;25(2):155–61.

122. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet. 2002 Dec 14;360(9349):1903–13.

123. Emerging Risk Factors Collaboration, Sarwar N, Gao P, Seshasai SRK, Gobin R, Kaptoge S, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet. 2010 Jun 26;375(9733):2215–22.

124. Bentzon JF, Otsuka F, Virmani R, Falk E. Mechanisms of Plaque Formation and Rupture. Circulation Research. 2014 Jun 6;114(12):1852–66.

125. Libby P, Theroux P. Pathophysiology of Coronary Artery Disease. Circulation. 2005 Jun 28;111(25):3481–8.

126. Bosco E, Hsueh L, McConeghy KW, Gravenstein S, Saade E. Major adverse cardiovascular event definitions used in observational analysis of administrative databases: a systematic review. BMC Medical Research Methodology. 2021 Nov 6;21(1):241.

127. Luo J, Yang H, Song BL. Mechanisms and regulation of cholesterol homeostasis. Nat Rev Mol Cell Biol. 2020 Apr;21(4):225–45.

128. Wijers M, Kuivenhoven JA, van de Sluis B. The life cycle of the low-density lipoprotein receptor: insights from cellular and in-vivo studies. Curr Opin Lipidol. 2015 Apr;26(2):82–7.

129. Tobert JA. Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. Nat Rev Drug Discov. 2003 Jul;2(7):517–26.

130. Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet. 2005 Oct 8;366(9493):1267–78.

131. Tsimikas S, Gordts PLSM, Nora C, Yeang C, Witztum JL. Statin therapy increases lipoprotein(a) levels. European Heart Journal. 2020 Jun 21;41(24):2275–84.

132. Abdul-Rahman T, Bukhari SMA, Herrera EC, Awuah WA, Lawrence J, de Andrade H, et al. Lipid Lowering Therapy: An Era Beyond Statins. Curr Probl Cardiol. 2022 Dec;47(12):101342.

133. Sabatine MS. PCSK9 inhibitors: clinical evidence and implementation. Nat Rev Cardiol. 2019 Mar;16(3):155–65.

134. Raal FJ, Rosenson RS, Reeskamp LF, Hovingh GK, Kastelein JJP, Rubba P, et al. Evinacumab for Homozygous Familial Hypercholesterolemia. New England Journal of Medicine. 2020 Aug 19;383(8):711–20.

135. Gaudet D, Greber-Platzer S, Reeskamp LF, Iannuzzo G, Rosenson RS, Saheb S, et al. Evinacumab in homozygous familial hypercholesterolaemia: long-term safety and efficacy. Eur Heart J. 2024 Jun 10;45(27):2422–34.

136. Tardif JC, Karwatowska-Prokopczuk E, Amour ES, Ballantyne CM, Shapiro MD, Moriarty PM, et al. Apolipoprotein C-III reduction in subjects with moderate hypertriglyceridaemia and at high cardiovascular risk. Eur Heart J. 2022 Jan 13;43(14):1401–12.

137. Gaudet D, Alexander VJ, Baker BF, Brisson D, Tremblay K, Singleton W, et al. Antisense Inhibition of Apolipoprotein C-III in Patients with Hypertriglyceridemia. New England Journal of Medicine. 2015 Jul 30;373(5):438–47.

138. Witztum JL, Gaudet D, Freedman SD, Alexander VJ, Digenio A, Williams KR, et al. Volanesorsen and Triglyceride Levels in Familial Chylomicronemia Syndrome. New England Journal of Medicine. 2019 Aug 8;381(6):531–42.

139. Dron JS, Hegele RA. Genetics of Lipid and Lipoprotein Disorders and Traits. Curr Genet Med Rep. 2016;4(3):130–41.

140. Khachadurian AK. The inheritance of essential familial hypercholesterolemia. The American Journal of Medicine. 1964 Sep 1;37(3):402–7.

141. Brown MS, Goldstein JL. Receptor-mediated control of cholesterol metabolism. Science. 1976 Jan 16;191(4223):150–4.

142. Sijbrands EJG, Westendorp RGJ, Defesche JC, de Meier PHEM, Smelt AHM, Kastelein JJP. Mortality over two centuries in large pedigree with familial

hypercholesterolaemia: family tree mortality study. BMJ. 2001 Apr 28;322(7293):1019–23.

143. Netea MG, Demacker PN, Kullberg BJ, Boerman OC, Verschueren I, Stalenhoef AF, et al. Low-density lipoprotein receptor-deficient mice are protected against lethal endotoxemia and severe gram-negative infections. J Clin Invest. 1996 Mar 15;97(6):1366–72.

