Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Evaluation of genetic and plasma markers of VEGF levels for prognosis in coronary heart disease

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Health Sciences

at Massey University, Wellington, New Zealand.

Juan Carlos Meza Alvarado

2024

Note for Examiners of Doctoral Theses
Explanation of Impacts on Research

Thank you for taking the time to examine this thesis, which has been impacted by the pandemic, extreme weather-related events, such as Cyclone Gabrielle, and/or changes in the university, including relocation, staffing and resources.

This information has been provided to allow you as an examiner, to take into account how certain factors such as delays and disruptions experienced by the PhD student will have impacted their research outcomes. This form has been completed by the supervisor and student, and outlines the extent that the student's research has been affected.

Please consider the factors listed below in your assessment of the work.

This statement has been prepared by the candidate's supervisor in consultation with the student and has been endorsed by the relevant Head of Academic Unit.

Student Name:	Juan Carlos Meza Alvarado	ID Number:	
Supervisor Name:	Dr. Barry Palmer	Date:	02-Jul-2024
Thesis title: Evaluation of genetic and plasma markers of VEGF levels for prognosis in coronary heart disease			

Considerations to be taken into account. Note: This statement will remain in the final copy of the thesis which will be available from the Massey University Library following the examination process.
[*Enter key considerations here for the examiners. This can include but is not limited to change of scope, scale, topic, focus; limitations in relation to data collection, access to necessary literature or archival materials, laboratories, field sites; disruptions as a result of lockdown and various alert levels, medical or health considerations etc*]

For a six-week period from March 26 to April 27, 2020, New Zealand was placed under very strict lockdown conditions (Level 4 – Lockdown), with students and staff unable to physically access University facilities, unless they were involved in essential research related to Covid-19. Additional Covid-19-related restrictions on activities were observed at lower emergency conditions, as observed by Alert Level 2 (Reduce) and Level 3 (Restrict) conditions. Alert Levels were changing as part of the pandemic management across the country. From 2020-September 2022 Massey University followed COVID-19 protection plans set up for the country. This meant that some research students had experimental, clinical, laboratory, field work, and/or data collection or analysis interrupted, and consequently may have had to adjust their research plans.

Carlos' ability to work on his thesis project was very significantly affected from August 2021 to end of 2022.

1) Carlos was not able to access the MU laboratory in Wellington from August 2021 to November 2021. Lack of access to the laboratory and COVID-19 symptoms delayed Carlos' initial establishment of the experimental laboratory pipeline and induction to the Massey Wellington research laboratory.

2) The length of the changes in the national lockdown measures affected contact and negotiations with University of Otago collaborators. Their input was required to refine project scale and focus based on preliminary results presented by Carlos using partial cohort data provided by Dr Barry Palmer. There were also concerns on the scope of work achievable considering COVID-19 alert level fluctuations and personal circumstances of research collaborators. Overall, this had the downstream effect of delaying and causing modifications of the collaboration to reach a final Material Transfer Agreement which granted the student use of the full study cohort datasets and access to available samples.

As a result, samples from the CDCS and HVOL cohorts were not able to be accessed until many months later than initially planned. The time available to perform laboratory investigations within the 3 years of Carlos' scholarship was restricted by approximately 6 months due to delays in accessing samples and lack of access to research laboratories. Additionally, research funding acquired by the student and his supervisory team in March 2022 to fund DNA sample genotyping and ELISA assays of plasma samples was only able to be used once lockdown and traffic light COVID-19 measures were removed. This resulted in 5 weeks delay in starting this lab work.

The work presented in this thesis was possible due to the research collaborators' support of Carlos accessing data and material. However, the legal process of an MTA between Universities is by nature a lengthy process which was initially complicated due to COVID-19 management from 2021 to 2022. Accessing the patient samples was delayed by 7 months because of this.

3) Carlos' household consists of 9-12 international and/or domestic flatmates. On 4 April 2022, a flatmate tested positive for COVID leading to a 2-week household isolation for Carlos. Additionally, Carlos was diagnosed with COVID on 12 May 2022 which led to a further 2-week isolation until 26 May 2022. Carlos was unable to work on his project for most of this time.

Thesis abstract

Cardiovascular disease (CVD) is the leading cause of death. CVD risk assessment is complicated by the multifactorial nature of disease onset and the lack of predictive biomarkers in patients. The vascular endothelial growth factor (VEGF-A), involved in blood vessel formation, could be considered a novel biomarker since increased levels have been observed in CVD aetiologies. Genetic variants, such as single nucleotide polymorphisms (SNPs), influence circulating VEGF-A levels and have been linked to CVD risk. SNPs are biomarkers that are independent of age that can be link to molecular mechanisms involved in CVD. The identification of clinically relevant SNPs will complement the existing CVD risk framework and further our understanding on the role of VEGF-A in CVD.

Imputed genotype data was obtained for 47 SNPs located at the *VEGFA* locus (human chromosome 6, n = 30), the VEGF receptor 2 (*VEGFR2*) locus (human chromosome 4, n = 13) and the very low-density lipoprotein receptor (*VLDLR*) locus (human chromosome 9, n = 4). Imputed genotype data for 1935 patients from the Coronary Disease Cohort Study (CDCS) and 1183 individuals from the Canterbury Healthy Volunteers Study (HVOL) was assessed. Association between genotype groups with cardiometabolic parameters was tested using one-way ANOVA tests. Association of genotypes with clinical endpoints was examined by Kaplan-Meir analyses and multivariate regression models. Candidate SNPs were selected from among all the imputed SNPs if data was $p < 0.1$ for any of the analyses. Manual genotyping, using predesigned TaqMan assays, was performed for SNPs with multiple significant associations ($p < 0.05$). Validation of imputed findings with manual genotyping data was possible for rs6921438, rs7767396, rs2305948 and rs1870377. VEGF-A levels for 227 HVOL participants were measured by an ELISA immunoassay to compare with previously reported levels from 549 CDCS patients.

SNPs identified that influence circulating VEGF-A levels included five at the *VEGFA* locus (rs6921438. rs7767396, rs45137773, rs7763440 and rs11757868), seven at the *VEGFR2* locus (rs2305948, rs1870377, rs1870378, rs1870379, rs7677779, rs13136007 and rs10016064) and four at the *VLDLR* locus (rs7043199, rs10738760. rs7030781 and rs2375981). The homozygote minor allele genotypes for each SNP were associated with lower VEGF-A levels. Manual genotyped data for *VEGFA* locus variants rs6921438 and rs7767396 showed: **a)** rs6921438 AA was associated with increased all-cause death ($p = 0.03$), non ST-elevated myocardial infarction (NSTEMI, $p = 0.0003$),

heart failure (HF, p = 0.035) and major adverse cardiovascular event (MACE, p = 0.032) risk **b)** rs7767396 GG was associated with increased NSTEMI (p = 0.001) HF (p = 0.023) risk **c)** rs6921438 AA (Hazard Ratio (HR) = 6.6 p = 0.016) and VEGF-A (HR = 2.64, p = 0.014) were independent HF admission risk predictors, along with established predictors. Manual genotyped data for *VEGFR2* locus variants rs2305948 and rs1870377 showed **a)** rs2305948 CC was associated with higher all-cause mortality (p = 0.045) and shorter time to first cardiovascular readmission risk (p = 0.045) **b)** rs1870377 was an independent predictor for cardiovascular death when adjusting for NTproBNP, hypertension, creatinine, and beta blocker treatment (TT vs TA+AA, p = 0.048, HR = 1.125). Lastly, analyses showed that *VLDLR* locus variant rs10738760 AA genotype was associated with increased risk of cardiovascular death (p = 0.047, HR = 1.50).

The use of imputation data can facilitate the identification of clinically relevant SNPs by observing the amount and statistical significance of associations. In total, 11 imputed variants were identified as expression quantitative trait loci (eQTL) SNPs that affect circulating VEGF-A levels. Validation by genotyping confirmed that rs6921438 and rs7767396, at the *VEGFA* locus, are associated with VEGF-A levels and CVD risk. Additional data supported that *VEGFR2* exonic variants rs2305948 and rs1870377 have an impact on increased outcome risk. Moreover, the *VLDLR* variant rs10738760 can interact with molecular mechanisms that exacerbate CVD risk. Overall, these five SNPs are promising genetic markers for CVD risk profiling assessments before outcome occurrence.

.

Acknowledgements

I would like to thank my supervisory team Dr. Barry Palmer, Associate Professor Rachel Page, Dr. Collette Bromhead and Dr Beth Mallard for their invaluable support. The opportunities, support, dynamic and positive reinforcement you all provided to me throughout my PhD have made it the best academic experience I have had in my life. I am deeply grateful to all of you.

I want to express my gratitude to thank Maurice and Phyllis Paykel Trust, New Zealand Heart Foundation, the Massey University School of Health Sciences Postgraduate Fund, the Massey University Doctoral Conference Grant, and Massey University Scholarship for their financial support.

Special thanks to our collaborators Prof. Vicky Cameron, Dr. Anna Pilbrow, Prof. Mark Richards, Prof. Rob Doughty, and Prof. Chris Frampton. for their support in accessing and using samples and data on the clinical cohorts for this project. I also want to thank Dr Vinicius Tragante for his bioinformatic support with the genotype imputation and extraction of specific SNPs. I also want to thank Dr. Anna Pilbrow and Miss Anna Newburn for their support on site at the University of Otago Christchurch, Christchurch Heart Institute Omics Laboratory when collecting samples and going over ELISA methodology.

No academic endeavour can go by without a support network. I would like to thank my two flat brothers Ruiteng (Roy) and Oscar. Roy your energy and the activities you promoted during your stay were amazing. Plus being in touch even after leaving and despite your own circumstances speaks volumes. Oscar, we get on like a house on fire. There are not enough words to describe the full depth of what you've meant to me and what you have taught me during my PhD. Between movies, series, MTG, philosophy, music, and morals you're one of a kind (Go Shintai!). Special mention to Evan and Zoe, you guys added a lot to the flat dynamics (Team Sherlock).

Next are the communities I've been fortunate to meet and enjoy. Massive thanks to all the members past and present of the GRS Steering group (Julia, Steph, Sharan, Alfredo) as well as fellow CoH PhD students Christine, Sa Weon, Jangrez and Vida. Honourable mention to Goblin Games (GG) and Two Right Feet Dance studio who allowed me to develop skills and many friendships. Francesca, your kindness, and passion for dancing are amazing. Lisa & Genna you managed to get me to actually dance and overcome my fears behind it. The community you've built is an achievement deserving of more recognition. From GG Cal, Dan, James, Lee, Vojden, Fraser, Naki, Toby, Theo, Matt, Halley, Matt, Rob, Morgan and so many more. Thursdays, Fridays, and Sundays became a welcome release from PhD work and mental MTG workouts were most welcome (Halloumi Best Girl).

Violeta and Risa, I have known you for the past 10 and 5 years, respectively. It is a blessing having long term friendships like ours where no matter the distance or time I can still rely on you. I hope life can allow us to meet again somewhere.

Dra. Rojas y Dr. Fausto, su apoyo y confianza en mis habilidades permitió que pudiera experimentar los últimos 3 años. Muchas gracias

Por último, gracias a mi familia Rosalinda, Rossy y Zoé. A pesar de los altibajos que nos han tocado en la vida hemos estado juntos. Mama, sé que estas orgullosa y espero Dios me dé chance de que ahora si consiga recursos para alivianar lo que nos queda juntos. Rossy, tienes una razón poderosa para seguir luchando. Animo carnal. Zoé, deseo de corazón que crezcas entendiendo y valorando a tu familia. Tu tío te quiere mucho. Espero que mis andanzas académicas y de vida puedan complementar tu crecimiento.

Abundat et pulchrum mundus est.

Belief makes me who I am.

Nuestro mañana brillara como la luz del Sol.

Table of contents

Note for Examiners of Doctoral Theses i

Thesis abstract iii

Acknowledgements v

Table of contents vii

List of figures xi

List of tables xii

Abbreviations xiii

Glossary of terms xv

Chapter 1.- Preface 1

1.1 Background 2

1.2 Research gap 3

1.3 Overall aim and specific research objectives 3

1.4 Thesis structure 4

1.5 Researcher contributions 6

1.6 References 9

Chapter 2.- Literature Review 11

Statement of contribution 12

2.1 Atherosclerosis pathophysiology 13

2.2 New Zealand CVD epidemiology 14

2.3 Cardiovascular disease risk algorithms 15

2.4 CVDs and single nucleotide polymorphisms 19

2.5 VEGF overview 21

2.6 VEGF-A gene and related SNPs in a cardiovascular context 23

2.7 Conclusion 44

2.8 References 45

Chapter 3.- Materials and Methods 52

3.1 Study populations 53

3.1.1 Coronary Disease Cohort Study (CDCS) 53

3.1.2 Canterbury Healthy Volunteers Study (HVOL) 54

3.1.3 Clinical events 54

3.2 Genetic Methods 55

3.2.1 Variant imputation 55

3.2.2 DNA extraction 56

3.2.3 TaqMan SNP genotyping 57

3.3 Plasma VEGF-A assay 58
3.4 Statistical analysis 60
3.5 Genomic context analysis 61
3.6 References 62
Chapter 4.- Study on *VEGFA* locus SNPs 63
Statement of contribution 64
4.1 Abstract 65
4.2 Introduction 66
4.3 Results 67
4.3.1 Baseline characteristics of study cohorts 67
4.3.2 SNP genotype analysis 68
4.3.3 rs6921438 and rs7767396 metabolite associations 69
4.3.4 Univariate survival association of selected SNPs in the CDCS cohort 73
4.3.5 Multivariate survival association of selected SNPs in the CDCS cohort 75
4.3.6 Genomic context for rs6921438 and rs7767396 78
4.4 Discussion 78
4.5 Conclusion 81
4.6 References 82
Chapter 5.- Study on *VEGFR2* locus SNPs 85
Statement of contribution 86
5.1 Abstract 87
5.2 Introduction 88
5.3 Results 89
5.3.1 Baseline characteristics of study cohorts 89
5.3.2 SNP genotype analysis 91
5.3.3 SNP associations within the CDCS cohort 92
5.3.4 SNP associations within the HVOL subset group 94
5.3.5 Association of genotypes with clinical outcomes in the CDCS cohort 96
5.3.6 Genomic context of *VEGFR2* SNPs 100
5.4 Discussion 101
5.5 Conclusion 104
5.6 References 105
Chapter 6.- Study on *VLDLR* locus SNPs 108
Statement of contribution 109
6.1 Abstract 110

6.2 Introduction 111
6.3 Results 112
6.3.1 Baseline characteristics of study cohorts 112
6.3.2 VLDLR SNP associations within the CDCS cohort 114
6.3.3 VLDLR SNP associations within the HVOL cohort 119
6.3.4 Association of genotypes with clinical outcomes in the CDCS cohort 124
6.3.5 Genomic context of *VLDLR* SNPs 126
6.4 Discussion 127
6.5 Conclusion 129
6.6 References 130
Chapter 7.-Discussion and project conclusions 133
7.1 Introduction 134
7.2 VEGF-A levels in the HVOL and CDCS cohorts 135
7.3 Proposed mechanisms for *VEGFA* locus SNPs 135
7.4 Proposed mechanisms for *VEGFR2* locus SNPs 139
7.5 Proposed mechanisms for *VLDLR* locus SNPs 144
7.6 VEGF-A impact on CVD risk pathways 150
7.7 Project applications and strengths 151
7.8 Project limitations 153
7.9 Future research 155
7.10 Concluding remarks 157
7.11 References 158
Appendices 162
Appendix 1.1 Research outputs 162
Appendix 1.2 Published minireview 163
Appendix 4.1 Imputed genotype comparison for five clinically relevant *VEGFA* locus SNPs in the CDCS cohort 172
Appendix 4.2 Variants associated with ANP levels in HVOL cohort 177
Appendix 4.3 Genotype frequencies for rs6921438 and rs7767396 178
Appendix 4.4. Multivariate HF readmissions model on the CDCS cohort using manual genotypes for rs6921438 and rs7767396 179
Appendix 4.5. Multivariate MACE readmissions model on the CDCS cohort using manual genotypes for rs6921438 and rs7767396 180
Appendix 4.6. Multivariate NSTEMI readmissions model on the CDCS cohort using manual genotypes for rs6921438 and rs7767396 181
Appendix 4.7. Multivariate unstable angina readmissions model on the CDCS cohort using manual genotypes for rs6921438 and rs7767396 182

Appendix 5.1 Imputed data for 13 VEGFR2 SNPs 183
Appendix 5.2 rs1870378 imputed genotype associations within the CDCS cohort 184
Appendix 5.3 rs1870379 imputed genotype associations within the CDCS cohort 185
Appendix 5.4. Variables with p<0.1 from one-way CDCS ANOVA analysis using rs2305948 imputed genotype group comparison. 186
Appendix 5.5. Kaplan-Meier survival plots using imputed genotype data from rs1870377, rs7677779 and rs13136007 for CABG events within the CDCS cohort. 187
Appendix 5.6. Multivariate all cause death model on the CDCS cohort using rs1870377 imputed genotype data 188
Appendix 5.7. Multivariate all cause death model on the CDCS cohort using rs7677779 imputed genotype data 189
Appendix 5.8 Multivariate all cause death model on the CDCS cohort using rs13136007 imputed genotype data 190
Appendix 6.1. Imputed frequencies for *VLDLR* locus SNPs in the CDCS cohort 191
Appendix 6.2. Imputed frequencies for *VLDLR* locus SNPs in the HVOL cohort 192
Appendix 6.3. Multivariate all cause death model on the CDCS cohort using rs2375981 imputed genotype data 193
Appendix 6.4. Multivariate NSTEMI readmissions model on the CDCS cohort using rs2375981 imputed genotype data 194
Appendix 6.5. Multivariate NSTEMI readmissions model on the CDCS cohort using rs10738760 imputed genotype data 195
Appendix 6.6. Multivariate cardiovascular death model on the CDCS cohort using rs10738760 imputed genotype data 196
Appendix 6.7. Multivariate cardiovascular death model on the CDCS cohort using rs2375981 imputed genotype data 197
Appendix 6.8. Multivariate NSTEMI readmissions model on the CDCS cohort using rs10738760 imputed genotype data 198
Appendix 6.9. Multivariate cardiovascular death model on the CDCS cohort using rs7030781 imputed genotype data 199

List of figures

Figure 2.1. Progression of atherosclerosis ... 14

Figure 2.2 New Zealand global disease burden 2021 data for non-communicable diseases ... 15

Figure 2.3 Potential effects of a SNP on VEGF-A gene expression ... 21

Figure 2.4. VEGFA gene and select SNPs on chromosome 6 ... 24

Figure 2.5. Schematic representation of key SNPs in key exons of VEGFR2 ... 37

Figure 3.1 Dilution series production from a standard solution. The initial volume of assay diluent in each tube and the final concentrations used in the dilution series are specified. ... 59

Figure 4.1. Comparison of mean VEGF-A levels for the CDCS and HVOL cohorts for manual genotypes of rs6921438 and rs7767396 ... 70

Figure 4.2. Kaplan-Meier survival plot for imputed genotypes of rs6921438 and rs7767396 in the CDCS cohort. ... 73

Figure 4.3. Kaplan-Meier survival plots of rs6921438 manual genotypes ... 74

Figure 4.4. Kaplan-Meier survival plots of rs7767396 manual genotypes. ... 74

Figure 4.5. Schematic representation of *VEGFA* locus SNPs ... 78

Figure 5.1. Significant associations of rs2305948 imputed genotypes within the CDCS cohort. ... 93

Figure 5.2. Relationship of rs1870377 manual genotypes within the CDCS cohort. ... 93

Figure 5.3. Imputed rs2305948 findings within the HVOL cohort ... 94

Figure 5.4. Imputed *VEGFR2* SNPs associated with VEGF-A levels in the HVOL cohort ... 95

Figure 5.5. Imputed VEGFR2 SNPs associated with triglyceride levels within the HVOL cohort ... 96

Figure 5.6. Kaplan-Meier survival plots of rs2305948 manual genotypes ... 98

Figure 5.7. Schematic representation of *VEGFR2* locus SNPs. ... 100

Figure 6.1. Schematic representation of a 312 kb region containing *VLDLR* locus SNPs ... 126

Figure 7.1. Proposed mechanisms of action for rs6921438 and rs7767396 ... 138

Figure 7.2. Associations of rs2305948 alleles with effector molecules and subsequent impact on risk factors that promote CVD risk or contribute to heart vascular homeostasis ... 141

Figure 7.3. Graphical summary of reported associations of rs1870377 A and mechanisms that contribute to CVD risk or low LDL levels. ... 143

Figure 7.4. Proposed links between rs10738760 A allele and CVD risk ... 146

Figure 7.5. Reported interactions for rs2375981 and potential effector functions in the context of risk and protective alleles ... 148

Figure 7.6. Potential interactions of rs7043199 the risk and protective alleles on VLDLR-AS1 with their effect on lipids and VEGF-A ... 149

Figure 7.7. Proposed interaction of elevated circulating VEGF-A levels on CVD risk factors ... 151

List of tables

Table 2.1. Cardiovascular Disease Risk Algorithms 17

Table 2.2. VEGF-A eQTLS on Chromosome 6........ 28

Table 2.3. Reports on trans-acting SNPs associated with VEGF-A or CVD risk......... 38

Table 3.1. TaqMan SNP Genotyping PCR Conditions 57

Table 4.1. Baseline characteristics of the CDCS and HVOL cohort 68

Table 4.2.- Manual genotype comparison for rs7767396 and rs6921438 the CDCS cohort 71

Table 4.3. Cox proportional hazards regression model for all-cause mortality in the CDCS cohort using imputed genotypes for rs6921438 and rs7767396........ 75

Table 4.4. Cox's proportional hazards regression model for all-cause mortality in the CDCS cohort using manual genotypes for rs6921438 and rs7767396........ 76

Table 4.5. Simplified Cox's proportional hazards regression model for all-cause mortality in the CDCS cohort using manual genotypes for rs6921438 and rs7767396 76

Table 4.6. Cox's proportional hazards regression model for heart failure readmissions in the CDCS cohort including manual genotypes for rs6921438 and rs7767396........ 77

Table 5.1. Baseline characteristics of the CDCS and HVOL cohort 90

Table 5.2. Frequency comparison of manual genotype vs imputed genotype data 91

Table 5.3. Cox's proportional hazards regression model for any cardiovascular readmission in the CDCS cohort using manual genotypes for rs1870377........ 99

Table 5.4. Cox's proportional hazards regression model for all cause death in the CDCS cohort using manual genotypes for rs2305438 99

Table 6.1 Baseline characteristics of the CDCS and HVOL cohort 113

Table 6.2 Genotype comparisons for *VLDLR* SNPs in the CDCS cohort 115

Table 6.3. Genotype comparison for *VLDLR* SNPS in the HVOL cohort........ 120

Table 6.4. Cox's proportional hazards regression model for cardiovascular death in the CDCS cohort using imputed genotypes for rs2375981 125

Table 6.5. Cox's proportional hazards regression model for cardiovascular death in the CDCS cohort using imputed genotypes for rs10738760 125

Abbreviations

°C	Celsius degrees
<	less than
µl	microlitre
µM	micromolar
ACS	Acute Coronary Syndrome
ANOVA	Analysis of Variance
ANP	Atrial Natriuretic Peptide
ApoE	Apolipoprotein E
BMI	Body Mass Index
BNP	B-type Natriuretic Peptide
BP	Blood Pressure
CAC	Coronary Artery Calcium
CAD	Coronary Artery Disease
CDCS	Coronary Disease Cohort Study
CHD	Coronary Heart Disease
Chr Resp	Chronic Respiratory diseases
Chr	Chromosome
CKD	Chronic Kidney Disease
CNP	C-type natriuretic peptide
CVD	Cardiovascular Disease
cvdGRS	CVD Genetic Risk Scores
dL	Decilitre
DBP	Diastolic Blood Pressure
DNA	Deoxyribonucleic Acid
EC	Endothelial Cell
ECG	Electrocardiogram
ECM	Extracellular Matrix
ELISA	Enzyme Linked Immunosorbent Assay
eQTL	Expression Quantitative Trait Loci
FRS	Framingham Risk Score
g	G-Force
GWAS	Genome Wide Association Studies
HbA1c	Haemoglobin A1C
HDL	High Density Lipoprotein
HF	Heart Failure
HR	Hazards Ratio
HVOL	Healthy Volunteer Cohort
ICAM	Intercellular Adhesion Molecule
IHD	Ischaemic Heart Disease
IL	Interleukin
IL-12	Interleukin 12
IL-6	Interleukin 6
LDL	Low Density Lipoprotein
LPL	Lipoprotein Lipase
lncRNA	Long Non-Coding RNA
MAF	Minor Allelic Frequency
MACE	Major Adverse Cardiovascular Events
MELAA	Middle Eastern/Latin American/African

MESA	Multi-Ethnic Study Of Atherosclerosis
MI	Myocardial Infarction
Min	Minutes
mg	Milligrams
mL	Millilitre
mM	Millimolar
mmHg	Millimetres of mercury
mRNA	Messenger RNA
MTA	Material Transfer Agreement
N/A	Not Applicable
NRP1	Co-Receptor Neuropilin-1
NSTEM	Non-ST Elevated Myocardial Infarction
NSTEMI	Non-ST-Elevation Myocardial Infarction
NT-ANP	Amino Terminal Atrial Natriuretic Peptide
NT-CNP	Amino Terminal C-Type Natriuretic Peptide
NTproBNP	Amino Terminal Pro B-Type Natriuretic Peptide
NVDPA	National Vascular Disease Prevention Alliance
NZ	New Zealand
NZPPE	New Zealand Primary Prevention Equation
OD	Optical Density
Ox-LDL	Oxidised LDL
PCE-ASCVD	Pooled Cohort Equations For Atherosclerotic Cardiovascular Disease
PCR	Polymerase Chain Reaction
PDGF	Platelet-Derived Growth Factor
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
S	Seconds
SBP	Systolic Blood Pressure
SCORE	Systematic Coronary Risk Evaluation
sFlt	Soluble Fms-like Tyrosine Kinase-1
sKDR	Soluble Kinase Insert Domain Receptor
SMC	Smooth Muscle Cells
SNP	Single Nucleotide Polymorphism
STEMI	ST-Elevation Myocardial Infarction
sVEGFR1	Soluble VEGF Receptor 1
sVEGFR2	Soluble VEGF Receptor 2
T2DM	Type 2 Diabetes Mellitus
TNF-α	Tumour Necrosis Factor Alpha
UA	Unstable Angina
USA	United States of America
VEGF	Vascular Endothelial Growth Factor
VEGF-A	Vascular Endothelial Growth Factor A
VEGFR1	Vascular Endothelial Growth Receptor 1
VEGFR2	Vascular Endothelial Growth Receptor 2
VLDL	Very Low-Density Lipoprotein
VLDLR	Very Low-Density Lipoprotein Receptor
WHO	World Health Organisation
χ^2	Chi-squared test

Glossary of terms

Cardiac death event: Any death due to a heart condition (myocardial infarction, arrythmia, heart failure)

Cardiovascular death event: Any death caused by cardiac death plus those that involve vascular conditions (e.g. stroke).

Enhancer: DNA region that can be bound by proteins to increase gene transcription.

eQTL: Genetic loci that explain variation in expression levels of mRNA.

Exon: Sections of DNA that code for proteins

Gene transcription: Process by which DNA is converted into RNA.

Genotype imputation: A process of estimating undetermined genotypes using a genotype reference panel.

Genotype: refers to the observed pair of alleles or variants that a person has.

Hazards Ratio: A measure of how often an event occurs in one group compared to how often the same event occurs in another group.

Hypercholesterolaemia: Lipid disorder where low-density lipoprotein cholesterol levels are above the normal threshold.

Hyperglycaemia: Condition where blood glucose levels are greater than 125 mg/dL.

Hypertension: Condition where blood vessel pressure is higher than 140 for systolic blood pressure and/or 90 mmHg for diastolic blood pressure

Hypoxia: Condition that occurs when there are insufficient blood supply or tissue oxygen levels.

Intron: A noncoding section of DNA that may be removed during DNA transcription.

Locus: A location within the genome where a particular gene is located.

Manual genotyping: Experimental procedure to identify the genetic variant(s) an individual carries.

Non-coding gene: Sequence of DNA nucleotides that do not produce a functional RNA but may interact with other genetic regions to modify gene transcription.

Nucleotide: Organic molecules that are the units of DNA or RNA.

Phenotype: Observable characteristics of an individual for a particular trait that result from the interaction between gene and environment

Promoter: A sequence of DNA where proteins bind to start transcription.

Protein-coding gene: Sequence of DNA nucleotides that are transcribed to produce a functional RNA.

Risk factor: Any variable associated with increased disease risk.

SNP: Genetic variants that cause a change of a single nucleotide at a specific position in the genome.

Transcription Factor: A protein that can control or modify DNA transcription.

VEGF-A eQTL SNP: Genetic variants that are associated with changes in expression levels of *VEGFA* .

Chapter 1.- Preface

This chapter provides a brief background on cardiovascular disease, vascular endothelial growth factor A (VEGF-A) and the rationale that justifies the PhD research work undertaken. The overall aim and specific research objectives are highlighted followed by an outline of the structure of the thesis and contributions of each researcher involved in this PhD project.

1.1 Background

Cardiovascular disease (CVD) is considered by the World Health Organisation (WHO) as the leading cause of death worldwide [1]. In New Zealand (NZ), CVDs represent a major health burden given they account for 14 deaths per day and result in high economic cost for the healthcare sector [2, 3]. Currently, primary prevention occurs through CVD risk algorithms which assess an individual's risk based on the measurement of well-established markers including age, gender, weight, cholesterol levels, blood pressure and metabolic dysfunction [4]. Preventative measures focus on risk factors that are age-dependent, but there is increasing evidence that genetic variants, which can be assayed at any age, influence cardiovascular risk phenotypes [5, 6]. Evidence on the clinical utility of single nucleotide polymorphisms (SNPs) has been observed in international studies where risk genotypes are used for treatment prioritisation, lifestyle recommendations and prediction of adverse events [7-11].

The vascular endothelial growth factor A (VEGF-A) participates in processes vital to the cardiovascular system, notably blood vessel formation [12-14]. VEGF-A can relieve arterial blockages by creating new blood vessels to recover blood circulation [12, 15]. However, elevated circulating levels of VEGF-A have been associated with risk factors that promote the onset of cardiovascular disease (CVD) [14, 16]. CVD onset mechanisms that increase VEGF-A production include high low-density lipoprotein (LDL) concentration, hypoxia, and interleukin activity [16, 17]. Research has shown VEGF-A is associated with inflammation, increased blood pressure and an elevated lipoprotein profile [18-21].

SNP alleles that are present more frequently in diagnosed individuals are considered genetic risk variants which can influence genes or disease phenotypes [22, 23]. In CVD research, SNPs have provided an insight into underlying pathological mechanisms and allow for patient risk profiling or stratification [22, 24, 25]. Studies have shown SNPs influencing circulating levels of VEGF-A, referred to as VEGF-A expression quantitative trait loci (eQTL) SNPs, are present on different human chromosomes [26, 27]. Some eQTL SNPs for VEGF-A have been associated with elevated circulating risk markers implicated in CVD pathogenesis including blood pressure, inflammatory cytokines, lipoprotein profile and natriuretic peptides [20, 28-33]. Studies have also identified eQTL SNPs for VEGF-A are associated with increased risk of metabolic syndrome, which can contribute to CVD onset [32, 34, 35]. Furthermore, there is evidence that specific VEGF-A eQTL SNPs have been associated with increased CVD risk and adverse coronary events [29, 36, 37]. Overall, international reports provide evidence on involvement of eQTL SNP for VEGF-A in CVD onset, development, and

progression that warrants further exploration.

1.2 Research gap

The study of genetic variants represents a developing research field that can bridge the gap between consolidating knowledge on a novel cardiovascular biomarker while identifying SNPs with clinical applications for the NZ population. The identification of SNPs associated with increased disease or clinical outcome risk can complement existing CVD risk profiling and stratification tools.

NZ based research focused on VEGF system SNPs managed to identify specific variants were associated with soluble VEGF receptor plasma levels in patients diagnosed with acute coronary syndromes [29]. Subsequent work also showed promising associations between eQTL SNPs for VEGF-A, natriuretic peptide levels and patient survival [38]. These studies employed a cohort of acute coronary syndrome patients with extensive inclusion parameters and follow-up data available. This resource represents an ideal opportunity that can support and benefit from the identification of additional eQTL SNPs for VEGF-A.

It should be possible to explore potential molecular mechanisms specific to any identified SNPs that relate them to VEGF-A activity, well-established CVD risk markers or CVD onset mechanisms. Conjointly, relevant eQTL SNPs for VEGF-A can further support a clinician's counsel. There is high likelihood that eQTL SNPs for VEGF-A could be prognostic markers for CVD risk. This may complement ongoing examination of a patient's condition following a therapeutic intervention or coronary event.

1.3 Overall aim and specific research objectives

The research work presented aims to provide an exploration on eQTL SNPs for VEGF-A of clinical relevance within established NZ cohorts: the Coronary Disease Cohort Study (CDCS) and the Canterbury Healthy Volunteers Study (HVOL).

Using these cohorts, this PhD project can contribute towards the development of knowledge on eQTL SNPs for VEGF-A while exploring their impact on CVD risk. Therefore the overall aim of this PhD thesis is to investigate VEGF-A eQTL SNPs to determine their association with circulating VEGF-A

plasma levels or circulating levels of metabolites routinely measured in coronary disease cohorts. Specific aims were:

A. To examine if eQTL SNPs for VEGF-A were associated with clinical outcome risk to identify novel CVD risk variant candidates and
B. To explore the genomic context for clinically relevant eQTL SNPs for VEGF-A to identify how they might be involved with CVD risk pathway mechanisms.

To achieve these aims, the project had the following research objectives:

1) Identify from the literature eQTL SNPs for VEGF-A that have been reported to influence circulating VEGF-A levels and show association with CVD or CVD risk factors.
2) Analyse imputed genotyping data on variants identified by Objective 1 for two NZ based cohorts: one of clinically diagnosed coronary patients (CDCS) and one of heart healthy individuals (HVOL).
3) Select eQTL SNPs for VEGF-A that have high likelihood of presenting clinical relevance based on their statistical associations with cardiometabolic and anthropometric variables measured in both cohorts.
4) Determine experimentally the genotype for SNPs selected as part of Objective 3 and VEGF-A plasma levels for a subset of the heart healthy cohort.
5) Identify if eQTL SNPs for VEGF-A from Objective 4 present associations with survival and/or risk of clinical outcomes in both the heart healthy cohort and the coronary disease cohort.
6) Propose potential pathways that would explain observed associations for SNPs from Objective 5.

1.4 Thesis structure

The thesis consists of seven chapters submitted in a Thesis with Publications format. The content of the PhD project has been disseminated throughout the years in various forms of research outputs (Appendix 1.1). Some chapters contain material that has or will be published in relevant peer-reviewed scientific journals following PhD examination. Prospective research outputs of the PhD project are also described in Appendix 1.1.

Due to overarching similarities in methodology, limitations, and implications for the results chapters (4-6) a specific methods chapter has been developed (Chapter 3) and the limitations of the studies have been discussed in Chapter 7 (Discussion and Conclusion). There are three individual results chapters, each focused on SNPs located on one of three human chromosomes (6, 4 and 9) that involve genes (*VEGFA*, *VEGFR2*, *VLDLR*) relevant to the research. Lastly, each result chapter contains a statement of contribution according to Massey University guidelines, an abstract and a brief introduction which are followed by results, specific discussion, and conclusion.

A brief description of each chapters' content is as follows:

<u>Chapter 1 – Preface</u>: This chapter contains a brief background on the PhD research topic. Additional sections provide an overview on the research gap, overall aim, and research objectives of the research work. Lastly, a description on the structure of the PhD Thesis is provided and researcher contributions to the PhD research studies.

<u>Chapter 2 – Literature review</u>: This chapter includes material describing CVD epidemiology (including NZ reports), a description of existing information on CVD risk algorithms and the biological process underlining atherosclerosis, a key feature leading to heart disease onset. This chapter also explores the link of SNPs to CVD and discusses the VEGF system, VEGF-A and eQTL SNPs for VEGF-A. The second portion of the chapter has been published during the PhD as a minireview in the journal *Frontiers in Cardiovascular Medicine* in the section *Cardiovascular Genetics and Systems Medicine* (Appendix 1.2).

<u>Chapter 3 – Material and Methods:</u> This chapter includes details on the methodological procedures conducted during the PhD project. The main sections involve cohort description, imputation genotyping, experimental analysis of DNA and plasma samples (where applicable) as well as the statistical methods employed.

<u>Chapter 4 – Study on *VEGFA* locus SNPs</u>: This chapter focuses on the work involving 30 SNPs mapping to the *VEGFA* locus (human chromosome 6) which had associations reported with cardiometabolic parameters or CVD risk in the existing literature. The chapter highlights the experimental findings surrounding two VEGF-A eQTL SNPs located in a non-coding region near *VEGFA*. The manuscript version of this chapter is under review by the external co-authors listed in

Section 1.5. This research work will be submitted to the BMC Cardiovascular Disorders journal by the time of thesis examination.

Chapter 5 – Study on *VEGFR2* locus SNPs: This chapter focuses on the work involving variants located at the gene for VEGF receptor 2 (*VEGFR2*), the main signal inducer of VEGF-A activity. The chapter highlights findings on four variants with clinical implications following the analysis of imputed genotypes for 13 SNPs mapping to the *VEGFR2* gene.

Chapter 6 – Study on *VLDLR* locus SNPs: This chapter focuses on genetic variants identified at the locus for the very low-density lipoprotein receptor (*VLDLR*). The chapter highlights findings on the analysis of 4 SNPs that have been associated with VEGF-A circulating levels and could be associated with CVD risk. The chapter shall be submitted to a relevant journal following the manual genotyping of the relevant clinically novel variants identified.

Chapter 7 – Discussion and Project Conclusions: This chapter aims to provide an overview on the difference in VEGF-A plasma levels between CVD and heart healthy cohorts and a discussion on potential mechanisms for specific alleles of eQTL SNPs for VEGF-A discussed in Chapters 4 to 6. This chapter also discusses the overarching limitations and strengths that are shared amongst Chapters 4 to 6. Lastly, the prospective future directions for this research, as well as concluding remarks, are included.

1.5 Researcher contributions

This PhD project has involved the collaboration of multiple researchers within the NZ academic community. Their contributions are summarised below.

Researcher	Contributions
Juan Carlos Meza Alvarado PhD researcher	Responsible for all aspects of the PhD project including grant writing, project design, SNP variant selection, drafting and updating inter-University material transfer agreement, data management, laboratory experiments (TaqMan SNP genotyping for variants on Chromosome 6, VEGF-A measurements on heart healthy plasma samples), statistical analyses using genotype data (imputed and manually generated), thesis writing. Responsible for major contributions pertaining manuscript writing including: conceptualisation, investigation, data curation and analysis, visualisation, draft writing, reviewing, editing, and submission. Responsible for major contributions to research dissemination including abstract writing and presentation of research findings
Dr Barry Palmer Primary Supervisor	Responsible for research project conceptualisation, methodology, funding acquisition, drafting inter-University material transfer agreement, provided laboratory and statistical support, reviewing of all research outputs (thesis chapters, poster abstracts, all manuscripts)
Associate Professor Rachel Page Co-supervisor	Supported funding acquisition, resource access, grant writing and reviewing of all research outputs (thesis chapters, poster abstracts, all manuscripts).
Dr Beth Mallard Co-supervisor	Supported funding acquisition, provided laboratory support and review of all research outputs (thesis chapters, poster abstracts, all manuscripts).
Dr Collette Bromhead Co-supervisor	Supported funding acquisition, provided laboratory support and review of all research outputs (thesis chapters, poster abstracts, all manuscripts).
Dr Anna Pilbrow Collaborator	Current research leader of the Christchurch Heart Institute Omics Laboratory in University of Otago Christchurch campus. Provided support in drafting and updating an inter-University material transfer agreement, management of both cohort samples and databases, reviewed conference abstracts and research article manuscripts.
Professor Mark Richards Collaborator	Current director of the Christchurch Heart Institute (University of Otago). Major role in conceptualising, funding acquisition, clinical oversight and reporting for CDCS. Steward of the CDCS cohort samples and databases. Provided approval for development of an inter-University material transfer agreement and use of

	material (samples, datasets). Provided support in the review of conference abstracts and research article manuscripts.
Dr Vicky Cameron Collaborator	Former research leader at the Christchurch Heart Institute (University of Otago). Provided approval and support of an inter-University material transfer agreement. Reviewed research article manuscripts.
Ms Anna Newburn	Research technician at the Christchurch Heart Institute Omics Laboratory in University of Otago Christchurch Campus. Provided laboratory support on VEGF-A measurement, access to plasma and DNA samples for both NZ cohorts.
Professor Rob Doughty Collaborator	Steward of the CDCS cohort samples and databases. Based at the University of Auckland. Provided approval for development of an inter-University material transfer agreement and use of material (samples, datasets). Provided support in the review of conference abstracts and research article manuscripts.
Professor Christopher Frampton Collaborator	Provided support on the methodology for statistical analysis of the cohort data sets. Based at the Christchurch Heart Institute (University of Otago).
Dr. Vinicius Tragante	Bioinformatician from deCODE genetics (Reykjavik, Iceland) responsible for providing imputed genotype data of selected SNP variants.

1.6 References

1. World Health Organisation. *Cardiovascular diseases (CVDs)*. 2021 11 June 2021; Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
2. Wilson, N., et al., *Prioritisation of intervention domains to prevent cardiovascular disease: a country-level case study using global burden of disease and local data.* Population Health Metrics, 2023. **21**(1): p. 1.
3. Martin Tobias and Maria Turley, *Health Loss in New Zealand: A report from the New Zealand Burden of Diseases, Injuries and Risk Factors Study, 2006-2016.* 2016, Ministry of Health: Wellington. p. 94.
4. Damen, J.A.A.G., et al., *Prediction models for cardiovascular disease risk in the general population: systematic review.* British Medical Journal, 2016. **353**: p. i2416.
5. McPherson, R. and A. Tybjaerg-Hansen, *Genetics of Coronary Artery Disease.* Circulation Research, 2016. **118**(4): p. 564-78.
6. Gladding, P.A., et al., *Polygenic Risk Scores in Coronary Artery Disease and Atrial Fibrillation.* Heart, Lung, and Circulation, 2020. **29**(4): p. 634-640.
7. Mega, J.L., 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. **385**(9984): p. 2264-2271.
8. Khera, A.V., et al., *Genetic risk, adherence to a healthy lifestyle, and coronary disease.* New England Journal of Medicine, 2016. **375**(24): p. 2349-2358.
9. Xu, Z., et al., *Association between a genetic risk score based on single nucleotide polymorphisms of coronary artery disease-related genes and left main coronary artery disease.* BioMed Research International, 2018. **2018**.
10. Xu, K., et al., *The combined effects of cardiovascular disease related SNPs on ischemic stroke.* Journal of the Neurological Sciences, 2018. **388**: p. 141-145.
11. Christiansen, M.K., et al., *A 45-SNP genetic risk score is increased in early-onset coronary artery disease but independent of familial disease clustering.* Atherosclerosis, 2017. **257**: p. 172-178.
12. Camaré, C., et al., *Angiogenesis in the atherosclerotic plaque.* Redox biology, 2017. **12**: p. 18-34.
13. Melincovici, C.S., et al., *Vascular endothelial growth factor (VEGF) – key factor in normal and pathological angiogenesis.* Romanian Journal of Morphology and Embryology, 2018. **59**(2): p. 455-467.
14. Smith, G.A., et al., *Vascular endothelial growth factors: multitasking functionality in metabolism, health, and disease.* Journal of Inherited Metabolic Disease, 2015. **38**(4): p. 753-763.
15. Zimarino, M., et al., *The dynamics of the coronary collateral circulation.* Nature Reviews Cardiology, 2014. **11**(4): p. 191-197.
16. Braile, M., et al., *VEGF-A in Cardiomyocytes and Heart Diseases.* International Journal of Molecular Sciences, 2020. **21**(15): p. 1-18.
17. Li, Z., et al., *Missense Variants in Hypoxia-Induced VEGFA/VEGFR2 Signaling Predict the Outcome of Large Artery Atherosclerotic Stroke.* Cellular and Molecular Neurobiology, 2021. **41**(6): p. 1217-1225.
18. Heinonen, S.E., et al., *The effects of VEGF-A on atherosclerosis, lipoprotein profile, and lipoprotein lipase in hyperlipidaemic mouse models.* Cardiovascular Research, 2013. **99**(4): p. 716-23.
19. Zhao, Z., et al., *Association between Single Nucleotide Polymorphisms in Cardiovascular Developmental Critical Genes and Hypertension: A Propensity Score Matching Analysis.* International Journal of Hypertension, 2020. **2020**: p. 1-8.
20. Ruotsalainen, S.E., et al., *An expanded analysis framework for multivariate GWAS connects inflammatory biomarkers to functional variants and disease.* European Journal of Human Genetics, 2021. **29**(2): p. 309-324.
21. Zhao, X., et al., *Vascular endothelial growth factor gene polymorphisms and coronary heart disease: a systematic review and meta-analysis.* Growth Factors, 2018. **36**(3-4): p. 153-163.
22. Roberts, R., C.C. Chang, and T. Hadley, *Genetic Risk Stratification: A Paradigm Shift in Prevention of Coronary Artery Disease.* Journal of the American College of Cardiology: Basic to Translational Science, 2021. **6**(3): p. 287-304.

23. Roberts, R., A. Campillo, and M. Schmitt, *Prediction and management of CAD risk based on genetic stratification.* Trends in Cardiovascular Medicine, 2020. **30**(6): p. 328-334.
24. Geiger, M., *Fundamentals of vascular biology*. Learning materials in biosciences. 2019, Vienna, Austria: Springer.
25. Fiatal, S. and R. Ádány, *Application of Single-Nucleotide Polymorphism-Related Risk Estimates in Identification of Increased Genetic Susceptibility to Cardiovascular Diseases: A Literature Review.* Frontiers in Public Health, 2018. **5**: p. 358-358.
26. Choi, S.H., et al., *Six Novel Loci Associated with Circulating VEGF Levels Identified by a Meta-analysis of Genome-Wide Association Studies.* Public Library of Science (PLOS) Genetics, 2016. **12**(2): p. e1005874.
27. Debette, S., et al., *Identification of cis- and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels.* Circulation Research, 2011. **109**(5): p. 554-63.
28. Skrypnik, D., et al., *Association of rs699947 (-2578 C/A) and rs2010963 (-634 G/C) Single Nucleotide Polymorphisms of the VEGF Gene, VEGF-A and Leptin Serum Level, and Cardiovascular Risk in Patients with Excess Body Mass: A Case-Control Study.* Journal of Clinical Medicine, 2020. **9**(2): p. 1-16.
29. Marks, E.C.A., et al., *Plasma levels of soluble VEGF receptor isoforms, circulating pterins and VEGF system SNPs as prognostic biomarkers in patients with acute coronary syndromes.* BMC Cardiovascular Disorders, 2018. **18**(1): p. 169.
30. Azimi-Nezhad, M., et al., *Associations of vascular endothelial growth factor (VEGF) with adhesion and inflammation molecules in a healthy population.* Cytokine, 2013. **61**(2): p. 602-7.
31. Stathopoulou, M.G., et al., *A common variant highly associated with plasma VEGFA levels also contributes to the variation of both LDL-C and HDL-C.* Journal of Lipid Research, 2013. **54**(2): p. 535-41.
32. Azimi-Nezhad, M., et al., *The Relationship Between Vascular Endothelial Growth Factor Cis- and Trans-Acting Genetic Variants and Metabolic Syndrome.* American Journal of the Medical Sciences, 2018. **355**(6): p. 559-565.
33. Ahola-Olli, A.V., et al., *Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors.* American Journal of Human Genetics, 2017. **100**(1): p. 40-50.
34. Hoseini, Z., et al., *VEGF gene polymorphism interactions with dietary trace elements intake in determining the risk of metabolic syndrome.* Journal of Cellular Biochemistry, 2019. **120**(2): p. 1398-1406.
35. Salami, A. and S. El Shamieh, *Association between SNPs of Circulating Vascular Endothelial Growth Factor Levels, Hypercholesterolemia and Metabolic Syndrome.* Medicina, 2019. **55**(8): p. 1-9.
36. Wang, Y., et al., *Vascular endothelial growth factor A polymorphisms are associated with increased risk of coronary heart disease: A meta-analysis.* Oncotarget, 2017. **8**(18): p. 30539-30551.
37. Li, L., et al., *Association of Genetic Polymorphisms on Vascular Endothelial Growth Factor and its Receptor Genes with Susceptibility to Coronary Heart Disease.* Medical Science Monitor, 2016. **22**: p. 31-40.
38. Palmer, B.R., et al., *Vascular endothelial growth factor-A promoter polymorphisms, circulating VEGF-A and survival in acute coronary syndromes.* PLOS One, 2021. **16**(7): p. 1-15.

Chapter 2.- Literature Review

The purpose of this chapter is to provide the reader with the background on genetics in cardiovascular disease (CVD), VEGF-A activity and current knowledge on VEGF-A related single nucleotide polymorphisms (SNPs) implicated in CVDs. This work served as a main focal point to identify genetic variants with implications in CVD or associations with cardiometabolic risk factors to fulfil Research Objective 1 (Chapter 1 Section 1.3) This content is included as a reformatted and restructured version of the published minireview article "VEGF-A related SNPs: a cardiovascular context", authored by Juan Carlos Meza Alvarado, Associate Professor Rachel Page, Dr Beth Mallard, Dr Collette Bromhead and Dr Barry Palmer (Sections 2.4 to 2.7). The article was published in Frontiers in Cardiovascular Medicine on May 23rd, 2023 (Appendix 1.2). Supplementary materials from the publication have been reformatted and introduced within the text of this chapter.

Additional content included in this chapter focuses on specific CVD data relevant to NZ, a description on existing CVD risk algorithm tools and a description on the molecular processes involved in atherosclerosis, a hallmark of CVD onset. This content is presented as sections 2.1 to 2.3 of this chapter.

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the student and the student's main supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the student's contribution as indicated below in the Statement of Originality.

Student name:	Juan Carlos Meza Alvarado
Name and title of main supervisor:	Dr. Barry Palmer, Senior Lecturer
In which chapter is the manuscript/published work?	Chapter 2

Describe the contribution that the student and members of the supervisory team have made to the manuscript/published work:[1]

The student was responsible for all aspects pertaining manuscript writing including: conceptualization, investigation, visualization, draft writing, reviewing, editing, and submission.

The supervisory team provided feedback of the manuscript drafts throughout the complete submission process.

Please select one of the following three options:

◉	The manuscript/published work is published or in press Please provide the full reference of the research output: Meza Alvarado, J.C., Page R.A, Mallard B., Bromhead C. Palmer, B.R. (2023). VEGF-A related SNPs: a cardiovascular context. Frontiers of Cardiovascular Medicine 10:1190513. https://doi.org/10.3389%2Ffcvm.2023.1190513
○	The manuscript is currently under review for publication Please provide the name of the journal:
○	It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Student's signature:	[signature]	Main supervisor's signature:	[signature]

This form should be placed at the beginning of each relevant thesis chapter.

Doctoral Research Committee

May 2023

2.1 Atherosclerosis pathophysiology

Atherosclerosis is a chronic inflammatory disease that affects arteries, with the most susceptible being coronary, iliofemoral and carotid arteries [1]. Atherosclerosis accounts for the majority of worldwide cardiovascular disease (CVD) mortality [2]. Some of the factors that promote the appearance of atherosclerosis include hypertension, lipid metabolism disruption, hyperglycaemia, and cigarette smoke toxins [2-4]. Atherogenesis can be divided into three stages: initiation, progression, and clinical complication [3]. The three stages are shown in Figure 2.1 and described below.

The initiation stage is characterised by endothelial cell (ECs) reacting to substances (tobacco toxins, aberrant cholesterol deposits) or hemodynamic stress. This promotes EC activation leading to increased permeability, higher inflammatory molecule production, increased expression of adhesion molecules on the cell surface, increased production of reactive oxygen species (ROS) and increased release of angiogenic growth factors, such as the vascular endothelial growth factor A (VEGF-A) [3, 5, 6]. During this first stage of atherosclerosis, LDL transcytosis occurs towards the subendothelial layer of the intima due to increased cell permeability [2]. Given the presence of ROS, LDLs are oxidised (ox-LDLs) which further triggers a proinflammatory response [2, 7]. The presence of chemoattracting cytokines and surface adhesion molecules, induce the recruitment of circulating monocytes and T lymphocytes that migrate into the intima [2, 7].

The second stage of atherogenesis involves the differentiation of infiltrated monocytes into macrophages. The macrophages upregulate scavenger receptors, which take up ox-LDLs to become lipid-laden foam cells [1-3]. Ongoing accumulation of lipids, recruitment of circulating monocytes and foam cell formation leads to the appearance of a fatty streak lesion in the artery. Cholesterol efflux mechanisms become destabilised causing endoplasmic reticulum stress on macrophages leading to their death. and an upregulation of growth factors (e.g., VEGF-A) which induce EC proliferation and smooth muscle cell (SMC) migration [3, 8]. Subsequently, the extracellular matrix (ECM) components interact with inflammatory cytokines causing calcification of the artery and increase lipid retention [9, 10]. Lastly, in response to the hypoxic and inflammatory conditions, the vessel wall undergoes angiogenesis mediated by VEGF-A. This neovascularisation enhances lesion progression, increases macrophage infiltration, lipid deposition and vessel wall thickening [1, 6]. Overall, the defective clearance of cellular debris, the heightened inflammatory response and ongoing monocyte infiltration give rise to atherosclerotic plaque [1, 2].

The final process associated with atherosclerosis involves clinical complications caused by plaque behaviour. The fibrous cap of the plaque can be catalogued as stable or unstable according to its composition [1]. In the case of a stable cap, plaque growth is slow and blood flow is decreased which impacts blood circulation [1, 8]. An unstable plaque occurs when proinflammatory cytokines thin out the fibrous cap by inducing SMC apoptosis, inhibiting ECM component synthesis, and overexpressing proteases that degrade existing ECM molecules [2, 3]. When the thin fibrous cap cannot resist physical stress, the plaque will break and release its contents triggering blood clotting in the vessel [3]. This can lead to occlusion (by interacting with circulating platelets) and a coronary event (e.g., myocardial infarction) [1, 8, 11]. Regardless of the plaque type, ischaemia and plaque growth are the main events that lead to CHD.

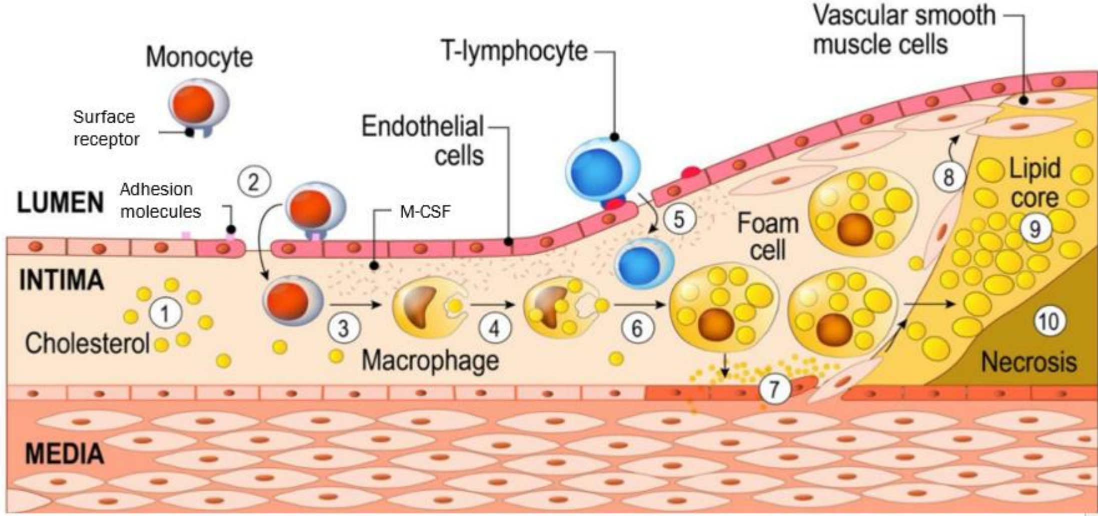


Figure 2.1. Progression of atherosclerosis Adapted from [7]. 1) Cholesterol molecules (e.g. oxidised-LDLs) induce the activation of endothelial cells and expression of adhesion molecules 2) circulating monocytes adhere and migrate to the intima 3) the macrophage colony stimulating factor (M-CSF) promotes monocyte differentiation into macrophages 4) the macrophages recognise ox-LDLs through scavenger receptors initiating a proinflammatory response including production of chemoattractant molecules 5) circulating T lymphocytes can also migrate and participate in lesion progression 6) the lipid-charged macrophages (foam cells) increase the proinflammatory environment 7) the destabilised cholesterol efflux causes macrophage apoptosis and there is increased endothelial cell proliferation 8) smooth muscle cell migration from the media to the intima occurs 9) the vessel thickens due to an accumulation of proinflammatory molecules, minerals (Ca^{2+}), lipids, endothelial cells and smooth muscle cells that may be supported by vasa vasorum formation. 10) There is an accumulation of necrotic bodies and cell debris that form the necrotic core

2.2 New Zealand CVD epidemiology

In NZ non-communicable diseases account for 90% of deaths, with CVDs contributing to a third of them (Figure 2.2) [12]. Specifically, CVDs have been identified as a major cause of mortality, and it is estimated that they cause 14 deaths per day [13]. Other NZ estimates show one out of every twenty individuals have CVD and Māori CVD mortality is twice that of non-Māori individuals [14]. Recent

data shows that CVD deaths in NZ are attributed to high systolic blood pressure, dietary intake and high LDL levels [15, 16]. Furthermore, it has been recently proposed that intervention prioritisation based on high systolic blood pressure presence would contribute to a reduction of CVD related health loss in NZ [15]. This agrees with previous evidence where increased health loss was observed in people with high blood cholesterol, high blood pressure, and tobacco use [13]. A national primary care study identified that patients of Māori, Pasifika, and Indian heritage had increased risk of presenting CVD than European patients [17]. Lastly, Māori and Pasifika ancestry individuals have higher probability of presenting CVD risk factors and were 1.5 times more likely to be hospitalised than patients of European ancestry [13, 14, 18]. The statistical data represents evidence on the negative impact of CVD in the NZ community, including high risk ethnic groups, and identifies risk factors critical for CVD prognosis.