144. Khovidhunkit W, Kim MS, Memon RA, Shigenaga JK, Moser AH, Feingold KR, et al. Thematic review series: The Pathogenesis of Atherosclerosis. Effects of infection and inflammation on lipid and lipoprotein metabolism mechanisms and consequences to the host1. Journal of Lipid Research. 2004 Jul 1;45(7):1169–96.

145. McGowan MP, Hosseini Dehkordi SH, Moriarty PM, Duell PB. Diagnosis and Treatment of Heterozygous Familial Hypercholesterolemia. Journal of the American Heart Association. 2019 Dec 17;8(24):e013225.

146. Beheshti SO, Madsen CM, Varbo A, Nordestgaard BG. Worldwide Prevalence of Familial Hypercholesterolemia: Meta-Analyses of 11 Million Subjects. JAmCollCardiol. 2020;75(20):2553–66.

147. Akioyamen LE, Genest J, Shan SD, Reel RL, Albaum JM, Chu A, et al. Estimating the prevalence of heterozygous familial hypercholesterolaemia: a systematic review and meta-analysis. BMJ Open. 2017 Sep 1;7(9):e016461.

148. Hu P, Dharmayat KI, Stevens CAT, Sharabiani MTA, Jones RS, Watts GF, et al. Prevalence of Familial Hypercholesterolemia Among the General Population and Patients With Atherosclerotic Cardiovascular Disease: A Systematic Review and Meta-Analysis. Circulation. 2020 Jun 2;141(22):1742–59.

149. Benn M, Watts GF, Tybjaerg-Hansen A, Nordestgaard BG. Mutations causative of familial hypercholesterolaemia: screening of 98 098 individuals from the Copenhagen General Population Study estimated a prevalence of 1 in 217. EurHeart J. 2016;37(17):1384–94.

150. Clarke SL, Tcheandjieu C, Hilliard AT, Lee KM, Lynch J, Chang KM, et al. Coronary Artery Disease Risk of Familial Hypercholesterolemia Genetic Variants Independent of Clinically Observed Longitudinal Cholesterol Exposure. Circ Genom Precis Med. 2022 Apr;15(2):e003501.

151. Trinder M, Francis GA, Brunham LR. Association of Monogenic vs Polygenic Hypercholesterolemia With Risk of Atherosclerotic Cardiovascular Disease. JAMA Cardiol. 2020 Apr 1;5(4):390–9.

152. Nurmohamed NS, Navar AM, Kastelein JJP. New and Emerging Therapies for Reduction of LDL-Cholesterol and Apolipoprotein B: JACC Focus Seminar 1/4. Journal of the American College of Cardiology. 2021 Mar 30;77(12):1564–75.

153. Humphries SE, Cooper JA, Seed M, Capps N, Durrington PN, Jones B, et al. Coronary heart disease mortality in treated familial hypercholesterolaemia:

Update of the UK Simon Broome FH register. Atherosclerosis. 2018 Jul;274:41–6.

154. Mortality in treated heterozygous familial hypercholesterolaemia: implications for clinical management. Scientific Steering Committee on behalf of the Simon Broome Register Group. Atherosclerosis. 1999 Jan;142(1):105–12.

155. Versmissen J, Oosterveer DM, Yazdanpanah M, Defesche JC, Basart DCG, Liem AH, et al. Efficacy of statins in familial hypercholesterolaemia: a long term cohort study. BMJ. 2008 Nov 11;337:a2423.

156. Schreuder MM, Hamkour S, Siegers KE, Holven KB, Johansen AK, van de Ree MA, et al. LDL cholesterol targets rarely achieved in familial hypercholesterolemia patients: A sex and gender-specific analysis. Atherosclerosis. 2023 Nov;384:117117.

157. Besseling J, Sjouke B, Kastelein JJP. Screening and treatment of familial hypercholesterolemia - Lessons from the past and opportunities for the future (based on the Anitschkow Lecture 2014). Atherosclerosis. 2015 Aug;241(2):597–606.

158. Sun YV, Damrauer SM, Hui Q, Assimes TL, Ho YL, Natarajan P, et al. Effects of Genetic Variants Associated with Familial Hypercholesterolemia on Low-Density Lipoprotein-Cholesterol Levels and Cardiovascular Outcomes in the Million Veteran Program. Circ Genom Precis Med. 2018 Dec;11(12):e002192.

159. Paththinige CS, Sirisena ND, Dissanayake V. Genetic determinants of inherited susceptibility to hypercholesterolemia - a comprehensive literature review. Lipids Health Dis. 2017 Jun 2;16(1):103.

160. Schneider WJ, Beisiegel U, Goldstein JL, Brown MS. Purification of the low density lipoprotein receptor, an acidic glycoprotein of 164,000 molecular weight. J Biol Chem. 1982 Mar 10;257(5):2664–73.

161. Unique variants in the LDLR gene - Global Variome shared LOVD [Internet]. [cited 2024 Apr 15]. Available from: https://databases.lovd.nl/shared/variants/LDLR/unique#object_id=VariantOnTranscriptUnique%2CVariantOnGenome&id=LDLR&order=VariantOnGenome%2FClinicalClassification%2CDESC&search_transcriptid=00011039&page_size=100&page=1

162. Leigh S, Futema M, Whittall R, Taylor-Beadling A, Williams M, den Dunnen JT, et al. The UCL low-density lipoprotein receptor gene variant database: pathogenicity update. J Med Genet. 2017 Apr;54(4):217–23.

163. Brautbar A, Leary E, Rasmussen K, Wilson DP, Steiner RD, Virani S. Genetics of familial hypercholesterolemia. Curr Atheroscler Rep. 2015 Apr;17(4):491.

164. Gene: LDLR (ENSG00000130164) - Summary - Homo_sapiens - Ensembl genome browser 107 [Internet]. [cited 2022 Oct 18]. Available from:

https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG0
0000130164;r=19:11089462-11133820

165. Gent J, Braakman I. Low-density lipoprotein receptor structure and folding. Cell Mol Life Sci. 2004 Oct;61(19–20):2461–70.

166. Whitfield AJ, Barrett PHR, van Bockxmeer FM, Burnett JR. Lipid disorders and mutations in the APOB gene. Clin Chem. 2004 Oct;50(10):1725–32.

167. Tybjaerg-Hansen A, Humphries SE. Familial defective apolipoprotein B-100: a single mutation that causes hypercholesterolemia and premature coronary artery disease. Atherosclerosis. 1992 Oct;96(2–3):91–107.

168. Myant NB, Forbes SA, Day IN, Gallagher J. Estimation of the age of the ancestral arginine3500-->glutamine mutation in human apoB-100. Genomics. 1997 Oct 1;45(1):78–87.

169. Teng YN, Pan JP, Chou SC, Tai DY, Lee-Chen GJ. Familial defective apolipoprotein B-100: detection and haplotype analysis of the Arg(3500)-->Gln mutation in hyperlipidemic Chinese. Atherosclerosis. 2000 Oct;152(2):385–90.

170. Abifadel M, Varret M, Rabès JP, Allard D, Ouguerram K, Devillers M, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet. 2003 Jun;34(2):154–6.

171. Gene: PCSK9 (ENSG00000169174) - Summary - Homo_sapiens - Ensembl genome browser 108 [Internet]. [cited 2022 Nov 2]. Available from: https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG0
0000169174;r=1:55039447-55064852

172. Lambert G, Sjouke B, Choque B, Kastelein JJP, Hovingh GK. The PCSK9 decade. J Lipid Res. 2012 Dec;53(12):2515–24.

173. Horton JD, Cohen JC, Hobbs HH. Molecular biology of PCSK9: its role in LDL metabolism. Trends Biochem Sci. 2007 Feb;32(2):71–7.

174. Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, Maioli M, et al. Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. Science. 2001 May 18;292(5520):1394–8.

175. D'Erasmo L, Di Costanzo A, Arca M. Autosomal recessive hypercholesterolemia: update for 2020. Curr Opin Lipidol. 2020 Apr;31(2):56–61.

176. Shaik NA, Al-Qahtani F, Nasser K, Jamil K, Alrayes NM, Elango R, et al. Molecular insights into the coding region mutations of low-density lipoprotein receptor adaptor protein 1 (LDLRAP1) linked to familial hypercholesterolemia. J Gene Med. 2020 Jun;22(6):e3176.

177. Nohara A, Tada H, Ogura M, Okazaki S, Ono K, Shimano H, et al. Homozygous Familial Hypercholesterolemia. J Atheroscler Thromb. 2021 Jul 1;28(7):665–78.

178. Bélanger AM, Akioyamen L, Alothman L, Genest J. Evidence for improved survival with treatment of homozygous familial hypercholesterolemia. Curr Opin Lipidol. 2020 Aug;31(4):176–81.