Figure 2.2 New Zealand global disease burden 2021 data for non-communicable diseases. Percentages of death for both sexes in 2021. Chr Resp: Chronic respiratory diseases, CKD: chronic kidney disease, CVD: cardiovascular diseases, Neuro: Neurological disorders, Oth NCD: other non-communicable diseases, Skin: Skin and subcutaneous diseases, Subs use: substance use disorders. Values not shown for causes with values below 0.5%. Image from [16]

2.3 Cardiovascular disease risk algorithms

CVD risk factors have been grouped into three categories by the WHO [19]: behavioural (tobacco use, physical activity, diet, alcohol use), metabolic (blood pressure levels, glycaemia, lipidaemia, weight) and others (age, gender, genetics) [19]. These categories can be further simplified to consider risk factors as modifiable or non-modifiable. The modifiable risk factors involve lifestyle circumstances that are present within the behavioural and metabolic risk factors [20]. Meanwhile, age, genetics and ethnicity at an individual level are non-modifiable. From all variables, the typical factors used in CVD

risk assessments include age, hypercholesterolaemia, HDL cholesterol, sex, smoking, diabetes, and systolic blood pressure [21].

Considering that risk factors can interact synergistically to promote CVDs, numerous studies around the world have developed risk algorithms. Their goal is to provide an insight into an individual's risk of presenting with a CVD, based on specific variables. Table 2.1 summarises examples of these algorithms, including their focus, limitations, populations from which they were derived, key variables considered and predicted outcomes.

The main goal of multivariate risk models is to estimate CVD risk in asymptomatic individuals based on multiple risk factors. Most of the models predict atherosclerotic CVDs or coronary events over a 10-year period except for the Australian (NVDPA) and New Zealand algorithms which are 5-year assessments [17]. Furthermore, each algorithm is tailored to their specific population, which limits their international use. The clinical implementation of these algorithms can benefit from including variables that are not routinely collected for the general population such as socioeconomic deprivation and health inequity [17]. Precision medicine using the risk scores should include these societal factors to avoid undertreating populations with high CVD risk [17].

In the 1990s, NZ introduced CVD risk prediction calculators, derived from the Framingham Risk Score (Table 2.1), to help clinicians identify high risk patients. However, in 2003, NZ guidelines recommended modifying the American equation to address the inequity of CVD risk management when estimating risk for ethnicities such as Māori, Pasifika and South Asians [37]. This led to the development of the PREDICT research study, which had the goal of deriving CVD risk prediction equations using local data [25]. The study used the PREDICT software tool implemented by the Ministry of Health, coupled to the nationwide health index, allowing GPs to continuously collate full data sets from NZ patients enrolled in primary health organisations [17]. In 2018, a guideline update modified the NZ Primary Prevention Equation (NZPPE) to include risk thresholds that prompt clinicians to start pharmacological treatments, provide advice on a recommended age when risk assessment should start and include mental illness as a risk factor [25]. Overall, the inclusion of these variables along with administrative data, would provide good levels of discrimination to be achieved at national, regional, and ethnic levels [38]. Furthermore, the five-year approach allows for a better follow up, given this time length is similar to randomised controlled trials of preventive medications and the period considers short term changes to risk management [25, 39].

Table 2.1. Cardiovascular Disease Risk Algorithms

Algorithm Name	Years updated	Focus & limitations	Population	Key variables	Predicted Outcomes
Framingham Risk Score (FRS)	1998, 2008	Multigenerational study that analyses patterns of CVD risk. It has been validated for mixed American populations. The FRS has been shown to overestimate risk in people of European heritage and underestimate risk in South Asian ethnicities [22].	North Americans	Age, BMI, Diabetes mellitus, HDL cholesterol, Smoking status, Systolic blood pressure, Total cholesterol	Angina, CHD death, Fatal or nonfatal ischaemic stroke, Heart failure, Nonfatal MI
Systematic Coronary Risk Evaluation (SCORE)	2003 2021	The initial algorithm involved a data set of different European populations from 1986. It underestimated CVD burden due to the inclusion of fatal outcomes and a restricted number of variables [23, 24]. Recently, a second algorithm was developed that adjusted for European region risk based on WHO incidence data [23]. The score does not account for other variables (socioeconomic status. family history, diabetes) nor has it been evaluated in non-European regions [23].	Europeans	Age, Gender, HDL cholesterol, Regions of Europe, Smoking status, Systolic blood pressure, Total cholesterol	Aortic aneurysm, Arrhythmia, CHD, CHD death, Heart failure, Peripheral vascular disease, Stroke
New Zealand Primary Prevention Equation (NZPPE)	2003 2018	New Zealand algorithm based on the FRS however it did not consider ethnic diversity [25]. To derive CVD risk prediction equations from local data, a study was done using primary care patients [17, 26]. This led to improved identification of high CVD risk noted in Māori, Pacific and Indian patients [17, 26]. In this manner, the NZ Primary Prevention Equation, which serves as a basis for current CVD risk assessments, was developed. The equation includes more variables than FRS, presents a revised definition of high CVD risk accounting for other risk factors (e.g. mental illness) and supports a clinician's decision on pharmacological treatment [25].	Multiple ethnicities	Age, Atrial fibrillation, Diabetes, Ethnicity, Family history, Gender, HDL cholesterol, Medication use (blood pressure, lipid-related, antithrombotic), Smoking status, Socioeconomic deprivation, Systolic blood pressure, Total cholesterol	Cardiomyopathy, Cardiovascular death (MI, stroke, or atherosclerotic aneurysm), CHD revascularisation, Heart failure, Nonfatal MI, Nonfatal stroke, Peripheral vascular disease, Transient ischaemic attack
QRISK Scores	2007, 2008, 2017	Scores derived using clinical data from England and Wales. The first version included traditional risk factors and social deprivation [27]. The second iteration (QRISK2) included more sociodemographic and medical variables (renal disease, rheumatoid arthritis) [28]. For its part, QRISK3 considers further chronic diseases and mental conditions [29].	Europeans (United Kingdom)	Age, Blood pressure treatment, BMI, Family history of CVD, Gender, HDL cholesterol, Smoking status, Steroid prescription, Systolic blood pressure, Total cholesterol, UK sociodemographic region	CHD death, Coronary insufficiency, Coronary revascularisation, Fatal or nonfatal stroke, Intermittent claudication, Nonfatal MI, Transient ischaemic attack

National Vascular Disease Prevention Alliance (NVDPA)	2012, 2020	This score has been compared and validated with the New Zealand PREDICT score. The differing healthcare practices within the United States and Europe complicate its use [30, 31]. Coronary artery calcium, score may be included to complement this risk algorithm [32].	Australia	Age, Alcohol intake, BMI, CAC, Diabetes, Family history of CVD, Gender, HDL cholesterol, Nutrition, Physical activity, Smoking status, social history, Systolic blood pressure, Total cholesterol, Waist circumference	Absolute risk of cardiovascular events including heart, strokes, and blood vessel diseases
Pooled Cohort Equations for atherosclerotic cardiovascular disease (PCE-ASCVD)	2013. 2018	Developed conjointly by the American College of Cardiology and the American Heart Association Task force. It has been validated in Caucasian and African American ethnicities. However, it under- or over-estimates CVD risk for Hispanics, Asians, and American Indian populations. [33].	White and Black Americans	Age, Blood pressure treatment, Diabetes mellitus, Gender, HDL cholesterol, Smoking, Systolic blood pressure, Total cholesterol	CHD death, Fatal stroke, Nonfatal MI, Nonfatal stroke
Multi-Ethnic Study of Atherosclerosis (MESA) risk score	2015	Ethnic comparison focusing on subclinical atherosclerosis progression and pre-emptive CVD diagnosis. This risk score has been compared with the FRS and PCE-ASCVD [34]. It is considered an improved risk score assessment in asymptomatic people compared to PCE. It may aid the decision-making process to start statin therapy in high-risk patients [35, 36].	African American, Chinese American, Hispanic American and Non-Hispanic Americans.	Age, Blood pressure treatment, Coronary artery calcium score, Diabetes mellitus, Gender, HDL cholesterol, Lipid lowering treatment, MI family history, Smoking status, Systolic blood pressure, Total cholesterol	CHD death, Coronary revascularisation, Nonfatal MI

BMI: Body Mass Index, CAC: Coronary artery calcium, CHD: Coronary Heart Disease, CVD: Cardiovascular disease, HDL: High density lipoprotein, MI: Myocardial Infarction, WHO: World Health Organisation

The PREDICT study and the NZPPE confirmed that Māori and Pasifika people have the most disease burden [18, 26]. Specifically, Māori have been reported to present increased rates of traditional risk factors including obesity, smoking and undiagnosed high blood pressure [18, 26]. NZ data also shows that Māori and Pasifika have an increased number of CVD-related hospitalisation events and low CVD treatment rates, which may be attributed to health inequity [14, 26]. Another factor that affects CVD risk is the degree of migrant acculturation into NZ given some ethnic groups (e.g. South Asians) appear to have increased obesity rates compared to Europeans [26]. This implies a need to create a CVD risk framework that increases equity along with avoiding under-estimation of risk due to ethnic or socioeconomic status [37]. Therefore, current CVD risk algorithms can benefit from including genetic factors that are associated with metabolites involved in CVD onset.

2.4 CVDs and single nucleotide polymorphisms

CVDs are defined by the WHO as a group of disorders that affect the heart and blood vessels in terms of structure or blood supply [40]. Notable examples of CVDs that are a leading cause of death globally include CHD, acute coronary syndrome (ACS) and congenital heart disease [41]. CHD involves inadequate coronary blood supply, which may arise from a blockage in the coronary arteries usually following progressive narrowing of the lumen of atherosclerotic blood vessels [42]. Given the multifactorial nature of CVDs, there are reviews available that explore in greater detail specific diseases such as coronary artery disease (CAD) [43], CHD [44], the underlying mechanism of atherosclerosis [2, 8] and the relationship of these diseases with specific variables [20, 45].

Genetic determinants are important non-modifiable risk factors for CVDs that have been studied intensively since the late 20th century [46-48]. The influence of genetic factors on CVD development was initially explored through family history studies focused on single gene disorders during the 1980s [43, 49]. Most CVDs are now considered to be polygenic disorders impacted by susceptibility and disease-linked genes, with major impacts from lifestyle and environmental factors [49]. Susceptibility genes are associated with an increase or reduction in the risk of developing a disease. Comparatively, disease-linked genes are those whose expression is linked to a pathological phenotype [43]. Both susceptibility and disease-linked genes can influence the regulation of other genes and/or factors that are directly involved in the pathobiology of different CVDs. The genetic basis for CVDs such as CAD and CHD has been reviewed in greater detail elsewhere [46, 50-52].

Considering this genetic complexity, numerous studies have focused on identifying associations between genetic variants and common cardiovascular disease traits [50, 53-56]. This has been supported with the establishment of genome wide association studies (GWAS), which employ technologies that detect many gene variants simultaneously [57]. The predominant variants identified through GWAS are SNPs [1, 50, 57, 58]. SNPs can be located within a protein-coding region, where they may display a functional effect, but they can also be in non-coding and regulatory areas of the genome (e.g. introns, enhancer, etc.). Moreover, SNPs can play a regulatory role by impacting gene expression and protein concentration if they are located within genetic elements such as transcription factor binding sites, splicing regions, enhancer, promoter, or silencer regions [1, 59]. These are often called expression quantitative trait loci (eQTLs) and explain a proportion of the genetic variance of a particular phenotype [60]. SNPs can also influence coding regions located within the same loci (cis-acting) or interact with coding regions of other chromosomes or distant loci on the same chromosome (trans-acting) [61, 62]. Specifically, SNP variants can influence CVD risk through traditional risk factors, such as plasma lipid levels and blood pressure [46, 61, 62]. Overall, SNP variants can have several potential effects on any given gene as summarised in Figure 2.3. One example covered in this review is *VEGFA,* which impacts the cardiovascular system through angiogenesis and increased EC activity.

Coupling our understanding of CVD pathogenesis with associations of regulatory SNPs with coronary biomarkers, there is potential for the combined use of CVD-relevant genetic risk scores (cvdGRS) in risk prevention [64]. This involves using multiple SNPs identified from GWAS studies in different populations and these SNP variants can be associated with clinical outcomes or risk factors [64, 65]. Overall, the goal of cvdGRS is to aid in patient risk stratification and treatment [57, 62, 65-67]. The functional effects of the SNP variants may provide evidence to underpin a clinical framework for prevention, treatment, and in severe cases, genetic counselling in primary care [57, 65, 68, 69].

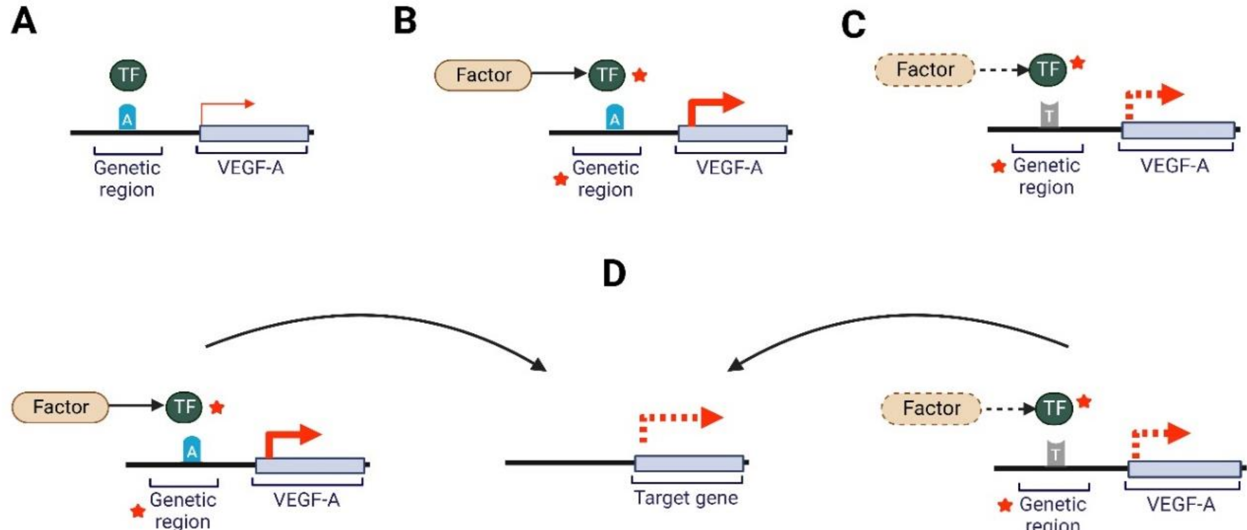


Figure 2.3 Potential effects of a SNP on VEGF-A gene expression. Gene regions where a SNP may be located include gene promoters, enhancers, introns, transcription factor binding motifs or non-coding RNAs. Dotted lines indicate hypothetical changes that could decrease or increase interactions or activity. Red arrows represent VEGF-A expression. Red stars indicate potential for an increase or decrease in activity through a variety of molecular mechanisms **(A)** An individual carrying the common allele "A" will show normally regulated levels of VEGF-A. **(B)** In the presence of a risk factor (e.g. elevated LDL-cholesterol plasma levels) and the common allele "A" there may be altered activity of genomic regions or transcription factors leading to altered VEGF-A expression and potentially impact circulating VEGF-A levels and/or activity. **(C)** Presence of a non-normal range risk factor (e.g. elevated plasma glucose levels) and the alternate allele "T" may impact activity of genomic regions or transcription factors leading to altered VEGF-A expression and potentially influence VEGF-A levels and/or activity **(D)** The changes illustrated in scenario B and C may impact expression of genes located at a distance. Trans-acting effects on a target gene is symbolised with black arrows. A potential pathway impacted could be lipid metabolism. Adapted from [63]. Created with BioRender.com.

2.5 VEGF overview

A molecule of interest in the development and progression of CVDs is VEGF-A, a member of the platelet-derived growth factor (PDGF)/VEGF family [70, 71]. This growth factor is involved in blood vessel formation, with reported impacts on the development of CVDs, as well as potential recovery [6, 72]. VEGF-A was originally referred to as a vascular permeability factor, with activity observed in tumour cells from rodents [73]. In 1989 several research groups identified that this factor selectively promoted the migration of vascular endothelium and induced angiogenesis *in vivo* [74-76]. Based on these findings, factors with this activity were renamed and classified as members of the VEGF family [76].

The VEGF family are glycoproteins expressed under the regulation of soluble mediators such as growth factors or cytokines [6, 77, 78]. They are involved in the regulation of blood vessel formation through endothelial cell differentiation or from existing blood vessels [77, 79]. Additionally, the VEGF family is involved in lymphangiogenesis, endothelial cell survival and vascular permeability

regulation, amongst other functions [77, 80]. However, alterations in their functionality have also been associated with the development of atherosclerosis, CHD, tumour formation, neovascularisation, and other pathologies including cancer, diabetic retinopathy, preeclampsia, and endometriosis [1, 6, 80, 81].

There are five VEGF family members that directly influence the human cardiovascular system. The archetype member is VEGF-A, a potent stimulator of vasculogenesis and angiogenesis [77, 81, 82]. VEGF-A production is influenced by oxygen tension, hormones (e.g., oestrogen) and proinflammatory cytokines [80, 82, 83]. VEGF-B induces the development of the cardiovascular system, embryonic angiogenesis and the formation of embryonic myocardium as well as participating in blood vessel survival [84]. VEGF-C and VEGF-D are primarily involved in lymphangiogenesis, while the placental growth factor (PIGF) participates in both angiogenesis and wound healing [6, 80, 82].

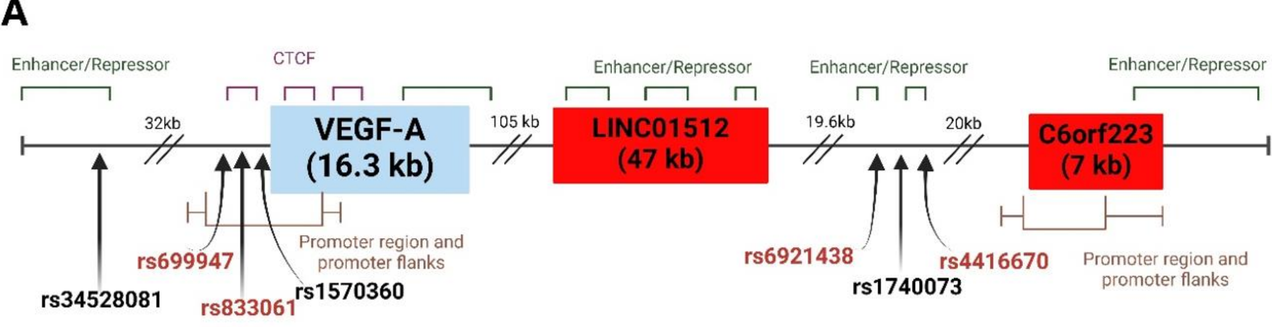
These VEGF proteins act through one or more of three tyrosine kinase VEGF receptors (VEGFRs) found on the surface of endothelial and non-endothelial cells [77]. VEGFR1 (Flt-1) and VEGFR2 (KDR) participate in angiogenesis. VEGFR2 is the primary inducer of VEGF-mediated blood vessel growth, while VEGFR3 is involved in lymphangiogenesis [80, 85, 86]. Additionally, VEGFR1 has the co-receptor neuropilin-1 (NRP1), which selectively potentiates VEGFR2-mediated vascular permeability, and endothelial cell motility in vascular development [82, 87]. Once activated, the signalling pathways of these receptors have the downstream effect of influencing vascular tone, blood vessel formation, endothelial cell proliferation and migration [80]. VEGFR signalling is reported to also be activated in a non-VEGF-dependent manner through receptor phosphorylation due to shear stress, or recognition of alternative ligands such as lactate and low-density lipoproteins (LDLs) [70, 86, 88].

Specifically, the VEGF-A canonical pathway occurs when it binds to either VEGFR1 or VEGFR2. This promotes receptor homodimerisation or heterodimerisation that leads to the phosphorylation of the receptor's intracellular domains [86, 88]. VEGFR1 has a soluble splice variant (sFlt-1) that acts as a decoy receptor, decreasing VEGF-A plasma concentration and limiting its binding to KDR [70, 77, 80]. Also, VEGF-A activity can be potentiated when PIGF displaces it from VEGFR1 to VEGFR2 [6]. These and other mechanisms surrounding the regulation of VEGF receptors have been reviewed in greater detail elsewhere [71, 88].

2.6 VEGF-A gene and related SNPs in a cardiovascular context

The *VEGFA* gene has a 16.3kb coding region located at 6p21.1 on the long arm of chromosome 6, including eight exons and seven introns [89, 90]. The first five exons are constitutively present among VEGF-A isoforms, since they encode the signal sequence for protein processing and residues that bind to VEGF receptors [87, 91]. Meanwhile, exons 6 and 7 contain the heparin binding domains that allow some isoforms to bind to cell surfaces and impact their activity or bioavailability depending on which are present [92, 93]. Lastly, exon 8 undergoes translational readthrough due to a non-canonical stop codon, leading to the production of sub-exons 8a and 8b, with the latter being reported to be present in a unique isoform with anti-angiogenic activity observed in bone disorders and brain diseases [87, 94-96]. So far, 16 distinct VEGF-A isoforms have been identified [80, 87]. The different isoforms depend on the presence or absence of exons 6 and 7, which affect the affinity for heparin or heparan sulphate proteoglycans. For example, the most prevalent VEGFA isoform is $VEGFA_{165}$, which lacks exon 6, but has moderate heparin affinity allowing the isoform to remain bound to cell surfaces [97]. Comparatively the isoform subtype $VEGFA_{121}$ lacks exon 6 and 7 so it is found only in free form [97]. Despite their size difference most of the VEGF-A isoforms act as endothelial cell mitogens, upregulate the endothelial expression of adhesion molecules and present pro-angiogenic activity [70, 84, 97]. Pathologies caused by increased angiogenesis include inflammatory diseases, cancers, retinopathy and atherosclerosis, while reduced angiogenesis has been observed in bone disorders and brain diseases [94]. The overall *VEGFA* gene structure including SNPs with reported influence on VEGF-A expression levels (discussed below and in Tables 2.2 and 2.3) is presented in Figure 2.4.

Altered plasma and tissue levels of VEGF-A have been observed in various conditions including ischaemic heart disease (IHD), CAD, strokes, heart failure, and myocardial infarction [72, 99-101]. Due to its impact on angiogenic processes, the effect of high VEGF-A circulating levels on CVD onset varies. High VEGF-A levels are associated with various CVD risk factors including smoking, hypercholesterolaemia, diabetes, hypertension, and hyperglycaemia [70]. Additionally, increased VEGF-A activity has been associated with inflammation, increased blood pressure and an increase in the formation of atherosclerotic lesions, leading to CHD [55, 102-104]. The impact of angiogenic molecules on atherosclerosis has been reviewed elsewhere [6].



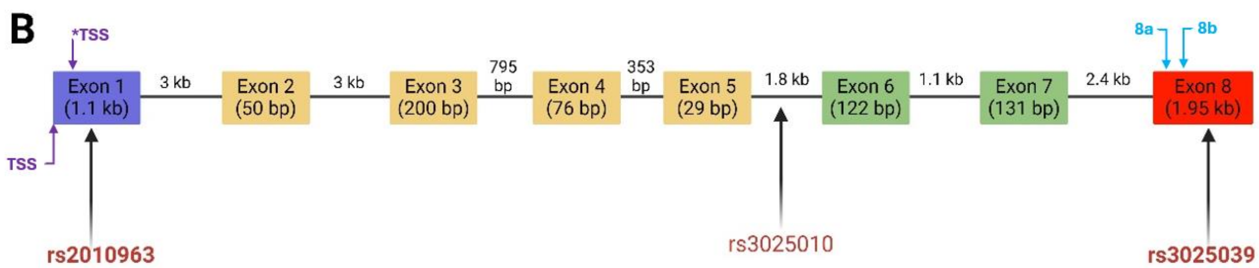


Figure 2.4. VEGFA gene and select SNPs on chromosome 6 **(A)** Location of VEGF-A related SNPs at chromosome 6 loci. SNPs in bold have been reported to influence VEGF-A levels. SNPs in red have reported associations to CVD risk factors or biomarkers. Red boxes indicate regions that represent lncRNA. Full gene length is indicated individually. Distances between SNPs and genes are indicated above dashes. **(B)** Schematic representation of SNPs located within exons and introns of *VEGFA*. Individual exon and intron lengths are shown [98]. Purple arrows and text show location of the main transcription starting site (TSS) and an alternative site (*TSS) located 633 nucleotides downstream. Light blue arrows and text show the location of splicing sites for the formation of exon 8a or exon 8b. The purple exon contains the signal sequence of the gene. The red exon confers anti-angiogenic capabilities when present. Green exons contain heparin-binding domains with extracellular matrix components. Yellow exons are involved in VEGF receptor binding. Location of SNPs and genomic regions determined using data from the UCSC genome browser and Ensembl databases (GRCh38-/hg38). Created with BioRender.com.

Expression of *VEGFA* can be upregulated by the hypoxia inducible factor, p53 allele polymorphisms, thyroid stimulating hormone, oestrogen levels and oxygen tension. [78, 80]. This matches studies that show VEGF-A production is influenced by elements associated with atherosclerosis including LDL concentration, hypoxia, and interleukin activity [72, 82, 105]. The increased production of VEGF-A can negatively impact human health by influencing the development of atherosclerotic plaques, by affecting vascular dilation, adhesion protein expression, monocyte migration, endothelium permeability and increased trans-endothelial lipid migration [6, 72]. High levels of VEGF-A in plasma have been associated with increased plaque growth and subsequent lesion vulnerability that can cause intraplaque haemorrhage [106]. There is evidence that proinflammatory cytokines (e.g., IL-1, IL-6, and IL-18) present during CVD onset can enhance VEGF-A production, thus exacerbating atherosclerotic lesion development [107-109].

Numerous variants related to the VEGF-A gene have been identified including 148 in the untranslated region (UTR), 209 in exons, 779 in introns, and 124 near the gene [110]. At least 30 SNPs within the

untranslated, exon, intron and promoter regions may have the potential to influence variation in VEGF-A expression [111, 112]. This genetic influence over VEGF-A circulating levels has been explored in various studies. Debette et al. (2011) investigated the heritability of VEGF-A levels in healthy individuals without a cancer diagnosis. This study identified four common variants (rs6921438, rs4416670, rs6993770 and rs10738760) distributed across three chromosomes that were independently associated with circulating VEGF-A levels and explained up to 48% of the heritability of serum VEGF-A levels [113]. A meta-analysis of GWAS data evaluated the association of variants with circulating VEGF-A levels [114]. Choi et al. (2016) found a total of ten SNPs contributed up to 52% of the variability in circulating VEGF-A levels with some SNPs associated with increased or decreased VEGF-A levels compared to median. Additional information on the study details of SNPs identified by these groups and other studies are presented in Tables 2.2 and 2.3. These tables also include SNPs that have been studied in relation to VEGF-A levels in healthy individuals, CVDs, or comorbidities related to the risk of CVD (e.g., diabetes, metabolic syndrome, hypertension).

Some of the SNPs that have been studied are located within exonic regions of *VEGFA* [115]. One noteworthy eQTL is rs2010963 from exon 1 of *VEGFA* (Figure 2.4B). The CC genotype has been associated with increased VEGF-A levels in type 2 diabetes mellitus (T2DM) [116, 117]. Furthermore, the rs2010963 CC genotype has been linked to risk factors including heart rate [116], blood glucose levels [110], blood pressure, cholesterol and HDL levels [116, 118]. There is also evidence for this variant influencing VEGF-A levels in non-CVDs such as glioma [119] and diabetic retinopathy [81, 120].The variant rs3025039, located within exon 8 of *VEGFA,* has similar effects (Figure 2.4B). Dong et al (2019) observed that patients diagnosed with gestational diabetes mellitus carrying the TT genotype had higher levels of VEGFA compared to healthy pregnant women [121]. Meanwhile Ruggiero et al (2011) [122] reported that the TT genotype was associated with lower median levels of VEGFA in healthy population samples from villages in Southern Italy. Some studies showed the CT genotype of rs3025039 is associated with reduced VEGFA levels as well as reducing risk of presenting with CHD and T2DM [110, 123]. The associations reported for the rs3025039 variant demonstrate its link to the cardiovascular system, but the variety of findings suggest additional studies are needed. An additional variant that has shown association to the cardiovascular system is rs3025010 located in intron 5 (Figure 2.4B). In a Chinese cohort diagnosed with hypertension the C allele of this variant was observed to be associated with lower systolic and diastolic blood pressure measurements [55]. Furthermore, in a Chinese case-control study, it was observed that the CC genotype of rs3025010 reduced risk of brain arteriovenous malformation [124]. This evidence shows a clear link to CVD risk

which could be further explored in additional ethnic groups to validate or identify other biomarker associations.

Other variants of interest can be found on Chromosome 6, but outside the intronic and exonic regions of *VEGFA* [59, 113]. rs69214328, is located within an enhancer region found between two long non-coding RNA genes (Figure 2.4A). The GWAS findings of Choi et al. (2016) and Debette et al. (2011) identified that the A allele of rs69214328 is associated with lower serum levels of the VEGF-A protein. Additionally, the same allele has also been reported to influence the variability of HDL and LDL in individuals of European ancestry [125]. The A allele of rs6921438 appears to have eQTL activities since increased serum levels of IL-6, TNF-α and VEGF-A were observed in interaction with SNPs rs6993770 (Chr8), rs4416670 (Chr6) and rs10738760 (Chr9), respectively [126]. Two additional variants (rs1740073 and rs34528081) located on chromosome 6 (Figure 2.4A) were identified by Choi et al. (2016) to be associated with serum levels of VEGF-A (Table 2.2). Furthermore, the T allele of rs34528081 was observed to be associated with increased VEGF-A serum levels in an additional GWAS study (Table 2.2). Meanwhile, the T allele of rs1740073 has been reported to associate with increased VEGF-A serum in a GWAS study while analysis of IHD using 1000 Genomes European data reported that the same allele could contribute to VEGF variance [99].

Another variant that has been studied is rs699947, which is located in the promoter region of *VEGFA* (Figure 2.4A). Various groups report that the AA genotype of rs699947 is associated with increased risk in cardiovascular pathologies including CAD, CHD, stroke, and congenital heart diseases (Table 2.2). The A allele of rs699947 has been associated with total cholesterol, LDL, and apolipoprotein B [110, 116, 127]. These associations have been observed across different ethnic groups, which further suggests rs699947 is a potential genetic risk marker for CVDs [123, 128, 129]. For its part, rs833061 is another variant that is located within the promoter region of *VEGFA* (Figure 2.4A) whose CT genotype has been observed to reduce VEGF-A levels in a T2DM cohort [110]. Other reports have also shown this variant is associated with hypertension and a meta-analysis of 3 cohorts implies this variant can influence congenital heart disease risk in individuals of Asian ancestry (Table 2.2). A variant located further from the promoter region that presents a similar array of findings related to lipid metabolism and inflammatory molecules is rs4416670 (Figure 2.4A). Both its alleles have been linked to CVD risk factors and biomarkers (Table 2.2). Specifically, the T allele was reported by Debette et al (2011) to increase VEGF-A serum levels while a study by Azimi-Nezhad et al (2013) reported the same allele could decrease IL-6 levels by interacting with rs6921438 (Chr6) and rs10738760 (Chr9). However, Azimi-Nezhad et al (2013) also report that the C allele of rs4416670 can increase TNF-α

and IL-6 levels by interacting with the A allele of rs6921438 thus implying a link between both VEGF-A related SNPs and inflammatory molecules. Additionally, the C allele has also been observed in other studies to be associated with apolipoprotein E levels, hypertension, and metabolic syndrome [125, 130]. These findings demonstrate links between VEGF-A related SNPs and lipid metabolism, inflammatory biomarkers, and CVD risk factors.

Some gene variants have findings of associations with molecules used in CVD risk assessment. For example, the rs1570360 variant located in the promoter region of VEGF-A (Figure 2.4A), was observed to contribute to an increased risk of congenital heart disease [131]. Some reports showed that the GA genotype of this variant is associated with a reduced left ventricular ejection fraction and extracranial internal carotid artery (ECICA) stenosis which are both risk factors for systemic hypertension and ischaemic stroke, respectively (Table 2.2). However, in a Chinese study the GG genotype was observed to increase susceptibility for coronary heart disease in patients with high smoking habits and diagnosed with hypertension. As such, this variant shows consistent links to CVD risk factors which, given its location, could be attributed to a potential influence on VEGF as observed in variants located within the promoter region (rs699947 and rs833061).

Similar studies have been reported for other SNPs located across the genome, often denoted as trans-acting SNPs (Table 2.3). Broadly, these eQTL SNPs have been associated with increased risk of CVDs (e.g., CAD, CHD, IHD) [99, 123, 134, 135] or metabolic syndrome [79]. One example rs1870377, located on chromosome 4 in exon 11 of the *VEGFR2* (*KDR*) gene (Figure 2.5) can influence cardiovascular outcomes. Li et al. (2021) reported that the AA genotype reduces risk of unfavourable CVD outcomes, particularly those related with disability, in an Asian ancestry cohort. Marks et al. (2018) also reported that the AA genotype associated with reduced risk of heart failure readmission and the A allele associated with high levels of VEGF system components, specifically sFlt-1 and KDR [136], and increased the risk of ischaemic stroke in a Korean cohort [135]. The TA and TT genotypes were both associated with increased CHD prevalence in Han Chinese populations [123, 134]. Location of additional SNPs influencing VEGF-A expression levels within the *VEGFR2* (*KDR*) gene is presented in Figure 2.5. Additional associations observed for trans-acting SNPs are presented in Table 2.3.

Table 2.2. VEGF-A eQTLS on Chromosome 6

SNP	Location*	Study type	Population (n)	Disease studied	Associated Biomarkers[$]	Associated risk factors[$]	Variant Effect or Association	Genotype or Allele	Source
rs34528081 (- →T)	Chr6: 43736681	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	VEGF-A	N/A	Increased serum VEGF-A levels (9 of 10 cohorts)	T allele	[114]
		GWAS community-based study	3 healthy cohorts from USA (n = 3,527), France (n = 859), and Sweden (n = 868)	Effect on VEGF levels	VEGF-A	N/A	Increased serum VEGF-A levels (2 study groups)	T allele	[113]
rs2010963 (G>C)	Chr6: 43770613 (+405 G/C)	Cross-sectional case-control study	Polish participants (n_{cases} = 265, $n_{controls}$ = 158)	Excess body mass	Estimated glomerular filtration rate	BMI	Negative correlation of VEGF-A serum levels with biomarkers	GG genotype	[116]
					Heart rate, DBP and HDL serum levels	N/A	Positive correlation of VEGF-A serum levels with biomarkers	CC genotype	
		Cohort study	Unrelated Han Chinese patients (n = 242)	CAD	N/A	N/A	Potential influence on CAD	CC genotype	[115]
		Case-control study	Iranian participants (n_{cases} = 347, $n_{controls}$ = 173)	CAD	N/A	N/A	Increased susceptibility to CAD	CC genotype	[132]
		Case-control study	Brazilian participants (n_{cases} = 169, $n_{controls}$ = 179)	Systemic hypertension	N/A	Left ventricular mass index	N/A	GC and CC genotypes	[133]

		Meta-analysis of case-control studies	10 cohorts: 6 Asian (n_{cases} = 1826, $n_{controls}$ = 1946) and 4 Caucasian (n_{cases} = 477, $n_{controls}$ = 916)	CAD and MI	N/A	N/A	Increased risk of CAD	CC genotype	[128]
		Cross-sectional study	Slovenian participants (n_{cases} = 595, $n_{controls}$ = 200)	T2DM	N/A	N/A	Higher VEGF-A serum levels in T2DM patients.	CC genotype	[117]
		Retrospective case-control study	Unrelated adults of Tunisian Arab origin (n_{cases} = 815, $n_{controls}$ = 805)	T2DM	Serum glucose, VEGF-A	Age	Reduced risk of T2DM Increased VEGF levels	GC genotype	[110]
		Cohort study	Unrelated Mexican mestizo participants (n = 415)	T2DM	Diastolic blood pressure, total cholesterol, and HDL	N/A	N/A	Both alleles	[118]
		Case-control study	Unrelated Chinese participants(n_{cases} = 533, $n_{controls}$ = 533)	CHD	N/A	Increased risk of CHD	N/A	GG genotype	[134]
		Case-control study	Unrelated Han Chinese participants (n_{cases} = 319, $n_{controls}$ = 333)	Haemorrhagic stroke	N/A	Increased risk of brain arterio-venous malformation	N/A	CC genotype	[124]

rs3025039 (C>T)	Chr 6:43784799 (+936 C/T)	Meta-analysis of case-control studies	8 cohorts: 4 Asian (n_{cases} = 1626, $n_{controls}$ = 1681) and 4 Caucasian (n_{cases} = 510, $n_{controls}$ = 796)	CAD and MI	N/A	Asian ancestry	Increased risk of CAD	TT genotype	[128]
		Cohort study	New Zealand patients including 3 ethnic groups: European, Māori and Pasifika (n = 1927)	ACS	Collateral vessel perfusion, BNP and NT-proBNP	N/A	Increased VEGF-A levels	TT genotype	[129]
		Retrospective case-control study	Unrelated adults of Tunisian Arab origin (n_{cases} = 815, $n_{controls}$ = 805)	T2DM	Diabetes duration, HbA1c, serum triglycerides, VEGF-A	N/A	Reduced T2DM risk and reduced VEGF levels	CT genotype	[110]
		Population-based cohort study	Italian (n = 1957)	Influence on VEGF levels	VEGF-A	N/A	Lower median serum VEGF levels	TT genotype	[122]
		Hospital-based case-control study	Korean (n_{cases} = 650, $n_{controls}$ = 308)	Ischaemic stroke	ApoB	Stroke ad Extracranial internal carotid artery stenosis	Increased risk of stroke	T allele	[127]
		Meta-analysis	9 Asian cohorts (n_{cases} = 1565, $n_{controls}$ = 2551)	Congenital heart diseases	N/A	N/A	Increased risk of tetralogy of Fallot	N/A	[131]

		Case-control study	Unrelated Han Chinese participants (n_{cases} = 810, $n_{controls}$ = 805)	Coronary Heart Disease	N/A	Smoking, alcohol intake, diabetes	Reduced risk of presenting CHD	CT genotype	[123]
		Case-control study	Chinese (n_{cases} = 239, $n_{controls}$ = 275)	Gestational diabetes mellitus	VEGF-A	Increased risk of GDM	Higher VEGF-A expression levels	CT and TT genotypes	[121]
rs1570360 (G>A)	Chr6:43770093 (-1154 G/A)	Case-control study	Brazilian participants (n_{cases} = 169, $n_{controls}$ = 179)	Systemic hypertension	N/A	Reduced ejection fraction	N/A	GA and AA genotypes	[133]
		Retrospective case-control study	Unrelated adults of Tunisian Arab origin (n_{cases} = 815, $n_{controls}$ = 805)	T2DM	Serum glucose	N/A	Increased risk of T2DM	A allele	[110]
		Hospital-based case-control study	Korean (n_{cases} = 650, $n_{controls}$ = 308)	Ischaemic stroke	Total homocysteine levels	Extracranial internal carotid artery (ECICA) stenosis	N/A	GA genotype	[127]
		Meta-analysis	5 cohorts: 3 Asian (n_{cases} = 455, $n_{controls}$ = 670) and 2 Caucasian (n_{cases} = 248, $n_{controls}$ = 368)	Congenital heart diseases	N/A	N/A	Increased risk of congenital heart disease	N/A	[131]
		Case-control study	Unrelated Han Chinese participants (n_{cases} = 810, $n_{controls}$ = 805)	Coronary Heart Disease	N/A	Smoking, Hypertension	Increased susceptibility to CHD	GG genotype	[123]

rs699947 (C>A)	Chr6:43768652 (−2578 C/A)	Cross-sectional case-control study	Polish participants (n_{cases} = 265, $n_{controls}$ = 158)	Excess body mass	HDL, VEGF-A	BMI	Positive correlation of VEGF-A serum levels with biomarkers	CC genotype	[116]
				Normal weight	Total cholesterol, LDL, VEGF-A	BMI	Negative correlation of VEGF-A serum levels with biomarkers	AA genotype	
		Cohort study	Unrelated Han Chinese patients (n = 242)	CAD	N/A	N/A	N/A	AA genotype	[115]
		Meta-analysis of case-control studies	8 cohorts: 5 Asian (n_{cases} = 2062, $n_{controls}$ = 2113) and 3 Caucasian (n_{cases} = 409, $n_{controls}$ = 698)	CAD and MI	N/A	Asian ancestry	Increased risk of CAD	AC genotype	[128]
		Cohort study	NZ patients including 3 ethnic groups: European, Māori and Pasifika (n = 2067)	Acute Coronary Syndrome	N/A	Lower physical activity	Predictor of mortality in male non-diabetic participants	AA genotype	[129]
		Retrospective case-control study	Unrelated adults of Tunisian Arab origin (n_{cases} = 815, $n_{controls}$ = 805)	T2DM	Serum glucose, LDL	N/A	Increased risk of T2DM	A allele	[110]
		Hospital-based case-control study	Korean (n_{cases} = 650, $n_{controls}$ = 308)	Ischaemic stroke	Total homocysteine levels, ApoB	Extracranial internal carotid artery stenosis	Increased risk of stroke	A allele	[127]

		Case-control study	Unrelated Han Chinese participants (n_{cases} = 810, $n_{controls}$ = 805)	CHD	N/A	Hypertension, Diabetes, alcohol intake	Increased risk of CHD	AA genotype	[123]
rs1740073 (T>C)	Chr6:43979661 (Intergenic between *LINC0512* and *C6orf223*)	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	VEGF-A	N/A	Increased serum VEGF-A levels (8 of 10 cohorts)	T allele	[114]
		Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	Ischaemic Heart Disease	VEGF-A	N/A	Potential contributor to VEGF phenotypic variance	T allele	[99]
rs6921438 (G>A)	Chr6:43957870 (Intergenic between *LINC0512* and *C6orf223*)	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	VEGF-A	N/A	Lower serum VEGF-A levels (10 cohorts)	A allele	[114]
		Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	IHD	HDL, LDL and VEGF-A	N/A	Increased serum VEGF levels	G allele	[99]
		Population-based phenome wide association study	Finnish participants (n = 6,890)	Inflammatory biomarker driver trait search	VEGF-A, IL-10. IL-12p70	N/A	Estimated effect on VEGF protein production	N/A	[103]

		GWAS community -based study	3 healthy cohorts from USA (n = 3,527), France (n = 859), and Sweden (n = 868)	Effect on VEGF levels	VEGF-A	N/A	Lower serum VEGF-A levels (2 study groups)	A allele	[113]
		Population study	2 groups of unrelated healthy European ancestry (n_1 = 1,006 n_2 = 1,145)	Influence on lipid metabolism	HDL, LDL	N/A	Contributes to 1% of HDL variability and 0.2% of LDL variability	A allele	[125]
		Case-control study	Iranian participants (n_{cases} = 248, $n_{controls}$ = 100)	Association between dietary intake and metabolic Syndrome	Zinc and manganese intake	Metabolic Syndrome	Association of SNP with low iron intake and high manganese intake.	AA genotype	[130]
		Population study	Healthy French individuals (n = 403)	Association between VEGF, adhesion, and inflammation molecules.	IL-6, TNF-α, VEGF-A	N/A	Epistatic interaction with two other SNPs is associated with increased IL-6 levels. Another interaction with two individual SNPs is associated with increased VEGF-A and TNF-α levels.	A allele	[126]
					IL-6	N/A	Epistatic interaction with two other SNPs is associated with decreased IL-6 levels.	G allele	[126]

rs4416670 (T>C)	Chr6:43982716 (Intergenic between *LINC0512* and *C6orf223*)	GWAS community-based study	3 healthy cohorts from USA (n = 3,527), France (n = 859), and Sweden (n = 868)	Effect on VEGF levels	VEGF-A	N/A	Increased serum VEGF-A levels (2 study groups)	T allele	[113]
		Population study	2 groups of unrelated healthy European ancestry (n_1 = 1,006 n_2 = 1,145)	Influence on lipid metabolism	Apolipoprotein E	Hypertension	Potential functional effect between variant, hypertension and ApoE levels.	C allele	[125]
		Case-control study	Iranian participants (n_{cases} = 248, $n_{controls}$ = 100)	Association between dietary intake and metabolic Syndrome	Iron, copper, zinc, manganese, and iodine intake	Metabolic Syndrome	Increased risk of Metabolic syndrome	CT and CC genotypes	[130]
		Population study	Healthy French individuals (n = 403)	Association between VEGF, adhesion, and inflammation molecules.	IL-6, TNFα	N/A	Epistatic interaction with one SNP is associated with increased TNF-α levels. Another interaction with two other SNPs is associated with increased IL-6 levels.	C allele	[126]
					IL-6		Epistatic interaction with two SNPs is associated with decreased IL-6 levels	T allele	[126]

rs3025010 (C>T)	Chr6:43779840 (Intron 5)	Cross-sectional population study	Chinese (n_{cases} = 258, $n_{controls}$ = 258)	Hypertension	N/A	Lower systolic and diastolic blood pressure	N/A	C allele	[55]
		Case-control study	Unrelated Han Chinese participants (n_{cases} = 319, $n_{controls}$ = 333)	Haemorrhagic stroke	N/A	Reduced risk of brain arterio-venous malformation	N/A	CC genotype	[124]
rs833061 (C>T)	Chr6:43769749-(460 T/C)	Retrospective case-control study	Unrelated adults of Tunisian Arab origin (n_{cases} = 815, $n_{controls}$ = 805)	T2DM	VEGF-A	N/A	Reduced VEGF-A serum levels	CT genotype	[110]
		Cross-sectional population study	Chinese (n_{cases} = 258, $n_{controls}$ = 258)	Hypertension	N/A	Hypertension	N/A	N/A	[55]
		Meta-analysis	3 cohorts: 2 Asian (n_{cases} = 233, $n_{controls}$ = 318) and 1 Caucasian (n_{cases} = 102, $n_{controls}$ = 112)	Congenital heart diseases	N/A	Asian ancestry	Influence on congenital heart disease risk	N/A	[131]

*Location reported with point location (Ensembl human database GRCh38.p13) specifying the variant's genomic context and where applicable indicating nomenclature in reference to the VEGF-A gene.

$ Biomarkers or risk factors associated to the allele or genotype reported by each study.

N/A: Not reported by the article.

ACS: Acute Coronary Syndrome, Apob: Apolipoprotein b, ApoE: Apolipoprotein E, BMI: Body Mass Index, BNP: B-type natriuretic peptide, CAD: Coronary Artery Disease, CHD: Coronary Heart Disease, eQTLs: expression quantitative trait loci, GWAS: Genome Wide Association Study, HbA1c: Haemoglobin A1C, HDL: High Density Lipoprotein, IL: Interleukin, IHD; Ischaemic Heart Disease, LDL: Low Density Lipoprotein, N/A: not applicable, NT-proBNP: N-terminal pro B-type natriuretic peptide, SNP: Single Nucleotide Polymorphism, T2DM: Type 2 Diabetes Mellitus, TNF: Tumour Necrosis Factor, USA: United States of America, VEGF: Vascular Endothelial Growth Factor.

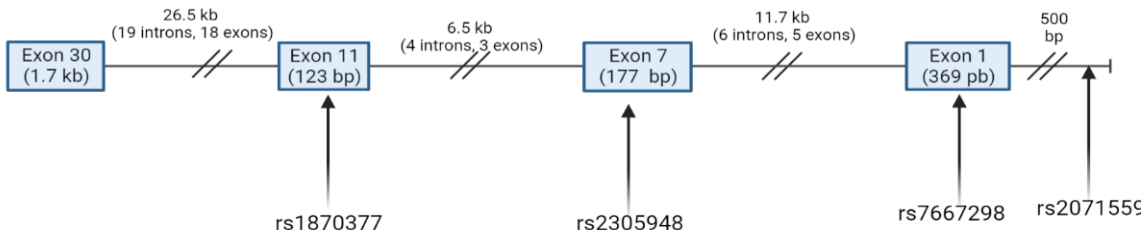


Figure 2.5. Schematic representation of key SNPs in key exons of VEGFR2 (Chromosome 4, reverse strand). Full exon length is indicated individually. Distances between exon and SNPs are indicated above dashes. Distance measured using NCBI GenBank Accession NC_000004

rs6993770 is located on chromosome 8 in intron 4 of the *ZPFM2* gene, which codes for a protein involved in heart morphogenesis and coronary vessel development. Broadly, studies on this variant have shown relationships with VEGF-A, CVD and CVD risk factors (Table 2.3). In the GWAS findings of Choi et al. [114] and the Mendelian Randomisation study done by Au Yeung [99], the A allele correlated with increased VEGF-A serum levels. GWAS findings of Debette et al. [113] showed the T allele was associated with increased VEGF-A serum levels. Other studies involving individuals of European and Iranian ancestry observed the T allele was also associated with risk biomarkers of CVD, particularly fasting blood glucose, triglyceride levels, systolic blood pressure and HDL levels [125, 137]. The TA genotype has been reported to increase the risk of metabolic syndrome [137], and impacts the expression of adhesion molecules (ICAM-1, E-selectin) as well as IL-6 levels [126]. Meanwhile, the TT genotype appears to contribute to metabolic syndrome risk in individuals with low iron intake [130]. This spectrum of reports demonstrates the range of associations that alleles and genotypes of trans-acting SNPs, such as rs1870377 or rs6993770, may have within the cardiovascular system.

Additional trans-acting SNPs (e.g. rs2071559, rs114694170, rs6993770, rs10738760, rs10761741, rs4782371) have been reported to be capable of influencing VEGF-A circulating levels [114] or soluble VEGFR levels (rs1870377) [136]. Specific study details and overall findings are presented in Table 2.3. Notably, two SNPs (rs2305948 and rs7667298) have associations with potential CVD risk, but their direct impact on VEGF system components was observed in cancer related studies [138, 139]. Interestingly, trans-acting SNPs most likely involve interactions with molecules or homeostatic mechanisms that have known roles in CVD onset, including inflammatory interleukins [103, 126], triglycerides, adhesion molecules, blood cell count and blood pressure [99, 137, 140, 141]. There are cases of specific variants that correlate with increased risk of presenting major adverse coronary events (rs2305948, rs7667298) [141], CHD (rs2305948, rs1870377, rs2071559, rs7667298) [123, 134], ischaemic stroke (rs1870377) [135] and MetS (rs6993770) [137]. As such, some SNPs appear to be potential contributors to phenotypes (IHD, CAD, CHD) while others may increase or reduce disease risk depending on the presence or absence of risk factors [105, 123, 136].

Table 2.3. Reports on trans-acting SNPs associated with VEGF-A or CVD risk.

SNP (change)*	Location (proximal gene/s) *	Gene function[$]	Study type	Population	Disease Studied	Notable findings	Source
rs2305948 (C>T)	Chr4: 55113391 *(VEGFR2)*	VEGF receptor	Case-control study	Unrelated Chinese participants (n_{cases} = 533, $n_{controls}$ = 533)	CHD	Increased CHD risk (TT genotype)	[134]
			Case-control study	Unrelated Han Chinese participants (n_{cases} = 810, $n_{controls}$ = 805)	CHD	Reduced risk to CHD when smoking, alcohol intake and diabetes were considered (TT genotype)	[123]
			Meta analysis of case-control studies	10 cohorts: 8 Asian and 2 Caucasian (n_{cases} = 5,474 $n_{controls}$ = 8,584)	CAD, MI, IS	This SNP is associated with coronary artery disease	[142]
			Cohort study	Han Chinese patients undergoing coronary intervention (n = 275)	CHD	Higher presence of main adverse cardiovascular events including angina pectoris and recurrent myocardial infarction (TT genotype)	[141]
rs1870377 (A>T)	Chr4: 55106807 (*VEGFR2*)	VEGF receptor	Case-control study	Unrelated Chinese participants (n_{cases} = 533, $n_{controls}$ = 533)	CHD	Increased CHD risk (TT genotype)	[134]
			Cohort Study	New Zealand patients including 3 ethnic groups: European, Māori and Pasifika (n = 2067)	ACS	Increased levels of sFlt-1 and sKDR (A allele). There is an association with reduced risk of heart failure remission (AA genotype)	[136]
			Meta analysis of case-control studies	10 cohorts: 8 Asian and 2 Caucasian (n_{cases} = 5,474 $n_{controls}$ = 8,584)	CAD, MI, IS	This SNP is associated with coronary artery disease and ischaemic stroke	[142]

			Case-control study	Unrelated Han Chinese participants (n_{cases} = 810, $n_{controls}$ = 805)	CHD	Increased CHD risk in non-hypertensive, nondiabetic, and non-smoking cases (TA genotype)	[123]
			Cohort study	Han Chinese patients (n = 1,016)	LAAS	Decreased risk of unfavourable outcomes (AA genotype)	[105]
			Case-control study	Korean participants (n_{cases} = 501, $n_{controls}$ = 478)	IS	Confers risk of presenting ischaemic stroke (A allele)	[135]
rs7667298 (A>G)	Chr4: 55125564 (*VEGFR2*)	VEGF receptor	Case-control study	Unrelated Han Chinese participants (n_{cases} = 810, $n_{controls}$ = 805)	CHD	Increased CHD risk in non-hypertensive, nondiabetic, and non-smoking patients (G allele)	[123]
			Cohort study	Han Chinese patients undergoing coronary intervention (n = 275)	CHD	Lower occurrence of target lesion revascularisation (GG genotype)	[141]
rs2071559 (T>C)	Chr4: 55126199 (*VEGFR2*)	VEGF receptor	Cross-sectional study	Slovenian participants (n_{cases} = 595, $n_{controls}$ = 200)	T2DM	Higher VEGF-A serum levels in T2DM patients (CC genotype). The SNP also shows association with carotid intima-media thickness in T2DM patients	[117]
			Case-control study	Unrelated Chinese participants (n_{cases} = 533, $n_{controls}$ = 533)	CHD	Increased CHD risk (CC genotype)	[134]
			Meta analysis of case-control studies	10 cohorts: 8 Asian and 2 Caucasian (n_{cases} = 5,474 $n_{controls}$ = 8,584)	CAD, MI, IS	This SNP is associated with coronary artery disease.	[142]
rs114694170 (T>C)	Chr5: 88884379 (*MEF2C*)	Myogenesis transcription enhancer in cardiac morphogenesis and vascular development.	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	Lower serum VEGF-A levels in 9 out of 10 cohorts (T allele)	[114]
			Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	IHD	Potential contributor to VEGF phenotypic variance and IHD risk (C allele)	[99]

rs6993770 (A>T)	Chr8: 105569300 (*ZFPM2*)	Transcription regulator involved in heart morphogenesis and development of coronary vessels	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	Increased serum VEGF-A levels in 9 out of 10 cohorts (A allele)	[114]
			Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	IHD	Contributor to VEGF phenotypic variance by Increasing VEGF-A serum levels (A allele). The SNP is also associated with platelet count.	[99]
			Cross sectional population-based study	Lebanese unrelated participants (n = 460)	Influence on circulating lipid levels.	This SNP is a contributor to VEGF-A circulating level heritability. It is also associated with total cholesterol, LDL, and hypercholesterolaemia	[140]
			Case-control study	Iranian participants (n_{cases} = 235, $n_{controls}$ = 101)	MetS	Association with fasting blood glucose, triglyceride levels and systolic blood pressure (T allele) Increased MetS risk (AT genotype)	[137]
			GWAS community-based study	3 healthy cohorts from USA (n = 3,527), France (n = 859), and Sweden (n = 868)	Effect on VEGF levels	Increased serum levels in 3 study groups (T allele)	[113]
			Population study	2 groups of unrelated healthy European ancestry (n_1 = 1,006 n_2 = 1,145)	Influence on lipid metabolism	Increased HDL levels (T allele)	[125]
			Case-control study	Iranian participants (n_{cases} = 248, $n_{controls}$ = 100)	Association between dietary intake and MetS	Association of metabolic syndrome with low dietary iron intake (TT genotype)	[130]
			Population study	Healthy French individuals (n = 403)	Association between VEGF, adhesion, and interleukins.	Epistatic interaction (TA genotype and T allele) with other SNPs associate with Increased ICAM-1, E-selectin, and IL-6 levels.	[126]

rs2375981 (C>G)	Chr9: 2692583 (*KCNV2*)	Codes for a potassium channel, regulates smooth muscle contraction and heart rate	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	Increased serum VEGF-A levels in 10 cohorts (C allele)	[114]
			Population based phenome wide association study	Finnish participants (n = 6,890)	Inflammatory biomarker driver trait search	This SNP has a potential effect on VEGF-A protein production. It has been associated with IFN-γ, IL-10, IL-12p70 and VEGF-A molecules.	[103]
			Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	IHD	Potential contributor to VEGF phenotypic variance and IHD (C allele)	[99]
			GWAS community-based study	3 healthy cohorts from USA (n = 3,527), France (n = 859), and Sweden (n = 868)	Effect on VEGF levels	Lower serum VEGF-A levels in 3 study groups (G allele)	[113]
rs7043199 (A>T)	Chr9: 2621145 (*VLDLR*)	Low density lipoprotein receptor involved in endocytosis	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	Lower serum VEGF-A levels in 10 cohorts (A allele)	[114]
			Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	IHD	Potential contributor to VEGF phenotypic variance and IHD (T allele)	[99]

rs10738760 (A>G)	Chr9: 2691186 (*VLDLR* and *KCNV2*)	*VLDLR*: Low density lipoprotein receptor involved in endocytosis. *KCNV2*: Potassium channel subunit that is predominantly expressed in the heart and retina	Cross sectional population-based study	Lebanese unrelated participants (n = 460)	Influence on circulating lipid levels.	This SNP is a contributor to VEGF-A circulating level heritability. It also has potential implications in metabolic syndrome	[140]
			GWAS community-based study	3 healthy cohorts from USA (n = 3,527), France (n = 859), and Sweden (n = 868)	Effect on VEGF levels	Increased VEGF-A serum levels in 3 study groups (A allele)	[113]
			Healthy cohort study	Lebanese general population (m = 460)	Effect of VEGF-A SNPs on iron levels	Association with low iron levels and obesity (GG genotype)	[143]
			Population study	Healthy French individuals (n = 403)	Association between VEGF, adhesion, and inflammation molecules.	An epistatic interaction of the G allele with two SNPs is associated with Increased ICAM-1 and E-selectin levels. An epistatic interaction of the A allele with two other SNPs is associated with Increased VEGF-A and decreased IL-6 levels	[126]
rs10761741 (T>G)	Chr10: 63306426 (*JMJD1C*)	Histone demethylase involved in transcription regulation, RNA processing and DNA repair	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	Increased serum VEGF-A levels in 10 cohorts (T allele)	[114]
			Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	IHD	Potential contributor to VEGF phenotypic variance and IHD (T allele)	[99]
			Healthy cohort study	Middle-aged men of European ancestry (n = 1,300)	Association with haemostatic factors	Increased mean platelet volume, decreased platelet count, and decreased platelet reactivity (G allele)	[144]

rs4782371 (G>T)	Chr16: 88502423 (*ZFPM1*)	Transcription regulator involved in cardiac development and platelet production	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	Lower serum VEGF-A levels in 10 cohorts (T allele)	[114]
			Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	IHD	This SNP is a potential contributor to VEGF phenotypic variance and IHD (G allele)	[99]
rs2639990 (C>T)	Chr18: 75203596 *(ZADH2)*	Protein that shows oxidoreductase and transferase activity predicted to be involved in negative regulation of fat cell differentiation.	GWAS Meta-analysis	10 cohorts of European ancestry (n = 16,112)	Effect on VEGF levels	Increased serum VEGF-A levels in 10 cohorts (T allele)	[114]
			GWAS community-based study	3 healthy cohorts from USA (n = 3,527), France (n = 859), and Sweden (n = 868)	Effect on VEGF levels	Increased VEGF-A serum levels in 3 study groups (T allele)	[113]
			Mendelian randomisation study	1000 Genomes data on adults of European ancestry (n_{cases} = 60,801 $n_{controls}$ = 123,504)	IHD	This SNP is a potential contributor to VEGF phenotypic variance and IHD (T allele)	[99]

*Location obtained from the Ensembl human database (GRCh38.p13).