179. Cuchel M, Raal FJ, Hegele RA, Al-Rasadi K, Arca M, Averna M, et al. 2023 Update on European Atherosclerosis Society Consensus Statement on Homozygous Familial Hypercholesterolaemia: new treatments and clinical guidance. Eur Heart J. 2023 May 2;44(25):2277–91.

180. Rosenson RS. Existing and emerging therapies for the treatment of familial hypercholesterolemia. Journal of Lipid Research [Internet]. 2021 Jan 1 [cited 2024 Aug 1];62. Available from: https://www.jlr.org/article/S0022-2275(21)00042-0/abstract

181. Aalto-Setälä K, Gylling H, Miettinen T, Kontula K. Identification of a deletion in the LDL receptor gene. A Finnish type of mutation. FEBS Lett. 1988 Mar 28;230(1–2):31–4.

182. Aalto-Setälä K, Helve E, Kovanen PT, Kontula K. Finnish type of low density lipoprotein receptor gene mutation (FH-Helsinki) deletes exons encoding the carboxy-terminal part of the receptor and creates an internalization-defective phenotype. J Clin Invest. 1989 Aug;84(2):499–505.

183. Koivisto UM, Turtola H, Aalto-Setälä K, Top B, Frants RR, Kovanen PT, et al. The familial hypercholesterolemia (FH)-North Karelia mutation of the low density lipoprotein receptor gene deletes seven nucleotides of exon 6 and is a common cause of FH in Finland. J Clin Invest. 1992 Jul;90(1):219–28.

184. Vuorio AF, Turtola H, Piilahti KM, Repo P, Kanninen T, Kontula K. Familial hypercholesterolemia in the Finnish north Karelia. A molecular, clinical, and genealogical study. Arterioscler Thromb Vasc Biol. 1997 Nov;17(11):3127–38.

185. Koivisto UM, Viikari JS, Kontula K. Molecular characterization of minor gene rearrangements in Finnish patients with heterozygous familial hypercholesterolemia: identification of two common missense mutations (Gly823-->Asp and Leu380-->His) and eight rare mutations of the LDL receptor gene. Am J Hum Genet. 1995 Oct;57(4):789–97.

186. Hämäläinen T, Palotie A, Aalto-Setälä K, Kontula K, Tikkanen MJ. Absence of familial defective apolipoprotein B-100 in Finnish patients with elevated serum cholesterol. Atherosclerosis. 1990 Jun;82(3):177–83.

187. Ilmonen M, Heliö T, Ebeling T, Pyörälä K, Uusitupa M, Palotie A, et al. Screening for mutations in the exon 26 of the apolipoprotein B gene in hypercholesterolemic Finnish families by the single-strand conformation polymorphism method. Hum Mutat. 1994;4(3):217–23.

188. Erdmann J, Stark K, Esslinger UB, Rumpf PM, Koesling D, de Wit C, et al. Dysfunctional nitric oxide signalling increases risk of myocardial infarction. Nature. 2013 Dec 19;504(7480):432–6.

189. Dang TA, Kessler T, Wobst J, Wierer M, Braenne I, Strom TM, et al. Identification of a Functional PDE5A Variant at the Chromosome 4q27 Coronary Artery Disease Locus in an Extended Myocardial Infarction Family. Circulation. 2021 Aug 24;144(8):662–5.

190. Graham SE, Clarke SL, Wu KHH, Kanoni S, Zajac GJM, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. Nature. 2021 Dec;600(7890):675–9.

191. Dron JS, Hegele RA. Polygenic influences on dyslipidemias. Curr Opin Lipidol. 2018 Apr;29(2):133–43.

192. D'Erasmo L, Minicocci I, Di Costanzo A, Pigna G, Commodari D, Ceci F, et al. Clinical Implications of Monogenic Versus Polygenic Hypercholesterolemia: Long-Term Response to Treatment, Coronary Atherosclerosis Burden, and Cardiovascular Events. J Am Heart Assoc. 2021 May 4;10(9):e018932.

193. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Atri DS, et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. Nat Genet. 2022 Dec;54(12):1803–15.

194. Fahed AC, Wang M, Homburger JR, Patel AP, Bick AG, Neben CL, et al. Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. Nat Commun. 2020 Aug 20;11(1):3635.

195. Erdmann J, Kessler T, Munoz Venegas L, Schunkert H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. Cardiovasc Res. 2018 Jul 15;114(9):1241–57.

196. Eales JM, Maan AA, Xu X, Michoel T, Hallast P, Batini C, et al. Human Y Chromosome Exerts Pleiotropic Effects on Susceptibility to Atherosclerosis. Arterioscler Thromb Vasc Biol. 2019 Nov;39(11):2386–401.

197. Tcheandjieu C, Zhu X, Hilliard AT, Clarke SL, Napolioni V, Ma S, et al. Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. Nat Med. 2022 Aug;28(8):1679–92.

198. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998 May 12;97(18):1837–47.

199. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General Cardiovascular Risk Profile for Use in Primary Care. Circulation. 2008 Feb 12;117(6):743–53.

200. SCORE2 working group and ESC Cardiovascular risk collaboration. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur Heart J. 2021 Jul 1;42(25):2439–54.

201. Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2. BMJ. 2012 Jun 21;344:e4181.

202. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. Heart. 2008 Jan 1;94(1):34–9.

203. Mehta PK, Wei J, Wenger NK. Ischemic heart disease in women: a focus on risk factors. Trends Cardiovasc Med. 2015 Feb;25(2):140–51.

204. Elliott J, Bodinier B, Bond TA, Chadeau-Hyam M, Evangelou E, Moons KGM, et al. Predictive Accuracy of a Polygenic Risk Score-Enhanced Prediction Model vs a Clinical Risk Score for Coronary Artery Disease. JAMA. 2020;323(7):636–45.

205. Jansen ACM, van Aalst-Cohen ES, Tanck MW, Trip MD, Lansberg PJ, Liem AH, et al. The contribution of classical risk factors to cardiovascular disease in familial hypercholesterolaemia: data in 2400 patients. J Intern Med. 2004 Dec;256(6):482–90.

206. Pérez de Isla L, Alonso R, Mata N, Saltijeral A, Muñiz O, Rubio-Marin P, et al. Coronary Heart Disease, Peripheral Arterial Disease, and Stroke in Familial Hypercholesterolaemia: Insights From the SAFEHEART Registry (Spanish Familial Hypercholesterolaemia Cohort Study). Arterioscler Thromb Vasc Biol. 2016 Sep;36(9):2004–10.

207. Fahed AC, Wang M, Patel AP, Ajufo E, Maamari DJ, Aragam KG, et al. Association of the Interaction Between Familial Hypercholesterolemia Variants and Adherence to a Healthy Lifestyle With Risk of Coronary Artery Disease. JAMA Network Open. 2022 Mar 16;5(3):e222687.

208. Arsenault BJ, Perrot N, Couture P. Does lifestyle contribute to disease severity in patients with inherited lipid disorders? Curr Opin Lipidol. 2017 Apr;28(2):177–85.

209. Khera AV, Emdin CA, Drake I, Natarajan P, Bick AG, Cook NR, et al. Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary Disease. N Engl J Med. 2016 Dec 15;375(24):2349–58.

210. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2019 Sep 10;140(11):e596–646.

211. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice: Developed by the Task Force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies With the special contribution of the European Association of

Preventive Cardiology (EAPC). European Heart Journal. 2021 Sep 7;42(34):3227–337.

212. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk: The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). European Heart Journal. 2020 Jan 1;41(1):111–88.

213. Law MR, Wald NJ, Thompson SG. By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? BMJ. 1994 Feb 5;308(6925):367–72.

214. Clarke R, Frost C, Collins R, Appleby P, Peto R. Dietary lipids and blood cholesterol: quantitative meta-analysis of metabolic ward studies. BMJ. 1997 Jan 11;314(7074):112–7.

215. Jenkins DJA, Kendall CWC, Faulkner DA, Nguyen T, Kemp T, Marchie A, et al. Assessment of the longer-term effects of a dietary portfolio of cholesterol-lowering foods in hypercholesterolemia. Am J Clin Nutr. 2006 Mar;83(3):582–91.

216. Ravera A, Carubelli V, Sciatti E, Bonadei I, Gorga E, Cani D, et al. Nutrition and Cardiovascular Disease: Finding the Perfect Recipe for Cardiovascular Health. Nutrients. 2016 Jun 14;8(6):363.

217. Rosato V, Temple NJ, La Vecchia C, Castellan G, Tavani A, Guercio V. Mediterranean diet and cardiovascular disease: a systematic review and meta-analysis of observational studies. Eur J Nutr. 2019 Feb;58(1):173–91.