$ Function of genes obtained from the Genecards database.

ACS: Acute Coronary Syndrome, CAD: Coronary Artery Disease. CHD : Coronary Heart Disease, GWAS: Genome Wide Association Study, HDL: High Density Lipoprotein, ICAM: Intercellular Adhesion Molecule, IHD: Ischaemic Heart Disease, IL: Interleukin, IS: Ischaemic Stroke, LAAS: Large Artery Atherosclerotic Stroke LDL: Low Density Lipoprotein, MetS: Metabolic Syndrome, MI: Myocardial Infarction, sFlt: soluble fms-like tyrosine kinase-1 (sVEGFR1), sKDR: soluble Kinase insert Domain Receptor (sVEGFR2), SNP: Single Nucleotide Polymorphism, T2DM: Type 2 Diabetes Mellitus, USA: United States of America, VEGF: Vascular Endothelial Growth Factor

2.7 Conclusion

Overall, the impact of VEGF-A related SNPs in various forms of heart disease has been explored in many different types of studies. The collective evidence reveals a critical subset of cis-acting SNPs mapping to the region of *VEGFA* (Figure 2.4 and Table 2.2), several trans-acting SNPs mapping in the region of the *VEGFR2* gene (Figure 2.5) and elsewhere on the human genome (Table 2.3), with repeatable associations with circulating levels of VEGF-A. A small group of SNPs reproducibly associate with established biomarkers and risk factors for heart disease (rs2010963, rs3025039, rs1570360, rs699947, rs6921438) or with increased susceptibility to common heart disease pathologies (rs2010963, rs3025039, rs1570360, rs699947, rs2305948, rs1870377). This review highlights that these SNPs can be potential markers for CVDs and may influence significant biological pathways that impact the cardiovascular system (e.g., lipid metabolism). The wide range of pathologies that VEGF-A and its related SNPs impact emphasizes the complexity of VEGF-A interactions within the cardiovascular system. Both cis- and trans-acting SNP eQTLs can affect expression levels, but there remain many unknowns around the specific mechanisms involved. There is a clear link between SNPs and VEGF-A levels, as well as established cardiovascular disease biomarkers (HDL, LDL, BNP, NTproBNP). Together these have the potential to act synergistically on the development of CVDs.

The complexity of SNP influences on CVD and CVD risk factors reinforces the importance of studying them in relation to VEGF-A. Particularly considering how altered levels of VEGF-A contribute to disease onset or exacerbate an individual's health depending on the risk factors they present with. Exploring the link between CVDs, SNPs, and VEGF-A may contribute to improved cardiovascular disease risk assessment, prevention, treatment, and prognosis.

2.8 References

1. Geiger, M., *Fundamentals of vascular biology*. Learning materials in biosciences. 2019, Vienna, Austria: Springer.
2. Libby, P., *The changing landscape of atherosclerosis.* Nature, 2021. **592**(7855): p. 524-533.
3. Ramji, D.P. and T.S. Davies, *Cytokines in atherosclerosis: Key players in all stages of disease and promising therapeutic targets.* Cytokine & Growth Factor Reviews, 2015. **26**(6): p. 673-685.
4. Libby, P., K.E. Bornfeldt, and A.R. Tall, *Atherosclerosis: Successes, Surprises, and Future Challenges.* Circulation Research, 2016. **118**(4): p. 531-534.
5. Aziz, M. and K.S. Yadav, *Pathogenesis of atherosclerosis.* Medical & Clinical Reviews, 2016. **2**(3): p. 1-6.
6. Camaré, C., et al., *Angiogenesis in the atherosclerotic plaque.* Redox biology, 2017. **12**: p. 18-34.
7. Minelli, S., P. Minelli, and M.R. Montinari, *Reflections on Atherosclerosis: Lesson from the Past and Future Research Directions.* Journal of Multidisciplinary Healthcare, 2020. **13**: p. 621-633.
8. Libby, P., et al., *Atherosclerosis.* Nature Reviews Disease Primers, 2019. **5**(1): p. 56.
9. Shi, X., et al., *Calcification in Atherosclerotic Plaque Vulnerability: Friend or Foe?* Frontiers in Physiology, 2020. **11**(56).
10. Andrews, J., et al., *Coronary arterial calcification: A review of mechanisms, promoters, and imaging.* Trends in Cardiovascular Medicine, 2018. **28**: p. 491-501.
11. Douglas, G. and K.M. Channon, *The pathogenesis of atherosclerosis.* Medicine, 2014. **42**(9): p. 480-484.
12. Ministry of Health. *Mortality Web Tool (2014-2018)*. 2018 30 June 2021; Available from: https://www.health.govt.nz/publication/mortality-web-tool.
13. Tobias, M. and Turley, M., *Health Loss in New Zealand: A report from the New Zealand Burden of Diseases, Injuries and Risk Factors Study, 2006-2016.* 2016, Ministry of Health: Wellington. p. 94.
14. Yeh, L.-C., et al., *Tatau kahukura : Māori health chart book 2015*. 3rd ed. 2015: Ministry of Health.
15. Wilson, N., et al., *Prioritisation of intervention domains to prevent cardiovascular disease: a country-level case study using global burden of disease and local data.* Population Health Metrics, 2023. **21**(1): p. 1.
16. Institute for Health Metrics and Evaluation. *Global Disease Burden*. 2019 [cited 2021; Available from: https://vizhub.healthdata.org/gbd-compare/.
17. Pylypchuk, R., et al., *Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study.* Lancet, 2018. **391**(10133): p. 1897-1907.
18. Barnes, L.A., et al., *The Prevalence of Cardiovascular Risk Factors in Different Occupational Groups in New Zealand.* Annals of Work Exposures & Health, 2020. **64**(6): p. 645-658.
19. World Health Organisation, *Global Atlas on cardiovascular disease prevention and control.* 2011.
20. Akyuz, A., *Exercise and Coronary Heart Disease.* Advances in Experimental Medicine and Biology, 2020. **1228**: p. 169-179.
21. Damen, J.A.A.G., et al., *Prediction models for cardiovascular disease risk in the general population: systematic review.* British Medical Journal, 2016. **353**: p. i2416.
22. Fitzgerald, X., et al., *Cardiovascular disease risk assessment in an Aboriginal community-controlled health service: comparing algorithms.* Australian Journal of Primary Health, 2020. **26**(4): p. 281-286.
23. Systematic Coronary Risk Evaluation working group and European Society of Cardiology, *SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe.* European Heart Journal, 2021. **42**(25): p. 2439-2454.
24. Tomasik, T., et al., *The systematic coronary risk evaluation (SCORE) for the prevention of cardiovascular diseases. Does evidence exist for its effectiveness? A systematic review.* Acta Cardiologica, 2017. **72**(4): p. 370-379.
25. BPAC, *What's new in cardiovascular disease risk assessment and management for primary care clinicians.* Best Practice Advocacy Centre New Zealand, 2018.

26. Selak, V., et al., *Ethnic differences in cardiovascular risk profiles among 475,241 adults in primary care in Aotearoa, New Zealand.* New Zealand Medical Journal, 2020. **133**(1521): p. 14-27.
27. Hippisley-Cox, J., et al., *Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study.* British Medical Journal, 2007. **335**(7611): p. 136.
28. Hippisley-Cox, J., et al., *Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2.* British Medical Journal, 2008. **336**(7659): p. 1475-82.
29. Hippisley-Cox, J., C. Coupland, and P. Brindle, *Development, and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study.* British Medical Journal, 2017. **357**: p. j2099.
30. National Vascular Disease Prevention Alliance, *Guidelines for the management of absolute cardiovascular disease risk*. 2012.
31. Australian Chronic Disease Prevention Alliance. *Australian Absolute Cardiovascular Disease Risk Calculator*. 2021; Available from: http://www.cvdcheck.org.au/australian-absolute-cardiovascular-disease-risk-calculator.
32. Jennings, G.L., et al., *National Heart Foundation of Australia: position statement on coronary artery calcium scoring for the primary prevention of cardiovascular disease in Australia.* Medical Journal of Australia, 2021. **214**(9): p. 434-439.
33. Grundy, S.M., et al., *2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines.* Circulation, 2019. **139**(25): p. e1082-e1143.
34. McClelland, R.L., et al., *10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors: Derivation in the MESA (Multi-Ethnic Study of Atherosclerosis) With Validation in the HNR (Heinz Nixdorf Recall) Study and the DHS (Dallas Heart Study).* Journal of the American College of Cardiology, 2015. **66**(15): p. 1643-1653.
35. Shemesh, J.S., et al., *Multi-Ethnic Study of Atherosclerosis (MESA) score compared to the Pooled Cohort Equations (PCE) score for cardiovascular risk assessment.* European Heart Journal, 2020. **41**(Supplement_2).
36. Blaha, M.J., et al., *Providing Evidence for Subclinical CVD in Risk Assessment.* Global Heart, 2016. **11**(3): p. 275-285.
37. Andrew J., K., et al., *A unified national cardiovascular disease (CVD) risk generator is required to address equity in the management of CVD risk in clinical practice in New Zealand.* New Zealand Medical Association, 2019. **132**.
38. Mehta, S., et al., *Development and validation of alternative cardiovascular risk prediction equations for population health planning: A routine health data linkage study of 1.7 million New Zealanders.* International Journal of Epidemiology, 2018. **47**(5): p. 1571-1584.
39. Ministry of Health, *Cardiovascular Disease Risk Assessment and Management for Primary Care*. 2018: Wellington, New Zealand.
40. World Health Organisation. *Cardiovascular diseases (CVDs)*. 2021 11 June 2021; Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
41. Joseph, P., et al., *Reducing the Global Burden of Cardiovascular Disease, Part 1.* Circulation Research, 2017. **121**(6): p. 677-694.
42. Sanchis-Gomar, F., et al., *Epidemiology of coronary heart disease and acute coronary syndrome.* Annals of Translational Medicine, 2016. **4**(13): p. 256-256.
43. Malakar, A.K., et al., *A review on coronary artery disease, its risk factors, and therapeutics.* Journal of Cellular Physiology, 2019. **234**(10): p. 16812-16823.
44. Khamis, R.Y., T. Ammari, and G.W. Mikhail, *Gender differences in coronary heart disease.* Heart, 2016. **102**(14): p. 1142-9.
45. Hartman, J. and W.H. Frishman, *Inflammation and Atherosclerosis.* Cardiology in Review, 2014. **22**(3): p. 147-151.

46. Assimes, T.L. and R. Roberts, *Genetics: Implications for Prevention and Management of Coronary Artery Disease.* Journal of the American College of Cardiology, 2016. **68**(25): p. 2797-2818.
47. Roberts, R., *Genetics-Current and Future Role in the Prevention and Management of Coronary Artery Disease.* Current Atherosclerosis Reports, 2016. **18**(12): p. 78.
48. Roberts, R., *A Breakthrough in Genetics, and its Relevance to Prevention of Coronary Artery Disease in LMIC.* Global Heart, 2017. **12**(3): p. 247-257.
49. Roberts, R. and J. Fair, *Genetics, its role in preventing the pandemic of coronary artery disease.* Clinical Cardiology, 2021. **44**(6): p. 771-779.
50. McPherson, R. and A. Tybjaerg-Hansen, *Genetics of Coronary Artery Disease.* Circulation Research, 2016. **118**(4): p. 564-78.
51. Khera, A.V., et al., *Genetic risk, adherence to a healthy lifestyle, and coronary disease.* New England Journal of Medicine, 2016. **375**(24): p. 2349-2358.
52. Roberts, R., *A genetic basis for coronary artery disease.* Trends in Cardiovascular Medicine, 2015. **25**(3): p. 171-8.
53. LeBlanc, M., et al., *Identifying Novel Gene Variants in Coronary Artery Disease and Shared Genes With Several Cardiovascular Risk Factors.* Circulation Research, 2016. **118**(1): p. 83-94.
54. Sun, B.B., et al., *Genomic atlas of the human plasma proteome.* Nature, 2018. **558**(7708): p. 73-79.
55. Zhao, Z., et al., *Association between Single Nucleotide Polymorphisms in Cardiovascular Developmental Critical Genes and Hypertension: A Propensity Score Matching Analysis.* International Journal of Hypertension, 2020. **2020**: p. 1-8.
56. Earle, N., et al., *Single nucleotide polymorphisms in arrhythmia genes modify the risk of cardiac events and sudden death in long QT syndrome.* Heart Rhythm, 2014. **11**(1): p. 76-82.
57. Roberts, R., C.C. Chang, and T. Hadley, *Genetic Risk Stratification: A Paradigm Shift in Prevention of Coronary Artery Disease.* Journal of the American College of Cardiology: Basic to Translational Science, 2021. **6**(3): p. 287-304.
58. Zhang, X., et al., *Genetic associations with expression for genes implicated in GWAS studies for atherosclerotic cardiovascular disease and blood phenotypes.* Human Molecular Genetics, 2014. **23**(3): p. 782-95.
59. Buroker, N.E., *VEGFA rSNPs, transcriptional factor binding sites and human disease.* Journal of Physiological Sciences, 2014. **64**(1): p. 73-6.
60. Nica, A.C. and E.T. Dermitzakis, *Expression quantitative trait loci: present and future.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1620): p. 20120362.
61. Dunham, I., et al., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.
62. Roberts, R., A. Campillo, and M. Schmitt, *Prediction and management of CAD risk based on genetic stratification.* Trends in Cardiovascular Medicine, 2020. **30**(6): p. 328-334.
63. Wang, K., et al., *A transcription-centric model of SNP-age interaction.* Public Library of Science (PLOS) Genetics, 2021. **17**(3): p. e1009427.
64. Yun, H., N.I. Noh, and E.Y. Lee, *Genetic risk scores used in cardiovascular disease prediction models: a systematic review.* Reviews in Cardiovascular Medicine, 2022. **23**(1): p. 1-9.
65. O'Sullivan, J., et al., *Polygenic Risk Scores for Cardiovascular Disease: A Scientific Statement From the American Heart Association.* Circulation, 2022. **146**(8): p. e93-e118.
66. Gladding, P.A., et al., *Polygenic Risk Scores in Coronary Artery Disease and Atrial Fibrillation.* Heart, Lung, and Circulation, 2020. **29**(4): p. 634-640.
67. Roberts, R. and A. Campillo, *Genetic stratification for primary prevention of coronary artery disease.* Current Opinion in Cardiology, 2018. **33**(5): p. 529-534.
68. Knowles, J.W. and E.A. Ashley, *Cardiovascular disease: The rise of the genetic risk score.* Public Library of Science (PLOS) Medicine, 2018. **15**(3): p. e1002546.
69. Rotter, J.I. and H.J. Lin, *An Outbreak of Polygenic Scores for Coronary Artery Disease.* Journal of the American College of Cardiology, 2020. **75**(22): p. 2781-2784.

70. Smith, G.A., et al., *Vascular endothelial growth factors: multitasking functionality in metabolism, health, and disease.* Journal of Inherited Metabolic Disease, 2015. **38**(4): p. 753-763.
71. Shibuya, M., *Vascular endothelial growth factor and its receptor system: physiological functions in angiogenesis and pathological roles in various diseases.* Journal of Biochemistry, 2013. **153**(1): p. 13-9.
72. Braile, M., et al., *VEGF-A in Cardiomyocytes and Heart Diseases.* International Journal of Molecular Sciences, 2020. **21**(15): p. 1-18.
73. Senger, D.R., et al., *Tumor Cells Secrete a Vascular Permeability Factor That Promotes Accumulation of Ascites Fluid.* Science, 1983. **219**(4587): p. 983-985.
74. Keck, P.J., et al., *Vascular Permeability Factor, an Endothelial Cell Mitogen Related to PDGF.* Science, 1989. **246**(4935): p. 1309-1312.
75. Connolly, D.T., et al., *Tumor vascular permeability factor stimulates endothelial cell growth and angiogenesis.* Journal of Clinical Investigation, 1989. **84**(5): p. 1470-1478.
76. Leung, D.W., et al., *Vascular Endothelial Growth Factor Is a Secreted Angiogenic Mitogen.* Science, 1989. **246**(4935): p. 1306-1309.
77. Niklander, S., et al., *Vascular Endothelial Growth Factor: A Translational View in Oral Non-Communicable Diseases.* Biomolecules, 2021. **11**(1): p. 1-27.
78. Singh, P., et al., *Genetic Scores of eNOS, ACE and VEGFA Genes Are Predictive of Endothelial Dysfunction Associated Osteoporosis in Postmenopausal Women.* International Journal of Environmental Research and Public Health, 2021. **18**(3): p. 1-18.
79. Holmes, D.I. and I. Zachary, *The vascular endothelial growth factor (VEGF) family: angiogenic factors in health and disease.* Genome Biology, 2005. **6**(2): p. 209.
80. Shaik, F., et al., *Structural Basis for Vascular Endothelial Growth Factor Receptor Activation and Implications for Disease Therapy.* Biomolecules, 2020. **10**(12): p. 1-25.
81. Hu, L., et al., *Associations between Vascular Endothelial Growth Factor Gene Polymorphisms and Different Types of Diabetic Retinopathy Susceptibility: A Systematic Review and Meta-Analysis.* Journal of Diabetes Research, 2021. **2021**: p. 7059139.
82. Le, T.H.V. and S.M. Kwon, *Vascular Endothelial Growth Factor Biology, and Its Potential as a Therapeutic Target in Rheumatic Diseases.* International Journal of Molecular Sciences, 2021. **22**(10): p. 1-17.
83. Gu, H., et al., *Methionine sulfoxide reductase A rs10903323 G/A polymorphism is associated with increased risk of coronary artery disease in a Chinese population.* Clinical Biochemistry, 2013. **46**(16-17): p. 1668-1672.
84. Zhou, Y., et al., *The Role of the VEGF Family in Coronary Heart Disease.* Frontiers in Cardiovascular Medicine, 2021. **8**: p. 1-16.
85. Oszajca, K., et al., *Single-nucleotide polymorphisms of VEGF-A and VEGFR-2 genes and risk of infantile hemangioma.* International Journal of Dermatology, 2018. **57**(10): p. 1201-1207.
86. Apte, R.S., D.S. Chen, and N. Ferrara, *VEGF in Signaling and Disease: Beyond Discovery and Development.* Cell, 2019. **176**(6): p. 1248-1264.
87. Peach, C.J., et al., *Molecular Pharmacology of VEGF-A Isoforms: Binding and Signalling at VEGFR2.* International Journal of Molecular Sciences, 2018. **19**(4): p. 1264.
88. Simons, M., E. Gordon, and L. Claesson-Welsh, *Mechanisms, and regulation of endothelial VEGF receptor signalling.* Nature Reviews Molecular Cell Biology, 2016. **17**(10): p. 611-625.
89. Pagès, G. and J. Pouysségur, *Transcriptional regulation of the Vascular Endothelial Growth Factor gene–a concert of activating factors.* Cardiovascular Research, 2005. **65**(3): p. 564-573.
90. Tischer, E., et al., *The human gene for vascular endothelial growth factor. Multiple protein forms are encoded through alternative exon splicing.* Journal of Biological Chemistry, 1991. **266**(18): p. 11947-11954.
91. Keyt, B.A., et al., *The carboxyl-terminal domain (111-165) of vascular endothelial growth factor is critical for its mitogenic potency.* Journal of Biological Chemistry, 1996. **271**(13): p. 7788-95.
92. Krilleke, D., et al., *Molecular Mapping and Functional Characterisation of the VEGF164 Heparin-binding Domain.* Journal of Biological Chemistry, 2007. **282**(38): p. 28045-28056.

93. Stevens, M. and S. Oltean, *Modulation of Receptor Tyrosine Kinase Activity through Alternative Splicing of Ligands and Receptors in the VEGF-A/VEGFR Axis.* Cells, 2019. **8**(4): p. 288.
94. Dehghanian, F., Z. Hojati, and M. Kay, *New Insights into VEGF-A Alternative Splicing: Key Regulatory Switching in the Pathological Process.* Avicenna Journal of Medical Biotechnology, 2014. **6**(4): p. 192-9.
95. Eswarappa, S.M., et al., *Programmed translational readthrough generates antiangiogenic VEGF-Ax.* Cell, 2014. **157**(7): p. 1605-18.
96. Bridgett, S., M. Dellett, and D.A. Simpson, *RNA-Sequencing data supports the existence of novel VEGFA splicing events but not of VEGFAxxxb isoforms.* Scientific Reports, 2017. **7**(1): p. 58.
97. Al Kawas, H., et al., *How VEGF-A and its splice variants affect breast cancer development – clinical implications.* Cellular Oncology, 2022. **45**(2): p. 227-239.
98. Arcondéguy, T., et al., *VEGF-A mRNA processing, stability, and translation: a paradigm for intricate regulation of gene expression at the post-transcriptional level.* Nucleic Acids Research, 2013. **41**(17): p. 7997-8010.
99. Au Yeung, S.L., H. Lam, and C.M. Schooling, *Vascular Endothelial Growth Factor, and Ischaemic Heart Disease Risk: A Mendelian Randomisation Study.* Journal of the American Heart Association, 2017. **6**(8): p. 1-8.
100. Zhang, Y., et al., *Serum vascular endothelial growth factor as a biomarker for prognosis of minor ischaemic stroke.* Clinical Neurology and Neurosurgery, 2020. **196**: p. 106060.
101. Rodríguez-Rodríguez, L., et al., *Vascular endothelial growth factor A and cardiovascular disease in rheumatoid arthritis patients.* Tissue Antigens, 2011. **77**(4): p. 291-297.
102. Heinonen, S.E., et al., *The effects of VEGF-A on atherosclerosis, lipoprotein profile, and lipoprotein lipase in hyperlipidaemic mouse models.* Cardiovascular Research, 2013. **99**(4): p. 716-23.
103. Ruotsalainen, S.E., et al., *An expanded analysis framework for multivariate GWAS connects inflammatory biomarkers to functional variants and disease.* European Journal of Human Genetics, 2021. **29**(2): p. 309-324.
104. Zhao, X., et al., *Vascular endothelial growth factor gene polymorphisms and coronary heart disease: a systematic review and meta-analysis.* Growth Factors, 2018. **36**(3-4): p. 153-163.
105. Li, Z., et al., *Missense Variants in Hypoxia-Induced VEGFA/VEGFR2 Signaling Predict the Outcome of Large Artery Atherosclerotic Stroke.* Cellular and Molecular Neurobiology, 2021. **41**(6): p. 1217-1225.
106. Michel, J.-B., et al., *Intraplaque haemorrhages as the trigger of plaque vulnerability.* European Heart Journal, 2011. **32**(16): p. 1977-1985.
107. Mitrokhin, V., et al., *Circulating interleukin-18: Association with IL-8, IL-10, and VEGF serum levels in patients with and without heart rhythm disorders.* International Journal of Cardiology, 2016. **215**: p. 105-109.
108. Tzeng, H.E., et al., *Interleukin-6 induces vascular endothelial growth factor expression and promotes angiogenesis through apoptosis signal-regulating kinase 1 in human osteosarcoma.* Biochaemical Pharmacology, 2013. **85**(4): p. 531-40.
109. Fahey, E. and S.L. Doyle, *IL-1 Family Cytokine Regulation of Vascular Permeability and Angiogenesis.* Frontiers in Immunology, 2019. **10**: p. 1-15.
110. Sellami, N., et al., *Association of VEGFA variants with altered VEGF secretion and type 2 diabetes: A case-control study.* Cytokine, 2018. **106**: p. 29-34.
111. Stevens, A., et al., *Haplotype analysis of the polymorphic human vascular endothelial growth factor gene promoter.* Cancer Research, 2003. **63**(4): p. 812-6.
112. Kim, Y.-J., et al., *Genetic polymorphisms of vascular endothelial growth factor (VEGF) associated with gastric cancer recurrence after curative resection with adjuvant chemotherapy.* BMC Cancer, 2019. **19**(1): p. 483.
113. Debette, S., et al., *Identification of cis- and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels.* Circulation Research, 2011. **109**(5): p. 554-63.
114. Choi, S.H., et al., *Six Novel Loci Associated with Circulating VEGF Levels Identified by a Meta-analysis of Genome-Wide Association Studies.* Public Library of Science (PLOS) Genetics, 2016. **12**(2): p. e1005874.

115. Cui, Q.T., et al., *Further evidence for the contribution of the vascular endothelial growth factor gene in coronary artery disease susceptibility.* Gene, 2013. **521**(2): p. 217-221.
116. Skrypnik, D., et al., *Association of rs699947 (-2578 C/A) and rs2010963 (-634 G/C) Single Nucleotide Polymorphisms of the VEGF Gene, VEGF-A and Leptin Serum Level, and Cardiovascular Risk in Patients with Excess Body Mass: A Case-Control Study.* Journal of Clinical Medicine, 2020. **9**(2): p. 1-16.
117. Merlo, S., et al., *Vascular Endothelial Growth Factor Gene Polymorphism (rs2010963) and Its Receptor, Kinase Insert Domain-Containing Receptor Gene Polymorphism (rs2071559), and Markers of Carotid Atherosclerosis in Patients with Type 2 Diabetes Mellitus.* Journal of Diabetes Research, 2016. **2016**: p. 1-6.
118. Totomoch-Serra, A., et al., *Association of common polymorphisms in the and genes with type 2 diabetes-related traits in Mexicans.* Archives of Medical Science, 2018. **14**(6): p. 1361-1373.
119. Linhares, P., et al., *Genetic variants of vascular endothelial growth factor predict risk and survival of gliomas.* Tumour Biology, 2018. **40**(3): p. 1-11.
120. Petrovič, M.G., et al., *Local, and genetic determinants of vascular endothelial growth factor expression in advanced proliferative diabetic retinopathy.* Molecular Vision, 2008. **14**: p. 1382-1387.
121. Dong, P.-P., *Association of vascular endothelial growth factor expression and polymorphisms with the risk of gestational diabetes mellitus.* Journal of Clinical Laboratory Analysis, 2019. **33**(2): p. e22686-e22686.
122. Ruggiero, D., et al., *Genetics of VEGF serum variation in human isolated populations of cilento: importance of VEGF polymorphisms.* PLOS One, 2011. **6**(2): p. e16982.
123. Liu, D., et al., *Association of Genetic Polymorphisms on VEGFA and VEGFR2 With Risk of Coronary Heart Disease.* Medicine, 2016. **95**(19): p. e3413.
124. Chen, H., et al., *Polymorphisms of the vascular endothelial growth factor A gene and susceptibility to sporadic brain arteriovenous malformation in a Chinese population.* Journal of Clinical Neuroscience, 2011. **18**(4): p. 549-553.
125. Stathopoulou, M.G., et al., *A common variant highly associated with plasma VEGFA levels also contributes to the variation of both LDL-C and HDL-C.* Journal of Lipid Research, 2013. **54**(2): p. 535-41.
126. Azimi-Nezhad, M., et al., *Associations of vascular endothelial growth factor (VEGF) with adhesion and inflammation molecules in a healthy population.* Cytokine, 2013. **61**(2): p. 602-7.
127. Yadav, B.K., et al., *Genetic Polymorphisms rs699947, rs1570360, and rs3025039 on the VEGF gene are correlated with extracranial internal carotid artery stenosis and ischaemic stroke.* Annals of Clinical and Laboratory Science, 2017. **47**(2): p. 144-155.
128. Ma, W.-Q., et al., *Association of genetic polymorphisms in vascular endothelial growth factor with susceptibility to coronary artery disease: a meta–analysis.* BMC Medical Genetics, 2018. **19**(1): p. 1-12.
129. Palmer, B.R., et al., *Vascular endothelial growth factor-A promoter polymorphisms, circulating VEGF-A and survival in acute coronary syndromes.* PLOS One, 2021. **16**(7): p. 1-15.
130. Hoseini, Z., et al., *VEGF gene polymorphism interactions with dietary trace elements intake in determining the risk of metabolic syndrome.* Journal of Cellular Biochemistry, 2019. **120**(2): p. 1398-1406.
131. Wang, W., A. Xu, and H. Xu, *The roles of vascular endothelial growth factor gene polymorphisms in congenital heart diseases: a meta-analysis.* Growth Factors, 2018. **36**(5-6): p. 232-238.
132. Nia, S.K., et al., *The impact of vascular endothelial growth factor+405 C/G polymorphism on long-term outcome and severity of coronary artery disease.* Journal of Clinical Laboratory Analysis, 2017. **31**(4).
133. Lacchini, R., et al., *Effect of Genetic Polymorphisms of Vascular Endothelial Growth Factor on Left Ventricular Hypertrophy in Patients With Systemic Hypertension.* The American Journal of Cardiology, 2014. **113**(3): p. 491-496.
134. Li, L., et al., *Association of Genetic Polymorphisms on Vascular Endothelial Growth Factor and its Receptor Genes with Susceptibility to Coronary Heart Disease.* Medical Science Monitor, 2016. **22**: p. 31-40.
135. Oh, S.-H., et al., *Association between kinase insert domain-containing receptor gene polymorphism and haplotypes and ischaemic stroke.* Journal of the Neurological Sciences, 2011. **308**(1): p. 62-66.

136. Marks, E.C.A., et al., *Plasma levels of soluble VEGF receptor isoforms, circulating pterins and VEGF system SNPs as prognostic biomarkers in patients with acute coronary syndromes.* BMC Cardiovascular Disorders, 2018. **18**(1): p. 169.
137. Azimi-Nezhad, M., et al., *The Relationship Between Vascular Endothelial Growth Factor Cis- and Trans-Acting Genetic Variants and Metabolic Syndrome.* American Journal of the Medical Sciences, 2018. **355**(6): p. 559-565.
138. Góra-Tybor, J., et al., *Clinical relevance of vascular endothelial growth factor type A (VEGFA) and VEGF receptor type 2 (VEGFR2) gene polymorphism in chronic lymphocytic leukemia.* Blood Cells, Molecules and Diseases, 2015. **54**(2): p. 139-143.
139. Mafu, T., A. September, and D. Shamley, *KDR inferred haplotype is associated with upper limb dysfunction in breast cancer survivors of mixed ancestry.* Cancer Management and Research, 2019. **11**: p. 3829-3845.
140. Salami, A. and S. El Shamieh, *Association between SNPs of Circulating Vascular Endothelial Growth Factor Levels, Hypercholesterolaemia and Metabolic Syndrome.* Medicina, 2019. **55**(8): p. 1-9.
141. Zhang, L.J., et al., *Association of VEGFR-2 Gene Polymorphisms With Clopidogrel Resistance in Patients With Coronary Heart Disease.* American Journal of Therapeutics, 2016. **23**(6): p. e1663-e1670.
142. Wang, L., et al., *A meta-analysis of the relationship between VEGFR2 polymorphisms and atherosclerotic cardiovascular diseases.* Clinical Cardiology, 2019. **42**(10): p. 860-865.
143. Chedid, P., et al., *The association of vascular endothelial growth factor related SNPs and circulating iron levels might depend on body mass index.* Frontiers in Bioscience-Landmark, 2022. **27**(1).
144. Eicher, J.D., et al., *Replication and hematological characterisation of human platelet reactivity genetic associations in men from the Caerphilly Perspective Study (CaPS).* Journal of Thrombosis and Thrombolysis, 2016. **41**(2): p. 343-50.

Chapter 3.- Materials and Methods

The main difference amongst individual results chapters of this thesis (Chapters 4 to 6) involves the number and location of genetic variants across specific chromosomes. The same material and methodology were used for each results chapter (4-6) and are collated in this chapter. The main goal of Chapter 3 is to provide a continuous flow in terms of thesis writing structure as well as avoiding content repetitiveness attributable to a thesis with publications format. A condensed version of this chapter shall be included when publishing the drafted manuscripts of this thesis following discussion and review by the manuscript authors.

3.1 Study populations

3.1.1 Coronary Disease Cohort Study (CDCS)

The Coronary Disease Cohort Study (CDCS) recruited 2140 patients with a diagnosis of acute coronary syndrome (ACS), admitted to Christchurch or Auckland City Hospitals, New Zealand, between July 2002 and January 2009. Inclusion criteria included ischaemic discomfort plus one or more of ECG change (ST-segment depression or elevation of ≥0.5 mm, T-wave inversion of ≥3 mm in ≥3 leads, or left bundle branch block), elevated levels of cardiac markers, a history of coronary disease, age ≥65 years, and a history of diabetes or vascular disease [1]. Patients with serious co-morbidity (e.g. end-stage renal failure, cancer) that limited their life expectancy to <3 years, were excluded. Recruitment included a wide spectrum of age, both genders and significant sub-groups with established risk factors for CHD including hypertension and type II diabetes. Demographic and clinical data were collected at baseline, including blood pressure, height, weight, electrocardiogram (ECG), echocardiography, family and personal medical history and medication regimes as described previously [2, 3]. The collection of data and samples was done by the corresponding baseline clinic staff coordinated by the CDCS cohort stewards specified in Section 1.5.

Plasma samples were assayed for natriuretic peptides (ANP, NT-ANP, BNP, NTproBNP, CNP, NT-CNP) and other cardiometabolic variables (total cholesterol, creatinine, urate, troponin I, aldosterone, endothelin and adrenomedullin). Measurement methods for these variables are described elsewhere [2, 3]. Patients were followed for a median of 5.04 (0.08–9.49) years. Patients attended follow-up clinics at 3–5 months and 12–14 months post-onset of ACS and participants completed questionnaires at 2- and 3-years post-discharge. Ethnicity was self-declared and categorised as European, Māori/Pasifika (Pacific Islander), Asian and Middle Eastern/Latin American/African (MELAA). Standardised transthoracic echocardiography was performed at baseline and at each follow-up clinic either at Christchurch Hospital or University of Auckland clinics, as described previously [1, 2, 4]. The study was approved by the New Zealand (NZ) Multi-Region Ethics Committee and registered with the Australian New Zealand Clinical Trials Registry (ACTRN12605000431628 on 16 September 2005). All participating patients provided written, informed consent. Access to the cohort dataset and DNA samples was established by agreement between Massey University and the University of Otago.

3.1.2 Canterbury Healthy Volunteers Study (HVOL)

A subset (n = 1183), age and gender matched with the CDCS cohort were selected from the Canterbury Healthy Volunteers Study (HVOL, n = 3,358) [5]. Participants in the HVOL were randomly selected from the electoral roll of Canterbury, New Zealand, excluding those with prior hospital admissions and cardiac diagnoses. CVD risk factors, anthropometric measures, personal health information and family history of cardiovascular events were recorded for each participant. Plasma sample aliquots were biobanked at -80 °C and assayed for different analytes as previously described [5]. Recruitment commenced in January 2002, and clinical events were documented from the NZHIS over a median follow-up of 9.2 years (range 20 days – 16.5 years). Sample collection and analyte measurements were carried out by researchers from the Christchurch Cardioendocrine Research Group (Department of Medicine, University of Otago) as described by Ellis et al. (2011) [5]. The study was approved by the New Zealand Health and Disability Ethics Committee (Reference CTY/01/05/062) and registered with the Australian New Zealand Clinical Trials Registry (ACTRN1260500448640). All participants gave written, informed consent. Access to the cohort dataset, DNA (n = 1183) and plasma (n = 223) samples was established in an agreement between Massey University and the University of Otago.

3.1.3 Clinical events

Clinical events in CDCS and HVOL were documented at scheduled follow-up clinic visits, through consulting patient notes and Hospital Patient Management System databases plus regular updates and corroboration acquired from the NZ Health Information Service.

Endpoints of interest included death from any cause, myocardial infarction (ST elevated, non-ST elevated or unspecified), stroke (ischaemic or haemorrhagic), heart failure and unstable angina. Additional death categories include cardiac, cardiovascular and MI related death. Cardiac death was defined as any death due to a heart condition (e.g. MI, arrythmia, heart failure) whereas cardiovascular death includes all cardiac deaths plus those that involve vascular conditions (e.g. stroke).

The composite variable major adverse cardiovascular event (MACE) was also used. MACE included patients presenting any of the following: any cause of death, any type of MI (ST elevated, non-ST elevated

or unspecified) or any type of stroke (ischaemic or haemorrhagic). Survival times and readmission events were calculated from the date of index admission.

3.2 Genetic Methods

3.2.1 Variant imputation

Genome wide imputation data was generated for the CDCS and HVOL by Dr Vinicius Tragante (deCODE Genetics, Reykjavik). This involved genotype calling using the Axiom Analysis Suite (ThermoFisher Scientific) in conjunction with the database Axiom_PMDA_Plus.r6 and mapping done with genome assembly hg38 (Dr Vinicius Tragante, personal communication). A total of 800,000 variants were called across 21 microplates each with capacity for 96 samples. Quality control was performed using Plink 1.9 [6] to remove SNP missingness (>3%), individual missingness (>2%), relatedness (PI_HAT > 0.1) and variants not in Hardy-Weinberg equilibrium ($p < 1 \times 10^{-6}$).

For this PhD project, Juan Carlos Meza Alvarado selected SNPs which had reported associations with cardiometabolic parameters, CVD risk or had reports as eQTL SNPs for VEGF-A levels (discussed in Chapter 2). Additional variants considered for the *VEGFR2* locus (Chapter 5) included those that overlapped with genomic regulatory regions (e.g. transcription factor binding motifs, enhancer signatures) based on genome assembly GRCh38/hg38 from the University of California, Santa Cruz Genome Browser (http://genome.ucsc.edu) and Ensembl databases [7, 8]. In total 47 SNPs were selected over three chromosomes with 30 mapping to the *VEGFA* locus on human chromosome 6 (Chapter 4), 13 mapping to the *VEGFR2* locus on chromosome 4 (Chapter 5) and 4 SNPs located between the *VLDLR* and *KCNV2* genes on chromosome 9 (Chapter 6).

Extraction of the imputed genotype data for the 47 selected variants from the CDCS and HVOL genome-wide imputation dataset was performed by Dr Vinicius Tragante (deCODE Genetics, Reykjavik). Specifically, imputation was carried out for the selected 47 SNPs using the Michigan Imputation Server, which implemented the 'minimac4' algorithm [9]. Strand-aligned genotype data were loaded into the server. The imputation was performed using the HRC r1.1 2016 reference panel and GRCh37/hg19 array build in 'Quality Control and Imputation' mode (population = other/mixed). Phasing was performed with Eagle v2.4. All biallelic variants with imputation quality threshold of INFO score ≥ 0.3 were extracted (Dr

Vinicius Tragante, personal communication). The quality of the imputation was assessed by r^2-values with $r^2=1$ meaning the strand-aligned and imputed genotypes match perfectly. Imputed genotypes for the selected SNPs were available for 1935 CDCS patients and 1183 HVOL individuals.

3.2.2 DNA extraction

Four saliva sample donations were provided by the supervisory team at Massey University Wellington Campus to be used as DNA controls for TaqMan SNP genotyping (Section 3.2.3). The donations were collected and processed by Juan Carlos Meza Alvardo. The extraction process consisted of the donor rinsing their mouth vigorously with 10 mL of a 4% sucrose solution for 1 minute. This volume was recollected into a 50 mL tube. Immediately, each saliva sample was centrifuged at 1,620 x g for 10 minutes at room temperature. The supernatant was discarded while the pellet was resuspended in 500µL of phosphate buffer saline (PBS) 1X. The resuspended buccal cell mixture was then transferred to two 1.5 mL microcentrifuge tubes, each containing a volume between 250 and 300 µL. Subsequently, both tubes were centrifuged at 16,100 x g for 1 minute at room temperature and the supernatant was discarded.

The DNeasy Blood & Tissue kit (Qiagen) was used following the manufacturers' instructions to purify DNA from the pellet of buccal cells. Briefly, the buccal cell pellet was resuspended in 200 µL of PBS and 20 µL of proteinase K (20 mg/mL) were added. Following this, 200 µL of lysis buffer AL were added and the suspension was mixed thoroughly by vortexing. Immediately, the samples were incubated at 56°C for 10 minutes. Afterwards, 200 µL of 100% ethanol were added and mixed thoroughly by vortexing. The contents were then transferred into DNeasy Mini spin column (Qiagen) in a 2 mL collection tube (Qiagen). The column and tube were centrifuged for 1 minute at 6000 g. The flowthrough and collection tube were discarded while the spin column was placed in a new 2 mL collection tube. An additional 500 µL of Buffer AW1 were added to the spin column which was then centrifuged for 1 minute at 6000 g. Similarly, the flowthrough and collection tube were discarded while the spin column was placed in a new 2 mL collection tube. An additional 500 µL of Buffer AW2 were added to the spin column which was then centrifuged for 3 minutes at 16,100 x g . The flowthrough and collection tube were discarded while the spin column was placed in a new 1.7 mL microtube (Axygen). The DNA was eluted from the column by adding 100 µL of Buffer AE. Afterwards, the column was incubated for 1 minute at room temperature and centrifuged for 1 minute at 16,100 x g . The purified genomic DNA was quantified using a DS-11 spectrophotometer

(DeNovix, Wilmington, USA) using the preprogramed double strand DNA (dsDNA) quantification function.

3.2.3 TaqMan SNP genotyping

Manual genotyping to validate imputed genotype data was performed for 2027 DNA samples from the CDCS cohort and 1177 samples from the CHVS cohort. The genotyping used the allelic (VIC and FAM-labelled) discrimination method. This involved the use of predesigned TaqMan assays (ThermoFisher Scientific) for two survival-associated SNPs in human chromosome 6 (Chapter 4). The assays were C__11542106_10 for rs6921438 and C__29965406_20 for rs7767396. Juan Carlos Meza Alvarado performed the genotyping for both Chromosome 6 SNPs (rs6921438 and rs7767396) using 2027 DNA samples from the CDCS cohort and 1177 samples from the CHVS cohort. Manual genotyping for rs6921438 and rs7767396 addresses Research Objective 4 (Chapter 1, Section 1.3).

For rs6921438 and rs7767396 genotyping reactions were performed using a Roche LightCycler LC96 system (Roche Diagnostics Ltd., Rotkreuz, Switzerland) in 96 well plates. Each 10 µL reaction volume included TaqMan Master Mix (Applied Biosystems), SNP-specific primers (TaqMan probe) and 1 ng of genomic DNA. As quality control, a random selection of 10% of samples (n = 210) from the CDCS cohort were re-genotyped with 99.5% concordance with the original genotypes. Each plate included three DNA samples (Section 3.2.2) as experimental controls for each genotype group (homozygous reference, heterozygous and homozygous for the minor allele) and a non-template control of nuclease free water (Invitrogen). Genotype groups were determined based on scatter plot and fluorescent measurements as determined by the LightCycler LC96 system "Endpoint Genotyping" Analysis Software. The experimental conditions for the polymerase chain reaction (PCR) of each plate were as follows:

Table 3.1. TaqMan SNP Genotyping PCR Conditions

Step	Temperature	Time	Cycles
Polymerase activation	95 °C	10 min	1
DNA Denaturation	95 °C	15 s	40
TaqMan Probe Annealing and DNA extension	60 °C	1 min	40

Pre-existing manual genotyping data was used to validate imputed genotype data on two SNPs in human chromosome 4 (Chapter 5). Manual genotype data available for CDCS patients included variants

rs2305948 (n = 1752) and rs1870377 (n = 1912). The DNA extraction and SNP genotyping conditions for both chromosome 4 variants were carried out by Marks et. al. (2018) as described previously [2].

3.3 Plasma VEGF-A assay

Plasma samples for the CDCS and HVOL participants were collected and stored at -80˚C by the Christchurch Heart Institute as described previously [1, 2]. Levels of VEGF-A at baseline from a subset of 549 CDCS participants were originally reported by Palmer, et al (2021) [3]. For this PhD project, Juan Carlos Meza Alvarado measured VEGF-A levels in a subset of 223 plasma samples from the HVOL cohort. VEGF-A plasma level data generated for the 223 HVOL plasma samples were compared with the VEGF-A data previously generated by Palmer et al. (2021) for the CDCS cohort. The HVOL plasma samples were determined using a predesigned quantitative sandwich enzyme linked immunosorbent assay (ELISA) kits to measure VEGF-A concentration (R&D Systems Europe, Abingdon, UK).

Each kit included a calibrator diluent (RD6U), a buffered protein base assay diluent (RD1W), a Wash Buffer Concentrate, a Substrate solution, a Stop solution, a human VEGF standard, a polyclonal antibody specific for human VEGF conjugated with horseradish peroxidase (VEGF Conjugate), adhesive strips and a 96-well microplate pre-coated with a monoclonal antibody specific for human VEGF. Each microplate included two technical replicates for each sample and a standard dilution series of VEGF-A standard. A standard curve was generated from the standard dilution series, described below, for each set of samples analysed per plate (n = 38 samples). The VEGF-A assay's detectable concentration range was 0 – 115 pg/mL.

To produce a VEGF dilution series, the human VEGF Standard was reconstituted to a 2000 pg/mL stock solution by using 1 mL of the R&D calibrator diluent RD6U. This solution was diluted to prepare an 8-point standard curve. This involved adding 500 µL of the RD6U Diluent to six empty 1.7 mL microtubes (Axygen). A total of 500 µL were taken from the stock solution and added to a microtube with RD6U Diluent to create a 1000 pg/mL solution. The diluted mixture was mixed thoroughly by vortexing. This process of adding 500 µL from a dilution to a tube with RD6U diluent was repeated five more time to produce dilutions at 500 pg/mL, 250 pg/mL, 125 pg/mL, 62.5 pg/mL, and 31.3 pg/mL. Stock RD6U was used as the 0 pg/mL dilution. The process to develop the dilutions is displayed in Figure 3.1.

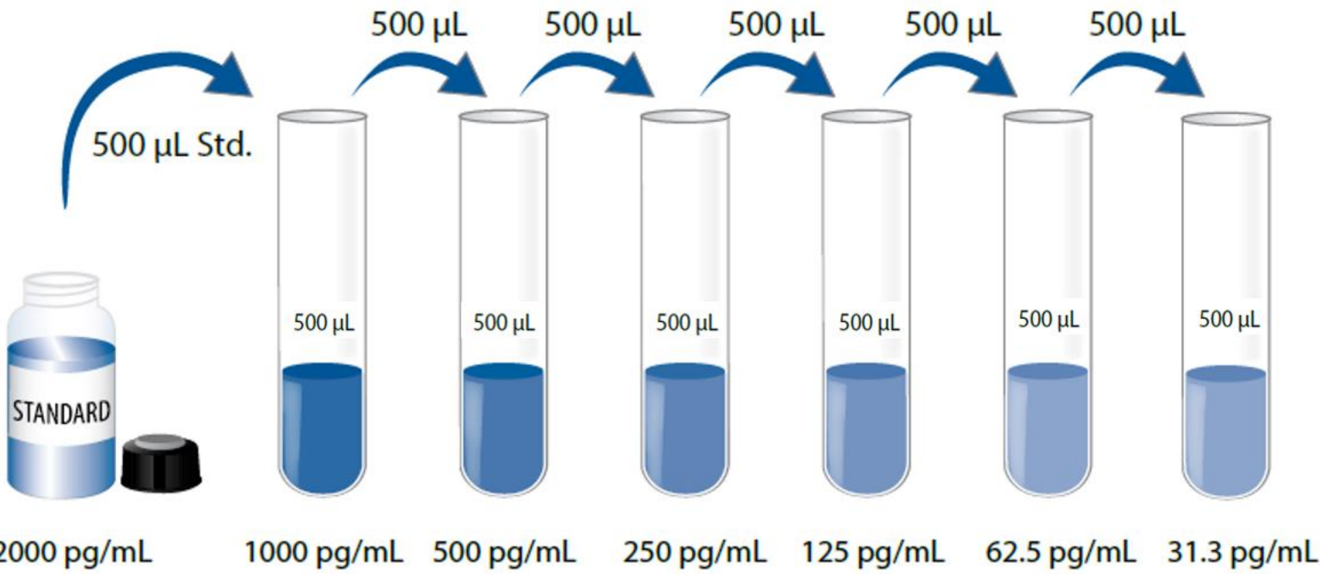


Figure 3.1 Dilution series production from a standard solution. The initial volume of assay diluent in each tube and the final concentrations used in the dilution series are specified.

Each 1 mL HVOL plasma aliquot was thawed and vortexed for 30 s followed by centrifugation at 2,400 x g for 15 s at 4°C to sediment cellular debris. A total of 100 µL of RD1W assay diluent were added to each well. Immediately, 100 µL of each sample were added to each well. The plate was then covered with an adhesive strip and incubated for 2 hours at room temperature (~20 °C). Following the incubation, the sample-RD1W mixture was removed through decantation. Then, each well was washed by adding 400 µL of 1X Wash Buffer and removing it through decantation. This process was repeated three times. On the last wash, removal of Wash Buffer was done by aspiration followed by plate inversion and blotting on clean paper towels. Subsequently, 200 µL of VEGF conjugate were added to each well and the plate was incubated for 2 hours. After this incubation, the plate was washed as described previously. The VEGFA-VEGF conjugate complexes formed were detected by adding 200 µL of Substrate Solution to each well. The reaction was developed in the dark for 25 minutes at room temperature (~20 °C) and stopped by adding 50 µL of stop solution to each well.

The reactions were then measured with the FLUOstar Omega microplate reader (BMG Labtech) at 570 nm and 450 nm. The optical density (OD) readings for each VEGF-A ELISA plate were exported in CSV format with the MARS 3.32 (BMG Labtech), and the calculations for VEGF-A concentration were done in Microsoft Office Excel. Data treatment involved generating OD corrected values for each well by

subtracting the 570 nm readings from their 450 nm counterparts. Afterwards, these OD corrected values were averaged for each technical duplicate of each sample and standard curve dilution. Lastly, the average of the OD corrected value for the blank well (0 pg/mL) was subtracted from all wells. A standard curve was generated using the resultant OD values for the standard dilution series and plotted against their respective concentrations. The sample resultant OD values were used to calculate VEGF-A concentration by using the standard curve. VEGF-A concentration was reported as pg/mL.

3.4 Statistical analysis

Univariate analyses were performed by the student to test associations between all 51 imputed SNP genotypes and baseline data for anthropomorphic measurements, analyte levels and echocardiographic measurements using Chi-squared (χ^2) and ANOVA tests, with age as a covariate with Bonferroni correction, where applicable. All SNP variants were analysed in a genotypic model. Due to low frequency (< 5 %) of the minor allele "a" for 8 of the 51 imputed SNPs, statistical comparisons between genotype groups were performed using a dominant model (AA vs Aa/aa). Skewed data were log-transformed before analysis and geometric means with 95% confidence intervals reported. The survival of groups was compared using Kaplan-Meier analysis and the log-rank test.

Candidate CVD risk variants were selected from among all the imputed SNPs if the ANOVA data presented significant association ($p<0.05$) or trended towards significance close to significant association ($p<0.1$) with cardiac risk markers or with VEGF-A levels. The selected variants were analysed for univariate association with survival using the Kaplan-Meier log-rank test and multivariate Cox proportional hazards to identify independent associations between genotype groups and all-cause mortality. The multivariate survival analyses included the following established predictors: age, gender, previous MI, beta blocker treatment, physical activity, and NTproBNP levels. This process will cover Research Objectives 2 and 3 (Chapter 1, Section 1.3).

The ANOVA, Kaplan-Meier and Cox proportional hazards multivariate analyses were repeated using manual genotyping data where available (Section 3.2.3), with survival tests expanded to include six clinical outcome endpoints (all-cause mortality, STEMI, NSTEMI, unstable angina, MACE and HF). Additional predictors included in the multivariate Cox models were statistically significant variables identified from the ANOVA tests for specific SNPs. Significant findings will cover Research Objective 5

(Chapter 1, Section 1.3) All analyses were performed using SPSS version 28.0.1.1 (IBM, Armonk, USA). Statistical significance was set at the 5% level ($p<0.05$).

3.5 Genomic context analysis

The genomic context for selected candidate CVD risk variants or survival associated variants (Section 3.4) was explored by the student to identify if the variants overlapped transcription factor binding motifs or regulatory regions involved with CVD risk. This assessment involved using assembly GRCh38/hg38 from the University of California, Santa Cruz Genome Browser (http://genome.ucsc.edu), the associated JASPAR database of transcription factor motifs, the Ensembl database and HaploReg databases [6, 10, 11]. This addresses Research Objective 6 (Chapter 1, Section 1.3). Relevant transcription factor binding motifs or regulatory regions are specified in the chapter corresponding to each variant's respective chromosome. This applied to n=13 SNPs with n=2 located in human chromosome 6 (Chapter 4), n = 7 in human chromosome 4 (Chapter 5) and n = 4 in human chromosome 9 (Chapter 6).