218. Salehi-Abargouei A, Maghsoudi Z, Shirani F, Azadbakht L. Effects of Dietary Approaches to Stop Hypertension (DASH)-style diet on fatal or nonfatal cardiovascular diseases--incidence: a systematic review and meta-analysis on observational prospective studies. Nutrition. 2013 Apr;29(4):611–8.

219. Sotos-Prieto M, Bhupathiraju SN, Mattei J, Fung TT, Li Y, Pan A, et al. Association of Changes in Diet Quality with Total and Cause-Specific Mortality. New England Journal of Medicine. 2017 Jul 13;377(2):143–53.

220. Koch CA, Kjeldsen EW, Frikke-Schmidt R. Vegetarian or vegan diets and blood lipids: a meta-analysis of randomized trials. Eur Heart J. 2023 Jul 21;44(28):2609–22.

221. Xu Z, Steffen LM, Selvin E, Rebholz CM. Diet quality, change in diet quality and risk of incident CVD and diabetes. Public Health Nutr. 2020 Feb;23(2):329–38.

222. Fiuza-Luces C, Santos-Lozano A, Joyner M, Carrera-Bastos P, Picazo O, Zugaza JL, et al. Exercise benefits in cardiovascular disease: beyond attenuation of traditional risk factors. Nat Rev Cardiol. 2018 Dec;15(12):731–43.

223. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet. 2012 Jul 21;380(9838):219–29.

224. Lear SA, Hu W, Rangarajan S, Gasevic D, Leong D, Iqbal R, et al. The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: the PURE study. Lancet. 2017 Dec 16;390(10113):2643–54.

225. Karjalainen JJ, Kiviniemi AM, Hautala AJ, Piira OP, Lepojärvi ES, Perkiömäki JS, et al. Effects of physical activity and exercise training on cardiovascular risk in coronary artery disease patients with and without type 2 diabetes. Diabetes Care. 2015 Apr;38(4):706–15.

226. Mann S, Beedie C, Jimenez A. Differential Effects of Aerobic Exercise, Resistance Training and Combined Exercise Modalities on Cholesterol and the Lipid Profile: Review, Synthesis and Recommendations. Sports Med. 2014;44(2):211–21.

227. Isath A, Koziol KJ, Martinez MW, Garber CE, Martinez MN, Emery MS, et al. Exercise and cardiovascular health: A state-of-the-art review. Progress in Cardiovascular Diseases. 2023 Jul 1;79:44–52.

228. Powell-Wiley TM, Poirier P, Burke LE, Després JP, Gordon-Larsen P, Lavie CJ, et al. Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation. 2021 May 25;143(21):e984–1010.

229. Sabaté E, World Health Organization, editors. Adherence to long-term therapies: evidence for action. Geneva: World Health Organization; 2003. 198 p.

230. Reach G, Benarbia L, Bruckert E, Kevorkian JP, Farnier M, Mourad JJ, et al. Intentionality in adherence to long-term therapies. Results from an online survey of 3,001 patients with cardio-metabolic pathologies in France. Patient Prefer Adherence. 2021;15:1739–53.

231. DiMatteo MR, Giordani PJ, Lepper HS, Croghan TW. Patient adherence and medical treatment outcomes: a meta-analysis. Med Care. 2002 Sep;40(9):794–811.

232. Lewey J, Shrank WH, Bowry ADK, Kilabuk E, Brennan TA, Choudhry NK. Gender and racial disparities in adherence to statin therapy: a meta-analysis. Am Heart J. 2013 May;165(5):665–78, 678.e1.

233. Bassi N, Karagodin I, Wang S, Vassallo P, Priyanath A, Massaro E, et al. Lifestyle modification for metabolic syndrome: a systematic review. AmJMed. 2014;127(12):1242.e1-1242.10.

234. Leung AWY, Chan RSM, Sea MMM, Woo J. An Overview of Factors Associated with Adherence to Lifestyle Modification Programs for Weight Management in Adults. IntJEnvironResPublicHealth. 2017;14(8):10.3390/ijerph14080922.

235. Lemstra M, Bird Y, Nwankwo C, Rogers M, Moraros J. Weight loss intervention adherence and factors promoting adherence: a meta-analysis. Patient Prefer Adherence. 2016;10:1547–59.

236. Burgess E, Hassmén P, Pumpa KL. Determinants of adherence to lifestyle intervention in adults with obesity: a systematic review. Clin Obes. 2017 Jun;7(3):123–35.