3.6 References

1. Ellis, K.L., et al., *Genomic risk variants at 1p13.3, 1q41, and 3q22.3 are associated with subsequent cardiovascular outcomes in healthy controls and in established coronary artery disease.* Circulation Cardiovascular Genetics, 2011. **4**(6): p. 636-46.
2. Marks, E.C.A., et al., *Plasma levels of soluble VEGF receptor isoforms, circulating pterins and VEGF system SNPs as prognostic biomarkers in patients with acute coronary syndromes.* BMC Cardiovascular Disorders, 2018. **18**(1): p. 169.
3. Palmer, B.R., et al., *Vascular endothelial growth factor-A promoter polymorphisms, circulating VEGF-A and survival in acute coronary syndromes.* PLOS One, 2021. **16**(7): p. 1-15.
4. Palmer, B.R., et al., *Angiotensin-converting enzyme 2 A1075G polymorphism is associated with survival in an acute coronary syndromes cohort.* American Heart Journal, 2008. **156**(4): p. 752-758.
5. Lassé, M., et al., *Fibrinogen and hemoglobin predict near future cardiovascular events in asymptomatic individuals.* Scientific Reports, 2021. **11**(1): p. 4605.
6. Chang, C.C., et al., *Second-generation PLINK: rising to the challenge of larger and richer datasets.* GigaScience, 2015. **4**(1).
7. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Research, 2002. **12**(6): p. 996-1006.
8. Martin, F.J., et al., *Ensembl 2023.* Nucleic Acids Research, 2022. **51**(D1): p. D933-D941.
9. Das, S., et al., *Next-generation genotype imputation service and methods.* Nature Genetics, 2016. **48**(10): p. 1284-1287.

Chapter 4.- Study on *VEGFA* locus SNPs

This chapter focuses on the work involving 30 SNPs mapping to the VEGFA locus (human chromosome 6) which had associations reported with cardiometabolic parameters or CVD risk in the existing literature (Chapter 2.- Literature review). This addresses project aim A as well as research objectives 2 and 3 (Chapter 1, Section 1.3).

The chapter highlights the experimental findings surrounding two VEGF-A eQTL SNPs located in a non-coding region near VEGFA. These results address project aim B as well as research objectives 4 and 5 (Chapter 1, Section 1.3).

STATEMENT OF CONTRIBUTION

DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the student and the student's main supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the student's contribution as indicated below in the Statement of Originality.

Student name:	Juan Carlos Meza Alvarado
Name and title of main supervisor:	Dr. Barry Palmer, Senior Lecturer
In which chapter is the manuscript/published work?	Chapter 4

Describe the contribution that the student and members of the supervisory team have made to the manuscript/published work:[1]

- The student's role involved project conceptualization, analysing the imputed genotype data. generating SNP data (rs6921438, rs7767396), generating data on HVOL VEGF-A levels, investigation of all variants, data curation and analysis, visualization, draft writing and editing.
- Data on CDCS VEGF-A levels was generated by others and described in Palmer et al (2021). Chapter 4 Ref #4
- Imputed genotype data analysed in this Chapter was generated by Dr. Anna Pilbrow and Dr. Vinicius Tragante
- The supervisory team provided feedback of the manuscript drafts.
- The student and main supervisor will be discussing further manuscript modifications with external project collaborators.

Please select one of the following three options:

○	The manuscript/published work is published or in press Please provide the full reference of the research output:
○	The manuscript is currently under review for publication Please provide the name of the journal:
●	It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Student's signature:	[signature]	Main supervisor's signature:	[signature]

This form should be placed at the beginning of each relevant thesis chapter.

Doctoral Research Committee May 2023

4.1 Abstract

Background: Cardiovascular disease (CVD) is the leading cause of death worldwide. Risk stratification of patients with CVD may be improved by predictive biomarkers, including genetic markers. Vascular endothelial growth factor A (VEGF-A) promotes angiogenesis, endothelial cell proliferation and vascular permeability regulation. Elevated circulating VEGF-A levels have been linked to CVD and single nucleotide polymorphisms (SNPs) have been previously associated with VEGF-A levels. We explored whether SNPs on the *VEGFA* locus are associated with VEGF-A levels and with clinical outcomes in patients with established coronary disease. VEGF-A levels were compared between coronary patients and heart healthy controls.

Methods: Imputed genotype data for 30 SNPs on human chromosome 6 from 1935 patients from the Coronary Disease Cohort Study (CDCS) and 1183 individuals from the Canterbury Healthy Volunteers Study (HVOL) were analysed for associations with cardiometabolic parameters using one-way ANOVA. Association with clinical endpoints was assessed using Kaplan-Meier analysis and multivariate regression models. To validate the findings from imputed data, DNA samples of 2027 CDCS patients and 227 HVOL participants were manually genotyped for variants rs6921438 and rs7767396. Plasma VEGF-A was measured by ELISA immunoassay in 227 HVOL participants and compared with plasma concentrations in 549 CDCS patients.

Results: Manual genotyping for rs6921438 and rs7767396 showed: **1)** In both cohorts, rs6921438 AA and rs7767396 GG genotype groups had lower VEGF-A levels at baseline (CDCS: rs6921438 AA (27.7 pg/mL) v AG (43.3 pg/mL), $p = 7.36 \times 10^{-11}$, AA v GG (63.2 pg/mL) $p = 1.49 \times 10^{-22}$, rs7767396: GG (27.4 pg/mL) v AG (42.8 pg/mL) $p = 1.98 \times 10^{-10}$, GG v AA (61.5 pg/mL) $p = 6 \times 10^{-6}$; HVOL rs6921438 AA (12.8 pg/mL) v GA (19.9 pg/mL) p = 0.258, AA v GG (26.4 pg/mL) p = 0.017; rs7767396 GG (12.6 pg/mL) v AG (19.6 pg/mL) p = 0.301; GG v AA (25.9 pg/mL) p = 0.023) **2)** rs6921438 AA was associated with increased all-cause death risk (p = 0.03), non ST-elevated myocardial infarction (NSTEMI, p = 0.0003), heart failure (HF, p = 0.035) and major adverse cardiovascular event (MACE, p = 0.032) risk **3)** rs7767396 GG was associated with increased NSTEMI (p = 0.001) HF (p = 0.023) risk **4)** rs6921438 AA (Hazard Ratio (HR) = 6.6 p = 0.016), rs7767396 GG (HR = 0.149, p = 0.017) and VEGF-A (HR = 2.64, p = 0.014) were independent HF readmission risk predictors

Conclusion: Variants rs6921438 and rs7767396 may influence plasma VEGF-A levels. Both SNPs and VEGF-A have potential to be used to aid in prognostic risk stratification for HF after acute coronary events.

4.2 Introduction

CVD is an important contributor to health deficits in New Zealand [1]. Prediction of later outcomes after CVD events may be aided by biomarkers [2]. Members of the vascular endothelial growth factor family have previously been proposed for CVD risk stratification and may be useful predictors in extant coronary disease [3-5]. Vascular endothelial growth factor (VEGF-A) is a key factor in blood vessel formation (angiogenesis) and collateral circulation (arteriogenesis), mediated by binding to the receptors VEGFR-1 (Flt-1) and VEGFR-2 (KDR) [6]. These receptors are found on the surface of endothelial and non-endothelial cells [7]. VEGFR1 participates in angiogenesis, while VEGFR2 is the primary inducer of VEGF-mediated blood vessel growth [8-10]. Additionally, VEGFR1 interacts with the co-receptor neuropilin-1 (NRP1), which selectively potentiates VEGFR2-mediated vascular permeability, and endothelial cell motility in vascular development [11, 12]. Furthermore, VEGFR1 has a soluble splice variant (sFlt-1) that acts as a decoy receptor, decreasing VEGF-A plasma concentration and limiting KDR binding [7, 8, 13].

Increased plasma and tissue levels of VEGF-A have been observed in various conditions including ischaemic heart disease (IHD), coronary artery disease (CAD), coronary heart disease (CHD), stroke, heart failure (HF), and myocardial infarction (MI) [14-17]. Additionally, high VEGF-A levels have been associated with various CVD risk factors including smoking, hypercholesterolaemia, diabetes, hypertension, metabolic syndrome, and hyperglycaemia [6, 13, 14]. Increased VEGF-A activity impacts the vascular system by increasing inflammatory molecule activity (e.g. TNF-α, IL-6) leading to increased vascular dilation, adhesion protein expression and trans-endothelial lipid migration which promote atherosclerotic lesion development [6, 14, 18-22].

The *VEGFA* gene has a 16.3kb coding region located at 6p21.1 on the short arm of chromosome 6, including eight exons and seven introns [23, 24]. Factors that upregulate VEGF-A gene expression include the hypoxia inducible factor, p53 allele polymorphisms, thyroid stimulating hormone, oestrogen levels and oxygen tension [8, 22, 23, 25]. VEGF-A circulating levels are influenced by SNPs. Four common variants (rs6921438, rs4416670, rs6993770 and rs10738760) distributed across three chromosomes have been independently associated with circulating VEGF-A levels and explain up to

48% of the heritability of serum VEGF-A levels [26]. A meta-analysis of genome wide association study (GWAS) data suggested ten SNPs contributed up to 52% of variance in total circulating VEGF-A [27]. Individual studies have also identified that SNPs at the *VEGFA* gene locus are associated with CVD [24-26] and CVD risk factors such as lipid metabolites [26–28] and coronary disease biomarkers [24-29]. Furthermore, elevated levels of VEGF-A may contribute to CVD onset or progression [5, 28, 29]. We explored the relationships of 30 *VEGFA* gene variants with cardiometabolic variables, including plasma concentrations of VEGF-A, in post-acute coronary syndrome patients and heart healthy matched controls [3, 4, 30-32]. SNPs were also analysed for associations with clinical endpoints.

4.3 Results

4.3.1 Baseline characteristics of study cohorts

Baseline characteristics of the CDCS cohort are given in Table 4.1. The population is predominantly of European ethnicity (90%) and male (71.7%). Throughout the follow-up period the clinical outcomes with highest rates of occurrence were major adverse cardiovascular events (MACE, n = 862, 40.3%), death (n = 500, 23.4%), NSTEMI (n = 488, 22.8%) and HF (n = 392, 8.59%). Additionally, discharge medications most reported by study participants were antithrombotic therapy (96.3%) followed by lipid lowering medication (86.1%) and vasoactive/anti-adrenergic agents. These include angiotensin converting enzyme inhibitors, angiotensin receptor blockers, beta blockers and mineralocorticoid receptor antagonists.

Baseline characteristics of the HVOL cohort are described in Table 4.1. The population was mostly European (98.4%) and male (64%). There were no significant differences for age and gender between the CDCS patients and HVOL cohort subset. As expected, compared to the CDCS participants this cohort had comparatively lower incidence of raised serum cholesterol (31%), hypertension (30.5%) and diabetes (5.2%). Additionally, low levels of medication usage were observed with blood pressure lowering medication (27.6%) being the most common.

Table 4.1. Baseline characteristics of the CDCS and HVOL cohort

Variables	CDCS		HVOL		p -value
	n	Mean ± SE or n (%)	n	Mean ± SE or n (%)	
Anthropometric					
Male gender	2026	1453 (71.7%)	250	160 (64%)	1.53×10^{-91}
Age	2026	66.7 ± 0.271	250	68.9 ± 0.522	3.3×10^{-3}
Ethnicity (European, Māori & Pasifika, Indian, MELAA)	2026	90%, 5.9%, 3.8%, 0.3%	250	98.4%, 1.6%, 0%, 0%	
BMI (kg/m^2)	1998	27.5 ± 0.112	250	26.4 ± 0.261	7×10^{-4}
MI Family history	2012	597 (29.5%)	248	105 (42%)	6.27×10^{-75}
Tobacco (smoker, ex-smoker, never smoked)	2026	6.3%, 57.4%, 36.3%	249	5%, 35.6%, 59.4%	1.38×10^{-211}
Alcohol (drinker, ex-drinker, non-drinker)	2023	63%, 12%, 25%	250	78.8%, 4.4%, 16.8%	2.17×10^{-240}
Cardiovascular Readmission after discharge	2026	1335 (65.9%)	N/A		
Previous disease history					
Hypertension	2009	1048 (51.7%)	249	76 (30.5%)	7.59×10^{-4}
High Cholesterol	2089	1135 (54.3%)	245	76 (31%)	0.068
Diabetes	2137	7 (0.3%)	250	13 (5.2%)	
Medication					
Antithrombotic	2026	1951 (96.3%)	250	31 (12.4%)	3.19×10^{-274}
Lipid lowering	2026	1792 (88.5%)	250	45 (18%)	9.33×10^{-189}
Blood pressure lowering	2022	1740 (86.1%)	250	69 (27.6%)	3.71×10^{-176}

Abbreviations: BMI: Body mass index, CDCS: Coronary Disease Cohort Study, HVOL: Canterbury Healthy Volunteers Study, MELAA: Middle Eastern Latin American African, MI: myocardial infarction, SE: Standard Error

4.3.2 SNP genotype analysis

Following initial ANOVA tests on the imputed genotype data obtained on 1935 CDCS patients for 30 SNPs, five variants (rs4513773, rs6921438, rs7763440, rs7767396, rs11757868) were associated with ten variables including VEGF-A and natriuretic peptides (Appendix 4.1). Imputed genotype data for 1183 HVOL individuals showed five variants (rs4513773, rs6921438, rs7763440, rs7767396, rs11757868) were associated with ANP levels (Appendix 4.2). To confirm genotype frequencies and imputed genotype findings, manual genotyping was performed for rs6921438 and rs7767396 in 2026 CDCS patients and 227 HVOL individuals. These variants were selected due to their reported involvement with CVD risk pathways.

The CDCS SNP frequencies for manual genotyped samples were rs6921438 GG 26.1%, GA 48.5%, and AA 25.4% (minor allele frequency [MAF] "A" = 0.4965); rs7767396 AA 28.4%, AG 48.6%, and GG 23% (MAF "G" = 0.472). Compared to the manually genotyped data, the imputed genotype SNP frequencies

were rs6921438 GG 28%, GA 49.5%, and AA 22.5% (MAF = 0.473); rs7767396 AA 30.7%, AG 48.3%, and GG 21% (MAF = 0.4515). Overall, manual genotyping for rs6921438 and rs7767396 was concordant with imputed genotypes from the CDCS cohort (rs6921438 – 60.5% $r^2 = 0.534$, rs7767396 – 60.8% $r^2 = 0.537$, $p < 0.001$). Genotype frequencies for this cohort are summarised in Appendix 4.3.

The HVOL imputed genotype SNP frequencies were rs6921438 GG 28.4%, GA 49.2%, and AA 22.4% (MAF = 0.47); rs7767396 AA 30.8%, AG 48.4%, and GG 20.8% (MAF = 0.45). Comparatively, the SNP frequencies for manually genotyped samples were rs6921438 GG 23%, GA 53%, and AA 23.9% (MAF = 0.495); rs7767396 AA 23.8%, AG 53.7%, and GG 22.5% (MAF = 0.4935). Overall, manual genotyping for rs6921438 and rs7767396 was concordant with imputed genotypes from the HVOL cohort subset (rs6921438 – 56.4% $r^2 = 0.529$, rs7767396 – 56% $r^2 = 0.535$, $p < 0.001$). Genotype frequencies for this cohort are summarised in Appendix 4.3.

4.3.3 rs6921438 and rs7767396 metabolite associations

A summary of the ANOVA analyses using imputed genotype data for rs4513773, rs6921438, rs7763440, rs7767396 and rs11757868 in the CDCS is shown in Appendix 4.1. Manual genotyping showed rs6921438 GG and rs7767396 AA were associated with lower systolic blood pressure (Table 4.2). Additionally, mean VEGF-A levels trended to progressively decrease with the addition of minor alleles for each SNP (Table 4.2, rs6921438 $p = 4.49 \times 10^{-22}$, rs7767396 $p = 3.47 \times 10^{-21}$). Specifically, CDCS patients with the rs6921438 AA genotype had lower plasma VEGF-A levels than those with the GA genotype (Figure 4.1A, $p = 7.36 \times 10^{-11}$) and GG (Figure 4.1A, $p = 1.49 \times 10^{-22}$) groups. Similarly, CDCS patients with the rs7767396 GG genotype had lower VEGF-A levels compared to the AG (Figure 4.1B, $p = 1.98 \times 10^{-10}$) and AA (Figure 4.1B. $p = 6 \times 10^{-6}$) groups.

CDCS VEGF-A levels were compared with VEGF-A levels measured in the HVOL subset (Figure 4.1). As expected from a heart healthy cohort, HVOL individuals presented with lower VEGF-A levels compared to the CDCS cohort (Figure 4.1). Mean VEGF-A levels trended to progressively decrease with the addition of minor alleles for each SNP (rs6921438, $p = 0.021$, rs7767396 $p = 0.029$). HVOL individuals with the rs6921438 AA genotype (12.8 pg/mL ± 1.61) showed statistically lower plasma VEGF-A than the GG (26.4 pg/mL ± 5.63, $p = 0.017$) but not the GA (19.9 pg/mL ± 1.83, $p = 0.258$) group. Likewise,

individuals carrying the rs7767396 GG genotype (12.6 pg/mL ± 1.74) had statistically lower mean VEGF-A levels than the AA (25.9 pg/mL ± 5.53, p = 0.023) but not the AG (19.6 pg/mL ± 1.81, AA, p = 0.301) group.

For the HVOL imputed genotype data, only ANP levels were significantly associated with five SNPs (rs4513773, rs6921438, rs7763440, rs7767396 and rs11757868). No variable besides VEGF-A was associated with manual genotyping data on rs6921438 and rs776736 for the HVOL subset.

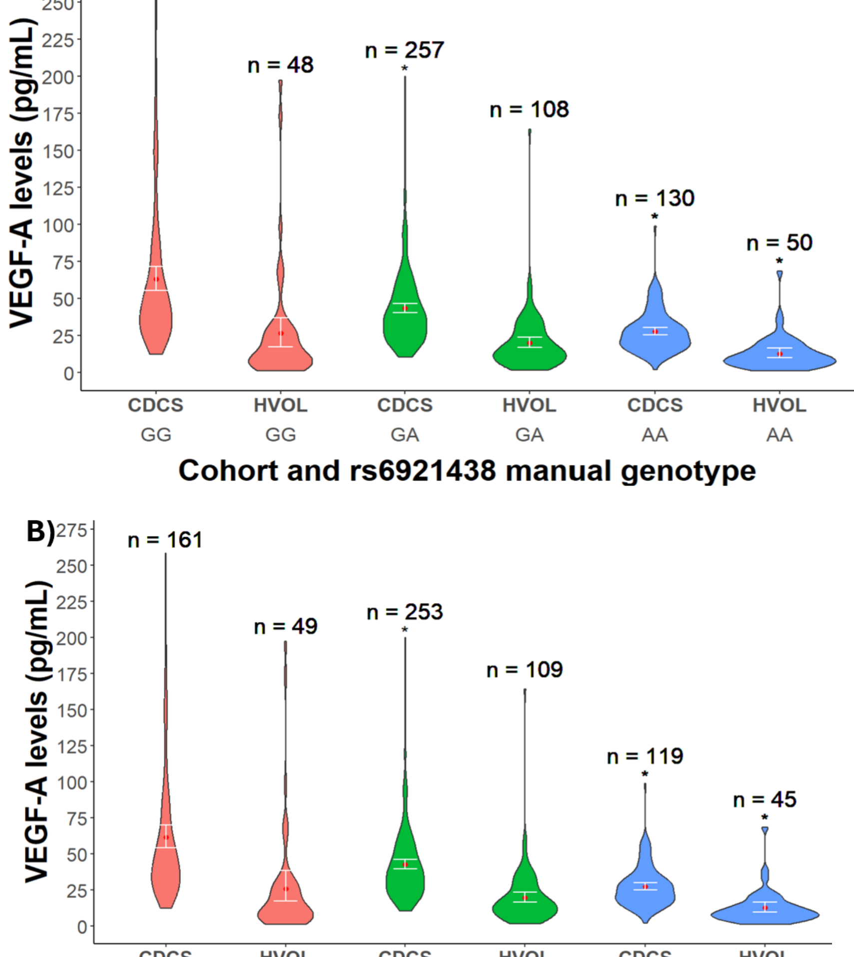


Figure 4.1. Comparison of mean VEGF-A levels for the CDCS and HVOL cohorts for manual genotypes of A) rs6921438 and B) rs7767396. Red represents the homozygous reference genotype, green heterozygous genotype, and blue homozygous for the minor allele genotype. Red dot indicates mean VEGF-A levels. Lines represent 95% CI SE error bars. * Indicates genotype groups that presented statistically significant difference when compared to the reference genotype group of the respective cohort ($p<0.05$). Sample size for each genotype group is showed.

Table 4.2.- Manual genotype comparison for a) rs7767396 and b) rs6921438 the CDCS cohort

			a) rs7767396 Genotypes				
Variable	**n**	**AA**	**n**	**AG**	**n**	**GG**	**p**
Anthropometric Variables							
Age (years)	577	65.7 ± 0.505	985	66.9 ± 0.389	465	67.2 ± 0.572	0.073
Activity (Sed, Mild, Mod, Act)	575	19.7%, 11%, 15%, 54.3%	983	21.6%, 13.6%, 13.3%, 51.5%	465	23.7%, 12.7%, 14.4%, 49.2%	0.443
Male Gender	577	435 (75.4%)	985	691 (70.2%)	465	328 (70.5%)	0.069
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	577	87.5%, 5.7%, 6.5%, 0.3%	985	90.4%, 6.1%, 3%, 0.5%	465	92%, 5.6%, 2.4%, 0%	**0.008**
Height (m)	571	1.71 ± 0.004	972	1.69 ± 0.003	460	1.70 ± 0.005	0.092
Weight (kg)	573	80.5 ± 0.672	973	79.3 ± 0.546	459	79.7 ± 0.808	0.428
BMI (kg/m^2)	571	27.6 ± 0.211	969	27.5 ± 0.161	459	27.4 ± 0.228	0.818
Systolic BP (mmHg)	570	126.9 ± 0.868	963	129.7 ± 0.712	453	129.7 ± 1.04	**0.031**
Diastolic BP (mmHg)	570	74.8 ± 0.488	963	74.9 ± 0.391	453	74.3 ± 0.541	0.696
Beta blocker use	577	497 (86.1%)	985	835 (84.8%)	465	411 (88.4%)	0.179
Previous MI	570	175 (30.7%)	982	275 (28%)	461	147 (31.9%)	0.261
Tobacco use (Current, Ex Smoker, Non-smoker)	577	6.9%, 57.7%, 35.4%	985	6%, 57.5%, 36.5%	465	6.2%, 56.8%, 37%	0.941
Analytes							
Cholesterol (mmol/L)	454	4.94 ± 0.056	774	4.89 ± 0.042	364	4.84 ± 0.062	0.462
Creatinine (mmol/L)	559	99.8 ± 2.06	958	97.9 ± 1.59	451	101.6 ± 2.64	0.434
Urate (mmol/L)	330	0.369 ± 0.006	583	0.372 ± 0.004	270	0.377 ± 0.006	0.646
Troponin I (ng/L) [$]	549	35.5 (23.5 – 47.4)	944	49.7 (20.6 – 78.9)	441	32.5 (20.2 – 44.8)	0.517
ANP (pg/mL) [$]	575	41.7 (39.2 – 44.1)	980	43.6 (41.6 – 45.6)	460	45.1 (42.3 – 47.9)	0.210
NT-ANP (pmol/L) [$]	575	1.29 (1.21 – 1.38)	979	1.33 (1.25 – 1.39)	460	1.43 (1.31 – 1.55)	0.205
BNP (pmol/L) [$]	575	24.1 (21.7 – 26.5)	980	26.7 (24.5 – 28.8)	460	28.5 (25.04 – 31.9)	0.157
NTproBNP (pg/ml) [$]	575	119.3 (108 – 130.7)	980	136.9 (124.9 – 149.1)	460	149.4 (129.3 – 169.4)	0.111
CNP (pmol/L) [$]	568	0.641 (0.605 – 0.678)	964	0.668 (0.639 – 0.697)	453	0.677 (0.629 – 0.725)	0.505
NT-CNP (pmol/L) [$]	567	23.4 (21.5 – 25.2)	962	22.5 (20.9 – 24.2)	453	23.3 (21.2 – 25.3)	0.223
Aldosterone (pmol/L) [$]	562	170.2 (145.2 – 195.3)	956	165.4 (158.4 – 172.5)	445	161.1 (150.9 – 171.2)	0.598
Endothelin(pmol/L) [$]	575	2.66 (2.58 – 2.74)	980	2.66 (2.59 – 2.73)	460	2.74 (2.65 – 2.83)	0.252
Adrenomedullin (pg/ml) [$]	558	8.56 (8.09 – 9.03)	942	8.48 (8.19 – 8.77)	450	8.68 (8.18 – 9.19)	0.693
VEGF-A (pg/mL) [$]	161	61.5 (53.9 – 69.1)	253	42.8 (39.6 – 46.1)	119	27.4 (24.9 – 29.9)	**3.47×10^{-21}**

Variable	n	b) rs6921438 genotypes GG	n	GA	n	AA	p
Anthropometric Variables							
Age (years)	529	65.7 ± 0.526	983	67.1 ± 0.387	514	66.8 ± 0.55	0.090
Activity (Sed, Mild, Mod, Act)	527	20.3%, 10.2%, 15.4%, 54.1%	981	21.1%, 13.5%, 13.3%, 52.1%	514	23.5%, 13.8%, 14.4%, 48.3%	0.253
Male Gender	529	400 (75.6%)	983	689 (70.1%)	514	364 (70.8%)	0.066
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	529	89.4%, 4 %, 6.2%, 0.4%	983	90.2%, 6.6%, 2.8%, 0.4%	514	90.3%, 6.4%, 3.1%, 0.2%	**0.013**
Height (m)	523	1.71 ± 0.004	972	1.69 ± 0.003	507	1.70 ± 0.004	0.064
Weight (kg)	525	80.5 ± 0.703	973	79.2 ± 0.538	506	80.0 ± 0786	0.320
BMI (kg/m^2)	523	27.6 ± 0.215	969	27.5 ± 0.159	506	27.6 ± 0.228	0.865
Systolic BP (mmHg)	522	126.4 ± 0.898	962	129.7 ± 0.707	501	129.9 ± 1.004	**0.010**
Diastolic BP (mmHg)	522	74.4 ± 0.502	962	75.1 ± 0.386	501	74.4 ± 0.526	0.465
Beta blocker use	529	455 (86%)	983	840 (85.5%)	514	449 (87.4%)	0.600
Previous MI	522	158 (30.3%)	980	287 (29.3%)	510	152 (29.8%)	0.922
Tobacco use (Current, Ex Smoker, Non-smoker)	529	7.4%, 57.6%, 35%	983	5.9%, 57.6%, 36.5%	514	6%, 56.8%, 37.2%	0.793
Analytes							
Cholesterol (mmol/L)	419	4.96 ± 0.058	767	4.87 ± 0.042	404	4.86 ± 0.059	0.363
Creatinine (mmol/L)	514	98.7 ± 2.003	956	97.9 ± 1.58	497	102.8 ± 2.64	0.214
Urate (mmol/L)	304	0.372 ± 0.006	585	0.37 ± 0.004	295	0.377 ± 0.006	0.610
Troponin I (ng/L) $	504	31.5 (22.3 – 40.7)	943	52.4 (22.8 – 81.9)	486	31.9 (20.7 – 43.2)	0.786
ANP (pg/mL) $	527	41.4 (38.8 – 43.9)	979	43.9 (41.9 – 45.9)	508	44.6 (41.9 – 47.2)	0.269
NT-ANP (pmol/L) $	527	1.28 (1.19 – 1.37)	978	1.34 (1.27 – 1.41)	508	1.41 (1.3 – 1.52)	0.293
BNP (pmol/L) $	527	23.5 (21.01 – 26.04)	979	26.8 (24.7 – 29.02)	508	28.4 (25.2 – 31.6)	0.074
NTproBNP (pg/ml) $	527	117.1 (105.3 – 128.8)	979	138 (125.9 – 150.1)	508	147.7 (129.1 – 166.2)	0.087
CNP (pmol/L) $	520	0.639 (0.600 – 0.678)	965	0.670 (0.641 – 0.699)	499	0.676 (0.631 – 0.720)	0.535
NT-CNP (pmol/L) $	519	22.7 (21.4 – 23.9)	963	22.5 (20.9 – 24.2)	499	24.03 (21.5 – 26.5)	0.246
Aldosterone (pmol/L) $	515	168.9 (141.7 – 196.1)	955	166.8 (159.7 – 173.9)	492	160.9 (151.3 – 170.5)	0.207
Endothelin (pmol/L) $	527	2.66 (2.58 – 2.74)	979	2.65 (2.59 – 2.72)	508	2.74 (2.66 – 2.83)	0.153
Adrenomedullin (pg/ml) $	510	8.68 (8.17 – 9.19)	944	8.48 (8.19 – 8.77)	495	8.56 (8.09 – 9.03)	0.970
VEGF-A (pg/mL) $	148	63.2 (55.1 – 71.3)	257	43.3 (40.1 – 46.6)	130	27.7 (25.3 – 30.2)	**4.49×10^{-22}**

$Log10 transformed p-values are reported. Mean ± standard error or Mean (95% CI range) or incidence (%) are reported. Significantly associated variables and their p-values are shown in **bold.** Abbreviations: Act: active (≥30 minutes on ≥3 days/week), ANP: atrial natriuretic peptide, BP: blood pressure, BMI: body mass index, BNP: B-type natriuretic peptide, CNP: C-type natriuretic peptide, MELAA: Middle Eastern/Latin American/African, MI: Myocardial infarction, Mod: moderate (≥30 minutes on 2 days/week), NT-ANP Amino terminal atrial natriuretic peptide, NT-CNP: Amino terminal C-type natriuretic peptide NTproBNP = amino-terminal pro-B type natriuretic peptide, Sed: Sedentary, VEGF-A: Vascular endothelial growth factor A.

4.3.4 Univariate survival association of selected SNPs in the CDCS cohort

Imputed rs6921438 and rs7767396 genotype data (Figure 4.2) were associated with risk of all-cause death. Those homozygous for the minor allele genotypes (rs6921438 AA and rs7767396 GG) were more likely to die over the follow up period (8 years). When repeating the analysis using manual genotyping data, patients carrying rs6921438 AA were more likely to die during the follow up period (Figure 4.3A), had higher likelihood of having a readmission for NSTEMI (Figure 4.3B), HF (Figure 4.3C) or MACE (Figure 4.3D). Patients with the rs7767396 GG genotype were more often readmitted for NSTEMI (Figure 4.4A) or HF (Figure 4.4B), but this genotype was not significantly associated with all-cause death or MACE.

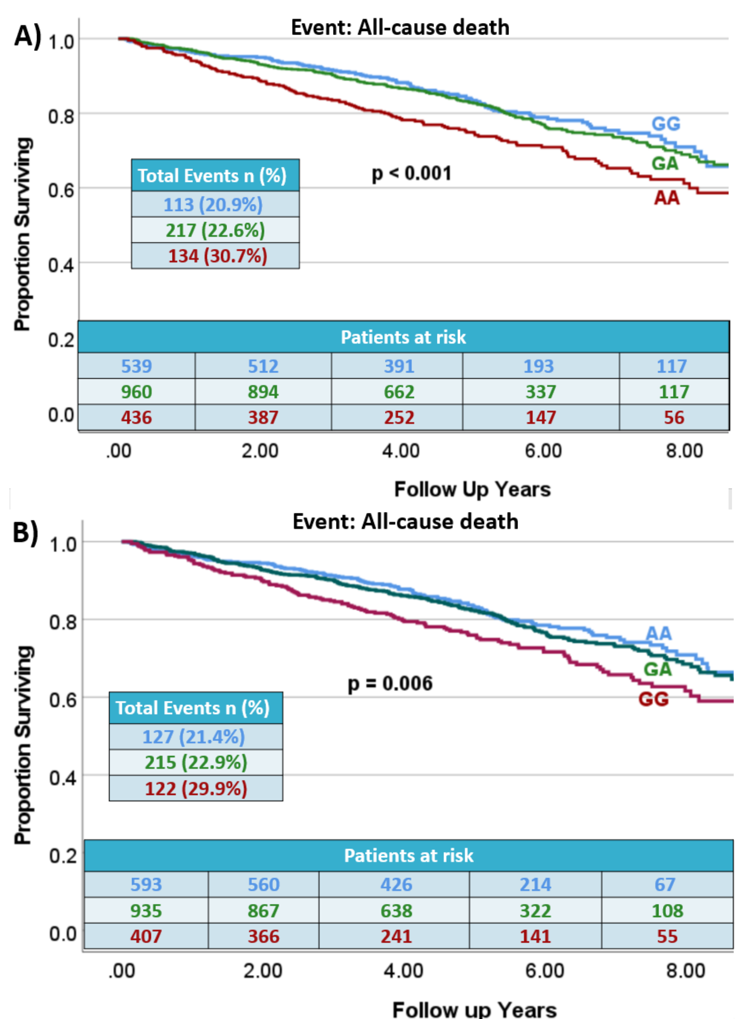


Figure 4.2. Kaplan-Meier survival plot for imputed genotypes of rs6921438 and rs7767396 in the CDCS cohort. Survival versus all-cause death stratified by genotypes of A) rs6921438 and B) rs7767396. Genotypes are colour coded blue for homozygous reference, green for heterozygous and red for homozygous for the minor allele. Patients are risk reported for every 2-year interval.

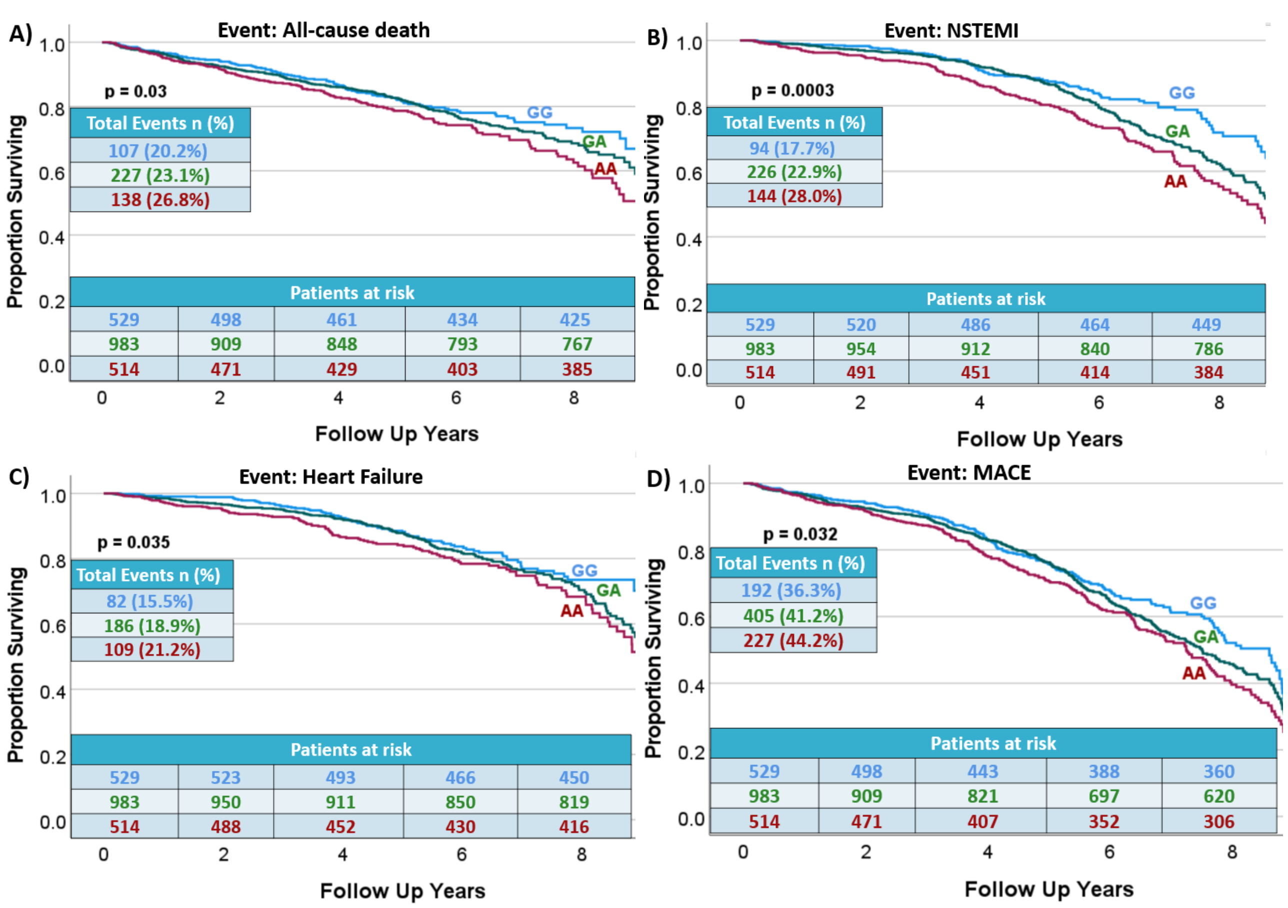


Figure 4.3. Kaplan-Meier survival plots of rs6921438 manual genotypes. Survival versus A) all-cause death, B) Non-ST-elevation myocardial infarction (NSTEMI) C) Heart Failure and D) Major adverse cardiovascular events (MACE) within the CDCS cohort. Genotypes are colour coded blue for GG, green for GA and red for AA.

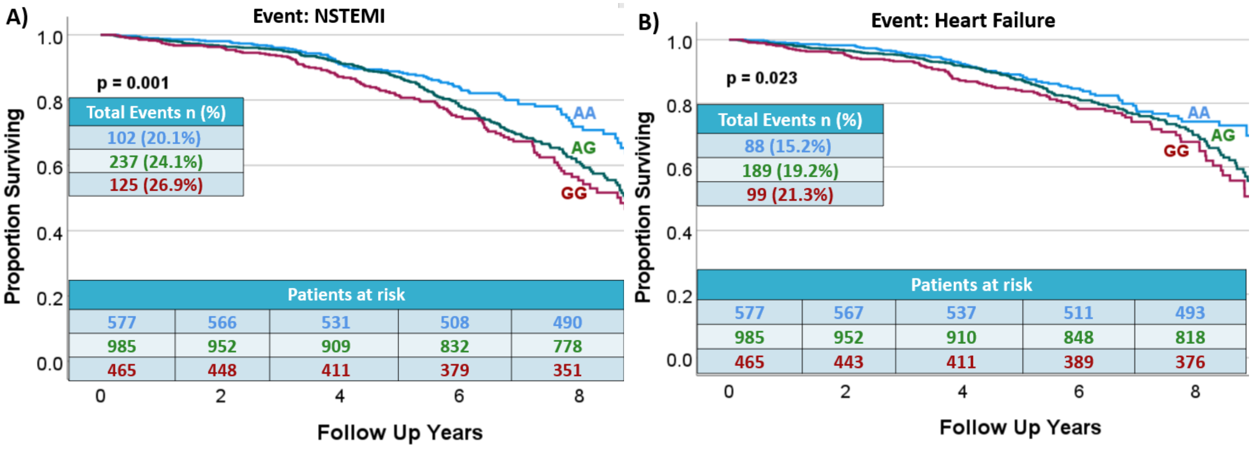


Figure 4.4. Kaplan-Meier survival plots of rs7767396 manual genotypes. Survival versus A) Non-ST-elevation myocardial infarction (NSTEMI) and B) Heart Failure within the CDCS cohort. Genotypes are colour coded blue for AA, green for AG and red for GG.

4.3.5 Multivariate survival association of selected SNPs in the CDCS cohort

Cox proportional hazards models for all-cause mortality showed imputed rs6921438 and rs7767396 genotypes were associated with increased risk (Table 4.3). When using manual genotype data in a similar model, neither SNP was significantly associated with increased risk of mortality (Table 4.4). We observed that baseline plasma VEGF-A was an independent predictor of all-cause mortality (Table 4.4, HR = 2.58, p = 0.003). A simplified model including NTproBNP, plasma VEGF-A and both variants showed VEGF-A (Table 4.5, HR = 2.22, p – 0.009) and NTproBNP (Table 4.5, HR = 6.62, p = 4.3 x 10^{-23}) were significantly associated with increased risk of mortality.

Table 4.3. Cox proportional hazards regression model for all-cause mortality in the CDCS cohort using imputed genotypes for rs6921438 and rs7767396. (n = 499, n = 161 (32.3%) events)

Predictor	Coeff.	SE	Wald	P-value	HR	95% CI for HR	
						Lower	Upper
Gender	0.073	0.193	0.143	0.705	1.08	0.737	1.57
Ethnicity			1.28	0.735			
European v Pasifika	-0.523	1.08	0.233	0.629	0.593	0.071	4.94
European v Asian	1.04	1.03	1.03	0.311	2.83	0.377	21.2
European v MELAA	-6.22	213.3	0.001	0.977	0.002	4.8 x10^{-185}	8.1 x10^{178}
***Physical Activity $^{\$\$}$**	-0.319	0.070	21.01	***4.5 x 10^{-6}**	0.727	0.634	0.833
***Previous MI**	0.571	0.169	11.4	***7.1 x10^{-4}**	1.77	1.27	2.46
***Age**	0.067	0.011	33.9	***5.6 x10^{-9}**	1.07	1.04	1.09
***Body mass index**	-0.052	0.022	5.75	***0.016**	0.949	0.910	0.991
Urate	1.56	0.850	3.38	0.066	4.78	0.903	25.2
***Creatinine**	0.005	0.001	10.4	***0.001**	1.005	1.002	1.007
Beta blocker	-0.180	0.229	0.615	0.433	0.835	0.533	1.31
***Log10 NTproBNP$^{\$}$**	0.867	0.257	11.4	***0.001**	2.37	1.43	3.93
***Log10 VEGF-A$^{\$}$**	0.906	0.298	9.25	***0.002**	2.47	1.38	4.43
***rs6921438 genotype**			9.87	***0.007**			
GG v GA	0.200	0.488	0.167	0.682	1.22	0.470	3.17
***GG v AA**	1.64	0.647	6.46	***0.011**	5.17	1.45	18.3
***rs7767396 genotype**			7.64	***0.022**			
AA v AG	-0.312	0.473	0.436	0.509	0.732	0.290	1.84
***AA v GG**	-1.58	0.649	5.93	***0.015**	0.206	0.058	0.735
Time to sampling	0.005	0.008	0.358	0.549	1.005	0.989	1.02

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or VEGF-A level.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Table 4.4. Cox's proportional hazards regression model for all-cause mortality in the CDCS cohort using manual genotypes for rs6921438 and rs7767396 (n = 506, n= 163 (32.2%) events)

Predictor	Coeff.	SE	Wald	P-value	HR	95% CI for HR	
						Lower	Upper
Gender	0.101	0.186	0.297	0.586	1.11	0.769	1.59
Ethnicity			1.14	0.766			
European v Pasifika	-0.500	1.07	0.220	0.639	0.607	0.075	4.91
European v Asian	0.978	1.03	0.910	0.340	2.66	0.357	19.8
European v MELAA	-6.18	201.2	0.001	0.976	0.002	1.24×10^{-174}	3.47×10^{168}
***Physical Activity [$$]**	-0.327	0.069	22.4	**$*2.2 \times 10^{-6}$**	0.721	0.630	0.826
***Previous MI**	0.516	0.166	9.64	***0.002**	1.68	1.21	2.32
***Age**	0.066	0.011	33.9	**$*5.6 \times 10^{-9}$**	1.07	1.05	1.09
***Body mass index**	-0.054	0.022	5.85	***0.016**	0.948	0.907	0.990
Urate	1.21	0.833	2.11	0.146	3.35	0.655	17.2
***Creatinine**	0.004	0.001	8.76	***0.003**	1.004	1.001	1.007
Beta blocker	-0.116	0.232	0.251	0.616	0.890	0.564	1.41
***Log10 NTproBNP[$]**	0.829	0.254	10.6	***0.001**	2.29	1.39	3.77
***Log10 VEGF-A[$]**	0.948	0.323	8.60	***0.003**	2.58	1.37	4.86
rs6921438 genotype			2.96	0.228			
GG v GA	0.088	0.505	0.030	0.862	1.09	0.406	2.94
GG v AA	0.860	0.648	1.76	0.185	2.36	0.663	8.42
rs7767396 genotype			2.27	0.321			
AA v AG	0.002	0.492	0.000	0.998	1.002	0.382	2.63
AA v GG	-0.701	0.646	1.17	0.278	0.496	0.140	1.76
Time to sampling	0.004	0.008	0.292	0.589	1.004	0.989	1.02

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or VEGF-A level.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Table 4.5. Simplified Cox's proportional hazards regression model for all-cause mortality in the CDCS cohort using manual genotypes for rs6921438 and rs7767396 (n = 532, n= 175 (32.9%) events)

Predictor	Coeff.	SE	Wald	P-value	HR	95% CI for HR	
						Lower	Upper
***Log10 NTproBNP[$]**	1.88	0.191	97.9	**$*4.3 \times 10^{-23}$**	6.62	4.55	9.62
***Log10 VEGF-A[$]**	0.798	0.306	6.81	***0.009**	2.22	1.22	4.05
rs6921438 genotype			3.33	0.189			
GG v GA	0.493	0.490	1.02	0.314	1.64	0.627	4.28
GG v AA	1.08	0.598	3.28	0.070	2.96	0.915	9.55
rs7767396 genotype			2.33	0.311			
AA v AG	-0.286	0.470	0.370	0.543	0.751	0.299	1.89
AA v GG	-0.883	0.596	2.19	0.139	0.414	0.129	1.33
Time to sampling	0.002	0.007	0.081	0.776	1.002	0.988	1.02

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or VEGF-A level.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

An additional model for heart failure readmission had rs6921438 AA associated with increased readmission risk (Table 4.6. HR = 6.60 p = 0.016) while rs7767396 GG associated with reduced readmission risk (Table 4.6, HR= 0.149, p = 0.017). We observed that elevated baseline plasma VEGF-A was an independent predictor of heart failure readmission (Table 4.6, HR = 2.64, p = 0.014). Additional multivariate models, without VEGF-A as a covariate, showed rs6921438 and rs7767396 were not independent predictors of clinical outcome (Appendices 4.4 to 4.7)

Table 4.6. Cox's proportional hazards regression model for heart failure readmissions in the CDCS cohort including manual genotypes for rs6921438 and rs7767396 (n = 462, 109 (23.6%) events)

Predictor	Coeff.	SE	Wald	P-value	HR	95% CI for HR	
						Lower	Upper
Gender	0.269	0.240	1.25	0.263	1.31	0.817	2.10
Ethnicity			6.25	0.100			
European v Pasifika	1.09	0.684	2.52	0.113	2.96	0.775	11.3
European v Asian	2.15	1.09	3.84	0.050	8.55	1.000	73.05
European v MELAA	-4.71	202.2	0.001	0.981	0.009	6.3×10^{-175}	1.3×10^{170}
***Physical Activity**	-0.396	0.086	21.3	**$*3.9 \times 10^{-6}$**	0.673	0.569	0.796
***Previous MI**	0.677	0.203	11.1	***0.001**	1.97	1.32	2.93
***Age**	0.062	0.014	19.8	**$*8.5 \times 10^{-6}$**	1.06	1.035	1.09
Body mass index	-0.013	0.030	0.187	0.666	0.987	0.930	1.05
***Urate**	2.96	1.08	7.47	***0.006**	19.3	2.31	161.4
Creatinine	0.002	0.002	1.23	0.267	1.002	0.998	1.006
Beta blocker	-0.573	0.276	4.31	***0.038**	0.564	0.328	0.969
***Log10 NTproBNP$**	1.73	0.357	23.5	**$*1.2 \times 10^{-6}$**	5.66	2.81	11.4
***Log10 VEGF-A$**	0.972	0.397	5.99	***0.014**	2.64	1.21	5.79
***rs6921438 genotype**			7.55	***0.023**			
GG v GA	0.589	0.678	0.756	0.385	1.803	0.477	6.81
***GG v AA**	1.89	0.787	5.75	***0.016**	6.60	1.41	30.8
***rs7767396 genotype**			7.78	***0.020**			
AA v AG	-0.448	0.655	0.469	0.493	0.639	0.177	2.30
***AA v GG**	-1.901	0.797	5.69	***0.017**	0.149	0.031	0.713
***Time to sampling**	0.021	0.010	4.60	***0.032**	1.02	1.002	1.04
LVEF at baseline	0.000	0.008	0.002	0.963	1.000	0.985	1.02

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or VEGF-A level.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF = left ventricular ejection fraction, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

4.3.6 Genomic context for rs6921438 and rs7767396

Both rs6921438 and rs7767396 are located downstream from *VEGFA* (Figure 4.5). The SNPs are approximately 1.4 kb from each other between long noncoding RNAs *LINC0512* and *C6orf23* (Figure 4.5), approximately 170 kb from the *VEGFA* locus. Two motifs that overlap with rs6921438 are binding sites for transcription factors FOXF2 and FOXH3. FOXF2 has the conserved function of participating in mural cell development and, according to the HaploReg database, the A allele of rs6921438 has increased FOXF2 binding activity compared to the reference G allele [33, 34]. Additionally, this variant is in a H3K4me2 site that allows access to a nearby promoter region, *LINC0512* being the potential target, but this was observed only in the liver [34]. Comparatively, rs7767396 does not overlap with an epigenetic region, but it can affect a binding motif for STAT3, a cardiac transcription factor [35]. The G allele can also affect the binding of NF-AT1 which is linked to lower VEGF-A levels [34, 36].

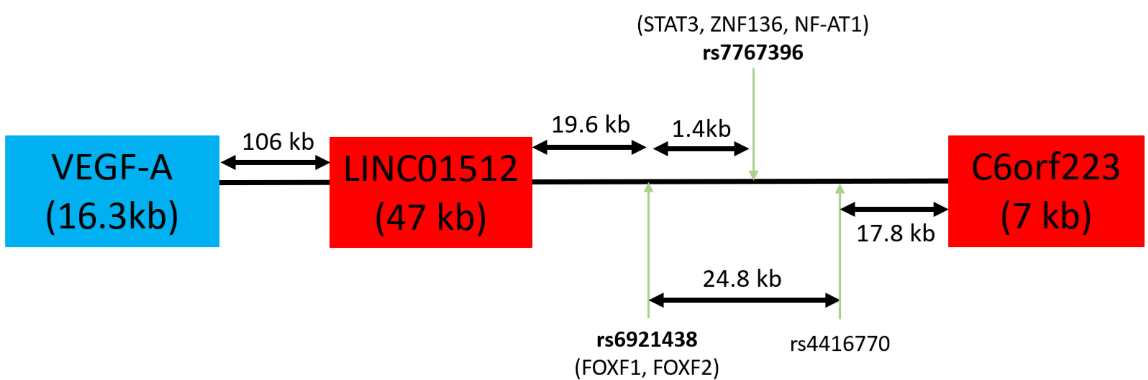


Figure 4.5. Schematic representation of *VEGFA* locus SNPs. SNPs analysed in the present study are shown in bold. Red boxes indicate lncRNAs. Blue box - coding gene for Vascular Endothelial Growth factor A. Green arrows point to individual SNP locations. Transcription Factor binding sites overlapping analysed SNPs are named. White arrow indicates SNPs with epistatic interaction described by Azimi-Nezhad et al [37]

4.4 Discussion

Lower plasma VEGF-A levels were univariately associated with rs6921438 AA and rs7767396 GG in patients diagnosed with ACS. In univariate survival analyses both SNPs were associated with clinical outcomes. rs6921438 AA was associated with increased mortality risk, NSTEMI, HF and MACE readmission risk (Figure 4.3). rs7767396 GG was associated with increased NSTEMI and HF readmission risk (Figure 4.4). In multivariate analysis higher NTproBNP and VEGF-A levels were independent

markers for increased risk of all-cause mortality and HF readmission (Tables 4.6 and 4.7). rs6921438 AA was associated with higher HF readmissions while rs7767396 GG was associated with lower risk (Table 4.6).

On comparing VEGF-A levels by genotype, we observed a trend of mean VEGF-A levels progressively decreasing with the addition of minor alleles for rs6921438 and rs7767396 in both cohorts. GWAS analyses using healthy individuals have observed that rs6921438 A is associated with lower serum VEGF-A levels [26, 27] and this variant can explain 41.2% of VEGF-A serum level variance [26]. A similar trend on the reduction of plasma VEGF levels was noted for rs7767396's minor allele in two oncological cohorts [36]. The impacts of specific alleles have been noted for other VEGF-A related SNPs on chromosome 6 (rs2010963, rs4416670, rs1740073, rs699947) whose major alleles have been associated with increased VEGF serum levels [26, 27, 38, 39]. Our results support lower VEGF-A plasma levels are associated with minor alleles of rs6921438 and rs7767396.

We observed an association between rs6921438 AA and rs7767396 GG (homozygous minor allele genotype) imputed data and mortality (Figure 4.2 and Table 4.3). Univariate survival analysis using manual genotyping data confirmed rs6921438 and rs7767396 were associated with 4 outcomes (death, NSTEMI, HF, MACE) or 2 outcomes (HF, MACE), respectively (Figures 4.3 and 4.4). A study on CAD patients identified that within a 5 year follow up period rs2010963 CC/GC genotype, located within exon 1 of the *VEGFA* gene, was associated with increased risk of CAD-related death [40]. The same study showed the minor allele "C" could be a high-risk allele for cardiac death. Other studies report the minor allele genotypes of *VEGFA* locus SNPs have been associated with increased stroke risk (rs699947 and rs3025039) [41] and coronary heart disease (rs699947 and rs1570360) [42, 43]. The current work used imputed genotype data to select SNP candidates for risk and explore further associations, prior to manual genotyping. Furthermore, assessment of multiple readmission risks allows us to establish the utility of rs6921438 and rs7767396 minor alleles in risk stratification. Our results, agree with studies on the influence of minor alleles of VEGF-A related SNPs on CVD related clinical outcomes.

To expand our survival analyses, multivariate regression models were generated for multiple clinical endpoints. These included manual genotypes of rs6921438 and rs7767396, established CVD risk predictors (age, physical activity, MI history, urate, creatinine, beta blocker usage and NTproBNP) and

VEGF-A levels. While neither variant was an independent predictor of death, we observed higher VEGF-A and NTproBNP were associated with increased mortality risk. In a HF multivariate model, VEGF-A and NTproBNP also behaved as independent predictors. These findings agree with current knowledge on NTproBNP's role as a cardiac biomarker [44]. Moreover. rs6921438 AA was associated with increased HF readmission risk while rs7767396 GG was a predictor for reduced HF readmission risk. A previous study identified a *VEGFA* promoter SNP (rs699947) as an independent predictor of mortality in male non-diabetic participants in the CDCS cohort [4]. The difference between analyses using imputed and manual genotyping data can be attributed to the statistical variation in genotype imputation (i.e. similar genotype frequency distribution but individuals may shift between genotype groups). However, the use of imputed genotyping data proved useful to identify SNPs likely to be disease relevant variants that were further explored. Furthermore, this work contributes by exploring associations at baseline and in the lead up to clinically relevant outcomes. Univariate results suggest rs6921438 AA and rs7767396 GG may be potential risk stratification markers for death, but they may impact other CVD onset mechanisms. The multivariate model for death including manual genotyping data supports that the minor alleles are not associated with elevated VEGF-A levels in our cohorts, but increased plasma VEGF-A has been shown to contribute to CVD risk in wider studies [14-17].

The multivariate model with HF as the endpoint showed the homozygous minor allele genotypes were associated with increased VEGF-A levels, but in a different risk effect direction for each SNP. In this scenario, their independent behaviour may be influenced by each minor allele interacting with VEGF-A, other metabolites, or different CVD risk molecular pathways. For instance, the rs6921438 A allele has been associated with increased inflammatory cytokine (TNF-α and IL-6) levels in a healthy population [37]. Inflammation is key in increasing VEGF-A expression and activity [22, 45, 46]. Additionally, lower plasma VEGF-A levels have been linked with rs6921438 A allele, which is also associated with decreased HDL and increased LDL levels in healthy European ancestry cohorts [47]. Notably, VEGF-A can reduce lipoprotein lipase activity, resulting in higher circulating LDL levels that confer CVD risk [18, 28]. Univariate findings correlate rs6921438 AA with lower baseline VEGF-A levels, but in the lead up to an HF event the AA group appears to confer increased HF risk while being linked to higher VEGF-A. It is possible that the changes of direction of AA's effect that we observed between survival analyses may be attributed to HF onset features (cardiac stress, hypoxia, and inflammation). These factors may have a more pronounced effect over VEGF-A prior to a clinical event which could alter rs6921438 AA's influence over

baseline VEGF-A. This suggests that over time rs6921438 AA may contribute to the pathophysiological microenvironment associated with HF onset.

For its part, rs7767396 AA is shown to be associated with higher VEGF-A levels than the AG and GG genotypes in cancer patients [36]. Additionally, rs7767396 G allele is associated with reduced NF-AT1 binding which is a transcription factor involved in heart development, heart failure and inflammatory pathways [36, 48]. No research, however, has explored this variant's involvement in CVD onset. This variant is located in a binding motif for STAT3 (Figure 4.5) which is a transcription factor that upregulates hypoxia responsive factor 1 alpha (HIF-1α), a well-established VEGF-A production inducer [35]. Our univariate survival data agrees that rs7767396 AA is associated with higher VEGF-A levels. When progressing towards HF onset rs7767396 GG was independently associated with higher VEGF-A levels, while also conferring reduced risk of disease. It is plausible that the reduced HF risk observed in our multivariate model is due to reduced NF-AT1 binding, an effect originally observed in cancer patients [36]. There are mechanisms by which rs6921438 and rs7767396 influence VEGF-A levels that differ according to the phenotype of a specific clinical outcome. It is possible that the minor alleles behave in one way at baseline, while in CVD onset, they may interact with known VEGF-A inducers (e.g. hypoxia, cytokines) to increase VEGF-A levels, thus contributing to the interaction loop between hypoxia, inflammation, and lipid metabolism pathways. Overall, rs7767396 G and rs6921438 A appear to be risk alleles of interest when HF onset biomarkers go above a certain threshold.

4.5 Conclusion

In summary we report an association between rs6921438 and rs7767396 genotypes with VEGF-A plasma levels in post-ACS patients and age- and gender-matched heart healthy controls. Our findings show the homozygous minor allele genotypes of both SNPs and VEGF-A levels may be associated with higher risk of heart failure readmissions. Further analysis of these variables in other longitudinal study cohorts may provide validation of the findings described in this chapter.

4.6 References

1. Wilson, N., et al., *Prioritisation of intervention domains to prevent cardiovascular disease: a country-level case study using global burden of disease and local data.* Population Health Metrics, 2023. **21**(1): p. 1.
2. Roberts, R., C.C. Chang, and T. Hadley, *Genetic Risk Stratification: A Paradigm Shift in Prevention of Coronary Artery Disease.* Journal of the American College of Cardiology: Basic to Translational Science, 2021. **6**(3): p. 287-304.
3. Marks, E.C.A., et al., *Plasma levels of soluble VEGF receptor isoforms, circulating pterins and VEGF system SNPs as prognostic biomarkers in patients with acute coronary syndromes.* BMC Cardiovascular Disorders, 2018. **18**(1): p. 169.
4. Palmer, B.R., et al., *Vascular endothelial growth factor-A promoter polymorphisms, circulating VEGF-A and survival in acute coronary syndromes.* PLOS One, 2021. **16**(7): p. 1-15.
5. Zhou, Y., et al., *The Role of the VEGF Family in Coronary Heart Disease.* Frontiers in Cardiovascular Medicine, 2021. **8**: p. 1-16.
6. Camaré, C., et al., *Angiogenesis in the atherosclerotic plaque.* Redox biology, 2017. **12**: p. 18-34.
7. Niklander, S., et al., *Vascular Endothelial Growth Factor: A Translational View in Oral Non-Communicable Diseases.* Biomolecules, 2021. **11**(1): p. 1-27.
8. Shaik, F., et al., *Structural Basis for Vascular Endothelial Growth Factor Receptor Activation and Implications for Disease Therapy.* Biomolecules, 2020. **10**(12): p. 1-25.
9. Oszajca, K., et al., *Single-nucleotide polymorphisms of VEGF-A and VEGFR-2 genes and risk of infantile hemangioma.* International Journal of Dermatology, 2018. **57**(10): p. 1201-1207.
10. Apte, R.S., D.S. Chen, and N. Ferrara, *VEGF in Signaling and Disease: Beyond Discovery and Development.* Cell, 2019. **176**(6): p. 1248-1264.
11. Le, T.H.V. and S.M. Kwon, *Vascular Endothelial Growth Factor Biology, and Its Potential as a Therapeutic Target in Rheumatic Diseases.* International Journal of Molecular Sciences, 2021. **22**(10): p. 1-17.
12. Peach, C.J., et al., *Molecular Pharmacology of VEGF-A Isoforms: Binding and Signalling at VEGFR2.* International Journal of Molecular Sciences, 2018. **19**(4): p. 1264.
13. Smith, G.A., et al., *Vascular endothelial growth factors: multitasking functionality in metabolism, health, and disease.* Journal of Inherited Metabolic Disease, 2015. **38**(4): p. 753-763.
14. Braile, M., et al., *VEGF-A in Cardiomyocytes and Heart Diseases.* International Journal of Molecular Sciences, 2020. **21**(15): p. 1-18.
15. Au Yeung, S.L., H. Lam, and C.M. Schooling, *Vascular Endothelial Growth Factor, and Ischemic Heart Disease Risk: A Mendelian Randomisation Study.* Journal of the American Heart Association, 2017. **6**(8): p. 1-8.
16. Zhang, Y., et al., *Serum vascular endothelial growth factor as a biomarker for prognosis of minor ischemic stroke.* Clinical Neurology and Neurosurgery, 2020. **196**: p. 106060.
17. Rodríguez-Rodríguez, L., et al., *Vascular endothelial growth factor A and cardiovascular disease in rheumatoid arthritis patients.* Tissue Antigens, 2011. **77**(4): p. 291-297.
18. Heinonen, S.E., et al., *The effects of VEGF-A on atherosclerosis, lipoprotein profile, and lipoprotein lipase in hyperlipidaemic mouse models.* Cardiovascular Research, 2013. **99**(4): p. 716-23.
19. Zhao, Z., et al., *Association between Single Nucleotide Polymorphisms in Cardiovascular Developmental Critical Genes and Hypertension: A Propensity Score Matching Analysis.* International Journal of Hypertension, 2020. **2020**: p. 1-8.
20. Ruotsalainen, S.E., et al., *An expanded analysis framework for multivariate GWAS connects inflammatory biomarkers to functional variants and disease.* European Journal of Human Genetics, 2021. **29**(2): p. 309-324.
21. Zhao, X., et al., *Vascular endothelial growth factor gene polymorphisms and coronary heart disease: a systematic review and meta-analysis.* Growth Factors, 2018. **36**(3-4): p. 153-163.

22. Ramakrishnan, S., V. Anand, and S. Roy, *Vascular endothelial growth factor signalling in hypoxia and inflammation.* Journal of Neuroimmune Pharmacology, 2014. **9**(2): p. 142-60.
23. Pagès, G. and J. Pouysségur, *Transcriptional regulation of the Vascular Endothelial Growth Factor gene–a concert of activating factors.* Cardiovascular Research, 2005. **65**(3): p. 564-573.
24. Tischer, E., et al., *The human gene for vascular endothelial growth factor. Multiple protein forms are encoded through alternative exon splicing.* Journal of Biological Chemistry, 1991. **266**(18): p. 11947-11954.
25. Singh, P., et al., *Genetic Scores of eNOS, ACE and VEGFA Genes Are Predictive of Endothelial Dysfunction Associated Osteoporosis in Postmenopausal Women.* International Journal of Environmental Research and Public Health, 2021. **18**(3): p. 1-18.
26. Debette, S., et al., *Identification of cis- and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels.* Circulation Research, 2011. **109**(5): p. 554-63.
27. Choi, S.H., et al., *Six Novel Loci Associated with Circulating VEGF Levels Identified by a Meta-analysis of Genome-Wide Association Studies.* Public Library of Science (PLOS) Genetics, 2016. **12**(2): p. e1005874.
28. Chen, H., et al., *Variation in VEGFA and risk of cardiovascular disease in the UK Biobank.* Frontiers in Cardiovascular Medicine, 2023. **10**: p. 1240288.
29. Scott, X.O., et al., *Cohort study on the differential expression of inflammatory and angiogenic factors in thrombi, cerebral and peripheral plasma following acute large vessel occlusion stroke.* Journal of Cerebral Blood Flow & Metabolism, 2022: p. 271678x221106956.
30. Palmer, B.R., et al., *KCNE5 polymorphism rs697829 is associated with QT interval and survival in acute coronary syndromes patients.* Journal of Cardiovascular Electrophysiology, 2012. **23**(3): p. 319-24.
31. Palmer, B.R., et al., *Angiotensin-converting enzyme 2 A1075G polymorphism is associated with survival in an acute coronary syndromes cohort.* American Heart Journal, 2008. **156**(4): p. 752-758.
32. Ellis, K.L., et al., *Genomic risk variants at 1p13.3, 1q41, and 3q22.3 are associated with subsequent cardiovascular outcomes in healthy controls and in established coronary artery disease.* Circulation Cardiovascular Genetics, 2011. **4**(6): p. 636-46.
33. Ryu, J.R., et al., *Stroke-associated intergenic variants modulate a human FOXF2 transcriptional enhancer.* Proceedings of the National Academy of Sciences of the United States of America (USA), 2022. **119**(35): p. e2121333119.
34. Ward, L.D. and M. Kellis, *HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease.* Nucleic Acids Research, 2016. **44**(D1): p. D877-81.
35. Harhous, Z., et al., *An Update on the Multifaceted Roles of STAT3 in the Heart.* Frontiers in Cardiovascular Medicine, 2019. **6**: p. 150.
36. Innocenti, F., et al., *Genetic variation determines VEGF-A plasma levels in cancer patients.* Scientific Reports, 2018. **8**(1): p. 16332.
37. Azimi-Nezhad, M., et al., *Associations of vascular endothelial growth factor (VEGF) with adhesion and inflammation molecules in a healthy population.* Cytokine, 2013. **61**(2): p. 602-7.
38. Merlo, S., et al., *Vascular Endothelial Growth Factor Gene Polymorphism (rs2010963) and Its Receptor, Kinase Insert Domain-Containing Receptor Gene Polymorphism (rs2071559), and Markers of Carotid Atherosclerosis in Patients with Type 2 Diabetes Mellitus.* Journal of Diabetes Research, 2016. **2016**: p. 1-6.
39. Skrypnik, D., et al., *Association of rs699947 (-2578 C/A) and rs2010963 (-634 G/C) Single Nucleotide Polymorphisms of the VEGF Gene, VEGF-A and Leptin Serum Level, and Cardiovascular Risk in Patients with Excess Body Mass: A Case-Control Study.* Journal of Clinical Medicine, 2020. **9**(2): p. 1-16.
40. Nia, S.K., et al., *The impact of vascular endothelial growth factor+405 C/G polymorphism on long-term outcome and severity of coronary artery disease.* Journal of Clinical Laboratory Analysis, 2017. **31**(4).
41. Yadav, B.K., et al., *Genetic Polymorphisms rs699947, rs1570360, and rs3025039 on the VEGF gene are correlated with extracranial internal carotid artery stenosis and ischemic stroke.* Annals of Clinical and Laboratory Science, 2017. **47**(2): p. 144-155.

42. Li, L., et al., *Association of Genetic Polymorphisms on Vascular Endothelial Growth Factor and its Receptor Genes with Susceptibility to Coronary Heart Disease.* Medical Science Monitor, 2016. **22**: p. 31-40.
43. Liu, D., et al., *Association of Genetic Polymorphisms on VEGFA and VEGFR2 With Risk of Coronary Heart Disease.* Medicine, 2016. **95**(19): p. e3413.
44. O'Donnell, C., et al., *N-Terminal Pro-B-Type Natriuretic Peptide as a Biomarker for the Severity and Outcomes With COVID-19 in a Nationwide Hospitalized Cohort.* Journal of the American Heart Association, 2021. **10**(24): p. e022913.
45. Fahey, E. and S.L. Doyle, *IL-1 Family Cytokine Regulation of Vascular Permeability and Angiogenesis.* Frontiers in Immunology, 2019. **10**: p. 1-15.
46. Tzeng, H.E., et al., *Interleukin-6 induces vascular endothelial growth factor expression and promotes angiogenesis through apoptosis signal-regulating kinase 1 in human osteosarcoma.* Biochemical Pharmacology, 2013. **85**(4): p. 531-40.
47. Stathopoulou, M.G., et al., *A common variant highly associated with plasma VEGFA levels also contributes to the variation of both LDL-C and HDL-C.* Journal of Lipid Research, 2013. **54**(2): p. 535-41.
48. Zhang, P., et al., *The mechanism of the NFAT transcription factor family involved in oxidative stress response.* Journal of Cardiology, 2024. **83**(1): p. 30-36.

Chapter 5.- Study on *VEGFR2* locus SNPs

This chapter focuses on the work involving 13 SNPs mapping to the VEGFR2 gene (human chromosome 4). The analysis of 13 variants addresses project aim A as well as research objectives 2 and 3 (Chapter 1, Section 1.3).

The chapter highlights findings on four variants with clinical implications following the analysis of the 13 imputed genotypes. These findings address project aim B and research objective 5 (Chapter 1, Section 1.3).

STATEMENT OF CONTRIBUTION DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the student and the student's main supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the student's contribution as indicated below in the Statement of Originality.

Student name:	Juan Carlos Meza Alvarado
Name and title of main supervisor:	Dr. Barry Palmer, Senior Lecturer
In which chapter is the manuscript/published work?	Chapter 5

Describe the contribution that the student and members of the supervisory team have made to the manuscript/published work:[1]

- The student took a major role in project conceptualization, statistical analyses on the imputed genotype data, investigation, data curation and analysis, visualization and draft writing.
-Data on CDCS VEGF-A levels was generated by others and described in Palmer et al (2021). Chapter 5 Reference #32
- Manual genotype data for rs1870377 and rs2305948 was generated by collaborators as described in Marks et al (2021). Chapter 5 Reference #29
--Imputed genotype data discussed in this Chapter was generated by Dr. Anna Pilbrow and Dr. Vinicius Tragante

Please select one of the following three options:

○	The manuscript/published work is published or in press Please provide the full reference of the research output:
○	The manuscript is currently under review for publication Please provide the name of the journal:
●	It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Student's signature:	[signature]	Main supervisor's signature:	[signature]

This form should be placed at the beginning of each relevant thesis chapter.

Doctoral Research Committee
May 2023

5.1 Abstract

Background: Vascular endothelial growth factor A (VEGF-A) influences the cardiovascular system by promoting endothelial cell proliferation and regulating vascular permeability. Vascular endothelial receptor 2 (VEGFR2) is the canonical receptor through which VEGF-A induces cellular signals to carry out its effector functions. Increased plasma levels of VEGF-A have been observed in various cardiovascular diseases (CVDs) and reports have associated diverse single nucleotide polymorphisms (SNPs) with increased VEGF-A levels. Specifically, two variants (rs2305948 and rs1870377) within exonic regions of *VEGFR2* influence VEGFR2 and VEGF-A activity. This may also apply to other SNPs located within the *VEGFR2* locus.

Methods: Imputed genotype data was obtained for 13 SNPs on chromosome 4 from 1935 patients from the Coronary Disease Cohort Study (CDCS) and 1183 individuals from the Canterbury Healthy Volunteers Study (HVOL). Imputed genotype data was analysed to assess the association of each genotype group with cardiometabolic parameters using one-way ANOVA tests and with clinical endpoints through Kaplan-Meir analyses and multivariate regression models. To validate the CDCS imputed data findings, previously collected manual genotyping data for rs2305948 and rs1870377 was used. VEGF-A levels for 227 HVOL participants were measured by an ELISA immunoassay to compare with previously reported levels from 549 CDCS patients.

Results: Imputed CDCS data showed rs2305948 was associated ($p < 0.05$) with higher levels of cystatin C, atrial natriuretic peptide (ANP), amino terminal pro-B-type natriuretic peptide (NTproBNP) and aldosterone levels. In the HVOL cohort, imputed data showed that rs2305948 was associated with high-density lipoprotein (HDL) levels ($p = 0.03$), while 7 SNPs (rs1870377, rs1870378, rs1870379, rs7677779, rs13136007, rs10016064) were associated with triglyceride and VEGF-A levels ($p < 0.05$).Imputed data identified individuals carrying one or two minor alleles of rs1870377, rs7677779 and rs13136007 had shorter time to first coronary artery bypass graft readmission risk. Manual genotyping data showed 1) rs1870377 trended towards associations with aldosterone and urate levels 2) rs2305948 CC was associated with higher all-cause mortality risk ($p = 0.045$) and cardiovascular readmission risk ($p = 0.045$) 3) rs1870377 is an independent predictor for cardiovascular death when adjusting for NTproBNP, hypertension, creatinine, and beta blocker treatment (TT vs TA+AA, $p = 0.048$, HR = 1.125).

Conclusion: Variants rs2305948 and rs1870377 represent potential genetic markers to assess CVD or mortality risk within a population of European ancestry.

5.2 Introduction

Cardiovascular diseases (CVDs) are an important contributor to health loss within New Zealand [1]. Preventive diagnosis and interventions are complicated by the lack of predictive biomarkers in patients prior to symptom onset [1, 2]. The vascular endothelial growth factor (VEGF-A), involved in blood vessel formation, could be considered a novel biomarker since increased levels have been observed in aetiologies such as ischaemic heart disease (IHD), coronary artery disease (CAD) heart failure (HF) and myocardial infarction (MI) [3, 4]. VEGF-A can increase inflammatory molecule activity and disrupt lipoprotein lipase leading to increased vascular dilation, adhesion protein expression and higher lipid levels which promote CVD development [3, 5-10].

VEGF-A's main signal inducer is the vascular endothelial growth factor receptor 2 (VEGFR2), also known as a tyrosine kinase receptor (KDR), which participates in vascular endothelial cell mitogenesis, vascular permeability and VEGF-mediated blood vessel growth [11-13]. VEGFR2 consists of seven extracellular immunoglobulin homology domain repeats, a transmembrane and a split tyrosine kinase domain [14]. Once VEGF-A binds to VEGFR2, homodimerisation or heterodimerisation with VEGFR1 occurs which causes intracellular domains to undergo phosphorylation at tyrosine residues to induce downstream intracellular signalling [5, 15]. Moreover, non-VEGF dependent VEGFR2 phosphorylation of intracellular tyrosine residues can occur due to binding of non-VEGF ligands such as low-density lipoproteins (LDLs), lactate, galectin and SRC tyrosine kinases [14-16]. Additional VEGFR2 activity may occur when VEGFR1 interacts with the co-receptor neuropilin-1 (NRP1) to selectively potentiate VEGFR2-mediated vascular permeability and endothelial cell motility [17, 18].

The *VEGFR2* gene is 47.11 kb long, comprised of 30 exons and 29 introns, located at 4q12 on the long arm of human chromosome 4 [19]. The gene is highly expressed in vascular endothelial cells and endothelial progenitor cells, and transcripts can be found in non-endothelial cells including haematopoietic cells [15, 20]. VEGFR2 is linked to survival since *VEGFR2* deficient mice died due to impaired development of haemopoietic and endothelial cells [15]. The major phosphorylation sites involved in VEGFR2 activity are Tyr1054 and Tyr1059 in the kinase domain activation loop although other notable sites include Tyr951, a binding site active in angiogenesis, and Tyr1175, a binding site involved in VEGF-

A mitogenic activity [15]. Mutational analyses revealed that a mutation of Y1175F blocked VEGF-A dependent migration in pathological angiogenesis [15, 21]

Studies have shown that circulating VEGF-A levels can be influenced by gene variants. There is evidence that expression quantitative trait loci (eQTL) are associated with circulating VEGF-A level variability, cardiometabolic risk markers and increased CVD risk [22-24]. There is existing knowledge on exonic variants that affect VEGFR2 function. Specifically, rs1870377 causes a missense mutation (Q472H) which increases VEGFR2 phosphorylation promoting increased VEGF-A binding efficiency [20, 25, 26]. Another example includes rs2305948 (missense V297I mutation), which affects VEGFR2's Ig-like domain dimerisation leading to reduced receptor affinity for VEGF-A [27, 28]. There is potential to analyse intronic and exonic variants mapping to *VEGFR2* to determine if they are associated with VEGF system activity, CVD risk factors and CVD risk in terms of both disease development and progression.

The aim of this study was to identify if 13 SNPs in *VEGFR2* are associated with cardiometabolic variables, including plasma VEGF-A, in post-acute coronary syndrome patients and heart healthy matched controls [29-33] using genome-wide association study (GWAS) imputation data for three exonic and ten intronic SNPs. These SNPs were also analysed for associations with clinical endpoints.

5.3 Results

5.3.1 Baseline characteristics of study cohorts

Baseline characteristics of the CDCS cohort are described in Table 5.1. The cohort is predominantly European (89.6%) and male (71.5%). Additionally, 65.6% of the patients have presented readmission events including HF, stroke, MI, unstable angina, coronary artery bypass graft (CABG), arrhythmia, and percutaneous coronary intervention. Study participants presented discharge medication with antithrombotic therapy being the highest (96.4%) followed by lipid lowering medication (88.3%) and blood pressure lowering medication (85.8%). Frequencies for the 13 imputed SNP genotype groups, their respective alleles and minor allele frequency (MAF) for the CDCS patients (n = 1935) are summarised in Appendix 5.1.

Table 5.1. Baseline characteristics of the CDCS and HVOL cohort

Variable	CDCS		HVOL		p - value
	n	Mean ± SE or n (%)	n	Mean ± SE or n (%)	
Anthropometric					
Age	2140	66.5 ± 0.27	1183	68.1 ± 0.216	3.76×10^{-5}
Male Gender	2140	1531 (71.5%)	1183	801 (67.7%)	1.71×10^{-119}
Weight (kg)	2116	79.9 ± 0.36	1183	77.4 ± 0.42	2.79×10^{-2}
Body Mass Index (kg/m^2)	2110	27.6 ± 0.109	1183	26.2 ± 0.11	8.33×10^{-15}
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	2140	89.6%, 6.2%, 3.9%, 0.3%	1183	98.8%, 1.1%, 0.1%, 0%	
Tobacco use (Current, Ex Smoker, Non-smoker)	2022	6.3%, 57.4%,36.3%	1183	4.8%, 39.6%,55.6%	5.66×10^{-253}
Alcohol (drinker, ex-drinker, non-drinker)	2137	62.8%, 12.1%, 25.1%	1183	80.7%, 4.1%, 15.2%	
Systolic blood pressure (mmHg)	2095	128.9 ± 0.476	1180	137 ± 0.545	1.09×10^{-24}
Diastolic blood pressure (mmHg)	2095	74.9 ± 0.26	1180	80.9 ± 0.306	2.06×10^{-46}
Left ventricular ejection fraction* (%)	1960	57.01 ± 0.27	372	64.6 ± 0.24	3.88×10^{-32}
Plasma analytes					
Creatinine (mmol/L)	2063	99.8 ± 1.17	1146	90.3 ± 11.7	3.64×10^{-9}
Cholesterol (mmol/L)	1681	4.89 ± 0.028	1146	5.65 ± 0.033	5.69×10^{-62}
Urate (mmol/L)	1238	0.37 ± 0.003	1146	0.347 ± 0.0022	1.54×10^{-11}
VEGF-A (pg/mL)	549	45.02 ± 1.50	223	19.5 ± 1.60	5.76×10^{-22}
Atrial natriuretic peptide (pmol/L)	2116	43.5 ± 0.68	1163	21.6 ± 0.325	7.24×10^{-9}
Amino terminal atrial natriuretic peptide (pmol/L)	2115	1.35 ± 0.026	1170	0.673 ± 0.325	4.99×10^{-76}
B-type natriuretic peptide (pmol/L)	2116	26.4 ± 0.74	1167	6.37 ± 0.119	3.08×10^{-84}
Amino terminal pro B-type natriuretic peptide (pmol/L)	2116	135.3 ± 4.07	1174	108 ± 3.98	4.35×10^{-5}
Medication					
Blood pressure lowering	2140	1837 (85.8%)	1183	293 (24.7%)	1.14×10^{-9}
Lipid lowering	2140	1890 (88.3%)	1183	189 (16%)	1.51×10^{-47}
Antithrombotic	2140	2063 (96.4%)	1183	120 (10.1%)	3.59×10^{-73}
Genetic data					
Imputed genotype	2140	1935 (90.4)	1183	1177 (99.5%)	

*Determined from echocardiographic measurements as previously described [31]

Abbreviations: CDCS: Coronary Disease Cohort Study, HVOL: Canterbury Healthy Volunteers Study, MELAA: Middle Eastern/Latin American/African, SE: Standard Error of Mean, VEGF-A: Vascular endothelial growth factor A.

Baseline characteristics for the 1183 individuals of the HVOL cohort subset are described in Table 5.1. The population was mostly European (98.8%) and male (67.7%). Compared to the CDCS, the HVOL cohort presented low levels of blood pressure lowering medication use (24.7%). Additionally, imputed genotype data was available for most of the HVOL subset group (99.5%). The frequencies for the 13

imputed SNP genotype groups, their respective alleles and MAF, for the HVOL subset group (n = 1177) are summarised in Appendix 5.1.

5.3.2 SNP genotype analysis

Imputed genotype data was available for 1935 CDCS patients and 1177 HVOL individuals for 13 SNPs on human chromosome 4 (Appendix 5.1). There was manual genotype data available only for variants rs2305948 and rs1870377 of CDCS patients (Table 5.2). Manual and imputed genotype frequencies for both variants in both cohorts are summarised in Table 5.2. Manual genotyping for rs2305948 and rs1870377 was concordant with imputed genotypes from the CDCS cohort (rs2305948 – 83.2% $r^2 = 0.564$, rs1870377 – 74.6 % $r^2 = 0.65$, $p < 0.001$).

Table 5.2. Frequency comparison of manual genotype vs imputed genotype data

SNP ID	Cohort	Genotype	Imputed frequencies n (%)	Physical Frequencies n (%)
rs2305948	CDCS	CC	1705 (88.1%)	1409 (80.4%)
		CT	222 (11.5%)	328 (18.7%)
		TT	8 (0.4%)	15 (0.9%)
		Total	1935	1752
	HVOL	CC	1037 (88.1%)	-
		CT	238 (11.7%)	-
		TT	2 (0.2%)	-
		Total	1177	-
rs1870377	CDCS	TT	1216 (62.8%)	1074 (56.2%)
		TA	657 (34%)	709 (37.1%)
		AA	62 (3.2%)	129 (6.7%)
		Total	1935	1912
	HVOL	TT	760 (64.6%)	-
		TA	371 (31.5%)	-
		AA	46 (3.9%)	-
		Total	1177	-

- Manual genotyping data not available for HVOL samples
Abbreviations: CDCS: Coronary Disease Cohort Study HVOL: Canterbury Healthy Volunteers Study

5.3.3 SNP associations within the CDCS cohort

A one-way ANOVA was performed to compare the association of the 13 imputed SNP genotypes on analyte measurements of the CDCS cohort. Imputed variants that had associations of $p < 0.1$ with CVD risk markers in the CDCS cohort were rs2305948, rs1870378 and rs1870379.

Trends towards significant associations for minor allele carriers for higher mean aldosterone levels and lower mean adrenomedullin levels compared to their respective reference genotype groups were observed for rs1870378 and rs1870379 (Appendix 5.2 and 5.3). Imputed genotype data for rs2305948 was significantly associated with cystatin C, atrial natriuretic peptide (ANP), amino terminal pro B-type natriuretic peptide (NTproBNP) and aldosterone levels. Data is summarised in Figure 5.1. Specifically, T allele carriers presented higher cystatin C (Figure 5.1A, n = 762, CC: 1.25 mg/dL CT/TT: 1.39 mg/dL, p = 0.03), ANP (Figure 5.1B, n = 1923, CC: 43.51 pg/mL, CT/TT: 46.2 pg/mL, p = 0.038), NTproBNP (Figure 5.1C, n = 1923, CC: 134.4 pg/mL, CT/TT: 148.4 pg/mL, p = 0.025) and aldosterone (Figure 5.1D, n = 1875, CC: 159.5 pmol/L, CC/TT: 217.3 pmol/L, p = 0.004) levels when compared to the rs2305948 CC genotype group. We also observed rs2305948 T allele carriers trended towards higher BNP (n = 1923, CC: 26.3 pmol/L, CT/TT: 29.7 pmol/L, p = 0.054) and creatinine levels (n = 1876, CC: 98.8 mmol/L, CT/TT: 105.5 mmol/L p=0.072) (Appendix 5.4). None of the imputed SNP data gave significant association with VEGF-A levels from the CDCS cohort.

When repeating the one-way ANOVA analysis in the CDCS with manual genotypes for rs2305948 and rs1870377, only rs1870377 A allele carriers trended towards higher aldosterone levels (Figure 5.2A, n = 1856, TT: 164.1.7 pmol/L, TA/AA: 166.4 pmol/L, p = 0.055) and lower urate levels (Figure 5.2B, n = 1174, TT: 0.377 mmol/L, TA/AA: 0.366 mmol/L, p = 0.083).

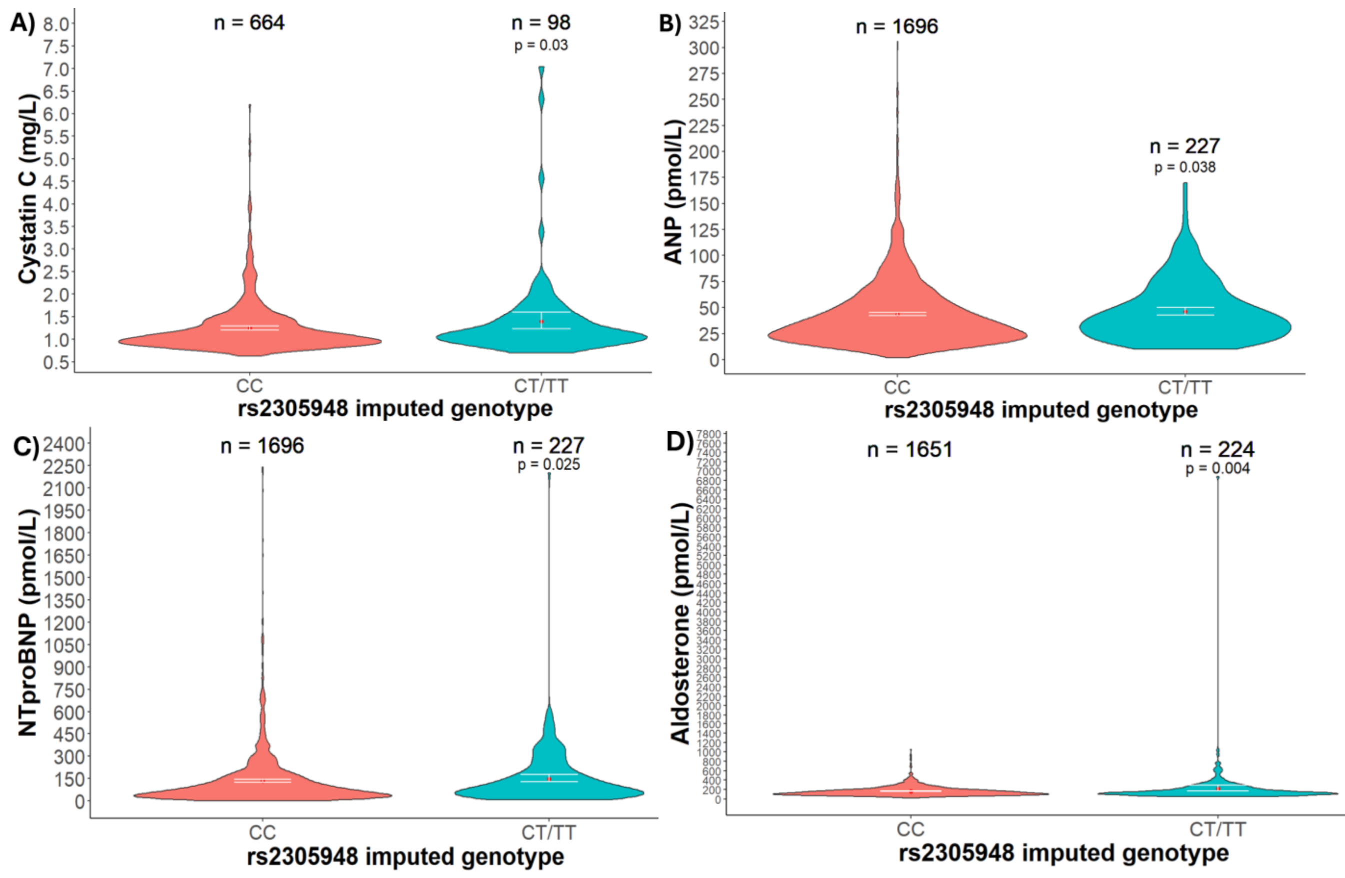


Figure 5.1. Significant associations of rs2305948 imputed genotypes within the CDCS cohort. Measurements of A) Cystatin C B) Atrial Natriuretic Peptide (ANP) C) Amino terminal pro B-type natriuretic peptide (NTproBNP) and D) Aldosterone mean levels. Red dot represents mean values of the variables. White lines represent ± SEM error bars.

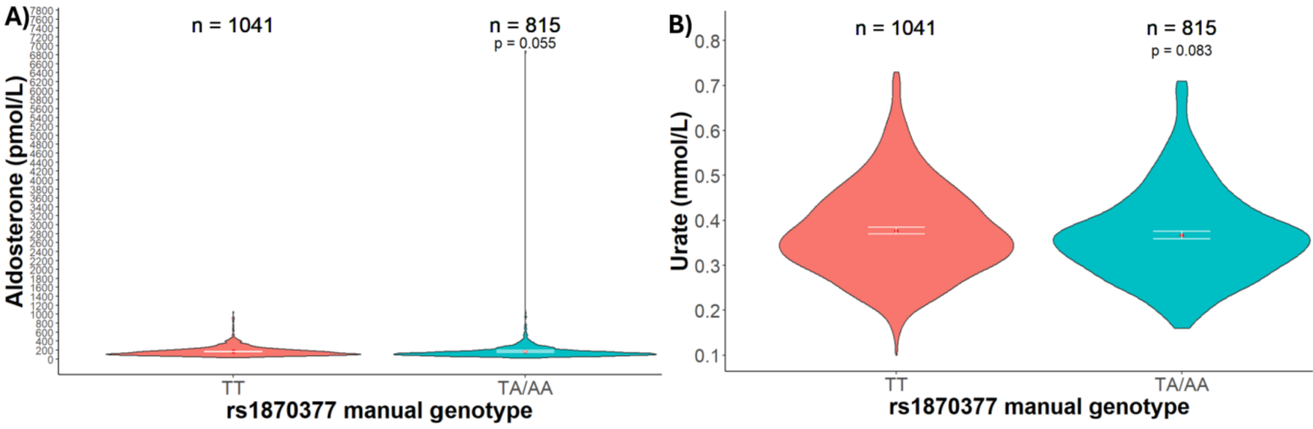


Figure 5.2. Relationship of rs1870377 manual genotypes within the CDCS cohort. Measurements of A) Aldosterone and B) Urate levels within the CDCS cohort. Red dot represents mean values of the variables. White lines represent ± SEM error bars.

5.3.4 SNP associations within the HVOL subset group

In the HVOL cohort, of the 13 imputed genotypes only 7 presented significant associations ($p < 0.05$) with at least one cardiometabolic parameter. Specifically, rs2305948 T allele carriers had higher levels of high-density lipoprotein (HDL) (Figure 5.3A, n = 1139, CC: 1.37 mmol/L, CT/TT: 1.44 mmol/L, p = 0.03) when compared to the reference genotype group. Additionally, the T allele carriers trended towards significantly higher endothelin levels (Figure 5.3B, n = 1146. CC: 2.13 pmol/L, CT/TT: 2.26 pmol/L, p = 0.08).

A group of six SNPs (rs1870377, rs1870378, rs1870379, rs7677779, rs13136007 and rs10016064) were associated with mean VEGF-A (Figure 5.4) and triglyceride (Figure 5.5) levels. Minor allele carriers for each SNP had lower mean levels of both analytes when compared to homozygous reference genotype. The VEGF-A measurements for minor allele carriers ranged from 14.3 pg/mL ± 1.19 SE (rs1870378 and rs1870379) to 15.0 pg/mL ± 1.25 SEM (rs10016064). The reference genotype group VEGF-A measurements ranged from 22.8 pg/mL ± 0.54 SEM (rs1870377 and rs7677779) to 23.1 pg/mL ± 2.52 SEM (rs1870378 and rs1870379). The triglyceride measurements for minor allele carriers ranged from 1.74 mmol/L ± 0.038 SEM (rs1870378 and rs1870379) to 1.76 mmol/L ± 0.038 SEM (rs10016064). The reference homozygote genotype group triglyceride measurements were similar for all six SNPS with values ranging from 1.89 to 1.90 mmol/L.

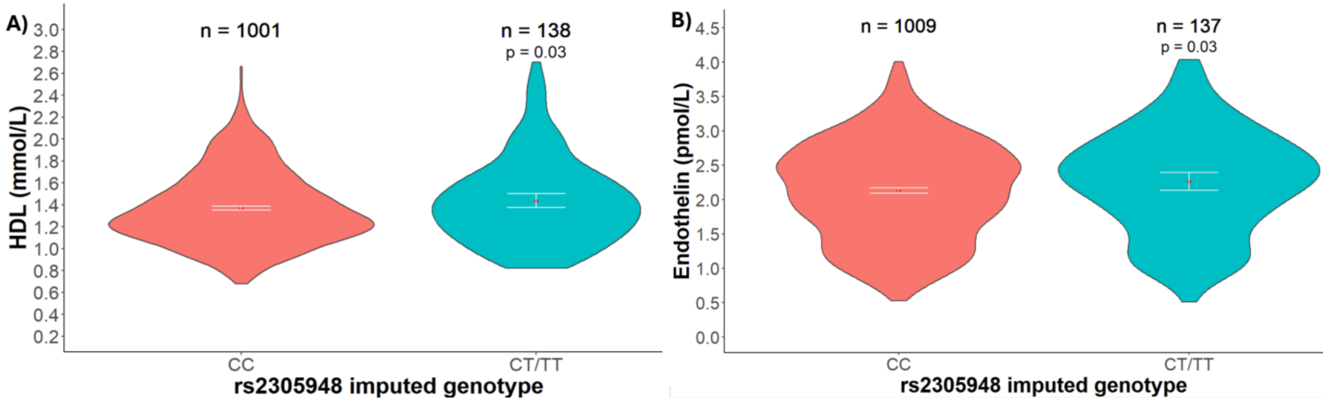


Figure 5.3. Imputed rs2305948 findings within the HVOL cohort relating to A) HDL and B) Endothelin levels. Red dot represents mean values of the variables. White lines represent ± SEM error bars.

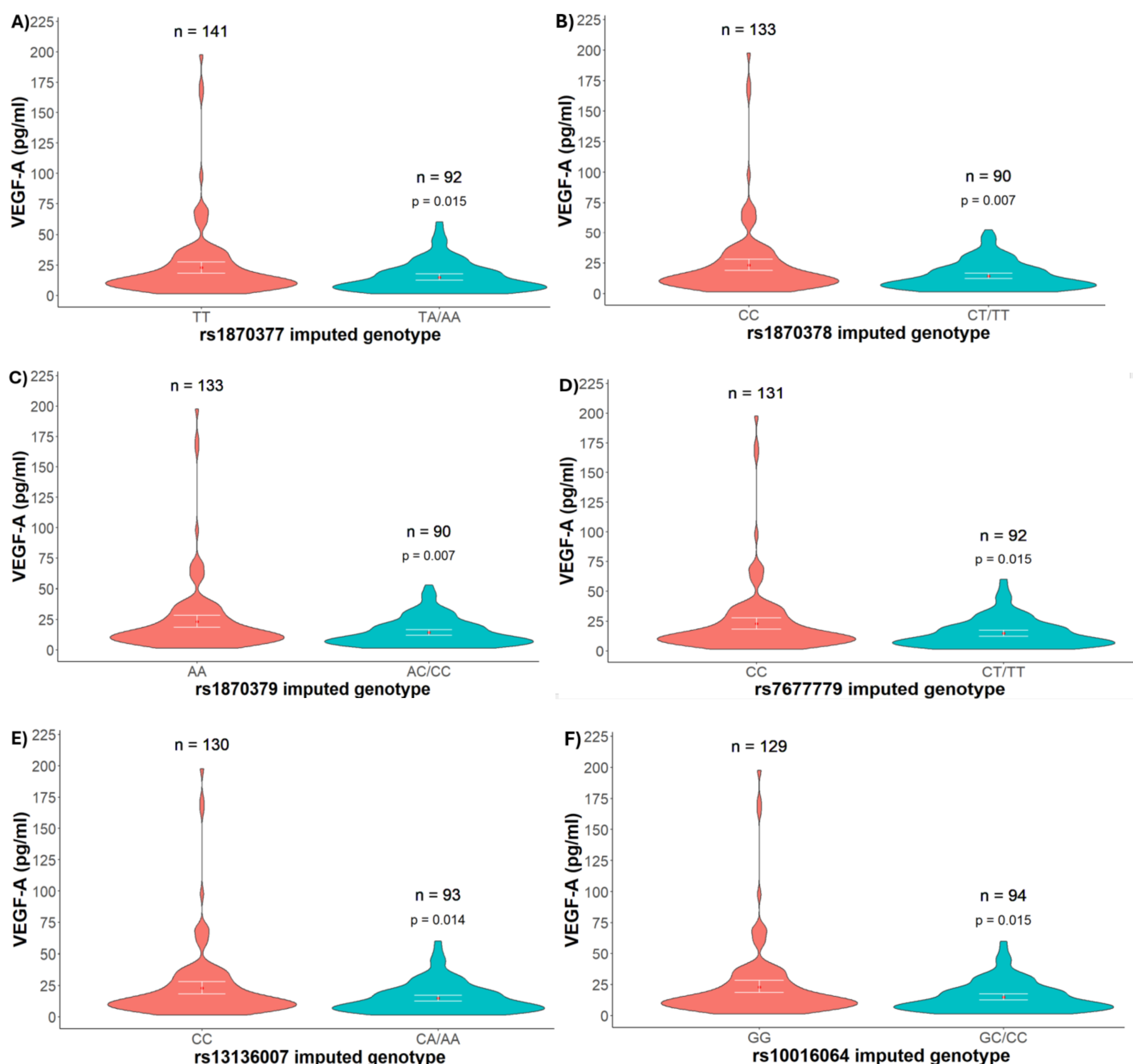


Figure 5.4. Imputed *VEGFR2* SNPs associated with VEGF-A levels in the HVOL cohort include A) rs1870377 B) rs1870378 C) rs1870379 D) rs7677779 E) rs13136007 and F) rs10016064. Red dot represents mean values of the variables. White lines represent ± SEM error bars.

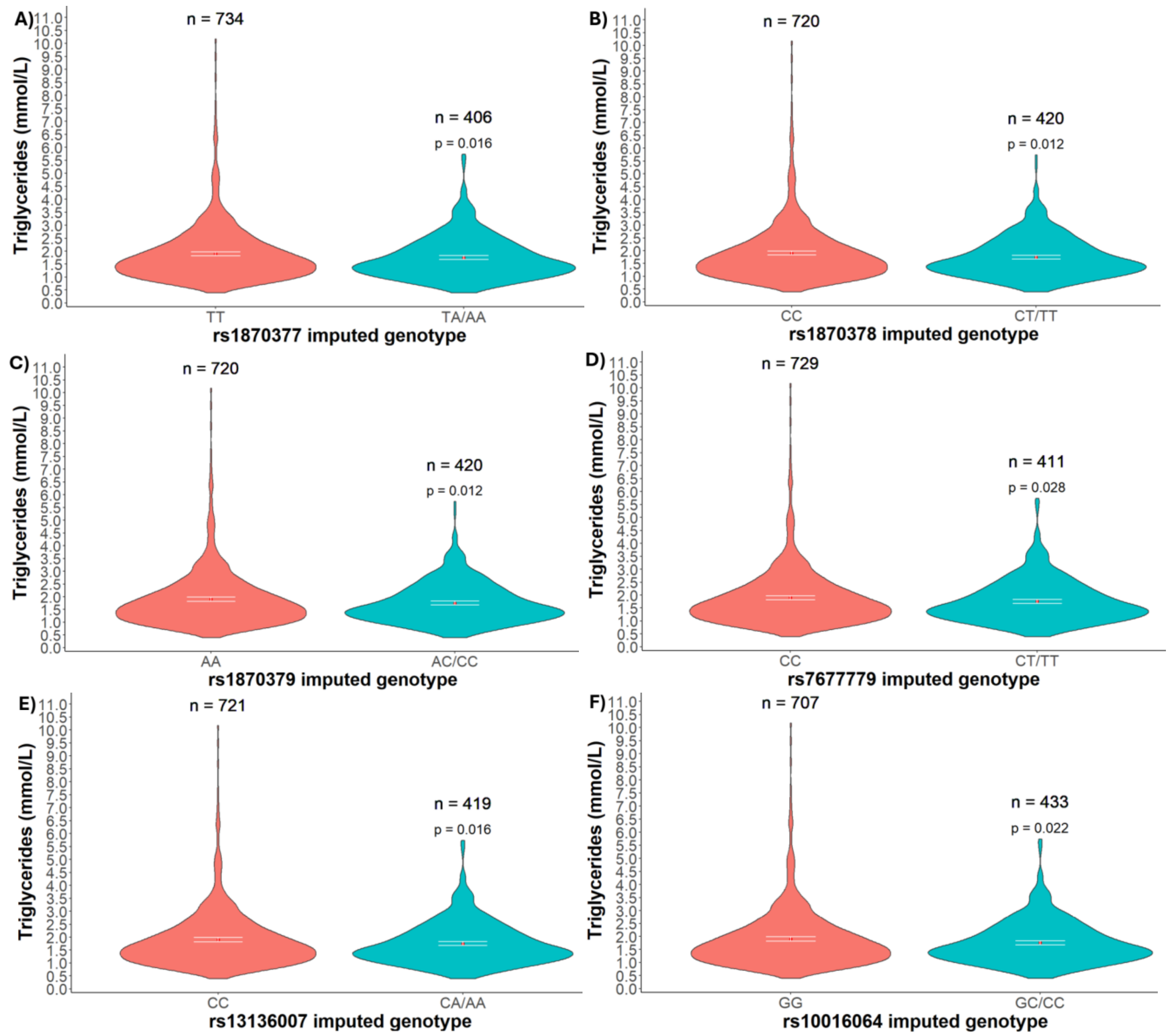


Figure 5.5. Imputed VEGFR2 SNPs associated with triglyceride levels within the HVOL cohort include A) rs1870377 B) rs1870378 C) rs1870379 D) rs7677779 E) rs13136007 and F) rs10016064. Red dot represents mean values of the variables. White lines represent ± SEM error bars.

5.3.5 Association of genotypes with clinical outcomes in the CDCS cohort

Univariate survival association analyses revealed that three SNPs (rs1870377, rs7677779 and rs13136007) presented significant associations with time to first CABG readmission event during the follow-up period (Appendix 5.5). Specifically, patients were more likely to have undergone CABG if they were carriers of the rs1870377 A, rs7677779 T or rs13136007 A alleles.

Imputed genotype data for rs23059438 showed no significant univariate survival associations with any clinical endpoints measured in the CDCS cohort. Manual genotype data showed rs2305948 CC was associated with higher cardiovascular readmission and all-cause mortality (Figure 5.6A and 5.6B). The same genotype was also significant for cardiac death (caused by events directly impacting heart structure) and cardiovascular death (caused by events involving blood vessels or blood flow) (Figure 5.6C and 5.6D). The association of the CC genotype was not significant for MI/CHD death and MACE events (Figure 5.6E and 5.6F). Manual genotype data for rs1870377 did not present statistically significant associations of univariate survival with any of the clinical endpoints assessed.

Given their association with survival, imputed data for SNPs rs1870377, rs767779 and rs13136007 was assessed in Cox proportional hazards regression models to identify their association with clinical endpoints. All three SNPs trended towards being significant independent predictors of all-cause mortality when using a selection of 9 covariates including age, gender, ethnicity, physical activity, previous MI, hypertension, beta blocker use, creatinine and NTproBNP levels (Appendix 5.6 to 5.8).

Multivariate regression models were generated using manual genotype data for rs2305948 and rs1870377 to determine if there was any change in the associations observed using imputed genotype data. rs1870377 A allele carriers had increased risk of cardiovascular disease readmissions (Table 5.3, $p = 0.048$, HR = 1.125). Meanwhile, the rs2305948 CC genotype came close to being a significant predictor for increased risk of all-cause death (Table 5.4, $p = 0.072$) in a similar model.

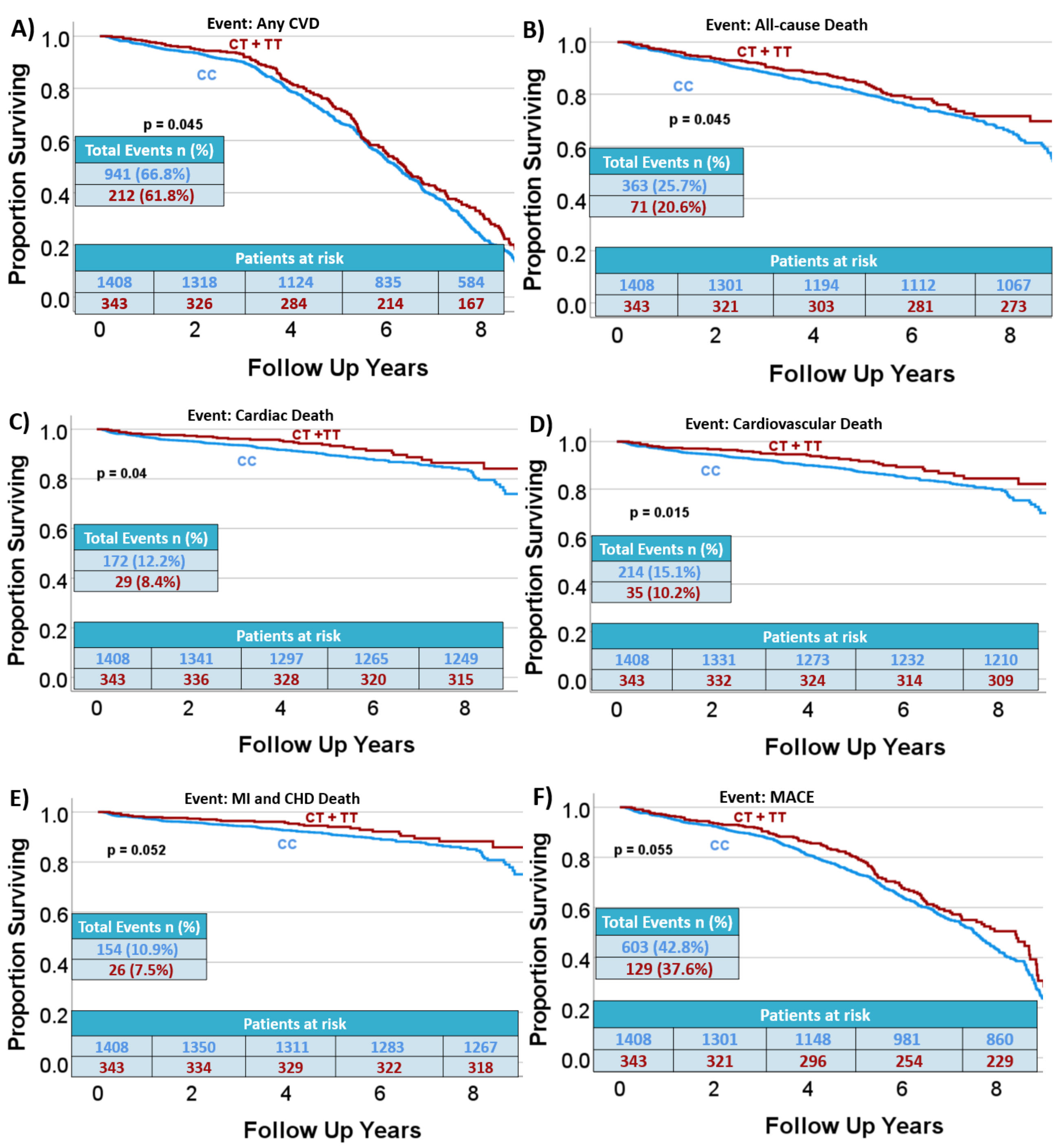


Figure 5.6. Kaplan-Meier survival plots of rs2305948 manual genotypes for survival versus A) first cardiovascular (CVD) readmission B) all-cause death, C) cardiac death FANPD) cardiovascular death E) myocardial infarction and coronary disease death and F) major adverse cardiovascular events (MACE) within the CDCS cohort. Genotypes are colour-coded blue for homozygous reference and red for homozygous for the minor allele genotypes. Patients at risk reported for every 2-year interval.

Table 5.3. Cox's proportional hazards regression model for any cardiovascular readmission in the CDCS cohort using manual genotypes for rs1870377 (n = 1821, 1186 (65.1%) events)

Predictor	Coeff.	SE	Wald	P – value	HR	95% CI for HR	
						Lower	Upper
Gender	0.099	0.069	2.07	0.15	1.10	0.965	1.27
Ethnicity			5.99	0.112			
European v Māori/Pasifika	0.236	0.137	2.99	0.084	1.27	0.969	1.65
European v Asian	0.327	0.179	3.35	0.067	1.39	0.977	1.97
European v MELAA	0.283	0.583	0.235	0.628	1.33	0.423	4.16
***Physical Activity (scale 1–4) $$**	-0.216	0.025	77.4	***1.3 x10^{-18}**	0.806	0.768	0.845
***Previous MI**	0.32	0.062	26.4	***2.8 x10^{-7}**	1.38	1.22	1.56
***Age**	0.017	0.003	29.3	***6 x 10^{-8}**	1.02	1.01	1.02
***Log10 NTproBNP$**	0.519	0.08	41.9	***9.1 x10^{-11}**	1.68	1.44	1.97
Beta blocker treatment at discharge	-0.158	0.086	3.39	0.065	0.854	0.721	1.01
***Creatinine**	0.002	0.0004	21.5	***3.4 x 10^{-6}**	1.002	1.001	1.003
Antecedent Hypertension	0.026	0.062	0.175	0.676	1.03	0.909	1.16
***rs1870377 genotype (TT v TA/AA)**	0.118	0.06	3.91	***0.048**	1.13	1.001	1.26

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B-type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Table 5.4. Cox's proportional hazards regression model for all cause death in the CDCS cohort using manual genotypes for rs2305438 (n = 1670, 408 (24.4%) deaths)

Predictor	Coeff	SE	Wald	P – value	HR	95% CI for HR	
						Lower	Upper
***Gender**	0.297	0.114	6.75	***0.009**	1.35	1.08	1.68
Ethnicity			2.56	0.466			
European v Māori/Pasifika	0.343	0.341	1.01	0.316	1.41	0.721	2.75
European v Asian	0.541	0.420	1.66	0.198	1.72	0.754	3.91
European v MELAA	-7.88	121.4	0.004	0.948	3.7 x 10^{-4}	0	8.1 x10^{99}
***Physical Activity**	-0.302	0.042	52.8	***3.7 x10^{-13}**	0.74	0.682	0.802
***Previous MI**	0.435	0.103	17.8	***2.4 x10^{-5}**	1.55	1.26	1.89
***Age**	0.065	0.006	106.3	***6.2 x10^{-25}**	1.07	1.05	1.08
***Log10 NTproBNP$**	1.37	0.144	91.3	***1.2 x10^{-21}**	3.95	2.98	5.23
***Beta blocker treatment at discharge**	-0.387	0.136	8.12	***0.004**	0.679	0.520	0.886
***Creatinine**	0.003	0.001	18.3	***1.8 x10^{-5}**	1.003	1.002	1.004
***Antecedent Hypertension**	0.251	0.107	5.52	***0.019**	1.29	1.04	1.59
rs2305438 genotype (CC v CT/TT)	0.244	0.136	3.23	0.072	1.28	0.978	1.67

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B-type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

5.3.6 Genomic context of *VEGFR2* SNPs

Using genome assembly GRCh38/hg38 from the University of California, Santa Cruz Genome Browser (http://genome.ucsc.edu), the Ensembl, JASPAR and HaploReg databases [19, 34-36], the genomic contexts for the 7 SNPs, associated with plasma analytes or clinical outcome risk, were explored to observe if DNA binding motifs for relevant effector molecules were found at these variants' loci. The location of the SNPs and the relevant transcription factor motifs are shown in Figure 5.7.

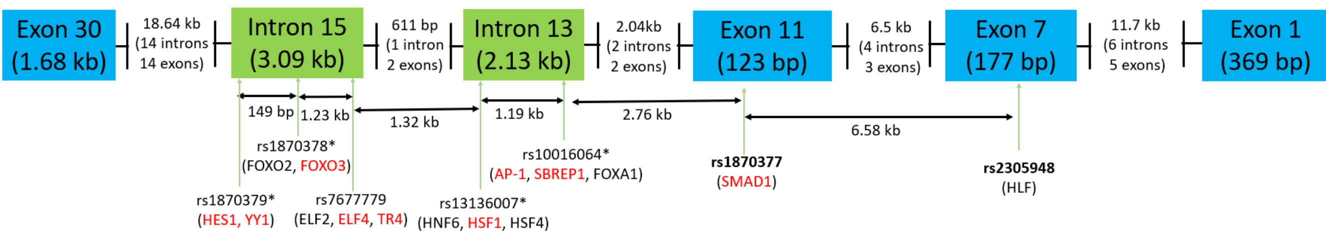


Figure 5.7. Schematic representation of *VEGFR2* locus SNPs. Functional SNPs with manual genotype data analysed in the present study are shown in bold. Green boxes represent intron regions. Blue box. Intron or exon length is indicated individually. Distance and number of exons/introns are specified between reported regions. Black double-ended arrows indicate distance between relevant locations (SNPs or genes). Green arrows point to individual SNP locations. Motifs overlapping SNPs are named between parentheses. Motifs in red are linked with cardiovascular risk pathways. *Indicates SNPs within enhancer regions active in the cardiovascular system, Distance measures calculated from NCBI GenBank Accession NC_000004.12

rs1870377, located in exon 11 induces an amino acid substitution that increases VEGFR2 phosphorylation when stimulated by VEGF-A leading to increased VEGF-A binding efficiency and increased capillary density [20, 25, 26]. This variant is located within a binding motif for the signal transducer protein SMAD1, which is predominantly expressed in endothelial cells and interacts with the cell surface activin receptor-like kinase-1 [37]. Similarly, the rs2305948 is a missense mutation located within exon 7 causing V297I, which is reported to affect the efficiency of dimerisation of two immunoglobulin-like domains on VEGFR2 which reduce VEGFR2's affinity for VEGF proteins [27, 28]. HaploReg data shows rs2305948 has the potential to disrupt VEGFR2 expression since it is located within enhancer marks that are active in adipose nuclei and human umbilical vein endothelial cells (HUVECs) [35].