237. Magnani JW, Mujahid MS, Aronow HD, Cené CW, Dickson VV, Havranek E, et al. Health Literacy and Cardiovascular Disease: Fundamental Relevance to Primary and Secondary Prevention: A Scientific Statement From the American Heart Association. Circulation. 2018 Jul 10;138(2):e48–74.

238. Jepson RG, Harris FM, Platt S, Tannahill C. The effectiveness of interventions to change six health behaviours: a review of reviews. BMC Public Health. 2010 Sep 8;10:538.

239. French DP, Cameron E, Benton JS, Deaton C, Harvie M. Can Communicating Personalised Disease Risk Promote Healthy Behaviour Change? A Systematic Review of Systematic Reviews. Ann Behav Med. 2017 Oct;51(5):718–29.

240. Umans-Eckenhausen MAW, Defesche JC, van Dam MJ, Kastelein JJP. Long-term compliance with lipid-lowering medication after genetic screening for familial hypercholesterolemia. Arch Intern Med. 2003 Jan 13;163(1):65–8.

241. McBride CM, Koehly LM, Sanderson SC, Kaphingst KA. The behavioral response to personalized genetic information: will genetic risk profiles motivate individuals and families to choose more healthful behaviors? Annu Rev Public Health. 2010;31:89–103.

242. Kullo IJ, Jouni H, Austin EE, Brown SA, Kruisselbrink TM, Isseh IN, et al. Incorporating a Genetic Risk Score Into Coronary Heart Disease Risk Estimates: Effect on Low-Density Lipoprotein Cholesterol Levels (the MI-GENES Clinical Trial). Circulation. 2016 Mar 22;133(12):1181–8.

243. Hollands GJ, French DP, Griffin SJ, Prevost AT, Sutton S, King S, et al. The impact of communicating genetic risks of disease on risk-reducing health behaviour: systematic review with meta-analysis. BMJ. 2016 Mar 15;352:i1102.

244. Robinson CL, Jouni H, Kruisselbrink TM, Austin EE, Christensen KD, Green RC, et al. Disclosing genetic risk for coronary heart disease: effects on perceived personal control and genetic counseling satisfaction. ClinGenet. 2016;89(2):251–7.

245. Marteau T, Senior V, Humphries SE, Bobrow M, Cranston T, Crook MA, et al. Psychological impact of genetic testing for familial hypercholesterolemia within a previously aware population: a randomized controlled trial. Am J Med Genet A. 2004 Jul 30;128A(3):285–93.

246. Beery TA, Williams JK. Risk reduction and health promotion behaviors following genetic testing for adult-onset disorders. Genet Test. 2007;11(2):111–23.

247. Brown SN, Jouni H, Marroush TS, Kullo IJ. Effect of Disclosing Genetic Risk for Coronary Heart Disease on Information Seeking and Sharing: The MI-GENES Study (Myocardial Infarction Genes). CircCardiovascGenet. 2017;10(4):10.1161/CIRCGENETICS.116.001613.

248. 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 Jun;18(6):499–502.

249. Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP, et al. Copy number variation detection and genotyping from exome sequence data. Genome Res. 2012;22(8):1525–32.

250. Fromer M, Moran JL, Chambert K, Banks E, Bergen SE, Ruderfer DM, et al. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. AmJHumGenet. 2012;91(4):597–607.

251. Surakka I, Sarin AP, Ruotsalainen SE, Durbin R, Salomaa V, Daly MJ, et al. The rate of false polymorphisms introduced when imputing genotypes from global imputation panels [Internet]. bioRxiv; 2016 [cited 2024 Mar 14]. p. 080770. Available from: https://www.biorxiv.org/content/10.1101/080770v1

252. Vilhjalmsson BJ, Yang J, Finucane HK, Gusev A, Lindstrom S, Ripke S, et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. AmJHumGenet. 2015;97(4):576–92.

253. Ripatti P, Rämö JT, Mars NJ, Fu Y, Lin J, Söderlund S, et al. Polygenic Hyperlipidemias and Coronary Artery Disease Risk. CircGenomPrecis Med. 2020;13(2):e002725.

254. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015 Oct;47(10):1121–30.

255. Ge T, Chen CY, Ni Y, Feng YCA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. 2019 Apr 16;10(1):1776.

256. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.

257. Surakka I, Horikoshi M, Mägi R, Sarin AP, Mahajan A, Lagou V, et al. The impact of low-frequency and rare variants on lipid levels. Nat Genet. 2015 Jun;47(6):589–97.