The remaining SNPs are in intronic regions of *VEGFR2* that contain motifs for multiple transcription factors, some of which have links to the cardiovascular system and CVD risk. rs10016064, from intron 13, is ~2.76 kb 3' of rs1870377 (Figure 5.7). This variant is inside an enhancer region that is active in the aorta, ventricles, and right atrium [35] Additionally, HaploReg reports that rs10016064 C increases activity for the binding motif of the activating protein-1 (AP-1) transcription factor while reducing activity

for the sterol regulatory element-binding protein 1 (SBREP1) transcription factor [34-36]. Also, rs10016064 is located within the binding motif for Forehead Box A1 (FOXA1), which is vital for development of several non-cardiac organs [38]. For its part, rs13136007 is ~1.19 kb 3' from rs10016067 and lies within an enhancer region active in adipose nuclei, aorta, and the right ventricle [35]. Furthermore, the rs13136007 C allele results in higher binding activity for the hepatocyte nuclear factor 6 (HNF6) compared to the A allele [35]. Additionally, the variant is within transcription factor binding motifs for heat shock transcription factors 1 and 4 (HSF1 and HSF4) [36].

Other variants analysed include rs7677779 in intron 15, which overlaps with binding motifs for transcription factors ELF2 and ELF4, that help regulate HUVEC quiescence, immune response regulation and in the case of ELF4, adipogenesis regulation [39, 40]. rs1870378 and rs1870379 are also located in intron 15 ~1.23 kb 3' from rs7677779 (Figure 5.7). These SNPs are located ~200 bp of each other within the same enhancer mark which is active in left ventricle [35]. Despite this proximity they impact different binding factor motifs. The rs1870378 C allele has markedly lower activity for FoxO2 and -3 transcription factors, compared to the T allele, and it can increase the activity of interferon regulatory factors [35]. rs18703789 is not located within transcription factor binding sites, but it overlaps enhancer marks active in foetal heart tissue and left ventricle [34-36].

5.4 Discussion

VEGF-induced homodimerisation of VEGFR-2 is the primary pathway of angiogenesis and the receptor participates in vascular permeability, inflammation as well as collateral vessel generation [41, 42]. Notable SNPs within *VEGFR2* implicated with detrimental effect on VEGF-A activity, involved in altered angiogenesis and CVD risk, include rs2305498 and rs1870377 [20, 25, 27, 28]. Multiple variants on human chromosome 6 have been associated with increased VEGF-A circulating levels [22, 23, 43, 44] and increased CVD risk [45-49]. However, there is little CVD-related research exploring SNPs located at the *VEGFR2* locus on chromosome 4. The present study focused on assessing imputed genotype data for 13 SNPs, including rs2305948 and rs1870377, within *VEGFR2,* to identify candidate risk variants based on the strength of their association with cardiometabolic parameters and clinical outcomes. Variants of interest identified in our study included rs2305948, rs1870377, rs1870378 and rs1870379. Manual genotype data for rs1870377 revealed there was no significant association with baseline aldosterone ($p = 0.055$) and urate ($p = 0.083$) levels. No significant associations were observed between imputed or manual

SNP genotype data with VEGF-A levels in the CDCS cohort. Comparatively, in the heart healthy HVOL cohort, imputed genotype analysis showed rs2305948 T allele carriers had higher levels of HDL cholesterol ($p = 0.03$). Moreover, HVOL individuals who had at least one copy of the minor allele for 6 SNPs (rs1870377, rs1870378, rs1870379, rs7677779, rs13136007 and rs10016064) had lower levels of VEGF-A ($p < 0.015$) and triglycerides ($p < 0.028$).

Imputed genotype data for rs2305948 showed association with two established CVD risk biomarkers and two natriuretic peptides in the CDCS cohort. Manual genotyping for rs2305948 did not validate these findings. In the HVOL imputed genotype rs2305948 T allele was associated with increased HDL levels and endothelin levels trended towards significance. This agrees with other reports where the T allele has been associated with circulating analyte levels. For instance, GWAS studies with European ancestry cohorts identified the rs2305948 T allele is associated with increased VEGFR2 bloodstream levels [50-52]. Additionally, Iraqi patients undergoing percutaneous coronary intervention had the rs2305948 T allele associated with increased LDL and total cholesterol levels [53]. A direct effect of the rs2305948 T allele has been noted in Chinese individuals, where a reduction of VEGFR2 binding efficiency has been implicated with increased VEGFR2 protein and expression levels [20, 53-55]. CDCS patients carrying the rs1870377 A allele had lower urate and aldosterone levels.

Meanwhile, HVOL carriers of rs1870377 A had lower VEGF-A and triglyceride levels ($p < 0.05$). This agrees with other reports where the rs1870377 A allele reportedly impacts VEGFR2 and metabolites involved in other pathways. An *in vitro* HUVEC study identified the rs1870377 A allele was associated with lower expression levels of endothelial adhesion molecules and the SNP could impact stress related endothelial dysfunction [56]. Furthermore, a case control study of CAD patients analysing antiplatelet treatment identified the rs1870377 A allele is linked with increased levels of LDL and serum VEGFR2 [57]. Similarly, a previous report on the CDCS cohort presented an association of rs1870377 with soluble plasma levels of VEGFR1 and VEGFR2 [29]. Overall, our findings agree with reports that associate rs2305948 T and rs1870377 A alleles with lipid metabolism molecules and VEGF system components.

Manual genotype data revealed CDCS patients with rs2305948 CC had higher likelihood of CVD readmission, all-cause death, cardiac and cardiovascular death. In a multivariate Cox hazards model CC genotype trended towards an increased risk of death ($p= 0.072$). Our results contrast with reports assessing

this SNP in the context of CVDs. For example, a small study on Arab ancestry participants (n = 439) observed the TT genotype had a slight association with CHD susceptibility when adjusting for age, gender, hypertension, and diabetes [58]. A meta-analysis of five case-control cohorts of Asian ancestry observed the rs2305948 T allele had a potential association with stroke risk [59]. Another Chinese study noticed the T allele could interact with *VEGFA* promoter SNP rs833061 (Chromosome 6) to increase haemorrhagic stroke risk [60]. However, a Han Chinese study observed rs2305948 TT genotype was associated with a reduced CHD risk when smoking, alcohol intake and diabetes were present [49]. Overall, rs2305948 presents a similar frequency distribution in Asian and European populations [19]. However, our differing results may be due cis- or transacting activity of rs2305948 T on other eQTL variants in addition to altering cardiovascular homeostasis via HDL, LDL and VEGFR2 pathways.

In the literature there are some contrasting findings in terms of the overall effect each rs1870377 allele has on clinical outcome. A study on South Korean stroke patients identified the rs1870377 T allele increases ischaemic stroke risk [61] while a Chinese study associated the A allele with increased CHD risk [20]. There is also evidence that the rs1870377 A allele may have a protective effect, while the T allele has negative implications. A case-control study on outcomes following large artery atherosclerotic stroke identified the rs1870377 AA genotype was significantly associated with decreased outcome risk ranging from moderate disability to death [42]. Research on a Han Chinese cohort also revealed rs1870377 AA is associated with reduced CHD risk [49]. Meanwhile, a study on Han Chinese CHD patients identified TT genotype carriers presented increased CHD risk [48]. Overall, most Asian ancestry studies show a protective effect of rs18703777 A allele, while the T allele increases risk of CVD. This agrees with our findings given that we identified rs1870377 TT increased risk of CVD related readmission (Table 5.3).

Compared to rs2305948 and rs1870377, minimal reports are available on rs7677779, rs1870378, rs1870379, rs13136007 and rs10016064; although the latter four SNPs were reported to be captured when analysing rs1870377 [20]. A published patent on macular degeneration treatment specifies rs1870377 and rs2305948 are genetic markers which can be complemented with any of the five other SNPs [62]. Additionally, genetic database tools show rs18703777 and 5 SNPs (rs1870378, rs1870379, rs7677779, rs13136007 and rs10016064) have high linkage disequilibrium (LD) [35, 63]. The similar analyte associations observed for these six SNPs in the HVOL subset reinforces their genetic behaviour. However, based on each SNP's location in *VEGFR2* it was possible to identify DNA motifs that may influence CVD

risk pathways. This may provide an explanation for the associations observed in this study and aid in identifying if any of these six additional SNPs have potential use as predictive markers.

5.5 Conclusion

In summary, the present study made use of imputation data to aid in the identification of potential VEGFR2 SNP candidates associated with cardiovascular biomarker levels and/or risk outcome. Manual genotyping data suggests that the rs1870377 A allele was linked with lower plasma VEGF-A levels in established ACS while in a dominant (TT vs TA/AA) model the variant is associated with increased cardiovascular readmission risk. These results agree with international reports on rs1870377's association with CVD [48, 49, 54]. rs2305948 CC genotype was associated with lower survival and trends towards being an independent predictor of mortality in a multivariate Cox proportional hazard model with established predictors. Additionally, we provided some contextual details on the positions of *VEGFR2* variants on chromosome 4 in relation to regulatory factor binding motifs that have not been previously reported, which suggest mechanisms for their impact on pathways relevant to CVD risk, onset, and progression.

5.6 References

1. Wilson, N., et al., *Prioritization of intervention domains to prevent cardiovascular disease: a country-level case study using global burden of disease and local data.* Population Health Metrics, 2023. **21**(1): p. 1.
2. Roberts, R., C.C. Chang, and T. Hadley, *Genetic Risk Stratification: A Paradigm Shift in Prevention of Coronary Artery Disease.* Journal of the American College of Cardiology: Basic to Translational Science, 2021. **6**(3): p. 287-304.
3. Braile, M., et al., *VEGF-A in Cardiomyocytes and Heart Diseases.* International Journal of Molecular Sciences, 2020. **21**(15): p. 1-18.
4. Au Yeung, S.L., H. Lam, and C.M. Schooling, *Vascular Endothelial Growth Factor, and Ischemic Heart Disease Risk: A Mendelian Randomization Study.* Journal of the American Heart Association, 2017. **6**(8): p. 1-8.
5. Camaré, C., et al., *Angiogenesis in the atherosclerotic plaque.* Redox biology, 2017. **12**: p. 18-34.
6. Heinonen, S.E., et al., *The effects of VEGF-A on atherosclerosis, lipoprotein profile, and lipoprotein lipase in hyperlipidaemic mouse models.* Cardiovascular Research, 2013. **99**(4): p. 716-23.
7. Zhao, Z., et al., *Association between Single Nucleotide Polymorphisms in Cardiovascular Developmental Critical Genes and Hypertension: A Propensity Score Matching Analysis.* International Journal of Hypertension, 2020. **2020**: p. 1-8.
8. Ruotsalainen, S.E., et al., *An expanded analysis framework for multivariate GWAS connects inflammatory biomarkers to functional variants and disease.* European Journal of Human Genetics, 2021. **29**(2): p. 309-324.
9. Zhao, X., et al., *Vascular endothelial growth factor gene polymorphisms and coronary heart disease: a systematic review and meta-analysis.* Growth Factors, 2018. **36**(3-4): p. 153-163.
10. Ramakrishnan, S., V. Anand, and S. Roy, *Vascular endothelial growth factor signalling in hypoxia and inflammation.* Journal of Neuroimmune Pharmacology, 2014. **9**(2): p. 142-60.
11. Shaik, F., et al., *Structural Basis for Vascular Endothelial Growth Factor Receptor Activation and Implications for Disease Therapy.* Biomolecules, 2020. **10**(12): p. 1-25.
12. Oszajca, K., et al., *Single-nucleotide polymorphisms of VEGF-A and VEGFR-2 genes and risk of infantile hemangioma.* International Journal of Dermatology, 2018. **57**(10): p. 1201-1207.
13. Apte, R.S., D.S. Chen, and N. Ferrara, *VEGF in Signaling and Disease: Beyond Discovery and Development.* Cell, 2019. **176**(6): p. 1248-1264.
14. Simons, M., E. Gordon, and L. Claesson-Welsh, *Mechanisms, and regulation of endothelial VEGF receptor signalling.* Nature Reviews Molecular Cell Biology, 2016. **17**(10): p. 611-625.
15. Koch, S., et al., *Signal transduction by vascular endothelial growth factor receptors.* Biochemical Journal, 2011. **437**(2): p. 169-83.
16. Ruan, G.X. and A. Kazlauskas, *Lactate engages receptor tyrosine kinases Axl, Tie2, and vascular endothelial growth factor receptor 2 to activate phosphoinositide 3-kinase/Akt and promote angiogenesis.* Journal of Biological Chemistry, 2013. **288**(29): p. 21161-21172.
17. Le, T.H.V. and S.M. Kwon, *Vascular Endothelial Growth Factor Biology, and Its Potential as a Therapeutic Target in Rheumatic Diseases.* International Journal of Molecular Sciences, 2021. **22**(10): p. 1-17.
18. Peach, C.J., et al., *Molecular Pharmacology of VEGF-A Isoforms: Binding and Signalling at VEGFR2.* International Journal of Molecular Sciences, 2018. **19**(4): p. 1264.
19. Martin, F.J., et al., *Ensembl 2023.* Nucleic Acids Research, 2022. **51**(D1): p. D933-D941.
20. Wang, Y., et al., *Polymorphisms of KDRGene Are Associated With Coronary Heart Disease.* Journal of the American College of Cardiology, 2007. **50**(8): p. 760-767.
21. Matsumoto, T., et al., *VEGF receptor-2 Y951 signalling and a role for the adapter molecule TSAd in tumor angiogenesis.* European Molecular Biology Organization (EMBO) Journal, 2005. **24**(13): p. 2342-53.
22. Debette, S., et al., *Identification of cis- and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels.* Circulation Research, 2011. **109**(5): p. 554-63.

23. Choi, S.H., et al., *Six Novel Loci Associated with Circulating VEGF Levels Identified by a Meta-analysis of Genome-Wide Association Studies.* Public Library of Science (PLOS) Genetics, 2016. **12**(2): p. e1005874.
24. Meza-Alvarado, J.C., et al., *VEGF-A related SNPs: a cardiovascular context.* Frontiers in Cardiovascular Medicine, 2023. **10**.
25. Glubb, D.M., et al., *Novel Functional Germline Variants in the VEGF Receptor 2 Gene and Their Effect on Gene Expression and Microvessel Density in Lung Cancer.* Clinical Cancer Research, 2011. **17**(16): p. 5257-5267.
26. Masago, K., et al., *Next-generation sequencing of tyrosine kinase inhibitor-resistant non-small-cell lung cancers in patients harbouring epidermal growth factor-activating mutations.* BMC Cancer, 2015. **15**(1): p. 908.
27. Matalliotaki, C., et al., *Implication of VEGFR2 in endometriosis: A structural biological and genetic approach.* World Academy of Sciences Journal, 2019. **1**(6): p. 283-289.
28. Yang, Y., et al., *Direct contacts between extracellular membrane-proximal domains are required for VEGF receptor activation and cell signalling.* Proceedings of the National Academy of Sciences of the United States of America (USA), 2010. **107**(5): p. 1906-11.
29. Marks, E.C.A., et al., *Plasma levels of soluble VEGF receptor isoforms, circulating pterins and VEGF system SNPs as prognostic biomarkers in patients with acute coronary syndromes.* BMC Cardiovascular Disorders, 2018. **18**(1): p. 169.
30. Palmer, B.R., et al., *KCNE5 polymorphism rs697829 is associated with QT interval and survival in acute coronary syndromes patients.* Journal of Cardiovascular Electrophysiology, 2012. **23**(3): p. 319-24.
31. Palmer, B.R., et al., *Angiotensin-converting enzyme 2 A1075G polymorphism is associated with survival in an acute coronary syndromes cohort.* American Heart Journal, 2008. **156**(4): p. 752-758.
32. Palmer, B.R., et al., *Vascular endothelial growth factor-A promoter polymorphisms, circulating VEGF-A and survival in acute coronary syndromes.* PLOS One, 2021. **16**(7): p. 1-15.
33. Ellis, K.L., et al., *Genomic risk variants at 1p13.3, 1q41, and 3q22.3 are associated with subsequent cardiovascular outcomes in healthy controls and in established coronary artery disease.* Circulation Cardiovascular Genetics, 2011. **4**(6): p. 636-46.
34. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Research, 2002. **12**(6): p. 996-1006.
35. Ward, L.D. and M. Kellis, *HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease.* Nucleic Acids Research, 2016. **44**(D1): p. D877-81.
36. Rauluseviciute, I., et al., *JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles.* Nucleic Acids Research, 2023. **52**(D1): p. D174-D182.
37. González-Núñez, M., J.M. Muñoz-Félix, and J.M. López-Novoa, *The ALK-1/Smad1 pathway in cardiovascular physiopathology. A new target for therapy?* Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2013. **1832**(10): p. 1492-1510.
38. Bernardo, G.M. and R.A. Keri, *FOXA1: a transcription factor with parallel functions in development and cancer.* Bioscience Reports, 2012. **32**(2): p. 113-30.
39. Sivina, M., et al., *The transcription factor E74-like factor controls quiescence of endothelial cells and their resistance to myeloablative treatments in bone marrow.* Arteriosclerosis, Thrombosis, and Vascular Biology, 2011. **31**(5): p. 1185-91.
40. Suico, M.A., T. Shuto, and H. Kai, *Roles, and regulations of the ETS transcription factor ELF4/MEF.* Journal of Molecular Cell Biology, 2017. **9**(3): p. 168-177.
41. Zhou, Y., et al., *The Role of the VEGF Family in Coronary Heart Disease.* Frontiers in Cardiovascular Medicine, 2021. **8**: p. 1-16.
42. Li, Z., et al., *Missense Variants in Hypoxia-Induced VEGFA/VEGFR2 Signaling Predict the Outcome of Large Artery Atherosclerotic Stroke.* Cellular and Molecular Neurobiology, 2021. **41**(6): p. 1217-1225.
43. Merlo, S., et al., *Vascular Endothelial Growth Factor Gene Polymorphism (rs2010963) and Its Receptor, Kinase Insert Domain-Containing Receptor Gene Polymorphism (rs2071559), and Markers of Carotid*

Atherosclerosis in Patients with Type 2 Diabetes Mellitus. Journal of Diabetes Research, 2016. **2016**: p. 1-6.
44. Skrypnik, D., et al., *Association of rs699947 (-2578 C/A) and rs2010963 (-634 G/C) Single Nucleotide Polymorphisms of the VEGF Gene, VEGF-A and Leptin Serum Level, and Cardiovascular Risk in Patients with Excess Body Mass: A Case-Control Study.* Journal of Clinical Medicine, 2020. **9**(2): p. 1-16.
45. Nia, S.K., et al., *The impact of vascular endothelial growth factor+405 C/G polymorphism on long-term outcome and severity of coronary artery disease.* Journal of Clinical Laboratory Analysis, 2017. **31**(4).
46. Lacchini, R., et al., *Effect of Genetic Polymorphisms of Vascular Endothelial Growth Factor on Left Ventricular Hypertrophy in Patients With Systemic Hypertension.* The American Journal of Cardiology, 2014. **113**(3): p. 491-496.
47. Yadav, B.K., et al., *Genetic Polymorphisms rs699947, rs1570360, and rs3025039 on the VEGF gene are correlated with extracranial internal carotid artery stenosis and ischemic stroke.* Annals of Clinical and Laboratory Science, 2017. **47**(2): p. 144-155.
48. Li, L., et al., *Association of Genetic Polymorphisms on Vascular Endothelial Growth Factor and its Receptor Genes with Susceptibility to Coronary Heart Disease.* Medical Science Monitor, 2016. **22**: p. 31-40.
49. Liu, D., et al., *Association of Genetic Polymorphisms on VEGFA and VEGFR2 With Risk of Coronary Heart Disease.* Medicine, 2016. **95**(19): p. e3413.
50. Sun, B.B., et al., *Genomic atlas of the human plasma proteome.* Nature, 2018. **558**(7708): p. 73-79.
51. Gudjonsson, A., et al. *A genome-wide association study of serum proteins reveals shared loci with common diseases*. Nature communications, 2022. **13**, 480 DOI: 10.1038/s41467-021-27850-z.
52. Pietzner, M., et al., *Mapping the proteo-genomic convergence of human diseases.* Science 2021. **374**(6569): p. eabj1541.
53. Ahmed, A.A., K.I. Amber, and N.R. Hadi, *The Impact of Kinase Insert Domain (KDR) Gene Polymorphism rs2305948 on Clopidogrel Resistance in Iraqi Patients Undergoing Elective Percutaneous Coronary Intervention (PCI).* Acta Informatica Medica, 2020. **28**(3): p. 202-208.
54. Wang, L., et al., *A meta-analysis of the relationship between VEGFR2 polymorphisms and atherosclerotic cardiovascular diseases.* Clinical Cardiology, 2019. **42**(10): p. 860-865.
55. Zhang, J., et al., *Genetic Variants of VEGF (rs201963 and rs3025039) and KDR (rs7667298, rs2305948, and rs1870377) Are Associated with Glioma Risk in a Han Chinese Population: a Case-Control Study.* Molecular Neurobiology, 2016. **53**(4): p. 2610-8.
56. Schacher, N.M., et al., *Impact of single nucleotide polymorphisms in the VEGFR2 gene on endothelial cell activation under non-uniform shear stress.* International Journal of Molecular Medicine, 2019. **44**(4): p. 1366-1376.
57. Al Awaida, W., et al., *Association of KDR rs1870377 genotype with clopidogrel resistance in patients with post percutaneous coronary intervention.* Heliyon, 2021. **7**(2): p. e06251.
58. Hajer, F., et al., *Genetic polymorphisms in VEGFA and VEGFR2 genes associated with coronary heart disease susceptibility and severity.* Molecular Biology Reports, 2023. **50**(12): p. 10169-10177.
59. Qiu, S., et al., *The Association between VEGFR Gene Polymorphisms and Stroke: A Meta-Analysis.* PLoS One, 2016. **11**(3): p. e0151371.
60. Feng, C., et al., *Effect of gene-gene and gene-environment interaction on the risk of first-ever stroke and poststroke death.* Molecular Genetics & Genomic Medicine, 2019. **7**(8): p. e846.
61. Oh, S.-H., et al., *Association between kinase insert domain-containing receptor gene polymorphism and haplotypes and ischemic stroke.* Journal of the Neurological Sciences, 2011. **308**(1): p. 62-66.
62. Gregory Hannum and K. Schmidt, *Genetic markers for macular degeneration disorder treatment*, World Intellectual Property Organization International Bureau, Editor. 2014: United States.
63. Machiela, M.J. and S.J. Chanock, *LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants.* Bioinformatics, 2015. **31**(21): p. 3555-7.

Chapter 6.- Study on *VLDLR* locus SNPs

This chapter focuses on the work involving 4 SNPs mapping to the very low-density lipoprotein receptor gene (VLDLR) on human chromosome 9. These variants have been previously associated with circulating VEGF-A levels. This addresses project aim A as well as research objectives 2 and 3 (Chapter 1, Section 1.3).

The chapter highlights the relationships of the VLDLR variants with cardiometabolic variables and specific outcome risk. These results address project aim B and research objectives 5 (Chapter 1, Section 1.3). .

STATEMENT OF CONTRIBUTION

DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the student and the student's main supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the student's contribution as indicated below in the Statement of Originality.

Student name:	Juan Carlos Meza Alvarado
Name and title of main supervisor:	Dr. Barry Palmer, Senior Lecturer
In which chapter is the manuscript/published work?	Chapter 6

Describe the contribution that the student and members of the supervisory team have made to the manuscript/published work:[1]

- The student took a major role in project conceptualization, statistical analyses of the imputed genotype data, , investigation, data curation and analysis, visualisation and draft writing.
- Data on CDCS VEGF-A levels was generated by others and described in Palmer et al (2021).
--Imputed genotype data analysed in this Chapter was generated by Dr. Anna Pilbrow and Dr. Vinicius Tragante
The supervisory team provided feedback on the manuscript drafts.

Please select one of the following three options:

◯	The manuscript/published work is published or in press Please provide the full reference of the research output:
◯	The manuscript is currently under review for publication Please provide the name of the journal:
◉	It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Student's signature:	[signature]	Main supervisor's signature:	[signature]

This form should be placed at the beginning of each relevant thesis chapter.

Doctoral Research Committee
May 2023

6.1 Abstract

Background: Vascular endothelial growth factor A (VEGF-A) is involved in endothelial cell proliferation, vascular permeability, and inflammatory activity. Increased VEGF-A activity has been linked to cardiovascular diseases (CVDs). Genome wide association studies (GWAS) have associated single nucleotide polymorphisms (SNPs) with VEGF-A levels, including variants rs7043199, rs10738760, rs7030781 and rs2375981 that map to the very low-density lipoprotein receptor (*VLDLR*) locus. The variants impact lipid metabolism and may be associated with CVD or outcome risk.

Methods: Imputed genotype data for 4 SNPs on human chromosome 9 from 1935 patients from the Coronary Disease Cohort Study (CDCS) and 1183 individuals from the Canterbury Healthy Volunteers Study (HVOL) was analysed. Statistical analyses focused on using the imputed genotype data to assess the influence of each genotype group on cardiometabolic parameters using one-way ANOVA tests and testing variant associations with clinical endpoints through Kaplan-Meir analyses as well as multivariate regression models. VEGF-A levels for 227 HVOL participants were measured by an ELISA immunoassay and compared with levels from 549 CDCS patients.

Results: Imputed CDCS data showed rs10738760 GG and rs2375981 GG genotype groups were associated with history of myocardial infarction (MI), lower left ventricular ejection fraction (LVEF), amino terminal C-type natriuretic peptide (NT-CNP) and blood pressure levels ($p<0.05$). The rs7043199 A allele was associated with lower VEGF-A levels ($p = 0.049$). In the HVOL, rs10738760 AG and rs2375981 CG genotype groups were associated with lower systolic blood pressure, urate, creatinine, and endothelin levels ($p < 0.05$). In this heart healthy group, rs10738760 GG and rs2375981 GG were associated with lower VEGF-A levels ($p = 0.027$; $p = 4.5 \times 10^{-4}$). We observed rs10738760 AA genotype was associated with increased risk of cardiovascular death ($p = 0.047$, HR = 1.50) when compared to the GG genotype. rs10738760 AG genotype was associated with reduced risk of non-ST elevation MI (NSTEMI) readmission ($p = 0.011$, HR = 0.747). The rs2375981 GC genotype was associated with reduced cardiovascular death risk ($p = 0.015$, HR = 0.7), but not incidence of NSTEMI during follow-up.

Conclusion: rs7043199 was associated with circulating plasma VEGF-A levels in the CDCS cohort. rs10738760. rs7030781 and rs2375981 were associated with circulating plasma VEGF-A levels in the HVOL subset. rs2375981 GC is associated with cardiovascular death risk. rs10738760 AA is implicated in higher cardiovascular death risk while rs10738760 AG is associated with lower NSTEMI readmission risk. The four *VLDLR* variants appear to be promising genetic markers for CVD risk profiling.

6.2 Introduction

Cardiovascular disease (CVD) is a leading cause of death worldwide. Diagnosis is complicated by the lack of strongly predictive biomarkers in individuals before the onset of symptoms [1]. According to a recent case study of New Zealand, the five major CVD risk factors in the country are high systolic blood pressure, high low-density lipoprotein (LDL) cholesterol, high body mass index (BMI), tobacco consumption and dietary risk [2]. A notable factor linked with CVD risk, is the vascular endothelial growth factor (VEGF-A). Elevated VEGF-A levels have been associated with ischaemic heart disease (IHD), coronary artery disease (CAD), stroke and myocardial infarction (MI) [3-6].

Factors that upregulate VEGF-A activity include hypoxia, oestrogen levels, thyroid stimulating hormone and inflammatory pathway molecules (e.g. interleukin 6 (IL-6) or signal transducer and activator of transcription proteins (STAT)) [7, 8]. Increased VEGF-A activity promotes a positive feedback loop for inflammation which causes increased vascular dilation, higher adhesion protein expression and increased trans-endothelial lipid migration collectively leading to CVD onset [3, 9-14]. VEGF-A has been implicated in reducing plasma lipoprotein lipase (LPL) activity, resulting in elevated very low-density lipoprotein (VLDL) levels, which are correlated with CVD risk mechanisms (e.g. atherosclerosis, insulin resistance and diabetes) [10, 15, 16]. Single nucleotide polymorphisms (SNPs) are a factor that influence VEGF-A levels and are associated with cardiometabolic risk markers or CVD risk [17]. Genome wide association studies (GWAS) identified that SNPs located on the VLDL receptor (*VLDLR*) locus, in human chromosome 9, are associated with VEGF-A circulating level variability [18-20].

VLDL functions as a triglyceride carrier, regulating blood pressure through nitric oxide signalling and aldosterone synthesis [16, 21, 22]. Additionally, VLDLs have surface apolipoproteins which function as ligands for VLDLR or cofactors for lipoprotein lipase (LPL) mediated hydrolysis [23]. Following hydrolysation, VLDL is transformed into VLDL remnants and intermediate density lipoprotein (IDL), with the latter becoming low density lipoprotein (LDL) [23]. For its part, VLDLR is expressed in different cell types (adipocytes, cardiomyocytes, macrophages, vascular endothelium) and it modulates triglyceride hydrolysis [16, 23]. Overall, VLDLR dysfunction can lead to high circulating VLDL and LDL levels, which are linked with CVD mortality by promoting cytotoxic proinflammatory atherogenesis [16, 23-25].

VLDLR variants associated with VEGF-A levels are rs7043199, rs10738760. rs7030781 and rs2375981 [18-20]. Specifically, rs10738760 along with three other SNPs associated with VEGF-A can explain up to 47.6% of variability of circulating VEGF-A levels [18]. Another study highlighted a group of 10 SNPs across six loci that could explain up to 52% of VEGF-A level variance, including VLDLR variants rs2375981 and rs7043199 [19]. Lastly, a study on circulating growth factors found the rs7030781 T allele was strongly associated with lower VEGF levels [20]. The current project used imputation data for these four VLDLR variants to explore their relationship with cardiometabolic variables, including plasma concentrations of VEGF-A, in post-acute coronary syndrome patients (Coronary Disease Cohort Study, CDCS) and heart healthy matched controls (Canterbury Healthy Volunteers Study, HVOL). This study also aimed to identify which VLDLR SNPs had significant associations with clinical outcome risk endpoints in the CDCS.

6.3 Results

6.3.1 Baseline characteristics of study cohorts

Baseline characteristics of the CDCS cohort are described in Table 6.1. The population is predominantly European (89.6%) and male (71.5%). Additionally, study participants' discharge medication included antithrombotic drug therapy being the highest (96.4%) followed by lipid lowering (88.3%) and blood pressure lowering medications (85.8%). Moreover, 65.6% of the patients had a cardiovascular event readmission including heart failure (HF), stroke, MI, unstable angina, coronary artery bypass graft (CABG), arrhythmia, and percutaneous coronary intervention. Imputed genotypes were obtained for 1935 patients (90.4%) of the CDCS cohort for four VLDLR SNPs located on human chromosome 9 (rs7043199, rs7030781, rs10738760 and rs2375981). The frequencies for each of the SNPs genotypes, their respective alleles, their MAF and imputation quality are summarised in Appendix 6.1.

Baseline characteristics for the 1183 individuals of the HVOL cohort subset are described in Table 6.1. The population was mostly European (98.8%) and male (67.7%). As expected from a heart healthy cohort, individuals presented low incidence of hypertension (28.9%), high cholesterol (28%) and diabetes (5.4%). Additionally, low levels of medication usage were observed with blood pressure lowering medication (24.8%) being the highest. Imputed genotypes were obtained for 1177 individuals (99.5%) of

the HVOL for all four VLDLR SNPs. The frequencies for each of the SNP genotypes, their respective alleles, their MAF and imputation quality are summarised in Appendix 6.2.

Table 6.1 Baseline characteristics of the CDCS and HVOL cohort

Variable	CDCS		HVOL		p - value
	n	Mean ± SE or n (%)	n	Mean ± SE or n (%)	
Anthropometric					
Age	2140	66.5 ± 0.27	1183	68.1 ± 0.216	3.76×10^{-5}
Male Gender	2140	1531 (71.5%)	1183	801 (67.7%)	1.71×10^{-119}
Weight (kg)	2116	79.9 ± 0.36	1183	77.4 ± 0.42	2.79×10^{-2}
Body Mass Index (kg/m^2)	2110	27.6 ± 0.109	1183	26.2 ± 0.11	8.33×10^{-15}
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	2140	89.6%, 6.2%, 3.9%, 0.3%	1183	98.8%, 1.1%, 0.1%, 0%	0
Tobacco use (Current, Ex Smoker, Non-smoker)	2022	6.3%, 57.4%,36.3%	1183	4.8%, 39.6%,55.6%	5.66×10^{-253}
Alcohol (drinker, ex-drinker, non-drinker)	2137	62.8%, 12.1%, 25.1%	1183	80.7%, 4.1%, 15.2%	0
Systolic blood pressure (mmHg)	2095	128.9 ± 0.476	1180	137 ± 0.545	1.09×10^{-24}
Diastolic blood pressure (mmHg)	2095	74.9 ± 0.26	1180	80.9 ± 0.306	2.06×10^{-46}
Cardiovascular readmission after discharge	2140	1404 (65.3%)	N/A		
Plasma analytes					
Creatinine (mmol/L)	2063	99.8 ± 1.17	1146	90.3 ± 11.7	3.64×10^{-9}
Cholesterol (mmol/L)	1681	4.89 ± 0.028	1146	5.65 ± 0.033	5.69×10^{-62}
Urate (mmol/L)	1238	0.37 ± 0.003	1146	0.347 ± 0.0022	1.54×10^{-11}
VEGF-A (pg/mL)	549	45.02 ± 1.50	223	19.5 ± 1.60	5.76×10^{-22}
Atrial natriuretic peptide (pmol/L)	2116	43.5 ± 0.68	1163	21.6 ± 0.325	7.24×10^{-9}
Amino terminal atrial natriuretic peptide (pmol/L)	2115	1.35 ± 0.026	1170	0.673 ± 0.325	4.99×10^{-76}
B-type natriuretic peptide (pmol/L)	2116	26.4 ± 0.74	1167	6.37 ± 0.119	3.08×10^{-84}
Amino terminal pro B-type natriuretic peptide (pmol/L)	2116	135.3 ± 4.07	1174	108 ± 3.98	4.35×10^{-5}
Medication					
Blood pressure lowering	2140	1837 (85.8%)	1183	293 (24.7%)	1.14×10^{-9}
Lipid lowering	2140	1890 (88.3%)	1183	189 (16%)	1.51×10^{-47}
Antithrombotic	2140	2063 (96.4%)	1183	120 (10.1%)	3.59×10^{-73}
Genetic data					
Imputed genotype	2140	1935 (90.4)	1183	1177 (99.5%)	

Abbreviations: CDCS: Coronary Disease Cohort Study, HVOL: Canterbury Healthy Volunteers Study, MELAA: Middle Eastern Latin American African, MI: myocardial infarction, SE: Standard Error

6.3.2 VLDLR SNP associations within the CDCS cohort

One-way ANOVA was performed to test the association of the four imputed SNP genotypes with anthropomorphic and plasma analyte data of the CDCS cohort patients (Table 6.2). rs10738760 GG and rs2375981 GG genotype groups had lower previous MI occurrence (rs10738760 $p = 0.032$; rs2375981 $p = 0.025$), lower systolic blood pressure (rs10738760 $p = 0.006$; rs2375981 $p = 0.024$), lower diastolic blood pressure (rs10738760 $p =0.005$; rs2375981 $p = 0.008$) and lower amino terminal C-type natriuretic peptide (NT-CNP) levels (rs10738760 $p = 0.048$; rs2375981 $p = 0.049$) compared to the other genotype groups. Additionally, the rs2375981 GG genotype group presented lower plasma CNP levels ($p = 0.049$). Meanwhile, the rs10738760 GG genotype group and the rs7030781 TT genotype group presented significantly lower LVEF (rs10738760 $p = 0.019$; rs7030781 $p = 0.038$). Lastly, the rs7043199 TT genotype group presented lower height and troponin I levels compared to A allele carriers (Table 6.2). rs7043199 A allele carriers had lower VEGF-A plasma levels ($p = 0.049$). The other three SNPs did not present significant associations with VEGF-A (rs10738760 $p = 0.366$, rs2375981 $p = 0.162$, rs7030781 $p = 0.623$), but trended to progressively decreasing mean levels with the addition of minor alleles.

Table 6.2 Genotype comparisons for *VLDLR* SNPs in the CDCS cohort

Variable	n	rs10738760 Genotypes AA	n	AG	n	GG	p
Anthropometric Variables							
Age (years)	574	66.7 ± 0.52	974	66.9 ± 0.39	387	66.3 ± 0.62	0.707
Male Gender	574	407 (70.9%)	947	707 (72.6%)	387	275 (71.1%)	0.730
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	574	89%, 5.7%, 4.7%, 0.6%	947	90.3%, 6.2%, 3.3%, 0.2%	387	91.2%, 4.4%, 3.9%, 0.5%	0.558
Height (m)	569	1.69 ± 0.004	960	1.7 ± 0.003	382	1.69 ± 0.004	0.108
Weight (kg)	570	79.2 ± 0.705	961	80.02 ± 0.557	382	79.4 ± 0.818	0.623
BMI (kg/m^2)	568	27.6 ± 0.21	957	27.5 ± 0.16	382	27.4 ± 0.25	0.860
Systolic BP (mmHg)	559	130 ±0.93	948	130 ± 0.706	386	126 ± 1.1	**0.006**
Diastolic BP (mmHg)	559	75.8 ± 0.48	948	74.8 ± 0.39	386	73.2 ± 0.614	**0.005**
LVEF (%)	527	57.2 ± 0.53	888	57.7 ± 0.39	352	55.5 ± 0.067	**0.019**
Beta blocker use	574	493 (85.9%)	947	852 (87.5%)	387	324 (83.7%)	0.184
Previous MI	571	181 (31.7%)	967	305 (31.5%)	384	95 (24.7%)	**0.032**
Tobacco use (Current, Ex Smoker, Non-smoker)	574	5.6%, 56.6%, 37.8%	947	6.3%, 58%, 35.7%	387	7.5%, 55.8%, 36.7%	0.725
Alcohol intake (current, ex-drinker, non-drinker)	572	60.3%, 11.9%, 27.8%	973	64%, 12.2%, 23.8%	387	63.6%, 11.6%, 24.8%	0.532
Plasma analytes							
Cholesterol (mmol/L)	434	4.92 ± 0.055	773	4.87 ± 0.043	302	4.92 ± 0.065	0.638
Creatinine (mmol/L)	553	100 (96.5 -104)	948	101 (97.1 – 105)	375	95.1 (91-99.1)	0.166
Urate (mmol/L)	337	0.38 (0.37-0.39)	596	0.37 (0.36 -0.37)	226	0.36 (0.35-0.37)	0.120
Troponin I (ng/L) $	545	55.2 (21-89.4)	931	42.2 (18.8-65.5)	372	28.1 (19.1-36.9)	0.739
ANP (pg/mL) $	570	44.4 (41.7 – 47.2)	967	43.6 (41.7 – 45.5)	385	43.5 (40.3 – 46.7)	0.793
NT-ANP (pmol/L)$	570	1.36 (1.27-1.45)	967	1.37 (1.29-1.45)	385	1.29 (1.18-1.4)	0.353
BNP (pmol/L) $	570	27.3 (24.4-30.1)	968	26.4 (24.2-28.6)	385	26.4 (23.1-29.5)	0.865
NTproBNP (pg/ml) $	570	133 (119-146)	968	138 (125-151)	385	136 (119-153)	0.950
CNP (pmol/L) $	562	0.66 (0.62-0.7)	954	0.66 (0.63-0.69)	379	0.66 (0.61-0.7)	0.111
NT-CNP (pmol/L)$	563	23.5 (21.6-25.4)	951	23.6 (21.7-25.4)	378	21.0 (19.8-22.2)	**0.048**
Aldosterone (pmol/L)$	558	177 (152-203)	944	161 (154-168)	373	164 (153-175)	0.482
Adrenomedullin (pg/ml)$	544	8.59 (8.12-9.07)	945	8.58 (8.26-8.89)	373	8.37 (7.9-8.84)	0.888
VEGF-A (pg/mL) $	160	47.9 (42-53.6)	273	45.0 (40.8-49.2)	104	41.9 (36-47.8)	0.366

Variable	n	CC	n	CG	n	GG	p
				rs2375981 genotypes			
Anthropometric Variables							
Age (years)	636	66.8 ± 0.49	957	67.03 ± 0.39	342	65.8 ± 0.67	0.261
Male Gender	636	453 (71.2 %)	957	694 (72.5%)	342	242 (70.8%)	0.767
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	636	88.2%, 6.9%, 4.2%, 0.7%	957	91%, 5.4%, 3.3%, 0.3%	342	91.2%, 4.1%, 4.4%, 0.3%	0.34
Height (m)	629	1.69 ±0.003	944	1.7 ± 0.003	338	1.7 ±0.005	0.109
Weight (kg)	632	79.7 ± 0.67	943	79.6 ± 0.56	338	79.7 ± 0.85	0.988
BMI (kg/m^2)	628	27.7 ± 0.2	941	27.3 ± 0.16	338	27.3 ± 0.16	0.303
Systolic BP (mmHg)	620	130 ± 0.89	932	129 ± 0.7	341	126 ± 1.17	**0.024**
Diastolic BP (mmHg)	620	75.9 ± 0.47	932	74.5 ± 0.385	341	73.5 ± 0.66	**0.008**
LVEF (%)	585	57.4 ± 0.5	869	57.5 ± 0.41	313	55.7 ± 0.71	0.077
Beta blocker use	574	493 (58.9%)	974	852 (87.5%)	387	324 (83.7%)	0.184
Previous MI	632	204 (32.3%)	951	295 (31%)	339	82 (24.2%)	**0.025**
Tobacco use (Current, Ex Smoker, Non-smoker)	574	5.6%, 56.6%, 37.8%	974	6.3%, 58%, 35.7%	387	7.5%, 55.8%, 36.7%	0.725
Alcohol intake (current, ex-drinker, non-drinker)	634	60.7%, 12%, 27.3%	956	64.4%, 12%, 23.6%	342	62.3%, 11.7%, 26%	0.538
Plasma analytes							
Cholesterol (mmol/L)	484	4.92 ± 0.05	759	4.87 ± 0.04	266	4.93 ± 0.07	0.701
Creatinine (mmol/L)	616	102 ± 2.03	928	99.9 ± 1.82	332	95.2 ± 2.32	0.182
Urate (mmol/L)	372	0.38 (0.37-0.39)	586	0.36 (0.36 -0.37)	201	0.36 (0.35 -0.38)	0.072
Troponin I (ng/L) $	606	51.9 (21-82.6)	915	43.3 (19.5-67.1)	327	26.7 (16.9-36.5)	0.371
ANP (pg/mL) $	632	44.2 (41.6 – 46.8)	951	43.9 (42.1 – 45.8)	340	42.9 (39.4 – 46.3)	0.416
NT-ANP (pmol/L)$	631	1.36 (1.27-1.44)	951	1.38 (1.3-1.46)	340	1.25 (1.14-1.36)	**0.042**
BNP (pmol/L) $	632	27.4 (24.6-30)	951	26.7 (24.4-29)	340	25.3 (22.1-25.5)	0.557
NTproBNP (pg/ml) $	632	135 (122-149)	951	138 (125-151)	340	131 (113-149)	0.718
CNP (pmol/L) $	622	0.67 (0.63-0.7)	940	0.66 (0.63-0.69)	333	0.65 (0.6-0.7)	**0.049**
NT-CNP (pmol/L)$	623	23.8 (22-25.6)	937	23.3 (21.4-25.1)	332	21.1 (19.7-22.4)	**0.049**
Aldosterone (pmol/L)$	620	177 (154-200)	926	161 (154-168)	329	161 (150-172)	0.411
Adrenomedullin (pg/ml)$	602	8.58 (8.13-9.02)	930	8.53 (8.22-8.84)	330	8.5 (7.98-9.03)	0.986
VEGF-A (pg/mL) $	178	49.5 (43.3-55.6)	269	43.9 (40.1-47.6)	90	41.1 (34.7-47.3)	0.162

Variable	n	AA	n	AT	n	TT	p
		rs7030781 Genotypes					
Anthropometric Variables							
Age (years)	765	67.04 ± 0.44	914	66.5 ± 0.4	256	66.6 ± 0.77	0.686
Male Gender	765	547 (71.5%)	914	663 (72.5%)	256	179 (69.9%)	0.696
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	765	90.1%, 5.8%, 3.5%, 0.6%	914	90.2%, 5.8%, 3.9%, 0.1%	256	90.2%, 5.1%, 4.3%, 0.4%	0.682
Height (m)	758	1.69 ± 0.003	900	1.7 ± 0.003	253	1.69 ± 0.005	0.288
Weight (kg)	760	79.4 ± 0.59	900	79.9 ± 0.58	253	79.6 ± 1	0.843
BMI (kg/m^2)	757	27.6 ± 0.18	897	27.4 ± 0.164	253	27.5 ±0.32	0.828
Systolic BP (mmHg)	747	129 ±0.8	891	130± 0.7	255	127 ± 1.3	0.192
Diastolic BP (mmHg)	747	75.2 ± 0.42	891	74.7 ± 0.39	255	73.8 ± 0.79	0.199
LVEF (%)	699	57.2 ± 0.46	833	57.5 ± 0.4	235	55.3 ± 0.86	**0.038**
Beta blocker use	765	664 (86.8%)	914	795 (87%)	256	210 (82%)	0.108
Previous MI	761	240 (31.5%)	907	273 (30.1%)	254	68 (26.8%)	0.356
Tobacco use (Current, Ex Smoker, Non-smoker)	765	5.4%, 58.1%, 36.5%	914	6.5%, 56.8%, 36.7%	256	8.6%, 55.5%, 35.9%	0.467
Alcohol intake (current, ex-drinker, non-drinker)	763	61.5%, 12.5%, 26%	913	64.3%, 11.5%, 24.2%	256	61.3%, 12.1%, 26.6%	0.805
Plasma analytes							
Cholesterol (mmol/L)	584	4.91 ± 0.04	722	4.87 ± 0.04	203	4.93 ± 0.08	0.742
Creatinine (mmol/L)	742	100 (96.8-103)	885	100 (96.4-104)	249	95.9 (90.2-101.5)	0.474
Urate (mmol/L)	462	0.38 (0.37-0.39)	552	0.36 (0.36-0.37)	145	0.36 (0.35-0.38)	0.102
Troponin I (ng/L) $	733	49 (23.2-74.7)	869	43.8 (18.8-68.7)	246	23.7 (15.4-31.9)	0.602
ANP (pg/mL) $	760	44.9 (42.5 – 47.2)	909	42.9 (41.1 – 44.9)	854	43.9 (39.7 – 48.1)	0.523
NT-ANP (pmol/L)$	759	1.36 (1.28-1.44)	909	1.35 (1.27-1.44)	254	1.29 (1.16-1.42)	0.363
BNP (pmol/L) $	760	28.0 (25.5-30.5)	909	25.8 (23.5-27.9)	254	26.1 (22.1-30)	0.263
NTproBNP (pg/ml) $	760	139 (126-152)	909	132.8 (120-146)	254	138.1 (116-160)	0.471
CNP (pmol/L) $	751	0.67 (0.64-0.71)	891	0.65 (0.62-0.68)	253	0.66 (0.61-0.72)	0.101
NT-CNP (pmol/L)$	752	23.2 (21.6-24.6)	888	23.6 (21.6-25.5)	252	21.1 (19.-22.5)	0.153
Aldosterone (pmol/L)$	745	178 (159-198)	881	157(150-164)	249	163 (149-177)	0.073
Adrenomedullin (pg/ml)$	730	8.56 (8.17-8.94)	884	8.55 (8.22-8.87)	248	8.47 (7.85-9.08)	0.999
VEGF-A (pg/mL) $	217	45.9 (41.3-50.5)	251	45.3 (40.8-49.8)	69	42.9 (35-50.8)	0.623

Variable	n	rs7043199 genotypes TT	n	TA + AA	p
Anthropometric Variables					
Age (years)	1244	66.8 ± 0.35	691	66.7 ± 0.46	0.946
Male Gender	1244	894 (71.9%)	691	495 (71.6%)	0.916
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	1244	87.1%, 7.1%, 5.3%, 0.5%	691	95.5%, 3.2%, 1.2%, 0.1%	**5.4×10^{-8}**
Height (m)	1233	1.69 ± 0.002	678	1.7 ± 0.003	**0.047**
Weight (kg)	1232	79.7 ± 0.49	681	79.6 ± 0.61	0.891
BMI (kg/m^2)	1230	27.6 ± 0.14	677	27.3 ± 0.18	0.212
Systolic BP (mmHg)	1219	129 ± 0.62	674	129 ± 0.837	0.952
Diastolic BP (mmHg)	1219	74.7 ± 0.5	674	74.8 ± 0.34	0.797
LVEF (%)	1128	57.1 ± 0.36	639	57.1 ± 0.49	0.912
Beta blocker use	1244	1067 (85.8%)	691	602 (87.1%)	0.449
Previous MI	1236	379 (30.7%)	686	202 (29.4%)	0.604
Tobacco use (Current, Ex Smoker, Non-smoker)	1244	6.1%, 57.2%, 36.7%	691	6.7%, 57.2%, 36.2%	0.883
Alcohol intake (current, ex-drinker, non-drinker)	1242	62.2%, 13.1%, 24.7%	690	64.1%, 9.9%, 26%	0.103
Plasma analytes					
Cholesterol (mmol/L)	981	4.89 ± 0.03	528	4.9 ± 0.05	0.816
Creatinine (mmol/L)	1201	101 (97.3 – 104)	675	97.8 (94.9-101)	0.265
Urate (mmol/L)	729	0.37 (0.36-0.38)	430	0.37 (0.36-0.38)	0.89
Troponin I (ng/L) $	1187	42.9 (26.5 – 59.1)	661	43.8 (11.4 – 76.1)	**0.048**
ANP (pg/mL) $	1237	43.5 (41.8 – 45.2)	686	44.4 (41.9 – 46.9)	0.966
NT-ANP (pmol/L)$	1237	1.33 (1.27 -1.4)	685	1.38 (1.28 – 1.47)	0.802
BNP (pmol/L) $	1237	26.6 (24.7-28.5)	686	26.8 (24.3-29.4)	0.978
NTproBNP (pg/ml) $	1237	134 (124-144)	686	140 (124-156)	0.979
CNP (pmol/L) $	1215	0.67 (0.64-0.69)	680	0.64 (0.61-0.68)	0.09
NT-CNP (pmol/L)$	1216	23.6 (22-25.1)	676	22.1 (20.7-23.5)	0.352
Aldosterone (pmol/L)$	1206	163 (157-170)	669	172 (150-193)	0.967
Adrenomedullin (pg/ml)$	1191	8.45 (8.18-8.73)	671	8.70 (8.28-9.11)	0.241
VEGF-A (pg/mL) $	341	46.9 (43.1-50.8)	196	42.3 (37.7-46.8)	**0.049**

$Log10 transformed p-values are reported.
Mean ± standard error or Mean (95% CI range) or incidence (%) are reported.
Significantly associated variables and their p-values are shown in **bold.**
Abbreviations: ANP: atrial natriuretic peptide , BP: blood pressure, BMI: body mass index, BNP: B-type natriuretic peptide, CNP: C-type natriuretic peptide, LVEF: left ventricular ejection fraction MELAA = Middle Eastern/Latin American/African, MI = Myocardial infarction, NT-ANP Amino terminal atrial natriuretic peptide, NT-CNP: Amino terminal C-type natriuretic peptide NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = Standard Error of Mean VEGF-A: Vascular endothelial growth factor A.

6.3.3 VLDLR SNP associations within the HVOL cohort

One-way ANOVA was performed to compare the effect of the four imputed SNP genotypes on anthropomorphic and plasma analyte data measurements from the HVOL cohort (Table 6.3). rs10738760 AG and rs2375981 CG genotypes had lower systolic blood pressure ($p = 0.006$; $p = 0.03$), urate ($p = 0.043$; $p = 0.015$), creatinine ($p = 0.012$; $p = 0.035$), and endothelin ($p = 0.012$; $p = 0.006$) levels compared to their respective reference genotype groups. Additionally, the homozygous minor allele genotype groups for both SNPs had lower VEGF-A levels (rs10738760 $p = 0.027$, rs2375981 $p = 4.5 \times 10^{-4}$). Only the rs10738760 GG group had lower levels of high sensitivity troponin I ($p = 0.036$). rs7030781 AT genotype was associated with lower creatinine levels ($p = 0.03$) while the TT genotype was associated with lower endothelin ($p = 0.035$) and VEGF-A levels ($p = 0.034$). The rs7043199 A allele was more frequent in males ($p = 0.037$) and individuals who never smoked ($p = 0.019$).

Table 6.3. Genotype comparison for *VLDLR* SNPS in the HVOL cohort

	rs10738760 Genotypes						
Variable	**n**	**AA**	**n**	**AG**	**n**	**GG**	**p**
Anthropometric variables							
Age (years)	343	68 ± 0.39	598	67.8 ±0.3	236	68.7 ± 0.49	0.264
Male Gender	460	326 (70.9%)	555	368 (66.3%)	162	102 (63%)	0.119
Ethnicity (European, Māori, Indian)	342	98.5%, 1.2%, 0.3%	598	99.2%, 0.8%, 0%	236	98.3%, 1.7%, 0%	0.463
Waist (cm)	343	94 ± 0.63	597	92.8 ± 0.46	236	92.2 ± 0.73	0.123
Weight (kg)	343	78.7 ± 0.76	597	77 ± 0.61	236	76.3 ± 0.87	0.105
BMI (kg/m^2)	342	26.6 ± 0.21	598	26.1 ± 0.16	236	26.1 ± 0.23	0.207
Systolic BP (mmHg)	341	138 ± 1.01	596	135 ± 0.76	236	139 ± 1.33	**0.006**
Diastolic BP (mmHg)	341	81.5 ± 0.57	596	80.2 ± 0.43	236	81.7 ± 0.65	0.083
LVEF (%)	100	64.2 ± 0.45	192	65.1 ± 0.32	79	64.1 ± 0.24	0.152
MI family history	342	29 (8.5%)	598	35 (5.9%)	236	13 (5.5%)	0.226
Tobacco use (Current, Ex Smoker, Never)	343	4.7%, 36.4%, 58.9%	598	4.7%, 40.5%, 54.8%	235	5.5%, 41.7%, 52.8%	0.628
Alcohol intake (current, ex-drinker, non-drinker)	343	80.8%, 4.6%, 14.6%	598	79.6%, 4.7%, 15.7%	236	82.6%, 2.1%, 15.3%	0.502
BP lowering medication	342	87 (25.4%)	598	145 (24.2%)	236	58 (24.6%)	0.92
Plasma analytes							
Cholesterol (mmol/L)	342	4.35 ± 0.06	598	4.26 ± 0.04	236	4.23 ± 0.07	0.339
HDL Cholesterol (mmol/L)	330	1.35 ± 0.01	581	1.38 ± 0.01	228	1.38 ± 0.02	0.364
LDL Cholesterol (mmol/L)	325	3.5 ± 0.05	567	3.46 ± 0.04	220	3.37 ± 0.06	0.331
Urate (mmol/L)	330	0.35 ± 0.004	582	0.34 ± 0.003	228	0.34 ± 0.004	**0.043**
Creatinine (µmol/L)$^\$$	330	91.7 (90.4-93.1)	582	89.4 (88.5-90.3)	228	90.7 (89.3-92.2)	**0.012**
NT-proBNP (pg/mL$^{)\$}$	285	110 (95.3-125)	461	102 (92.6-111)	189	118 (96.4-140)	0.878
Hs TnI (pg/mL$^{)\$}$	285	3.59 (2.93-4.24)	461	4.6 (3.42-5.78)	189	2.97 (2.57-3.36)	**0.036**
ST2 (mg/mL$^{)\$}$	177	33.5 (32-34.8)	296	31.8 (30.7-32.9)	117	31.6 (29.8-33.3)	0.107
BNP (pmol/L)$^\$$	339	6.28 (5.86-6.69)	589	6.26 (5.96-6.57)	234	6.69 (6.07-7.31)	0.865
Endothelin (pmol/L$^{)\$}$	332	2.23 (2.16-2.3)	584	2.09 (2.03-2.15)	230	2.14 (2.05-2.23)	**0.012**
Angiotensin II (pmol/L)$^\$$	315	12.1 (11.4-12.9)	568	12.2 (11.4-13)	217	11.7 (10.8-12.6)	0.553
VEGF-A (pg/mL$^{)\$}$	76	23.5 (16.7-30.2)	101	20.2 (15.6-24.8)	46	11.6 (9.07-14.0)	**0.027**

Variable	n	CC	n	CG	n	GG	p
			rs2375981 genotypes				
Anthropometric variables							
Age (years)	378	67.9 ± 0.37	591	67.9 ± 0.3	208	68.8 ± 0.54	0.304
Male Gender	378	268 (70.9%)	591	398 (67.3%)	208	130 (62.5%)	0.113
Ethnicity (European, Māori, Indian)	377	98.7%, 1.1%, 0.2%	591	99%, 1%, 0%	208	98.6%, 1.4%, 0%	0.665
Waist (cm)	378	94.4 ± 0.59	590	92.7 ± 0.47	208	91.8 ± 0.8	**0.016**
Weight (kg)	378	79.0 ± 0.72	591	77.0 ± 0.61	208	75.7 ± 0.94	**0.018**
BMI (kg/m^2)	377	26.6 ± 0.2	591	26.1 ± 0.16	208	25.9 ± 0.25	**0.048**
Systolic BP (mmHg)	376	137 ± 0.96	589	135 ± 0.76	208	139 ± 1.45	**0.03**
Diastolic BP (mmHg)	376	81.6 ± 0.54	589	80.5 ± 0.43	208	80.9 ± 0.71	0.303
LVEF (%)	113	64.5 ± 0.42	190	65.1 ± 0.32	68	63.9 ± 0.65	0.188
MI family history	377	30 (8%)	591	35 (5.9%)	208	12 (5.8%)	0.405
Tobacco use (Current, Ex Smoker, Never)	378	4.8%, 36.2%, 59%	590	4.4%, 41.4%, 54.2%	208	6.2%, 40.4%, 53.4%	0.431
Alcohol intake (current, ex-drinker, non-drinker)	378	80.7%, 4.8%,14.5%	591	80%, 4.6%, 15.4%	208	81.7%, 1.9%, 16.3%	0.491
BP lowering medication	377	92 (24.4%)	591	149 (25.2%)	208	49 (23.6%)	0.884
Plasma analytes							
Cholesterol (mmol/L)	377	4.39 ± 0.05	591	4.24 ± 0.04	208	4.19 ± 0.08	0.078
HDL Cholesterol (mmol/L)	364	1.35 ± 0.01	573	1.38 ± 0.01	202	1.39 ± 0.02	0.187
LDL Cholesterol (mmol/L)	357	3.52 ± 0.04	559	3.44 ± 0.04	196	3.37 ± 0.06	0.185
Urate (mmol/L)	364	0.35 ± 0.003	574	0.34 ± 0.003	202	0.34 ± 0.005	**0.015**
Creatinine (µmol/L)$^{\$}$	364	91.6 (90.4-92.9)	574	89.6 (88.6-90.5)	202	90.2 (88.7-91.8)	**0.035**
NT-proBNP (pg/mL)$^{\$}$	314	108 (94.7-122)	455	102 (92.3-111)	166	123 (98.5-147)	0.651
Hs TnI (pg/mL)$^{\$}$	320	3.91 (2.9-4.91)	468	4.37 (3.3-5.43)	168	2.96 (2.54-3.39)	0.081
ST2 (mg/mL)$^{\$}$	198	33.5 (32.1-34.9)	289	31.7 (30.6-32.8)	103	31.5 (29.6-33.4)	0.055
BNP (pmol/L)$^{\$}$	374	6.24 (5.85-6.63)	582	6.28 (5.96-6.6)	206	6.77 (6.12-7.42)	0.678
Endothelin (pmol/L)$^{\$}$	366	2.23 (2.16-2.3)	578	2.09 (2.03-2.15)	202	2.14 (2.04 - 2.24)	**0.006**
Angiotensin II (pmol/L)$^{\$}$	347	11.9 (11.3-12.6)	564	12.3 (11.5-13.1)	189	11.7 (10.8 -12.6)	0.784
VEGF-A (pg/mL)$^{\$}$	80	23.6 (17.1-30.1)	104	19.9 (15.4-24.3)	39	10.1 (7.71 – 12.5)	**4.5 x10^{-4}**

rs7030781 Genotypes							
Variable	**n**	**AA**	**n**	**AT**	**n**	**TT**	**p**
Anthropometric variables							
Age (years)	460	67.8 ± 0.34	555	68.3 ± 0.32	162	68.4 ± 0.58	0.475
Male Gender	460	326 (70.9%)	555	368 (66.3%)	162	102 (63%)	0.119
Ethnicity (European, Māori, Indian)	459	98.9%, 0.9%, 0.2%	555	98.7%, 1.3%, 0%	162	98.8%, 1.2%, 0%	0.747
Waist (cm)	460	78.6 ± 0.64	555	76.9 ± 0.64	162	75.9 ± 1.03	0.065
Weight (kg)	460	93.8 ± 0.52	554	92.6 ± 0.5	162	92.4 ± 0.89	0.186
BMI (kg/m^2)	459	26.6 ± 0.18	555	26.0 ± 0.17	162	26.0 ± 0.28	0.092
Systolic BP (mmHg)	457	137 ± 0.86	554	136 ± 0.81	162	139 ± 1.66	0.419
Diastolic BP (mmHg)	457	81.2 ± 0.48	554	80.4 ± 0.45	162	81.7 ± 0.8	0.277
LVEF (%)	141	64.7 ± 0.37	180	64.9 ± 0.33	50	63.9 ± 0.84	0.415
MI family history	459	36 (7.8%)	555	33 (5.9%)	162	8 (4.9%)	0.321
Tobacco use (Current, Ex Smoker, Never)	460	4.8%, 36.1%, 59.1%	554	4.7%, 41.5%, 53.8%	162	5.6%, 42.7%, 51.9%	0.369
Alcohol intake (current, ex-drinker, non-drinker)	460	80.9%, 4.8%, 14.3%	555	79.8%, 4.2%, 16%	162	82.1%, 2.5%, 15.4%	0.717
BP lowering medication	459	113 (24.6%)	555	139 (25%)	162	38 (23.5%)	0.918
Plasma analytes							
Cholesterol (mmol/L)	459	4.34 ± 0.05	555	4.25 ± 0.04	162	4.22 ± 0.08	0.374
HDL Cholesterol (mmol/L)	445	1.35 ± 0.01	537	1.39 ± 0.01	157	1.37 ± 0.02	0.262
LDL Cholesterol (mmol/L)	436	3.48 ± 0.04	523	3.47 ± 0.04	153	3.34 ± 0.07	0.292
Urate (mmol/L)	445	0.35 ± 0.003	538	0.34 ± 0.003	157	0.34 ± 0.005	0.343
Creatinine (μmol/L)[$]	445	91.3 (90.2-92.5)	538	89.4 (88.4-90.4)	157	90.7 (88.9-92.4)	**0.03**
NT-proBNP (pg/mL)[$]	378	110 (97.4-122)	432	104 (93.5-115)	125	115 (88.3-146)	0.733
Hs TNI (pg/mL)[$]	385	4.57 (3.18-5.97)	443	3.78 (3.2-4.35)	128	2.79 (2.37-3.21)	0.053
ST2 (mg/mL)[$]	244	33 (31.8-34.2)	272	31.84 (30.7-33.0)	74	31.4 (28.9-33.8)	0.189
BNP (pmol/L)[$]	454	6.21 (5.86-6.56)	547	6.44 (6.09-6.79)	161	6.46 (5.75-7.16)	0.518
Endothelin (pmol/L)[$]	444	2.21 (2.14-2.27)	545	2.1 (2.04-2.16)	157	2.1 (1.98-2.21)	**0.035**
Angiotensin II (pmol/L)[$]	426	11.9 (11.3-12.5)	527	12.4 (1.51-13.3)	147	11.4 (10.3-12.4)	0.397
VEGF-A (pg/mL)[$]	98	22.7 (17.2-28.2)	99	18.3 (13.8-22.8)	26	12.1 (8.49-15.7)	**0.034**

Variable	n	rs7043199 Genotypes TT	n	TA + AA	p
Anthropometric variables					
Age (years)	763	68.4 ± 0.26	414	67.2 ± 0.36	0.107
Male Gender	763	500 (65.5%)	414	296 (71.5%)	**0.037**
Ethnicity (European, Māori, Indian)	763	99%, 1%, 0%	413	98.5%, 1.2%, 0.3%	0.384
Waist (cm)	763	93.1 ± 0.42	414	93.02 ± 0.55	0.926
Weight (kg)	763	77.4 ± 0.53	414	77.5 ± 0.69	0.927
BMI (kg/m^2)	763	26.3 ± 0.14	413	26.1 ± 0.19	0.367
Systolic BP (mmHg)	760	137 ± 0.68	413	136 ± 0.96	0.374
Diastolic BP (mmHg)	760	81.1 ± 0.37	413	80.7 ± 0.52	0.523
LVEF (%)	248	64.7 ± 0.28	123	64.6 ± 0.46	0.816
MI family history	763	47 (6.2%)	414	30 (7.3%)	0.461
Tobacco use (Current, Ex Smoker, Never)	762	3.9%, 42%, 54.1%	414	6.5%, 35%, 58.5%	**0.019**
Alcohol intake (current, ex-drinker, non-drinker)	763	80.5%, 4.2%, 15.3%	414	80.7%, 4.1%, 15.2%	0.996
BP lowering medication	763	192 (25.2%)	413	98 (23.7%)	0.620
Plasma analytes					
Cholesterol (mmol/L)	737	5.7 ± 0.03	403	5.5 ± 0.06	0.053
HDL Cholesterol (mmol/L)	737	1.38 ± 0.01	402	1.36 ± 0.01	0.275
LDL Cholesterol (mmol/L)	718	3.5 ± 0.03	394	3.38 ± 0.05	0.053
Urate (mmol/L)	737	0.34 ± 0.002	403	0.34 ± 0.003	0.521
Creatinine (µmol/L)$	737	90.3 (89.5-91.1)	403	90.4 (89.3-91.5)	0.855
NT-proBNP (pg/mL)$	609	109 (99.7-118)	326	106 (91.3-121)	0.193
Hs TnI (pg/mL)$	623	4.01 (3.18-4.84)	333	3.89 (2.99-4.79)	0.379
ST2 (mg/mL)$	398	32.3 (31.3 – 33.2)	192	32.3 (30.8-33.8)	0.786
BNP (pmol/L)$	752	6.48 (6.18-6.78)	410	6.12 (5.75 – 6.49)	0.245
Endothelin (pmol/L)$	743	2.15 (2.1-2.2)	403	2.12 (2.05-2.19)	0.513
Angiotensin II (pmol/L)$	715	12.1 (11.6-12.6)	385	12.1 (10.9-13.2)	0.45
VEGF-A (pg/mL)$	133	19.9 (15.5-24.3)	90	18.9 (14.4-23.4)	0.793

$Log10 transformed p-values are reported. Mean ± standard error or Mean (95% CI range) or incidence (%) are reported. Significantly associated variables and their p-values are shown in **bold.** Abbreviations: BP: blood pressure, BMI: body mass index, BNP: B-type natriuretic peptide, HDL: high density lipoprotein, HsTnI: High sensitivity troponin I, LDL: low density lipoprotein LVEF: left ventricular ejection fraction MELAA = Middle Eastern/Latin American/African, MI = Myocardial infarction, , NTproBNP = amino-terminal pro-B type natriuretic peptide, ST2: Suppression of tumorigenicity 2, VEGF-A: Vascular endothelial growth factor A.

6.3.4 Association of genotypes with clinical outcomes in the CDCS cohort

Univariate Kaplan-Meier analyses revealed that none of the SNPs had significant association with any of the clinical endpoints assessed within the CDCS cohort. Based on the genotype associations with analytes measured, multivariate regression models incorporating established predictors (age, gender, ethnicity, previous MI, beta blocker treatment, NTproBNP, creatinine), and the covariates associated with genotype (LVEF, systolic and diastolic blood pressure), were generated for rs2375981 and rs10738760.

Multivariate Cox hazard regression analyses showed that the rs2375981 CG genotype was associated with reduced risk of cardiovascular death (Table 6.4, p = 0.015, HR = 0.7). Models for other outcome endpoints revealed the rs2375981 CG trended towards being a significant predictor of reduced risk of all-cause mortality and STEMI readmission events (Appendix 6.3 and 6.4). rs10738760 AG genotype was associated with reduced risk of presenting cardiovascular death (Table 6.5, p = 0.022, HR = 0.701) and reduced risk of NSTEMI readmission events (Appendix 6.5, p = 0.011, HR = 0.747) when compared to the AA genotype In a model using the GG genotype as reference the rs10738760 AA genotype was associated with increased risk of cardiovascular death (Appendix 6.6, p = 0.047, HR = 1.49). When including blood pressure variables (systolic and diastolic) in the models, the rs2375981 genotype on cardiovascular death (Appendix 6.7) and rs10738760 genotype on NSTEMI (Appendix 6.8) trended towards significance implying potential interactions with either type of blood pressure. The rs7030781 TT genotype was close to being a statistically significant predictor of reduced risk of cardiovascular death (Appendix 6.9).

Table 6.4. Cox's proportional hazards regression model for cardiovascular death in the CDCS cohort using imputed genotypes for rs2375981 (n = 1701, 223 (13.1%) events)

Predictor	Coeff.	SE	Wald	P – value	HR	95% CI for HR	
						Lower	Upper
Gender	0.126	0.162	0.608	0.435	1.13	0.826	1.56
Ethnicity			6.52	0.089			
European v Pasifika	0.555	0.382	2.12	0.146	1.74	0.825	3.68
***European v Asian**	0.885	0.397	4.98	***0.026**	2.42	1.11	5.27
European v MELAA	-7.32	120.9	0.004	0.952	0.001	6.6×10^{-107}	6.5×10^{99}
***Physical Activity (scale 1–4) $$**	-0.259	0.056	21.48	***3.58×10^{-6}**	0.772	0.691	0.861
***Previous MI**	0.741	0.142	27.2	***1.8×10^{-07}**	2.09	1.59	2.77
***Age**	0.052	0.008	37.9	***7.1×10^{-10}**	1.05	1.04	1.07
***Log10 NTproBNP$**	1.79	0.229	61.2	***5.1×10^{-15}**	5.99	3.83	9.38
Beta blocker	-0.168	0.193	0.762	0.383	0.845	0.579	1.23
***Creatinine**	0.002	0.001	5.22	***0.022**	1.002	1.000	1.004
***rs2375981 genotype**			6.82	***0.033**			
***CC v CG**	-0.356	0.147	5.90	***0.015**	0.700	0.526	0.933
CC v GG	-0.375	0.210	3.19	0.074	0.687	0.455	1.037
LVEF	-0.006	0.005	1.27	0.259	0.994	0.984	1.004

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF= left ventricular ejection fraction, MI: Myocardial infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error

Table 6.5. Cox's proportional hazards regression model for cardiovascular death in the CDCS cohort using imputed genotypes for rs10738760 (n = 1667, 211 (12.7%) events)

Predictor	Coeff.	SE	Wald	P – value	HR	95% CI for HR	
						Lower	Upper
Gender	0.134	0.166	0.649	0.421	1.14	0.826	1.58
Ethnicity			6.10	0.107			
European v Pasifika	0.467	0.409	1.31	0.253	1.59	0.716	3.55
***European v Asian**	0.915	0.399	5.25	***0.022**	2.49	1.14	5.45
European v MELAA	-7.47	133	0.003	0.955	0.001	2.9×10^{-117}	1.1×10^{110}
***Physical Activity (scale 1–4) $$**	-0.231	0.058	15.7	***7.1×10^{-6}**	0.794	0.708	0.890
***Previous MI**	0.766	0.147	27.1	***1.9×10^{-7}**	2.15	1.61	2.87
***Age**	0.049	0.009	29.4	***5.8×10^{-8}**	1.05	1.03	1.07
***Log10 NTproBNP$**	1.86	0.239	60.8	***6.1×10^{-15}**	6.46	4.05	10.3
Beta blocker	-0.215	0.195	1.22	0.269	0.806	0.550	1.18
***Creatinine**	0.002	0.001	4.502	***0.034**	1.002	1.000	1.004
***rs10738760 genotype**			6.04	***0.049**			
***AA v AG**	-0.356	0.155	5.26	***0.022**	0.701	0.517	0.950
AA v GG	-0.365	0.207	3.10	0.078	0.694	0.463	1.04
Systolic blood pressure	0.005	0.004	1.53	0.215	1.005	0.997	1.01
Diastolic blood pressure	-0.004	0.007	0.271	0.603	0.996	0.982	1.01

LVEF	-0.008	0.006	1.69	0.192	0.992	0.981	1.004

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF= left ventricular ejection fraction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error

6.3.5 Genomic context of *VLDLR* SNPs

Based on the genome assembly GRCh38/hg38 from the University of California, Santa Cruz Genome Browser (http://genome.ucsc.edu), the Ensembl, JASPAR and HaploReg databases [26-29], the genomic contexts of the 4 VLDLR SNPs were explored to see if DNA binding motifs for relevant effector molecules could be found. The location of the VLDLR SNPs and the relevant transcription factor motifs are shown in Figure 6.1.

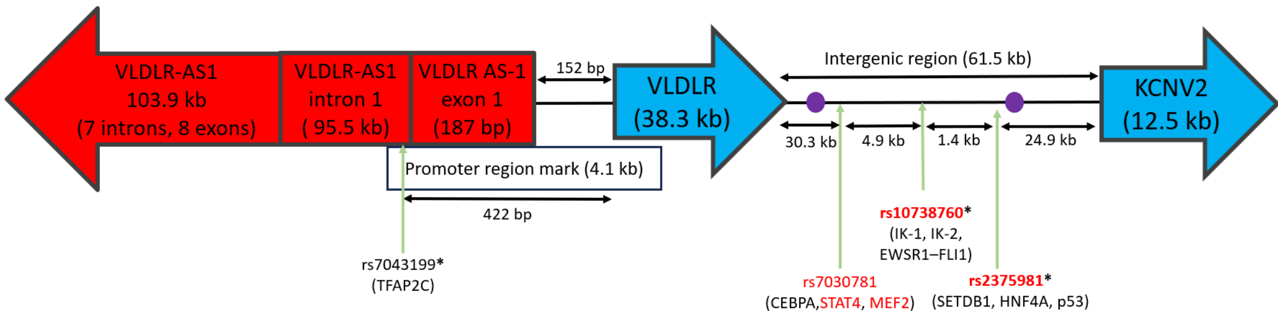


Figure 6.1. Schematic representation of a 312 kb region containing *VLDLR* locus SNPs. rs7043199 was associated with VEGF-A levels in the CDCS cohort, while SNPs shown in red were associated with VEGF-A levels in the heart healthy cohort. Bold SNPs were associated with clinical outcome in the CDCS cohort. Blue arrows represent coding genes for potassium voltage-gated channel modifier subfamily V member 2 (KCNV2) and very low-density lipoprotein receptor (VLDLR) genes. Red arrow indicates VLDLR antisense lncRNA regions. Blue dots represent approximate location of 220 bp CCCTC-binding factor (CTCF) sites. Transcription Factor (TF) motifs overlapping SNPs are named in parentheses. TF motifs in red are linked with cardiovascular risk pathways *Indicates SNPs that map in enhancer regions active in the cardiovascular system. Black double-ended arrows indicate distance between relevant locations (SNPs and/or genes). Distances calculated from NCBI GenBank Accession NC_000009.12.

rs7043199 is located within intron 1 of the long non-coding very low-density lipoprotein receptor antisense RNA 1 (*VLDLR-AS1*) gene on the negative strand of chromosome 9, while *VLDLR* is located on the positive strand ~152 bp from *VLDLR-AS1* (Figure 6.1). Specifically, rs7043199 is located within promoter histone marks that are active in 17 different tissues including the ventricles and atria [28]. Both promoter and enhancer signatures have been linked to influencing gene expression of the *VLDLR*, *VLDLR AS-1* and *KCNV2* genes [26, 30].

rs7030781 is located approximately 30.3 kb 3' from *VLDLR* and 31.2 kb 5' from *KCNV2*. The A allele increases the binding activity of the CCAAT/enhancer binding protein alpha (CEBPA) transcription factor [28]. This factor is expressed in liver and adipocytes, where it regulates adipogenesis which

impacts glucose and lipid homeostasis [31]. Additionally, the rs7030781 A allele is reported to increase the binding of STAT4 [28]. Lastly, the T allele is implicated in increased activity of the myocyte enhancer factor 2A [28].

Both rs10738760 and rs2375981 are located approximately 25 kb 5' from *KCNV2* and 35 kb 3' from *VLDLR,* 1.4 kb apart (Figure 6.1). rs10738760 is within an enhancer region that undergoes epigenetic control via Histone H3 lysine K4 mono-methylation (H3K4me1), which is active in both ventricles, right atrium, adipose nuclei, and foetal heart cells [28]. rs2375981 is associated with enhancer activity and H3K4me1modifications in right atrium and skeletal muscle [28]. The C allele reduces SETDB1 binding[28], a factor involved in structural cardiac defects [32]. The rs2375981 T allele can also reduce hepatocyte nuclear factor 4 alpha (HNF4A) binding activity [28],

6.4 Discussion

There is evidence in the literature that VEGF-A levels are influenced by cis- and trans acting SNPs with some of the variants being associated with increased CVD risk [17]. The present study focused on assessing imputed genotype data for 4 eQTL SNPs (rs7043199, rs7030781, rs10738760 and rs2375981) on the *VLDLR* locus to identify candidate risk variants associated with cardiometabolic parameters and clinical outcomes. In the CDCS cohort rs10738760 GG and rs2375981 GG genotypes were associated with less history of MI, lower blood pressure measurements and lower CNP levels. Moreover, rs10738760 GG and rs7030781 TT genotypes were associated with lower LVEF. Also, rs7043199 A allele carriers had lower VEGF-A levels. Comparatively, in the HVOL cohort, imputed genotype analysis revealed heterozygous genotypes for rs10738760 and rs2375981 were associated with lower systolic blood pressure, urate, creatinine, and endothelin levels. For rs10738760 GG and rs2375981 GG genotypes an association with lower VEGF-A levels was observed.

Regarding circulating VEGF-A levels, a GWAS meta-analysis using ten European cohorts identified the rs7043199 A allele was associated with decreased VEGF-A levels, while the rs2375981 C allele was associated with increased VEGF-A levels [19]. Other GWAS reports on European cohorts identified the rs2375981 G allele is associated with decreased VEGF-A levels [20, 33] In a GWAS study using three healthy cohorts the rs10738760 A allele was associated with increased VEGF-A [18]. However, rs7030781 has contrasting reports on association with VEGF-A levels. One Finnish study identified the rs7030781 T allele was associated with lower VEGF-A levels [20] while another study observed the same allele was associated with higher VEGF-A levels [34]. Our results agree with most

GWAS reports, that chromosome 9 variants are associated with VEGF-A levels. Specifically, we noticed mean VEGF-A levels progressively decreased with the addition of minor alleles in both the clinical (CDCS) and heart healthy (HVOL) cohorts.

Besides VEGF-A, the four chromosome 9 SNPs have been associated with lipid metabolites, blood pressure measurements, inflammatory and growth factors. European GWAS data has shown rs7030781, rs10738760 and rs2375981 are associated with increased circulating interleukin levels including IL-6, IL-10, IL-11, and IL-12 [12, 35, 36]. Additionally, rs10738760 and rs7030781 were associated with increased levels of fibroblast growth factor (FGF) and granulocyte colony-stimulating factor (G-CSF) [35, 36]. Lastly, a study of healthy Greek teenagers associated the rs7043199 A allele with higher systolic and diastolic blood pressure [37]. Our study did not observe an association of rs7043199 with blood pressure in either CDCS or HVOL cohorts. However, we did notice that in the HVOL cohort, SNPs rs7030781, rs2375981 and rs10738760 were associated with VEGF-A, endothelin and (exclusive to rs10738760) troponin levels. The variables we observed have been linked with increased inflammatory molecule activity, blood pressure regulation and lipid profile changes which are CVD pathogenesis factors [3, 10, 38-40]. This study's findings suggest VLDLR SNPs may impact CVD risk by affecting VEGF-A activity, VLDLR's lipid regulation, blood pressure regulation and inflammation.

In terms of risk, rs10738760 has been consistently associated with metabolic syndrome (MetS) biomarkers [41-45]. A population-based study on healthy Lebanese individuals identified the rs10738760 A allele was associated with lower levels of total cholesterol [41]. This study combined their data with a healthy Iranian cohort and, via multiple regression analysis, both studies demonstrated the rs10738760 GA genotype was associated with increased MetS risk [41]. Further studies on the same group of Iranian participants demonstrated rs10738760 AA genotype carriers had higher risk of developing MetS, with the effect being greater in individuals with high fat and sugar intake [42-45]. Additionally, MetS patients carrying the rs10738760 A allele had higher VEGF-A levels [43, 44]. Our data showed the rs10738760 GA and rs2375981 CG may confer reduced risk of cardiovascular death. Additionally, the rs10738760 GA could also be implicated in reduced NSTEMI readmission risk. However, we observed the rs10738760 AA genotype might confer increased risk of cardiovascular death. Both rs10738760 and rs2375981 present a degree of linkage disequilibrium ($r^2 = 0.88$) [28, 46] which can explain the similar associations we observed for metabolites and cardiovascular death associations. Our data reinforces existing findings that rs10738760 is a disease risk variant and, based on our data, it may be a prognostic marker for cardiovascular death.

The SNPs' impact on CVDs may be tied to epigenetic marks considering we observed rs10738760 and rs7043199 were within enhancer or promoter histone marks. Both variants have been linked to DNA methylation activity in healthy Caucasian families [47]. This implies the variants may display an effect on gene regulation that is cis- or trans-acting. In other contexts, some of the chromosome 9 variants may have trans-acting properties with other eQTL VEGF-A variants. A study focusing on Alzheimer disease (AD) highlighted that rs2375981 and rs7043199 present epistatic interaction with VEGF-A related SNPs on chromosome 6 (rs34528081) or chromosome 8 (rs6993770), respectively, that are associated with reduced AD risk [48]. Transcription factors may also be affected considering rs7030781 lies within motifs for STAT4 and MEF2A. STAT4 is actively expressed in heart tissue and can influence CVD onset through regulation of proinflammatory signals [49]. MEF2A is vital for cardiac development, cardiomyocyte differentiation and morphogenesis [50]. Meanwhile, rs2375981 is within the motif for HNF4A, a regulator of hepatic lipid metabolism genes [51] Overall, the individual SNPs may influence other regulatory mechanisms which can impact VEGF-A influence over inflammatory and lipid regulation pathways.

6.5 Conclusion

In summary, the present study made use of imputation data to identify candidate SNPs on chromosome 9 that may influence VEGF-A levels, associate with cardiovascular biomarker levels and/or risk outcome. Our results complement international findings from European cohorts by identifying the minor alleles for 3 SNPs (rs7030781, rs10738760 and rs2375981) in a heart healthy cohort and rs7043199 in a post-ACS cohort are associated with lower VEGF-A levels. Our results suggest there is a high likelihood that the heterozygous genotypes of rs10738760 and rs2375981 can present a protective effect for cardiovascular death and, exclusively for rs10738760, NSTEMI readmission. The rs10738760 AA genotype has a likelihood of increasing mortality risk post-ACS. Our models and the associations with biomarkers and risk factors reported lead us to speculate these VLDLR chromosome 9 variants can impact VEGF-A circulating levels, blood pressure regulation and lipid metabolism. Furthermore, variants rs10738760 rs2375981 and rs7043199 may have trans-acting activities that warrant exploration in further cohorts focusing on CVD onset and progression.

6.6 References

1. Roberts, R., C.C. Chang, and T. Hadley, *Genetic Risk Stratification: A Paradigm Shift in Prevention of Coronary Artery Disease.* Journal of the American College of Cardiology: Basic to Translational Science, 2021. **6**(3): p. 287-304.
2. Wilson, N., et al., *Prioritization of intervention domains to prevent cardiovascular disease: a country-level case study using global burden of disease and local data.* Population Health Metrics, 2023. **21**(1): p. 1.
3. Braile, M., et al., *VEGF-A in Cardiomyocytes and Heart Diseases.* International Journal of Molecular Sciences, 2020. **21**(15): p. 1-18.
4. Au Yeung, S.L., H. Lam, and C.M. Schooling, *Vascular Endothelial Growth Factor, and Ischemic Heart Disease Risk: A Mendelian Randomization Study.* Journal of the American Heart Association, 2017. **6**(8): p. 1-8.
5. Zhang, Y., et al., *Serum vascular endothelial growth factor as a biomarker for prognosis of minor ischemic stroke.* Clinical Neurology and Neurosurgery, 2020. **196**: p. 106060.
6. Rodríguez-Rodríguez, L., et al., *Vascular endothelial growth factor A and cardiovascular disease in rheumatoid arthritis patients.* Tissue Antigens, 2011. **77**(4): p. 291-297.
7. Pagès, G. and J. Pouysségur, *Transcriptional regulation of the Vascular Endothelial Growth Factor gene–a concert of activating factors.* Cardiovascular Research, 2005. **65**(3): p. 564-573.
8. Shaik, F., et al., *Structural Basis for Vascular Endothelial Growth Factor Receptor Activation and Implications for Disease Therapy.* Biomolecules, 2020. **10**(12): p. 1-25.
9. Camaré, C., et al., *Angiogenesis in the atherosclerotic plaque.* Redox biology, 2017. **12**: p. 18-34.
10. Heinonen, S.E., et al., *The effects of VEGF-A on atherosclerosis, lipoprotein profile, and lipoprotein lipase in hyperlipidaemic mouse models.* Cardiovascular Research, 2013. **99**(4): p. 716-23.
11. Zhao, Z., et al., *Association between Single Nucleotide Polymorphisms in Cardiovascular Developmental Critical Genes and Hypertension: A Propensity Score Matching Analysis.* International Journal of Hypertension, 2020. **2020**: p. 1-8.
12. Ruotsalainen, S.E., et al., *An expanded analysis framework for multivariate GWAS connects inflammatory biomarkers to functional variants and disease.* European Journal of Human Genetics, 2021. **29**(2): p. 309-324.
13. Zhao, X., et al., *Vascular endothelial growth factor gene polymorphisms and coronary heart disease: a systematic review and meta-analysis.* Growth Factors, 2018. **36**(3-4): p. 153-163.
14. Ramakrishnan, S., V. Anand, and S. Roy, *Vascular endothelial growth factor signalling in hypoxia and inflammation.* Journal of Neuroimmune Pharmacology, 2014. **9**(2): p. 142-60.
15. Chen, H., et al., *Variation in VEGFA and risk of cardiovascular disease in the UK Biobank.* Frontiers in Cardiovascular Medicine, 2023. **10**: p. 1240288.
16. Lee, H.C., A. Akhmedov, and C.H. Chen, *Spotlight on very-low-density lipoprotein as a driver of cardiometabolic disorders: Implications for disease progression and mechanistic insights.* Frontiers in Cardiovascular Medicine, 2022. **9**: p. 993633.
17. Meza-Alvarado, J.C., et al., *VEGF-A related SNPs: a cardiovascular context.* Frontiers in Cardiovascular Medicine, 2023. **10**.
18. Debette, S., et al., *Identification of cis- and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels.* Circulation Research, 2011. **109**(5): p. 554-63.
19. Choi, S.H., et al., *Six Novel Loci Associated with Circulating VEGF Levels Identified by a Meta-analysis of Genome-Wide Association Studies.* Public Library of Science (PLOS) Genetics, 2016. **12**(2): p. e1005874.
20. Ahola-Olli, A.V., et al., *Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors.* American Journal of Human Genetics, 2017. **100**(1): p. 40-50.
21. Magnifico, M.C., et al., *VLDL Induced Modulation of Nitric Oxide Signalling and Cell Redox Homeostasis in HUVEC.* Oxidative Medicine and Cellular Longevity 2017. **2017**: p. 2697364.
22. Tsai, Y.Y., W.E. Rainey, and W.B. Bollag, *Very low-density lipoprotein (VLDL)-induced signals mediating aldosterone production.* Journal of Endocrinology, 2017. **232**(2): p. R115-r129.

23. Huang, J.K. and H.C. Lee, *Emerging Evidence of Pathological Roles of Very-Low-Density Lipoprotein (VLDL).* Int J Mol Sci, 2022. **23**(8).
24. Varbo, A., J.J. Freiberg, and B.G. Nordestgaard, *Extreme nonfasting remnant cholesterol vs extreme LDL cholesterol as contributors to cardiovascular disease and all-cause mortality in 90000 individuals from the general population.* Clinical Chemistry, 2015. **61**(3): p. 533-43.
25. Perman, J.C., et al., *The VLDL receptor promotes lipotoxicity and increases mortality in mice following an acute myocardial infarction.* Journal of Clinical Investigation, 2011. **121**(7): p. 2625-40.
26. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Research, 2002. **12**(6): p. 996-1006.
27. Martin, F.J., et al., *Ensembl 2023.* Nucleic Acids Research, 2022. **51**(D1): p. D933-D941.
28. Ward, L.D. and M. Kellis, *HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease.* Nucleic Acids Research, 2016. **44**(D1): p. D877-81.
29. Rauluseviciute, I., et al., *JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles.* Nucleic Acids Research, 2023. **52**(D1): p. D174-D182.
30. Fishilevich, S., et al., *GeneHancer: genome-wide integration of enhancers and target genes in GeneCards.* Database (Oxford), 2017. **2017**.
31. Olofsson, L.E., et al., *CCAAT/Enhancer Binding Protein α (C/EBPα) in Adipose Tissue Regulates Genes in Lipid and Glucose Metabolism and a Genetic Variation in C/EBPα Is Associated with Serum Levels of Triglycerides.* The Journal of Clinical Endocrinology & Metabolism, 2008. **93**(12): p. 4880-4886.
32. Markouli, M., D. Strepkos, and C. Piperi, *Structure, Activity and Function of the SETDB1 Protein Methyltransferase.* Life 2021. **11**(8).
33. Gudjonsson, A., et al. *A genome-wide association study of serum proteins reveals shared loci with common diseases*. Nature communications, 2022. **13**, 480 DOI: 10.1038/s41467-021-27850-z.
34. Sliz, E., et al., *Genome-wide association study identifies seven novel loci associating with circulating cytokines and cell adhesion molecules in Finns.* Journal of Medical Genetics, 2019. **56**(9): p. 607-616.
35. Sun, W., et al. *Common Genetic Polymorphisms Influence Blood Biomarker Measurements in COPD*. PLoS genetics, 2016. **12**, e1006011 DOI: 10.1371/journal.pgen.1006011.
36. Nath, A.P., et al., *Multivariate Genome-wide Association Analysis of a Cytokine Network Reveals Variants with Widespread Immune, Haematological, and Cardiometabolic Pleiotropy.* American journal of human genetics, 2019. **105**(6): p. 1076-1090.
37. Kafyra, M., et al., *Associations of VEGF-A-Related Variants with Adolescent Cardiometabolic and Dietary Parameters.* Nutrients, 2023. **15**(8).
38. Pi, X., L. Xie, and C. Patterson, *Emerging Roles of Vascular Endothelium in Metabolic Homeostasis.* Circulation Research, 2018. **123**(4): p. 477-494.
39. Kowalczyk, A., et al., *The role of endothelin-1 and endothelin receptor antagonists in inflammatory response and sepsis.* Archivum Immunologiae et Therapiae Experimentalis (AITE), 2015. **63**(1): p. 41-52.
40. Klinkhammer, B., et al., *TROPONIN CORRELATES WITH INFLAMMATORY MARKERS IN COVID-19.* Journal of the American College of Cardiology, 2021. **77**(18, Supplement 1): p. 3029.
41. Salami, A. and S. El Shamieh, *Association between SNPs of Circulating Vascular Endothelial Growth Factor Levels, Hypercholesterolemia and Metabolic Syndrome.* Medicina, 2019. **55**(8): p. 1-9.
42. Ghazizadeh, H., et al., *Association of rs6921438 A<G with serum vascular endothelial growth factor concentrations in patients with metabolic syndrome.* Gene, 2018. **667**: p. 70-75.
43. Ghazizadeh, H., et al., *Interaction between a genetic variant in vascular endothelial growth factor with dietary intakes in association with the main factors of metabolic syndrome.* Gene Reports, 2020. **21**: p. 100813.
44. Ghazizadeh, H., et al., *Association of a Vascular Endothelial Growth Factor genetic variant with Serum VEGF level in subjects with Metabolic Syndrome.* Gene, 2017. **598**: p. 27-31.
45. Ghazizadeh, H., et al., *Association between serum cell adhesion molecules with hs-CRP, uric acid, and VEGF genetic polymorphisms in subjects with metabolic syndrome.* Mol Biol Rep, 2020. **47**(2): p. 867-875.
46. Machiela, M.J. and S.J. Chanock, *LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants.* Bioinformatics, 2015. **31**(21): p. 3555-7.

47. Gorenjak, V., et al., *Epigenome-wide association study in healthy individuals identifies significant associations with DNA methylation and PBMC extract VEGF-A concentration.* Clinical Epigenetics, 2020. **12**(1): p. 79.
48. Petrelis, A.M., et al., *VEGF-A-related genetic variants protect against Alzheimer's disease.* Aging 2022. **14**(6): p. 2524-2536.
49. Meng, Z.Z., et al., *The pro-inflammatory signalling regulator Stat4 promotes vasculogenesis of great vessels derived from endothelial precursors.* Nature Communications, 2017. **8**: p. 14640.
50. Desjardins, C.A. and F.J. Naya, *The Function of the MEF2 Family of Transcription Factors in Cardiac Development, Cardiogenomics, and Direct Reprogramming.* Journal of Cardiovascular Development and Disease, 2016. **3**(3).
51. Chen, L., et al., *HNF4 Regulates Fatty Acid Oxidation and Is Required for Renewal of Intestinal Stem Cells in Mice.* Gastroenterology, 2020. **158**(4): p. 985-999.e9.

Chapter 7.-Discussion and project conclusions

This chapter aims to discuss overarching aspects of the studies detailed in Chapters 4-6. This includes an overview on the difference in VEGF-A plasma levels between the CVD diagnosed and heart healthy cohorts. Furthermore, this chapter discusses the associations observed for the SNPs in the context of CVD onset and proposes links with potential molecular mechanisms that relate to CVD risk. This addresses research objective 6 (Chapter 1, Section 1.3). To reduce repetition of material due to a Thesis with publications format, the overarching limitations and strengths shared amongst Chapters 4 – 6 are discussed in this chapter. Lastly, the prospective future directions of the conducted research as well as concluding remarks, are included.

7.1 Introduction

The aetiology of CVDs is multifactorial given multiple risk factors are involved in their progression and risk assessment (Chapter 2, Sections 2.1 to 2.3). The typical risk factors that are used in CVD risk assessment are lifestyle and metabolic risk factors which include triglyceride levels, HDL cholesterol, sex, smoking, diabetes, and systolic blood pressure [1]. There is a clinical need for the identification of early predictive biomarkers prior to the onset of CVD symptoms [2]. Particularly, SNPs represent predictive biomarkers that correlate with specific CVD risk factors (Chapter 2, Sections 2.5 and 2.7). SNPs are a novel tool that can support CVD risk prevention and risk stratification [2-5].

A novel candidate risk marker is VEGF-A, a molecule involved in blood vessel formation that can interact with inflammatory and lipid pathways. There is evidence that increased VEGF-A activity can promote inflammation leading to increased vascular dilation, adhesion protein expression and trans-endothelial lipid migration which can promote atherosclerotic lesion development [6-12]. Additionally, elevated plasma and tissue levels of VEGF-A have been observed in CVDs including IHD, CAD, stroke, HF, and MI [6, 13-15]. Moreover, genome wide association studies have identified SNPs, spread across the human genome, that influence circulating VEGF-A level variation [16, 17]. Lastly, some of these SNPs may influence CVD onset or associated risk factors thus offering potential markers for CVD risk [18].

As mentioned in “Chapter 2 - Literature review”, genetic studies of CVD have shifted from focusing on variants close to candidate genes to GWAS data for the purpose of identifying trait associations. Specifically, GWAS data have made it possible to identify VEGF-A expression quantitative trait loci (eQTL) SNPs spread over different human chromosomes. Furthermore, some eQTL VEGF-A SNPs have been linked to CVD risk. This PhD project aimed to investigate associations of SNP candidates with cardiometabolic parameters in New Zealand individuals. The cohorts included in this study are one of post-acute coronary syndrome patients (CDCS) and a heart healthy cohort (HVOL). The project employed statistically generated imputed SNP genotypes (Chapter 3 Material and Methods) for 47 SNPs over three chromosomes. In this manner, imputation-based findings aided in the identification of risk variants that were associated with VEGF-A levels, CVD traits or clinical outcomes (Chapter 4 to Chapter 6). Afterwards, imputation-based results were contrasted with available physically generated SNP data to confirm findings (Chapter 4 and Chapter 5). SNPs that were associated with readmission risk for specific aetiologies were identified (Chapter 4 to Chapter 6). Lastly, each variant’s

genomic context was explored to elucidate potential molecular mechanisms that impact CVD risk. The impact of key CVD-associated variants shall be discussed considering the findings observed and relating them to the available literature.

7.2 VEGF-A levels in the HVOL and CDCS cohorts

As discussed, CVD patients tend to present with elevated levels of plasma VEGF-A and this biomarker had not previously been measured in the HVOL cohort. HVOL participants had mean plasma VEGF-A levels that ranged from ~26.4 pg/mL (rs6921438 AA, Chapter 4) to 10.1 pg/mL (rs2375981 GG, Chapter 6). This range is lower than data previously reported on three longitudinal cohorts studies on VEGF-A levels, where individuals free of chronic disorders had mean plasma VEGF-A levels of 27.4 pg/mL [17]. Comparatively, VEGF-A levels in heart disease cohorts are expected to be higher. For instance, the mean VEGF-A level for the CDCS cohort was 45.0 pg/mL which was previously reported by Palmer et al [19]. Another study focusing on patients with cerebrovascular atherosclerotic stenoses reported mean plasma VEGF levels ranging from 38.1 to 45.0 pg/ml [20]. A trend observed in the research findings was that mean VEGF-A levels were lowest in the homozygous for the minor allele genotype groups. This applied to two *VEGFA locus* variants (rs691438 and rs7767396) for both study cohorts, six *VEGFR2* locus variants (rs1870377, rs1870378, rs1870379, rs7677779, rs13136007 and rs10016064) for the HVOL cohort and four *VLDLR* locus variants in the CDCS (rs7043199) or HVOL (rs10738760, rs2375981 and rs7030781) cohorts.

Overall, the VEGF-A levels observed in this PhD project are within the range of those from studies that have measured circulating VEGF-A levels in healthy and CVD diagnosed individuals. This project's results confirm the variants discussed in Chapters 4 to 6 are associated with VEGF-A levels. Furthermore, specific variants were linked with increased or decreased risk of clinical outcomes. As such, considering their individual molecular mechanisms it is important to discuss how each key variant is linked to VEGF-A activity and/or CVD risk mechanisms.

7.3 Proposed mechanisms for *VEGFA* locus SNPs

As discussed in Chapter 4, *VEGFA* locus variants (rs6921438 and rs7767396) presented association with VEGF-A levels and four clinical outcomes (death, NSTEMI, heart failure, MACE). Specifically, the homozygous minor allele genotypes (rs6921438 AA and rs7767396 GG) are linked to lower VEGF-A levels and higher outcome risk. Furthermore, the research work identified rs6921438 AA,

rs7767396 GG and VEGF-A were associated with increased heart failure readmission risk (Chapter 4, Table 4.6). As part of understanding how each variant affects CVD, their influence can be explored based on their reported allelic effects and tying them to CVD effector molecules.

As discussed in Chapter 2 Section 2.6 (Table 2.2), existing literature shows that rs6921438 A allele is linked to high proinflammatory cytokines (TNF-α, IL-6), low HDL, high LDL, and low VEGF-A levels [10, 13, 16, 21, 22]. The allele's influence over these biomarkers could be due to its effect on overlapping transcription factor binding sites. As discussed in Chapter 4 (section 4.3.6) the HaploReg database showed the A allele of rs6921438 has increased FOXF2 binding activity compared to the reference G allele [23, 24]. FOXF2's wild type function involves cardiac morphogenesis, heart valve formation and endothelial integrity [25]. Loss of FOXF2 function is associated with CVD pathogenesis, however an increase in FOXF2 binding may result in increased inflammatory signalling [23, 25]. This suggests that increased activity of FOXF2 could influence downstream transcription or regulation of inflammatory genes (IL-6, TNF-α), with the latter being well established drivers of immune cell mediated vascular inflammation [12, 26, 27]. From an allelic perspective, rs6921438 A's effect on signalling molecules (e.g. interleukins), may be more complex to the point of involving other SNPs. There is evidence of this given Azimi-Nezhad et al. (2013) identified that the A allele of rs6921438 could interact with a chromosome 8 SNP (rs6993770) and a chromosome 6 SNP (rs4416670,~24.8 kb downstream of rs6921438) to increase IL-6 levels and TNF-α levels [22]. This would imply that inflammation may be a common pathway through which SNPs might influence VEGF-A levels. In turn, elevated VEGF-A levels are known to affect lipoprotein activity [8, 28, 29]. Specifically, VEGF-A is capable of binding to HDL, which would lower its circulating levels and disrupt lipid transport allowing the accumulation of LDL and triglycerides [30]. Additionally, VEGF-A has been identified to lower lipoprotein lipase (LPL) activity, which results in higher circulating LDL levels [30, 31]. Considering low VEGF-A levels in the presence of rs6921438 A, it is plausible that VEGF-A activity is more oriented towards interacting with LPL and HDL. This could explain why circulating VEGF-A levels are lower despite increased production due to proinflammatory cytokines. Given no significant associations between rs6921438 and natriuretic peptides or disease risk were observed in 'heart healthy' individuals, rs6921438 A may require other factors (e.g. hypoxia, dyslipidaemia) to promote CVD onset. In a healthy group. there may be low background activity of IL-6 or TNF-α as well as adequate blood oxygen levels which implies there would not be increased VEGF-A production. If hypoxia or alterations to lipid levels were to occur these could be reduced due to HDL antioxidant activity. Overall, this project identified that rs6921438 A is associated with higher heart failure readmission. However, there is evidence on the crosstalk between VEGF-A, inflammation

and lipoprotein homeostasis that link them to CVD risk. Therefore, inflammatory biomarkers may directly impact VEGF-A levels which can promote a disturbed lipoprotein balance, conjointly leading to CVD onset. The proposed mechanism of action for rs6921438 A, as well as current evidence on VEGF-A, inflammation and lipoprotein interactions, discussed in this paragraph, are summarised in Figure 7.1A.

In this project, the rs7767396 G allele was associated with lower VEGF-A levels and increased disease risk (Chapter 4). In the literature, the rs7767396 G allele has been reported to be associated with lower VEGF-A levels and reducing NF-AT1 binding [24, 32]. NF-AT1 is a transcription factor involved in signalling that induces proinflammatory cytokine expression [33]. A potential explanation for lower VEGF-A levels could be due to the G allele reducing proinflammatory cytokine production. Another transcription factor with binding affected by rs7767396 is STAT3 (Chapter 4, Section 4.3.6). STAT3 upregulates hypoxia responsive factor 1 alpha (HIF-1α), a well-established inducer of VEGF-A production [34]. The rs7767396 G allele overlaps with two motifs that regulate pathways (inflammation and hypoxia) involved in VEGF-A production. Theoretically, reduced binding of both motifs could decrease hypoxia and inflammation which would imply lower VEGF-A production. Therefore, rs7767396 G allele's association with low VEGF-A levels may be due to reduced STAT3 activity leading to lower HIF-1α activity and lower NF-AT1 binding. Based on its location rs7767396 appears to have more direct connections to VEGF-A levels, when compared to rs6921438. Due to its links to NF-AT1 and STAT3, rs7767396 may have a more robust effect on disease risk. Additionally, NF-AT1 and STAT3 are involved in heart development gene regulation [33, 34]. The proposed mechanism of action for rs7767396 and potential connections of the transcription factors with effector mechanisms are summarised in Figure 7.1B.

Although both variants have plausible links to VEGF-A regulation and downstream effects, due to the nature of this project, it is difficult to identify which variant has a more pronounced driving effect. Based on the pathways that it might connect with, rs6921438 could have an influence on CVD developmental mechanisms. Meanwhile rs7767396 may be tied more to an immediate susceptibility of disease outcome, in addition to altering VEGF-A levels. Although both variants are relatively close, from a genetic perspective, they interact with different regulators which affect overlapping effector molecules that contribute to CVD risk. Notably, the work presented agrees with reports on rs6921438's influence on VEGF-A levels, while displaying its complex interaction with inflammatory molecules and lipid regulation. Moreover, rs7767396 represents a novel VEGF-A eQTL SNP with potentially a more direct effect on VEGF-A. Additionally, its impact on disease risk could be due to interaction with

transcription factors that regulate inflammation and heart development genes. Overall, rs6921438 and rs7767396 are intergenic variants displaying trans-acting properties that impact CVD risk. Both variants should be considered as novel genetic markers to aid in NZ based CVD risk assessment.

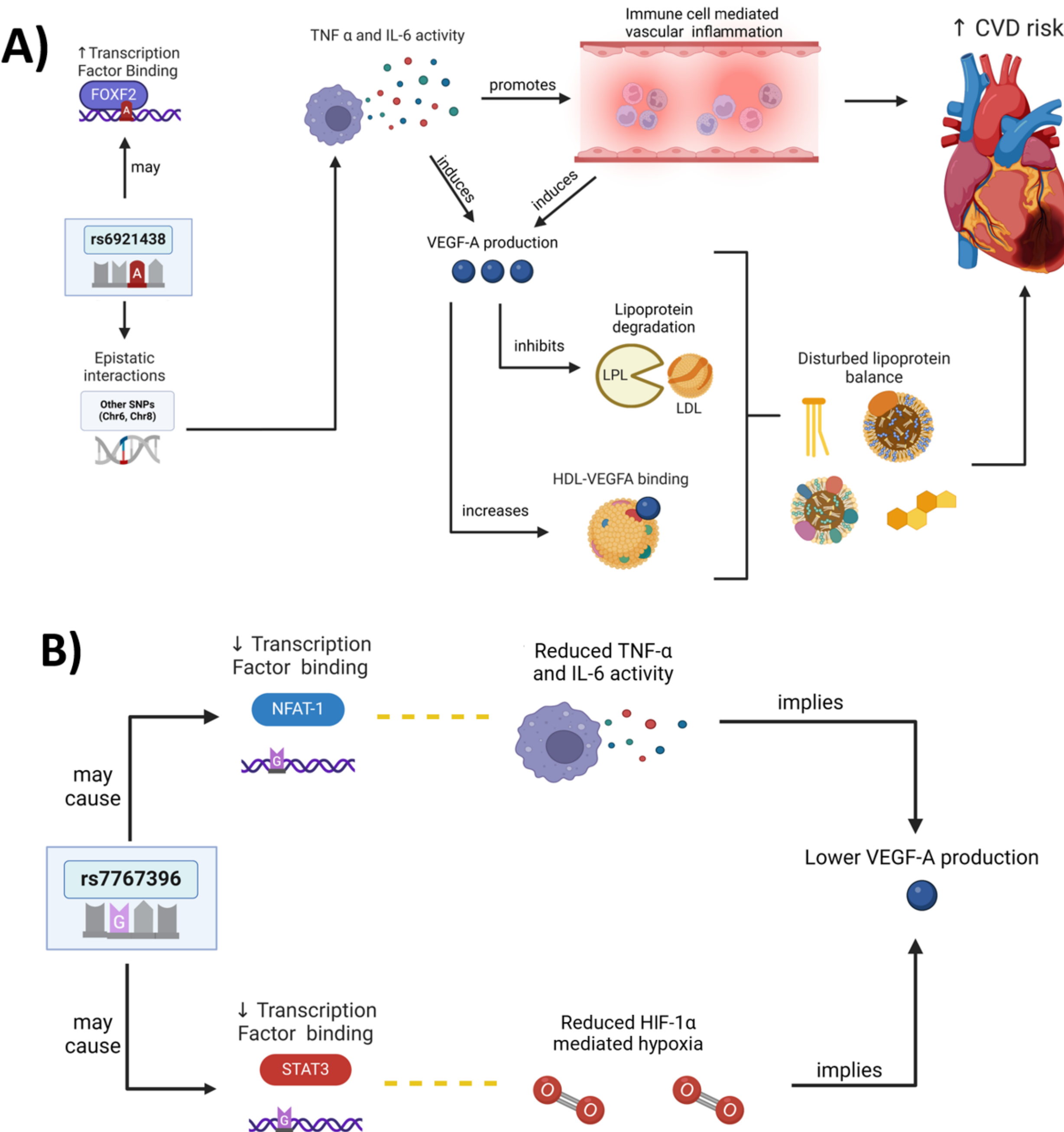


Figure 7.1. Proposed mechanisms of action for A) rs6921438 and B) rs7767396. Dashed yellow lines represent potential connections of the SNP effect with effector mechanisms regulated by the specified transcription factors. Black arrows represent effects referenced in text. Abbreviations: Chr: Chromosome, CVD: Cardiovascular Disease, HDL: High density lipoprotein, HIF-1α: Hypoxia inducible factor-1 alpha, IL-6: Interleukin 6, LDL: Low density lipoprotein, LPL: Lipoprotein lipase, SNP: Single nucleotide polymorphism, TNF-α: Tumour necrosis factor alpha. VEGF-A: Vascular endothelial growth factor A. Created with BioRender.com.

7.4 Proposed mechanisms for *VEGFR2* locus SNPs

VEGFR2 represents one of the main receptors that when activated, via VEGF-A or other non-VEGF agonists, promotes vascular permeability and angiogenesis (Chapter 5). Upon canvassing the literature (Chapter 2), VEGF-A eQTL SNPs reported on human chromosome 4 include two exonic SNPs in *VEGFR2* (rs1870377 and rs2305948) and two *VEGFR2* promoter region SNPs (rs7667298 and rs2071559). The work presented in Chapter 5 aimed to assess these SNPs and other variants in the *VEGFR2* gene to determine their involvement in CVD.

CDCS patients with the rs2305948 CC genotype had higher incidence of CVD readmissions and mortality events (Chapter 5). Evidence in the literature showed that the rs2305948 T allele is associated with high LDL and cholesterol levels [35]. Moreover, the rs2305948 T allele participates in decreased VEGF-A binding to VEGFR2 which can cause an increase of sVEGFR2 circulating levels [36, 37]. This project did not observe association between rs2305948 and VEGF-A levels, but the T allele could be a disease risk variant candidate (Chapter 5). This highlights rs2305948's role in CVD risk with potential influence on VEGF-A effector functions which indirectly give it eQTL properties. Following an ACS event, the lack of VEGF-A binding caused by rs2305948 would allow for an increase in circulating VEGF-A levels that could induce other pathways that lead to increased CVD risk. An affected pathway includes lipid catabolism caused by VEGF-A reducing plasma LPL activity [8, 28, 29] leading to circulating LDL levels, which agrees with rs2305948's T allele associations with higher lipid content. However, the role for the observed risk genotype (rs2305948 CC) following ACS should be elucidated further since the C allele has been associated with reduced circulating sVEGFR2 levels [38, 39]. Notably, VEGF binding to VEGFR2 stimulates signalling pathways that cause vascular dilation [6, 12]. Therefore, if there is reduced VEGFR2 activity and/or low circulating VEGFR2 there may be impaired vasodilation. This would then lead to hypertension which promotes CVD risk. The pathways and molecules associated with rs2305948, VEGF-A and VEGFR-2 function in ACS patients are summarised in Figure 7.2A.

When looking deeper into how each rs2305948 allele could affect healthy individuals, the SNP may play a role in CVD risk factors via HDL or endothelin. This project noted that rs2305948 T allele carriers had higher HDL and endothelin levels (Chapter 5, Section 5.3.4). It is possible that low VEGF-A binding to VEGFR2 due to rs2305948 T would activate compensatory mechanisms to maintain homeostasis. This could affect downstream effects such as VEGFR2 mediated in vasodilation [6, 12]

while unbound circulating VEGF-A could promote a localised inflammatory response [22]. HDL is well established for presenting anti-inflammatory and antioxidant properties which would reduce the possibility of circulating LDLs being oxidised [40, 41]. Moreover, HDL also displays vasodilation properties [41] which could complement VEGFR2-mediated vasodilation. Additionally, since endothelin is a vasoconstrictor [42], it may be involved along with HDL in modulating endothelial function. Overall, endothelin, HDL and VEGFR2 would modulate vasodilation in healthy individuals with rs2305948 being associated with increased HDL and endothelin levels. However, the variant's direct effect on blood pressure regulation may be superseded by the consequences of VEGF-A dysfunction. This could be further complicated considering the rs2305948 T allele has been reported to present trans-acting properties with other SNPs. For example, it can increase haemorrhagic stroke risk by interacting with rs833061 that is in the promoter region of VEGF-A on chromosome 6 [43]. Additionally, it has been reported that rs2305948 TT, rs3025039 CT and r1870377 AA are associated with reduced CHD risk when adjusting for risk factors such as smoking, alcohol intake and diabetes [44]. Overall, rs2305948 can be considered a VEGF-A eQTL variant, but it may have more complicated interactions that relate to downstream mechanisms involved in blood pressure regulation or inflammation influenced by VEGFR2 or VEGF-A, respectively. Associations observed in this project for rs2305948 in healthy individuals and effector mechanisms discussed are summarised in Figure 7.2B.