258. Kanerva N, Kaartinen NE, Ovaskainen ML, Konttinen H, Kontto J, Mannisto S. A diet following Finnish nutrition recommendations does not contribute to the current epidemic of obesity. Public Health Nutr. 2013;16(5):786–94.

259. Dyslipidemiat [Internet]. [cited 2023 Jul 12]. Available from: https://www.kaypahoito.fi/hoi50025

260. Tichy L, Freiberger T, Zapletalova P, Soska V, Ravcukova B, Fajkusova L. The molecular basis of familial hypercholesterolemia in the Czech Republic: spectrum of LDLR mutations and genotype-phenotype correlations. Atherosclerosis. 2012;223(2):401–8.

261. Benito-Vicente A, Uribe KB, Jebari S, Galicia-Garcia U, Ostolaza H, Martin C. Validation of LDLr Activity as a Tool to Improve Genetic Diagnosis of Familial Hypercholesterolemia: A Retrospective on Functional Characterization of LDLr Variants. Int J Mol Sci. 2018 Jun 5;19(6):1676.

262. Pfisterer SG, Brock I, Kanerva K, Hlushchenko I, Paavolainen L, Ripatti P, et al. Multiparametric platform for profiling lipid trafficking in human leukocytes. Cell Rep Methods. 2022 Feb 8;2(2):100166.

263. Norman D, Sun XM, Bourbon M, Knight BL, Naoumova RP, Soutar AK. Characterization of a novel cellular defect in patients with phenotypic homozygous familial hypercholesterolemia. J Clin Invest. 1999 Sep;104(5):619–28.

264. Pisciotta L, Priore Oliva C, Pes GM, Di Scala L, Bellocchio A, Fresa R, et al. Autosomal recessive hypercholesterolemia (ARH) and homozygous familial hypercholesterolemia (FH): a phenotypic comparison. Atherosclerosis. 2006 Oct;188(2):398–405.

265. Filigheddu F, Quagliarini F, Campagna F, Secci T, Degortes S, Zaninello R, et al. Prevalence and clinical features of heterozygous carriers of autosomal recessive hypercholesterolemia in Sardinia. Atherosclerosis. 2009 Nov;207(1):162–7.

266. Khera AV, Won HH, Peloso GM, Lawson KS, Bartz TM, Deng X, et al. Diagnostic Yield and Clinical Utility of Sequencing Familial Hypercholesterolemia Genes in Patients With Severe Hypercholesterolemia. J Am Coll Cardiol. 2016 Jun 7;67(22):2578–89.

267. Saadatagah S, Alhalabi L, Farwati M, Zordok M, Bhat A, Smith CY, et al. The burden of severe hypercholesterolemia and familial hypercholesterolemia in a population-based setting in the US. Am J Prev Cardiol. 2022 Sep 25;12:100393.

268. Dedoussis GVZ, Schmidt H, Genschel J. LDL-receptor mutations in Europe. Hum Mutat. 2004 Dec;24(6):443–59.

269. Bertolini S, Cantafora A, Averna M, Cortese C, Motti C, Martini S, et al. Clinical expression of familial hypercholesterolemia in clusters of mutations of the LDL receptor gene that cause a receptor-defective or receptor-negative phenotype. Arterioscler Thromb Vasc Biol. 2000 Sep;20(9):E41-52.

270. Kerminen S, Cerioli N, Pacauskas D, Havulinna AS, Perola M, Jousilahti P, et al. Changes in the fine-scale genetic structure of Finland through the 20th century. PLOS Genetics. 2021 Mar 4;17(3):e1009347.

271. Aragam KG, Dobbyn A, Judy R, Chaffin M, Chaudhary K, Hindy G, et al. Limitations of Contemporary Guidelines for Managing Patients at High Genetic Risk of Coronary Artery Disease. J Am Coll Cardiol. 2020 Jun 9;75(22):2769–80.

272. Fuat A, Adlen E, Monane M, Coll R, Groves S, Little E, et al. A polygenic risk score added to a QRISK®2 cardiovascular disease risk calculator demonstrated robust clinical acceptance and clinical utility in the primary care setting. European Journal of Preventive Cardiology. 2024 Jan 19;zwae004.

273. Patel AP, Khera AV. Advances and Applications of Polygenic Scores for Coronary Artery Disease. Annu Rev Med. 2023 Jan 27;74:141–54.

274. Roth GA, Johnson C, Abajobir A, Abd-Allah F, Abera SF, Abyu G, et al. Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. JAmCollCardiol. 2017;70(1):1–25.

# ORIGINAL PUBLICATIONS