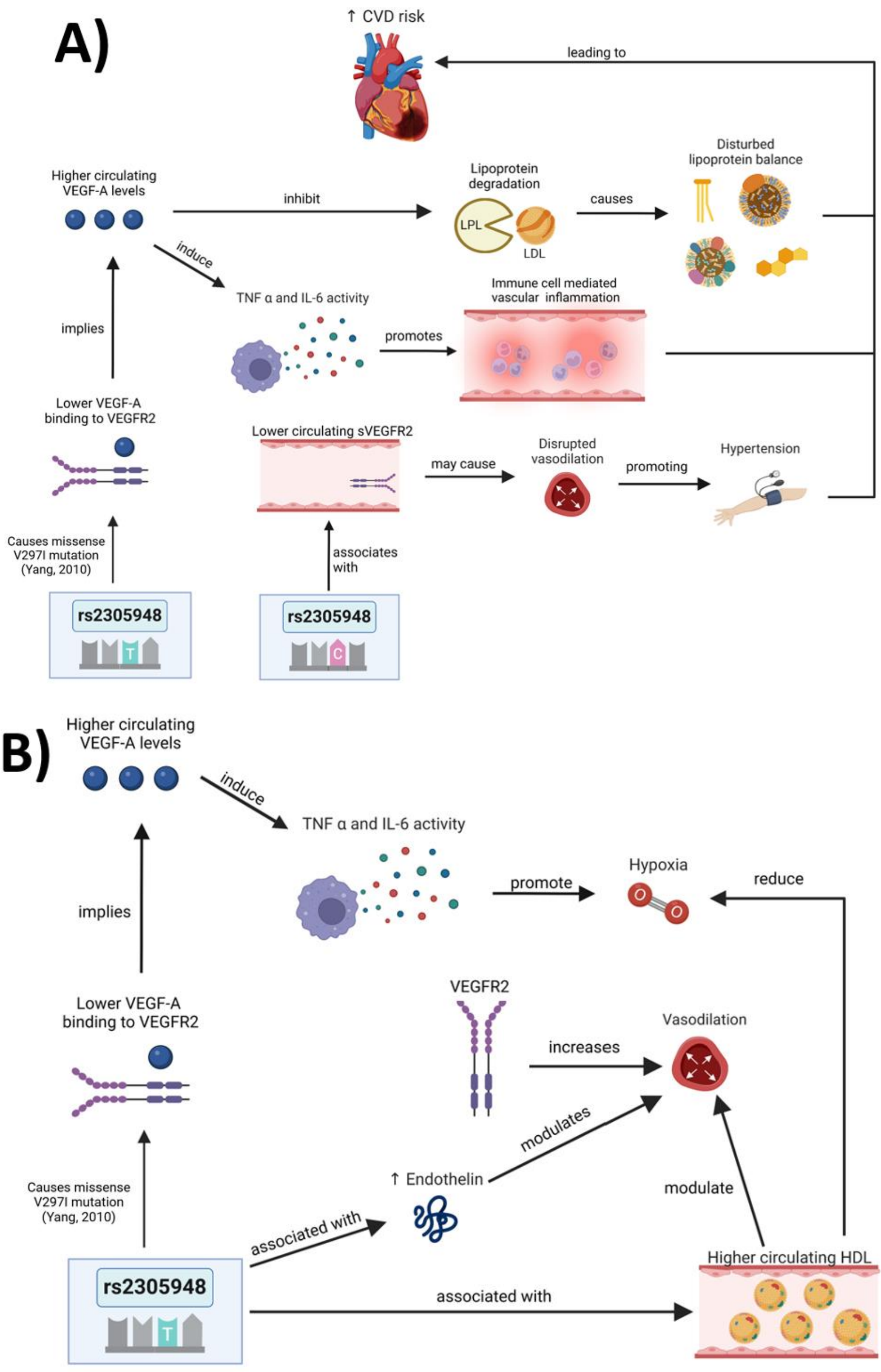


Figure 7.2. Associations of rs2305948 alleles with effector molecules and subsequent impact on risk factors that A) promote CVD risk or B) contribute to heart vascular homeostasis. Abbreviations: Chr: Chromosome, CVD: Cardiovascular Disease, HDL: High density lipoprotein, IL-6: Interleukin 6, LDL: Low density lipoprotein, LPL: Lipoprotein lipase, sVEGFR2: Soluble: vascular endothelial growth factor receptor 2, TNF-α: Tumour necrosis factor alpha. VEGF-A: Vascular endothelial growth factor A, VEGFR2: vascular endothelial growth factor receptor 2. Created with BioRender.com.

The second variant in the study of *VEGFR2* locus SNPs with notable associations to biomarkers was rs1870377 (Chapter 5). Specifically, in the CDCS cohort, rs1870377 A allele carriers had significantly higher risk of CVD readmissions (Section 5.3.5). Existing evidence shows that the rs1870377 A allele is associated with low circulating VEGFR2, higher circulating LDL and sVEGFR1 levels [45-47]. Particularly, sVEGFR1 inhibits VEGF-A angiogenic activity by binding to circulating VEGF-A and this can attenuate VEGFR2 expression and signalling [48]. Additionally, rs1870377 causes a missense mutation (Q472H) which increases VEGFR2 phosphorylation, promoting increased VEGF-A binding efficiency [49-51]. This implies that rs1870377 is connected to both VEGF receptors and their subsequent impact on VEGF-A activity or levels. However, the hypoxic and/or inflammatory conditions during CVD onset could induce greater signalling towards CVD development than the VEGF receptors. Additionally, any unbound VEGF-A could also contribute to reduced lipoprotein lipase activity leading to increased LDL levels or induction of cytokine signalling that promote CVD development. Overall, rs1870377 A appears to influence CVD onset through its mechanisms on sVEGFR1 and VEGFR2-mediated VEGF-A activity. This interaction may explain the disease risk behaviour observed in this project for the rs1870377 AA genotype (Chapter 5). The associations for rs1870377 A and interactions of molecules that contribute to CVD risk are summarised in Figure 7.3A.

In the heart healthy cohort, the rs1870377 A allele was associated with lower triglyceride and VEGF-A levels (Chapter 5). For healthy individuals, the lack of an ongoing hypoxic and inflammatory environment would allow VEGF-A to bind to VEGFR2 to carry out its effector functions. Additionally, it could be possible that regulation of VEGF-A activity can be modulated by background production of sVEGFR1. This implies low circulating VEGF-A levels in healthy individuals would be observed due to VEGF-A binding to either sVEGFR1 or VEGFR2. Moreover, VEGFA activity would be primarily focused on angiogenesis or artery maintenance implying little to no interference on lipoprotein lipase activity, which could contribute to lower triglyceride levels. Associations observed in this study and potential links to low LDL levels are summarised in Figure 7.3B

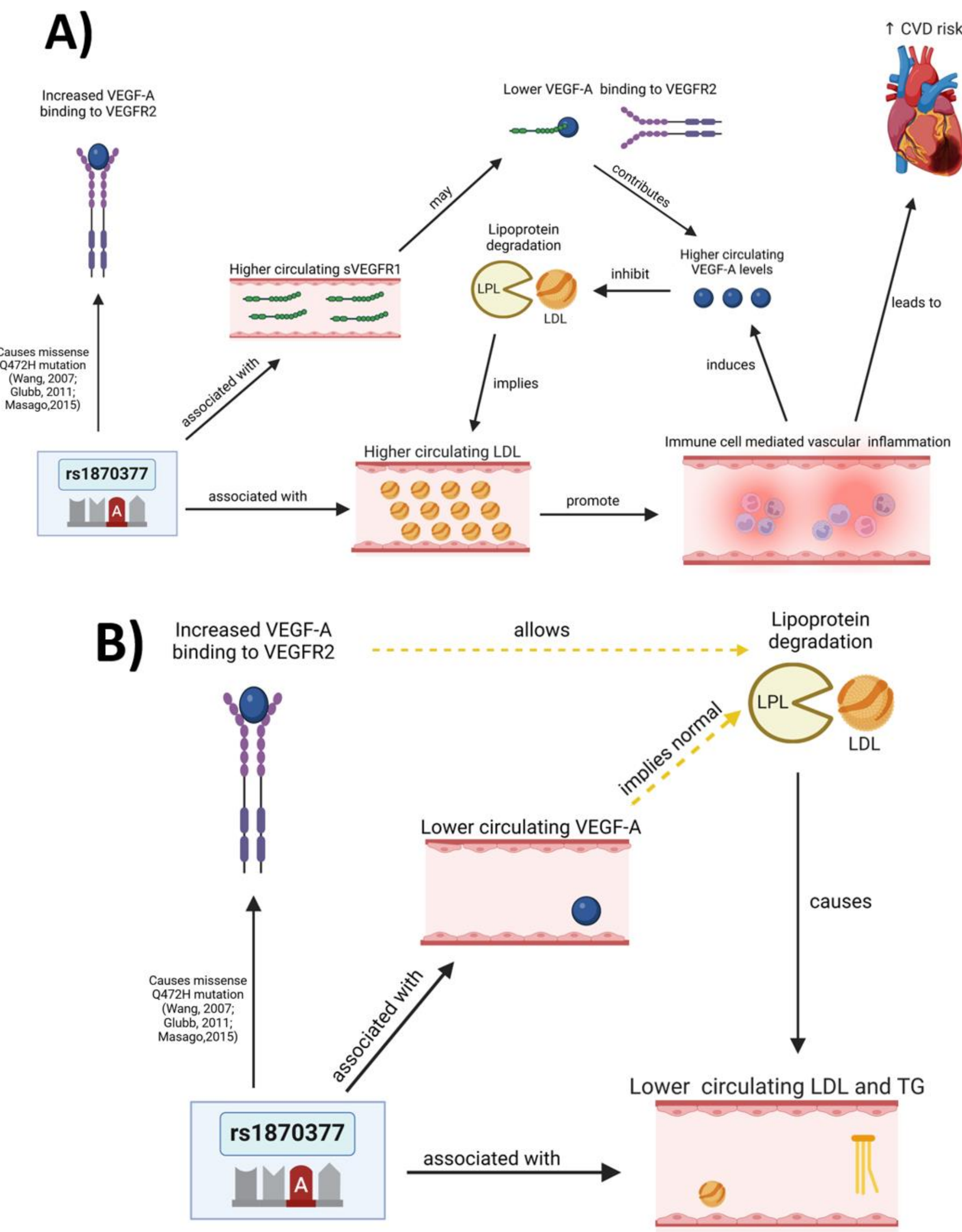


Figure 7.3. Graphical summary of reported associations of rs1870377 A and mechanisms that contribute to A) CVD risk or B) low LDL levels. Yellow arrows represent potential connections between VEGF-A and lipoprotein degradation. Black arrows represent effects referenced in text. Abbreviations: CVD: Cardiovascular Disease, LDL: Low density lipoprotein, LPL: Lipoprotein lipase, sVEGFR1: Soluble vascular endothelial growth factor receptor 1, sVEGFR2: Soluble: vascular endothelial growth factor receptor 2, TG: triglycerides, VEGF-A: Vascular endothelial growth factor A, VEGFR2: vascular endothelial growth factor receptor 2. Created with BioRender.com.

This project also identified two intronic SNPs (rs767779 and rs13136007) that have the potential of being candidates to assess mortality risk (Appendices 5.7 and 5.8). For rs7677779, a potential mechanism may involve interaction with overlapping transcription factor binding motifs. This includes transcription factors ELF2 and ELF4 which are involved in immune response regulation. Although no significant associations were observed with plasma analytes, rs7677779 T could promote immune cell migration and signalling that could exacerbate early CVD onset. Additionally, ELF4 is involved in adipogenesis regulation [52] making it possible that rs7677779 T's involvement in CVD could be due to increased lipid production and subsequent foam cell formation. Meanwhile rs13136007 overlaps with the binding motif for HNF6, a factor that contributes to the expression of genes involved in glucose and lipid metabolism [53, 54]. This would also allow it to impact the lipidic landscape to contribute to CVD onset. Considering rs7677779 T and rs13136007 A alleles were associated with triglyceride levels in the heart healthy cohort, the transcription factors could be downregulated in this context. As for their association with VEGF-A levels, it may be due to the interplay between VEGF-A and lipids. In CVD onset, they might play a role through lipid-VEGF-A interaction, but their overall effect could be overshadowed by the exonic variants discussed.

The main findings of this study, in conjunction with the genomic context mentioned previously, agree that the most influential SNPs on the VEGF system and cardiovascular risk are rs2305948 and rs1870377. These variants may also contribute to the relationship between VEGF-A, blood pressure regulation and lipid catabolism. Additionally, there is potential for two intronic SNPs at the *VEGFR2* locus to be candidates for future studies involving lipid regulation. Given the additive, and cis- and trans-acting nature of SNPs, the study of additional VEGFR2 SNPs is vital to identify non-coding variants to complement research on coding variants. Altogether this project may complement a wider understanding on different VEGFR2 SNPs by looking into VEGF-A regulation and the overall effect on established CVD risk factors.

7.5 Proposed mechanisms for *VLDLR* locus SNPs

Different research groups have identified four VEGF-A eQTL variants at the *VLDLR* locus of chromosome 9 [16, 17, 35]. The findings on these variants (Chapter 6) for the CDCS cohort showed associations between homozygous minor allele genotypes and cardiometabolic parameters. rs10738760 GG and rs2375981 GG were associated with lower previous MI occurrence, lower systolic BP, lower diastolic BP and lower NT-CNP levels. Meanwhile rs10738760 GG and rs7030781 TT were associated with lower LVEF. Lastly, rs7043199 A allele carriers had lower VEGF-A levels and

rs7043199 TT was associated with slightly lower height and troponin I levels. Furthermore, rs2375981 CG and 10738760 AG were associated with reduced risk of cardiovascular death. As for the heart healthy cohort, rs10738760 GG, rs2375981 GG and rs7030781 TT were associated with lower VEGF-A levels.

Based on the associations observed with both study cohorts, there is a possibility that the G alleles for rs2375981 and 10738760 have protective properties. The variants' associations with urate and creatinine in healthy individuals might suggest there is no apparent metabolic dysfunction attributed to CVD. Additionally, by being associated with low systolic BP and low endothelin, vasodilation may occur which could contribute to minimal blood flow stress. Based on the CDCS associations, the effect on reduced death risk by either variant's G allele appears to confirm the variants' association with blood pressure regulation. This coincides with the observed low occurrence of previous MI events for the rs2375981 GG and 10738760 GG genotype groups. Following an ACS event, it would be expected to have elevated levels of natriuretic peptides due to cardiac stress. The variants' association with low NT-CNP could imply minimal cardiac stress, but this could be attributed to individuals having recovered by the time samples were taken in the clinic.

The rs10738760 AA genotype was associated with increased risk of cardiovascular death in the CDCS cohort (Chapter 6). The AA genotype group showed elevated BP and MI occurrence. Contrastingly the GG genotype presented lower levels of the same measurements. Moreover, the multivariate hazards models showed that beta blocker treatment was not independent from the SNP, further implying an effect through blood pressure regulation. Given rs10738760's intergenic location, the variant might have trans-acting properties on neighbouring genes. A gene of interest could be KCNV2 (located ~25kb downstream from rs10738760) which codes for a potassium channel that is predominantly expressed in the heart and retina [56, 57]. Research has primarily focused on KCNV2's role in retinopathy where it has been clearly linked with altered electrophysiological changes [58]. In heart tissue, potassium voltage-gated ion channels are critical for mediating cardiac electrophysiological signals that maintain heart contraction [57, 59]. As such, rs10738760 A could contribute to KCNV2 expression or activity which would link it with observed associations with systolic function measurements (LVEF and SBP) and cardiovascular death risk (Chapter 6). Notably, cardiovascular death involves having a heart condition (e.g. MI, arrythmia, heart failure) or vascular conditions (e.g. stroke, coronary artery disease, peripheral vascular disease). These scenarios can be, in part, caused by heart dysfunction or blood pressure changes, respectively. Overall, this potential interaction with KCNV2 can complement reports on this variant's association with metabolic syndrome (MetS) risk

[60-64]. A summary on the proposed link of rs1073870 A with KCNV2 and the variant's association with MetS, in the context of CVD risk is presented in Figure 7.4.

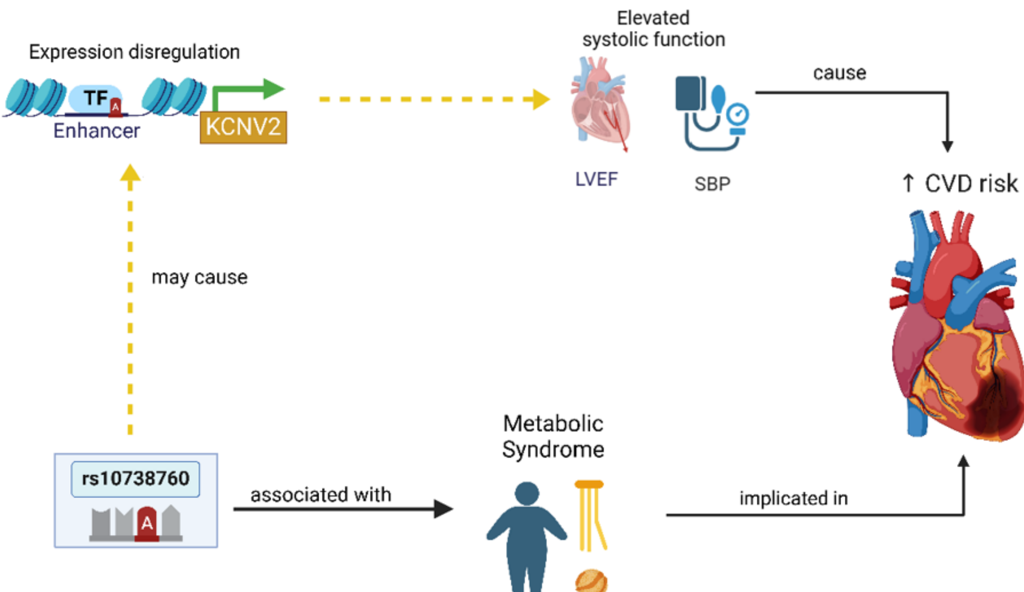


Figure 7.4. Proposed links between rs10738760 A allele and CVD risk. Yellow arrows represent potential connections between KCNV2 and systolic function. Black arrows represent effects referenced in text. Abbreviations: CVD: Cardiovascular Disease, LVEF: left ventricular ejection fraction, SBP: systolic blood pressure. Created with BioRender.com.

As part of this project, rs2375981 CG was observed to be associated with reduced cardiovascular death risk and lower BP in the CDCS cohort (Chapter 6). In a heart healthy cohort, the rs2375981 GG genotype was associated with lower VEGF-A levels, while the CG group presented low urate, creatinine, endothelin, and systolic BP levels (Chapter 6, Section 6.3.3). These findings suggest the allelic variation between rs2375981 G and C alleles may have different effects. Specifically, HaploReg data reports that the rs2375981 C allele can reduce SETDB1 binding compared to the G allele [24]. SETDB1 participates in gene silencing through histidine residue methylation, and it can suppress expression of proinflammatory mediators including IL-6 and IL-12 [65]. There is a possibility that the protective effect observed for rs2375981 G may be due to SETDB1 binding reducing inflammatory molecule activity via methylation. In turn, a decrease of inflammation would not increase VEGF-A levels. The G allele could also modify KCNV2 expression due to chromatin modification, potentially mediated by SETBDI, with systolic function being possibly affected given the role of potassium voltage-gated ion channels in the heart [59].

Another potential target is VLDLR upregulation which would assist in the clearance of circulating LDLs thereby contributing to reducing CVD risk [66, 67]. There are additional examples that relate rs2375981 with trans-acting properties by interacting with other variants affecting lipid metabolism genes. For example, a case-control study on Alzheimer's disease (AD) identified that rs2375981. the

rs2375981 GC genotype was associated with reduced AD risk by interacting with chromosome 18 variant rs2639990 [68]. Specifically, rs2639990 is located in intron 1 of prostaglandin reductase 3, a gene that negatively modulates adipogenesis [69]. However, the rs2639990 T allele is associated with lower VEGF-A levels [69], implying it could interact with rs2375981 to increase VEGF-A levels and subsequent negative effects. Overall, rs2375981 G may contribute to reducing CVD risk by interacting with lipid metabolism and blood pressure regulation. Meanwhile, it is plausible rs2375981 may affect circulating VEGF-A levels by interacting with other SNPs through other unknown mechanisms. A summary of the molecular mechanisms and potential links of each rs2375981 allele are summarised in Figure 7.5.

An additional finding of this study was rs7043199 A allele carriers presenting low VEGF-A levels in the CDCS cohort (Chapter 6, Section 6.3.2). Furthermore, rs7043199 overlaps with promoter, enhancer and histone marks that have been linked to the *VLDLR*, *VLDLR AS-1* and *KCNV2* genes [70, 71]. Specifically, the non-coding RNA for very low-density lipoprotein receptor antisense RNA 1 (*VLDLR-AS1*) has unknown functionality but it can regulate gene expression through different mechanisms including chromatin remodelling, transcriptional or post-transcriptional regulation [72, 73]. This suggests the rs7043199 has an impact on regulatory mechanisms that affect other pathways (e.g. lipid metabolism) which could modify VEGF-A levels. In a CVD context, the histone marks could be active due to rs7043199 A which could allow VLDLR-AS1 expression. This antisense RNA could influence the established interaction between VEGF-A and LDL [8, 28, 29]. Additionally, lipoprotein changes could induce subsequent inflammation and hypoxia that may increase VEGF-A levels. VEGF-A activity could then focus on disrupting LPL activity thus contributing to CVD onset. Oppositely, the rs7043199 T allele could promote VLDLR-AS1 activity allowing proper VLDLR functionality which, in theory, would not promote CVD onset mechanisms (e.g. cytokine signalling) or cause an increase of VEGF-A levels. These are potential pathways that would link the variant with lipoprotein regulation that may affect circulating VEGF-A levels. Figure 7.6 summarises potential links between each rs7043199 allele with VLDLR-AS1 regulation and downstream involvement with VEGF-A – lipid interaction.

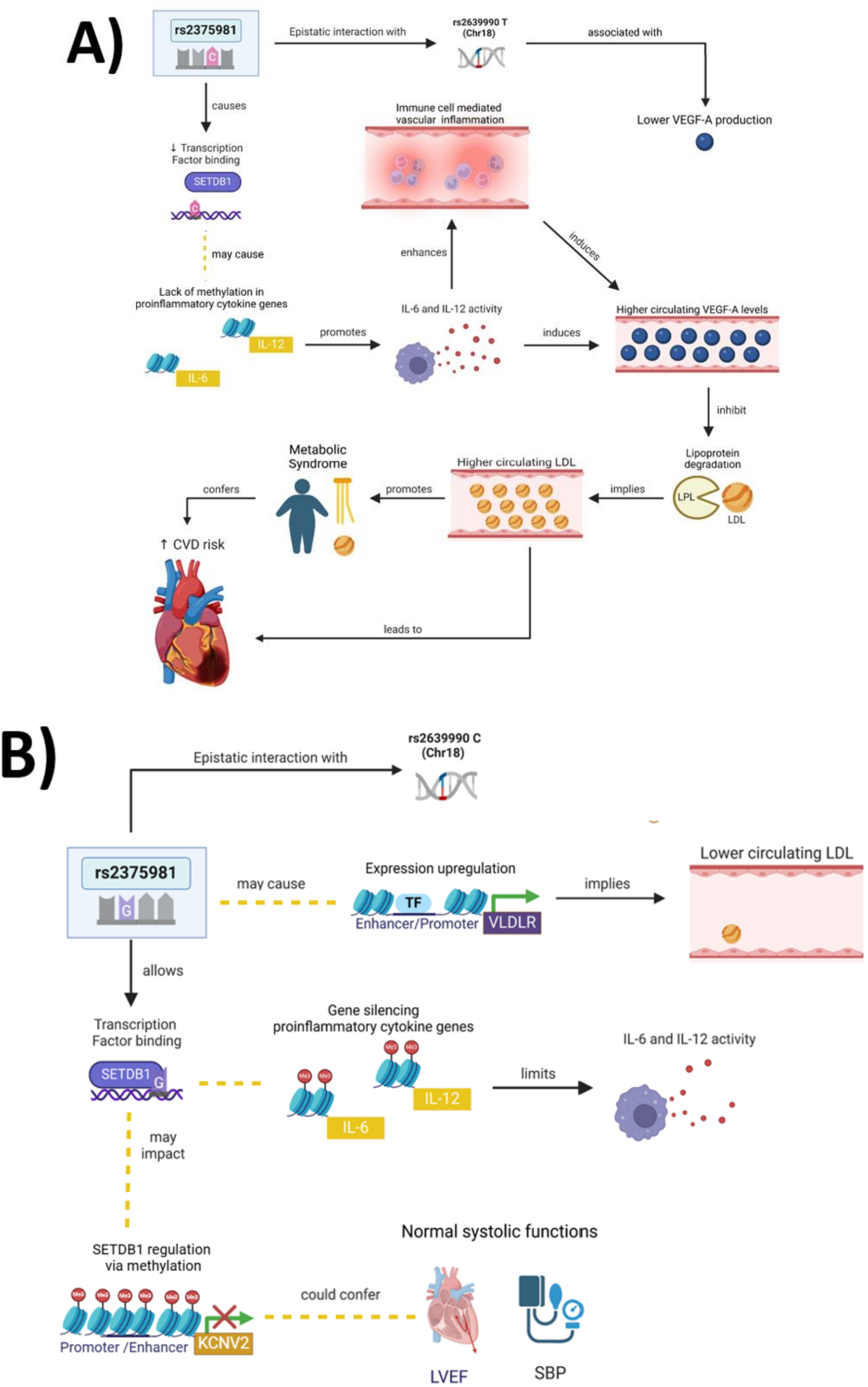


Figure 7.5. Reported interactions for rs2375981 and potential effector functions in the context of A) risk and B) protective alleles. Yellow lines represent potential connections of the SNP's altered binding with other molecular mechanisms. Black arrows represent processes referenced in text. Abbreviations: Chr: Chromosome, CVD: Cardiovascular Disease, IL-6: Interleukin 6, IL-12: Interleukin 12, KCNV2: Potassium Voltage-Gated Channel Modifier Subfamily V Member 2, LDL: Low density lipoprotein, LPL: Lipoprotein lipase, LVEF: left ventricular ejection fraction, SBP: systolic blood pressure, VEGF-A: Vascular endothelial growth factor A, VLDLR: very low density lipoprotein receptor. Created with BioRender.com.

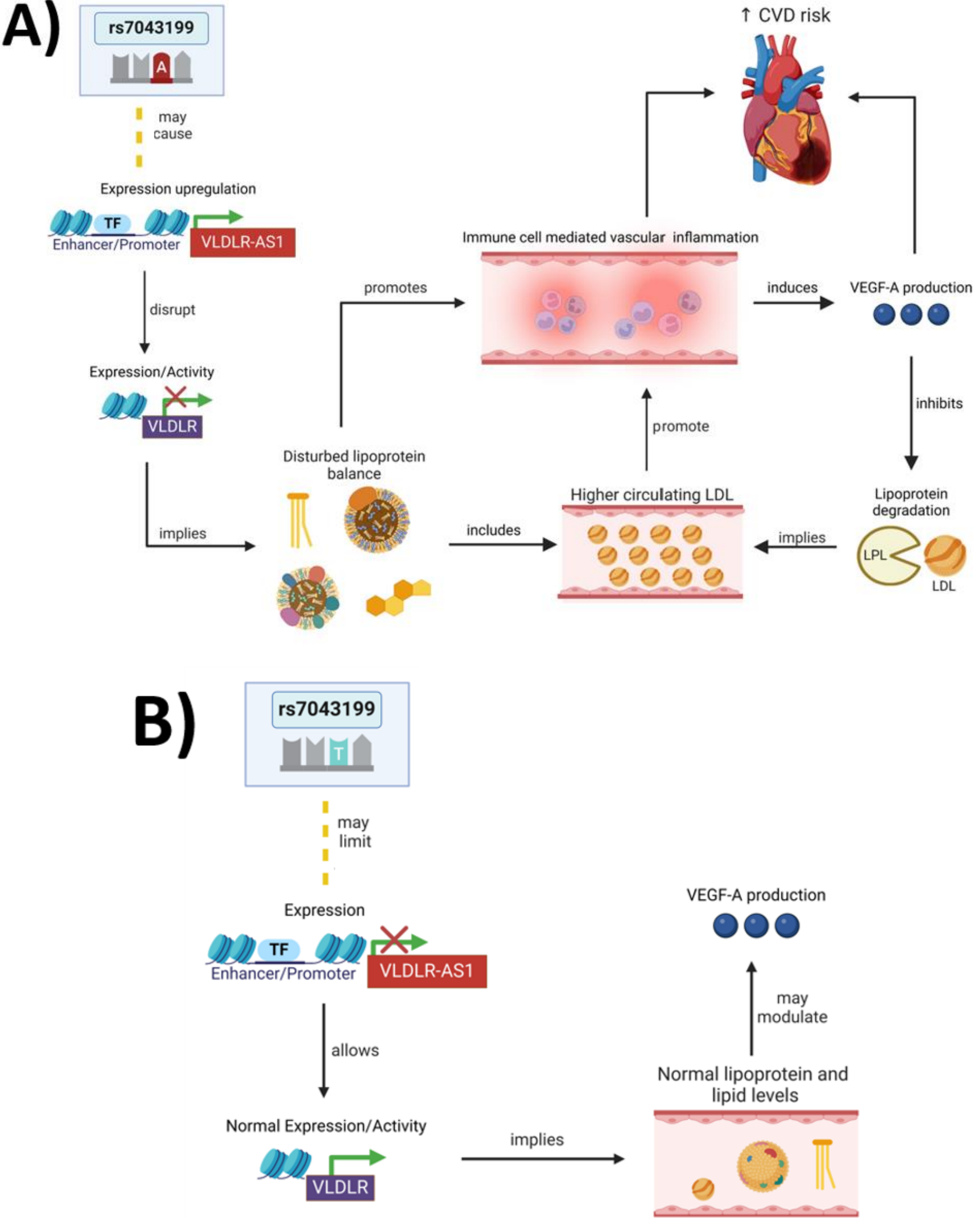


Figure 7.6. Potential interactions of rs7043199 the A) risk and B) protective alleles on VLDLR-AS1 with their effect on lipids and VEGF-A. Yellow lines represent potential connections of the SNP's with VLDLR-AS! expression. Black arrows represent processes referenced in text. Abbreviations: CVD: Cardiovascular Disease, LDL: Low density lipoprotein, LPL: Lipoprotein lipase, VEGF-A: Vascular endothelial growth factor A, VLDLR: very low-density lipoprotein receptor, VLDLR-AS1: very low-density lipoprotein receptor antisense RNA 1. Created with BioRender.com.

Additionally, there is a high likelihood that a network of interactions exists between the chromosome 9 variants described and elements on proximal or distal loci, including other SNPs. For example, a Greek cohort study using 9 VEGF-A eQTLs (including rs7043199, rs2375981 and rs10738760) found the variants were associated with elevated VEGF-A, BMI, and systolic BP and lower HDL levels [74].

This agrees with current knowledge on the synergistic and additive effect of multiple SNPs towards a specific disease phenotype [2, 75]. As such, multiple human chromosome 9 variants may impact CVD onset via blood pressure and VLDLR regulation.

7.6 VEGF-A impact on CVD risk pathways

Broadly, elevated circulating VEGF-A levels can have an impact on CVD risk factors. VEGF-A has been shown to disrupt HDL and LDL activity. HDL particles transport cholesterol through the bloodstream to the liver as well as playing a role in endothelial repair and inflammation inhibition. However, VEGF-A is capable of binding to HDL, and it can disrupt LPL activity [30, 31].. This leads to a decrease in HDL activity and increased VLDL levels, respectively, which would contribute to CVD risk by impacting anthropometric measurements (e.g. weight and BMI). Other risk factors that may be impacted include increased vasoconstriction due to reduced HDL activity promoting hypertension, as such following an ACS event, elevated VEGF-A levels may play a role by contributing to ongoing development which may be exacerbated based on which allelic variant an individual has. Overall, these represent additional mechanisms through which VEGF-A can impact risk factors involved in CVD risk. The effects of elevated circulating VEGF-A levels on CVD risk factors are summarised in Figure 7.7.

From what has been discussed, the genetic variants analysed in this project can be associated with VEGF-A circulating levels and other cardiometabolic parameters routinely assessed in healthcare. By exploring the genomic context of these variants, it is possible to propose potential mechanisms of action to explain their effect in CVDs. In the case of human chromosome 6 variants, they may increase VEGF-A expression through transcription factor activity that would promote expression of VEGF-A inducers such as IL-6, TNF-α and HIF-1α. Although these inflammatory and hypoxic markers occur at a systemic level as part of CVD onset, the variants can complement their production which contribute to CVD progression. Moreover, chromosome 4 variants may affect VEGF-A's interaction with its receptors VEGFR1 and VEGFR2, which can also impact other pathways including BP and LPL activity. This provides evidence that variants on the *VEGFA* and *VEGFR2* loci affect VEGF-A activity with the potential to exacerbate molecular mechanisms occurring during CVD onset. In addition, variants rs69214368, rs7767396, rs2305948 and rs1870377 (Chapters 4 and 5) were associated with clinical outcome. Between the theorised interactions discussed and associations observed, these four variants can be implied to present a systemic effect that warrants their use in CVD

risk profiling. This project noted chromosome 9 variants on the *VLDLR* locus had VEGF-A eQTL properties, primarily in healthy individuals.

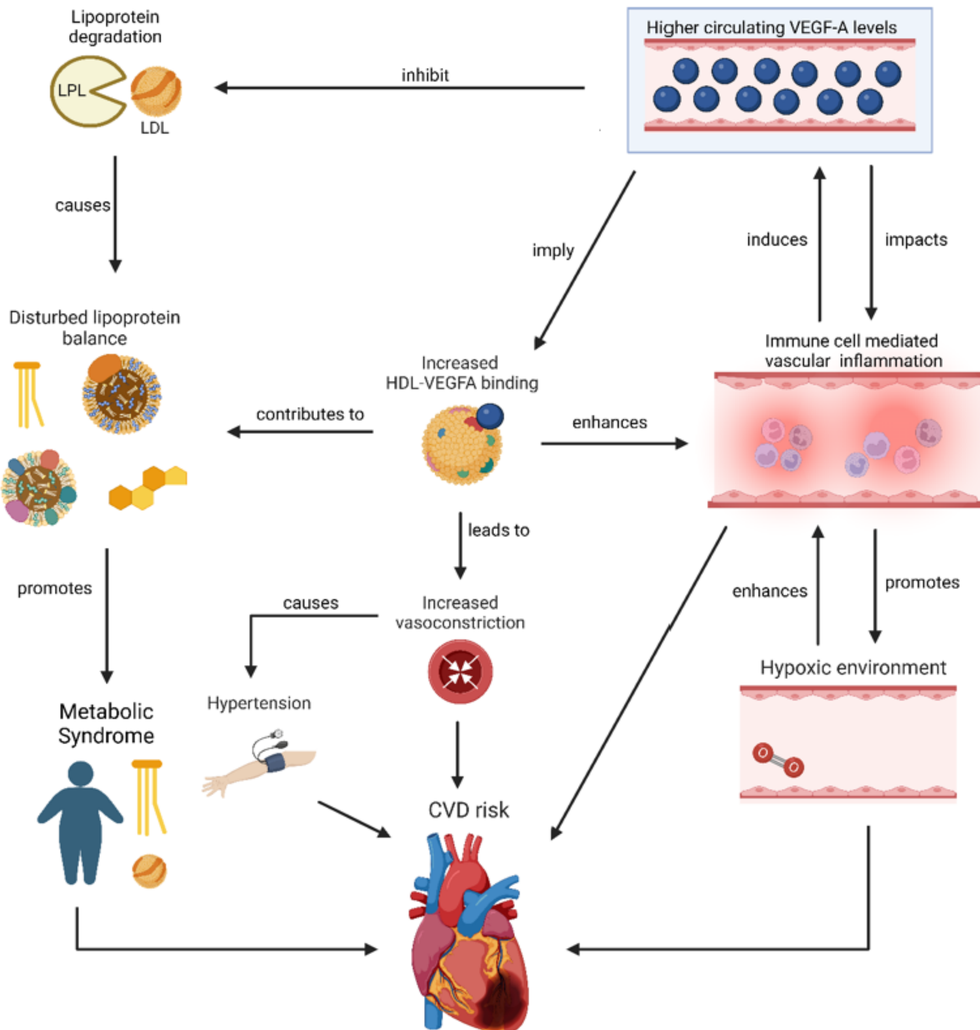


Figure 7.7. Proposed interaction of elevated circulating VEGF-A levels on CVD risk factors. Abbreviations: CVD: Cardiovascular Disease, HDL: High density lipoprotein, LDL: Low density lipoprotein, LPL: Lipoprotein lipase, VEGF-A: Vascular endothelial growth factor A, VLDLR: very low-density lipoprotein receptor. Created with BioRender.com.

7.7 Project applications and strengths

In addition to providing insights on their mechanism of action, this project could allow the risk variants identified to be used in risk profiling algorithms. Due to the polygenic nature of heart disease, there is a need to continue improving existing CVD risk algorithms. This involves the use of a GRS which allows the estimation of an individual's genetic predisposition to CVDs [76]. Existing evidence highlights that GRSs' predictive potential depends on the type of associations of SNPs. Particularly,

GRSs that would improve discrimination would include SNPs associated with intermediate risk factors (e.g. elevated blood lipid levels, diabetes, blood pressure) as well as CVD-associated SNPs [77]. Based on this project's findings, which are supported by other reports, the variants discussed in this chapter represent notable candidates for a GRS. The use of variants in GRSs requires several considerations related to the additive and synergistic properties of SNPs. Considering SNPs display pleiotropic effects, identifying associations with CVD or CVD risk factors requires a robust study design. Furthermore, variants may present trans acting mechanisms that could modify their individual associations and involve pathways with additional pleiotropic effects in the human body. Particularly, the direction and effect size of a SNP on any given trait or disease may vary. This implies that some SNPs may have a more pronounced effect on immediate pathological processes while others enhance CVD susceptibility via risk factor modification. Therefore, some SNPs could behave as confounding factors when trying to identify a combination of variants ideal for a GRS. As observed in this project, SNPs on the same locus displayed similar cardiometabolic associations that may occur through different or overlapping mechanisms. Overall, this project highlighted those SNPs with clinical outcome associations which could be considered priority SNPs for preliminary GRS use in New Zealand.

Another key application of these findings involves the prioritisation of healthcare. As mentioned previously, the genetic variants identified have associations with BP measurements, lipid markers and natriuretic peptides. Therefore, they could be used as preliminary proxies, when plasma measurements are unavailable, to prioritise health counselling. Additionally, this project employed patient follow up data to identify clinical outcomes with high readmission risk. Variants identified to confer increased or reduced risk of outcomes may complement a patient's clinical background and lifestyle choices to allow risk stratification. Lastly, those variants with potential links to inflammation, lipid biomarkers and blood pressure regulators could pave the way for a personalised treatment as well as offering the opportunity to explore novel drug treatments. At this stage, the main strength of the project involved identifying VEGF-A related variants in clinical and heart healthy cohorts from New Zealand. The two cohorts utilised in this study (CDCS and HVOL) had well characterised participants, with robust selection criteria and extensive follow up and clinical end point data. This project highlights the benefit of employing national data that is derived from a collaboration of New Zealand heart health researchers.

The strength of the project findings stems from an initial assessment of the literature to identify eQTL VEGF-A SNPs with CVD risk associations (Chapter 2 Literature review). From here, the project

focused on variants relevant to VEGF-A which were VEGFA, VEGFR2 and VLDLR. To expand coverage, additional proxy variants in the VEGFA and VEGFR2 loci were considered based on their linkage disequilibrium. This allowed the inclusion of reported variants and exploring potentially novel SNP candidates. Where possible, findings were validated using laboratory generated genotype data and exploring each variants genomic influence. Specifically, this approach combines bioinformatic aspects (imputation and genetic database data) with experimental data (SNP genotyping and VEGF-A level measurement) to bridge the gap with existing literature. This project's approach differs from studies that employ multiple international datasets to provide a thorough statistical analysis to identify statistically significant SNPs. Despite this, the project highlights the use of imputed genotype data to provide a cost-effective approach towards selecting individual SNPs for further analysis. Notably, the project findings agree with international research in CVD related risk variants. Overall, it was possible to identify novel clinically relevant SNPs (rs6921438, rs7767396, rs10738760, rs2375981, and rs7043199) while complementing existing research on others (rs2305948 and rs1870377) so they may serve as a basis for genetic risk stratification within the NZ CVD risk framework.

7.8 Project limitations

The project design made use of imputation data and genetic databases to explore associations of the variants with different variables and clinical outcomes. Where possible, data was generated to validate the initial associations. Due to this design, instead of proving causative properties of SNPs, the project focused on identifying associations and searching for novel risk variants candidates that agreed with the existing literature. Given the VEGF-A network is complicated and to acknowledge the need for information on potential causation, additional effort was done to explore genetic motifs for key variants (Section 7.3 to 7.5). Although, the proposed links will require validation, it was possible to identify metabolic pathways or candidate molecular mechanisms that can be pursued by other groups. Moreover, these links build up from reported evidence based on the established functions and research done within multiple fields.

In terms of findings, this thesis' results agree with studies conducting research on VEGF-A eQTLs. However, most work has assessed European or Asian populations through case-control, meta-analyses and GWAS studies (Chapter 2 Literature Review). The findings of this project are limited to cohorts of predominantly European ancestry individuals. This limits the extrapolation of the results to other populations. Therefore, this project's findings may not be conserved in non-European cohorts or cohorts focusing on different heart aetiologies (e.g. heart failure, structural defects). Although other

populations might share similar SNP associations, the cohorts studied in this project have a small number of non-European participants. In New Zealand individuals of Māori and Pasifika ancestry have been reported to present higher CVD risk factors compared to people of European heritage [78-80]. This highlights that current health inequities represent an overarching limiting factor on being able to apply this study's findings for all New Zealanders.

Imputed genotype data was analysed to identify associations of candidate SNPs with cardiometabolic parameters or clinical risk. When repeating the statistical analyses using manual genotype data for specific variants (rs6921438, rs7767396, rs1870377 and rs2305948), some associations did not remain. Although genotype group frequencies were similar between data types, the fact that fewer associations in the manual data were observed was most likely due to individuals shifting between genotype groups. Furthermore, imputed genotype data makes use of an imputation parameter (r^2) that estimates the correlation between imputed genotypes and unobserved genotypes. If all imputed genotype callings were concordant with unobserved genotype data, the r^2 value for each imputed SNP would be 1. The imputed genotypes for the 47 SNPs used in this thesis had an r^2 value between 0.5 to 0.8. This implies there is an inherent degree of error in the imputed genotype callings. Although imputed genotypes served as a tool to identify SNP candidates for CVD risk, imputed findings must be validated with manual genotype data. This was possible for some variants discussed in Chapter 4 (rs6921438, rs77673960) and Chapter 5 (rs1870377 and rs2305948). Manual genotyping validation for *VLDLR* locus variants (Chapter 6) was not possible due to funding limitations, but the strength and number of observed associations suggest specific variants (rs10738760, rs2375981, and rs7043199) warrant further analysis.

Other studies that used GWAS imputed data to identify associations between biomarkers (e.g. cytokines, VEGF-A) and SNPs make use of large cohorts ($n \geq 5000$) which would reduce the margin of error caused by genotype imputation [10, 28, 81]. These sample sizes would also allow higher degrees of data stratification, particularly in multivariate model development, with minimal loss of statistical power. Additionally, genotype imputation carried out in some of these studies used the same population as their reference panel which would reduce population specific variability [10, 28]. These aspects contrast with the research project since the reference panel of European ancestry used may include other groups (e.g. Irish, Finnish, Iberian) that may differ to the ancestry of some of the European New Zealander participants. Overall, lack of validation of imputed genotype findings, population specificity and smaller cohort sizes (CDCS n = 1935, HVOL n = 1183) represent some of the limitations on the use of imputed genotype data for this project.

Additional limitations can be observed considering the resources available for each cohort used in the project. Both cohorts are a national resource that has been amply used by different collaborators to study cardiac biomarkers, with plasma samples being quite limited compared to DNA samples. Plasma samples were collected at varying times after the index event and the available volume varies for each participant. This can affect subsequent laboratory testing. This factor contributes to the project having to account for the different amounts of data on various molecules which limit variables that were included in each multivariate hazards regression model. The main impact reflects on the sample size for any given clinical outcomes with key results depending on the behaviour of well-established biomarkers included. Inherent to this, there may be other pathways that affect SNP activity over VEGF-A and clinical outcomes to the point that the nature of disease complexity forces the project to focus on key well-established variables.

Other limitations involve funding prioritisation. Specifically, HVOL samples had no previously available VEGF-A data which was needed for comparison with the CDCS data. Furthermore, this involved the use of expensive optimised ELISA kits to match experimental procedures used for VEGF-A analysis which led to a minimal number of plasma samples being assayed. This also extends to SNP genotyping since variants with clinical outcome association were prioritised. This led to assessing two SNPs completely for both cohorts. Some alternative methods to circumvent these aspects include the development of an in-house ELISA protocol for VEGF-A, development of high-resolution melting SNP genotyping probes or restriction fragment length polymorphism genotyping. However, these options represent avenues outside the aims of this project. Additionally, funding prioritisation was also important in ensuring the project stayed focused on the aim of identifying and validating key associations rather than demonstrating potential causative links that were identified upon analysing the genomic context for the variants. Overall, multiple limitations are present, but alternatives were adequately considered to maximise research outputs, while not causing an increase in resource investment that would directly impact the aim, robustness and timeframe of the PhD project.

7.9 Future research

Given the extent of research done for this project, there are many aspects to explore as immediate future goals. Imputed findings differed for variants on chromosomes 6 and 4 so an immediate priority would involve confirming through manual genotyping additional findings for other variants of interest. This includes the VEGFR2 intronic SNP variants with potential links to CVD risk (rs767779 and rs13136007), as well as the four *VLDLR* locus variants. By genotyping other variants, various

complementary research questions could be explored, and this could confirm or deny the proposed mechanisms of action for the chromosome 9 variants. Additionally, with more genotyping data, an initial test of a uGRS could be explored. This would allow to test a new hypothesis regarding risk stratification based on the recorded clinical outcomes. Another future direction involves developing composite variables for different allelic combinations of the 7 clinically relevant SNPs identified. This would guide further elucidations on potential pathways to increase understanding of how multiple variants interact with other covariates. In this manner, some SNPs links to cardiovascular homeostasis could be revealed while other novel interactions may be identified even amongst SNPs. Lastly, if manual genotype data were available for neighbouring SNPs, it would be possible to expand the scope of this study's aims to explore haplotype interactions on each locus. Some candidates to consider would be variants mentioned throughout this project that did not present cardiometabolic associations but overlap potentially functional genomic regions (e.g. transcription factor binding motifs, enhancer marks).

Further mid-term goals would involve examining other cohorts. Project findings are limited to European ancestry individuals, the main ethnic group present in both cohorts studied (CDCS and HVOL). Other populations might not share the associations found. Exploring other groups is vital since in New Zealand, individuals of Māori and Pasifika ancestry present higher CVD risk factors than people of European heritage [78-80]. Another possibility involves testing the variants in a third cohort diagnosed for other specific cardiovascular events such as HF or subclinical atherosclerosis. This would also allow the exploration of variant interactions within other disease phenotypes with the potential to analyses cardiac markers. Other cohorts of interests could involve patients undergoing therapy following a cardiac intervention. This would allow the assessment of SNPs in the context of patient treatment and recovery. Moreover, doing routine measurements of VEGF-A levels in longitudinal cohorts would help ascertain how genetics affect baseline VEGF-A over time. Overall, this would allow the exploration of molecular targets for cardiovascular therapy and use genetic variants to assess current treatment prioritisation guidelines as well as VEGF-A behaviour over time.

Additional future research that can stem from this project involves validating the mechanistic proposals for some variants (Sections 7.3 to 7.5). In this regard, it would be needed to develop other projects that focus on gene expression of targets that may influenced by binding of transcription factors or empirically exploring how each SNP can affect transcription factor binding efficiency. Additional consideration should be taken to account for novel underlying mechanisms that have changed the Central Dogma or Sequence Hypothesis of DNA → RNA → protein [82]. This includes regulatory

mechanisms involving protein-RNA interactions, epigenetics, chromatin regulation and pharmacogenetics amongst others [82, 83]. By focusing on these different aspects involving genetic regulation, the use of multiple novel molecular tools and techniques can support the study of causative mechanisms SNPs may have in a specific cardiovascular context.

.

7.10 Concluding remarks

This project focused on assessing eQTL VEGF-A SNPs that have been reported in different European and Asian populations. The use of well-established NZ based cohorts for diagnosed heart disease and heart healthy individuals allowed the assessment of variants associated with cardiometabolic parameters and clinical outcome. The project focused on variants at loci relevant to the VEGF system, specifically *VEGFA*, *VEGFR2* and *VLDLR*. It was possible to identify SNPs that are associated with VEGF-A, blood pressure and triglyceride levels. Furthermore, specific SNPs were associated with increased or reduced disease readmission risk. Overall, variants on human chromosome 6 and 4 represented novel risk profile candidates that are associated with other cardiovascular parameters. SNP candidates on human chromosome 9 are additional candidates for future studies involving VEGF-A and lipid metabolism. The PhD project also explored the links between a SNP's location with other molecular mechanisms that may influence CVD risk pathways. This represents an additional insight on molecular mechanisms that contributes to improving our understanding of VEGF-A in healthy and CVD diagnosed individuals. The presented findings may contribute to enhancing current risk profiling tools by including some of the genetic variants analysed.

7.11 References

1. Damen, J.A.A.G., et al., *Prediction models for cardiovascular disease risk in the general population: systematic review.* British Medical Journal, 2016. **353**: p. i2416.
2. Roberts, R., C.C. Chang, and T. Hadley, *Genetic Risk Stratification: A Paradigm Shift in Prevention of Coronary Artery Disease.* Journal of the American College of Cardiology: Basic to Translational Science, 2021. **6**(3): p. 287-304.
3. Roberts, R. and A. Campillo, *Genetic stratification for primary prevention of coronary artery disease.* Current Opinion in Cardiology, 2018. **33**(5): p. 529-534.
4. Roberts, R. and J. Fair, *Genetics, its role in preventing the pandemic of coronary artery disease.* Clinical Cardiology, 2021. **44**(6): p. 771-779.
5. Assimes, T.L. and R. Roberts, *Genetics: Implications for Prevention and Management of Coronary Artery Disease.* Journal of the American College of Cardiology, 2016. **68**(25): p. 2797-2818.
6. Braile, M., et al., *VEGF-A in Cardiomyocytes and Heart Diseases.* International Journal of Molecular Sciences, 2020. **21**(15): p. 1-18.
7. Camaré, C., et al., *Angiogenesis in the atherosclerotic plaque.* Redox biology, 2017. **12**: p. 18-34.
8. Heinonen, S.E., et al., *The effects of VEGF-A on atherosclerosis, lipoprotein profile, and lipoprotein lipase in hyperlipidaemic mouse models.* Cardiovascular Research, 2013. **99**(4): p. 716-23.
9. Zhao, Z., et al., *Association between Single Nucleotide Polymorphisms in Cardiovascular Developmental Critical Genes and Hypertension: A Propensity Score Matching Analysis.* International Journal of Hypertension, 2020. **2020**: p. 1-8.
10. Ruotsalainen, S.E., et al., *An expanded analysis framework for multivariate GWAS connects inflammatory biomarkers to functional variants and disease.* European Journal of Human Genetics, 2021. **29**(2): p. 309-324.
11. Zhao, X., et al., *Vascular endothelial growth factor gene polymorphisms and coronary heart disease: a systematic review and meta-analysis.* Growth Factors, 2018. **36**(3-4): p. 153-163.
12. Ramakrishnan, S., V. Anand, and S. Roy, *Vascular endothelial growth factor signaling in hypoxia and inflammation.* Journal of Neuroimmune Pharmacology, 2014. **9**(2): p. 142-60.
13. Au Yeung, S.L., H. Lam, and C.M. Schooling, *Vascular Endothelial Growth Factor and Ischemic Heart Disease Risk: A Mendelian Randomization Study.* Journal of the American Heart Association, 2017. **6**(8): p. 1-8.
14. Zhang, Y., et al., *Serum vascular endothelial growth factor as a biomarker for prognosis of minor ischemic stroke.* Clinical Neurology and Neurosurgery, 2020. **196**: p. 106060.
15. Rodríguez-Rodríguez, L., et al., *Vascular endothelial growth factor A and cardiovascular disease in rheumatoid arthritis patients.* Tissue Antigens, 2011. **77**(4): p. 291-297.
16. Choi, S.H., et al., *Six Novel Loci Associated with Circulating VEGF Levels Identified by a Meta-analysis of Genome-Wide Association Studies.* Public Library of Science (PLOS) Genetics, 2016. **12**(2): p. e1005874.
17. Debette, S., et al., *Identification of cis- and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels.* Circulation Research, 2011. **109**(5): p. 554-63.
18. Meza-Alvarado, J.C., et al., *VEGF-A related SNPs: a cardiovascular context.* Frontiers in Cardiovascular Medicine, 2023. **10**.
19. Palmer, B.R., et al., *Vascular endothelial growth factor-A promoter polymorphisms, circulating VEGF-A and survival in acute coronary syndromes.* PLOS One, 2021. **16**(7): p. 1-15.
20. Yu, F., et al., *Correlation of Plasma Vascular Endothelial Growth Factor and Endostatin Levels with Symptomatic Intra- and Extracranial Atherosclerotic Stenosis in a Chinese Han Population.* J Stroke Cerebrovasc Dis, 2017. **26**(5): p. 1061-1070.
21. Stathopoulou, M.G., et al., *A common variant highly associated with plasma VEGFA levels also contributes to the variation of both LDL-C and HDL-C.* Journal of Lipid Research, 2013. **54**(2): p. 535-41.
22. Azimi-Nezhad, M., et al., *Associations of vascular endothelial growth factor (VEGF) with adhesion and inflammation molecules in a healthy population.* Cytokine, 2013. **61**(2): p. 602-7.

23. Ryu, J.R., et al., *Stroke-associated intergenic variants modulate a human FOXF2 transcriptional enhancer.* Proceedings of the National Academy of Sciences of the United States of America (USA), 2022. **119**(35): p. e2121333119.
24. Ward, L.D. and M. Kellis, *HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease.* Nucleic Acids Research, 2016. **44**(D1): p. D877-81.
25. Wu, Q., W. Li, and C. You, *The regulatory roles and mechanisms of the transcription factor FOXF2 in human diseases.* PeerJ, 2021. **9**: p. e10845.
26. Fahey, E. and S.L. Doyle, *IL-1 Family Cytokine Regulation of Vascular Permeability and Angiogenesis.* Frontiers in Immunology, 2019. **10**: p. 1-15.
27. Ramji, D.P. and T.S. Davies, *Cytokines in atherosclerosis: Key players in all stages of disease and promising therapeutic targets.* Cytokine & Growth Factor Reviews, 2015. **26**(6): p. 673-685.
28. Chen, H., et al., *Variation in VEGFA and risk of cardiovascular disease in the UK Biobank.* Frontiers in Cardiovascular Medicine, 2023. **10**: p. 1240288.
29. Lee, H.C., A. Akhmedov, and C.H. Chen, *Spotlight on very-low-density lipoprotein as a driver of cardiometabolic disorders: Implications for disease progression and mechanistic insights.* Frontiers in Cardiovascular Medicine, 2022. **9**: p. 993633.
30. Velagapudi, S., et al., *VEGF-A Regulates Cellular Localization of SR-BI as Well as Transendothelial Transport of HDL but Not LDL.* Arterioscler Thromb Vasc Biol, 2017. **37**(5): p. 794-803.
31. Kivelä, A.M., et al., *Regulation of endothelial lipase and systemic HDL cholesterol levels by SREBPs and VEGF-A.* Atherosclerosis, 2012. **225**(2): p. 335-340.
32. Innocenti, F., et al., *Genetic variation determines VEGF-A plasma levels in cancer patients.* Scientific Reports, 2018. **8**(1): p. 16332.
33. Zhang, P., et al., *The mechanism of the NFAT transcription factor family involved in oxidative stress response.* Journal of Cardiology, 2024. **83**(1): p. 30-36.
34. Harhous, Z., et al., *An Update on the Multifaceted Roles of STAT3 in the Heart.* Frontiers in Cardiovascular Medicine, 2019. **6**: p. 150.
35. Ahmed, A.A., K.I. Amber, and N.R. Hadi, *The Impact of Kinase Insert Domain (KDR) Gene Polymorphism rs2305948 on Clopidogrel Resistance in Iraqi Patients Undergoing Elective Percutaneous Coronary Intervention (PCI).* Acta Informatica Medica, 2020. **28**(3): p. 202-208.
36. Matalliotaki, C., et al., *Implication of VEGFR2 in endometriosis: A structural biological and genetic approach.* World Academy of Sciences Journal, 2019. **1**(6): p. 283-289.
37. Yang, Y., et al., *Direct contacts between extracellular membrane-proximal domains are required for VEGF receptor activation and cell signaling.* Proceedings of the National Academy of Sciences of the United States of America (USA), 2010. **107**(5): p. 1906-11.
38. Caron, B., et al. *Integrative genetic and immune cell analysis of plasma proteins in healthy donors identifies novel associations involving primary immune deficiency genes*. Genome medicine, 2022. **14**, 28 DOI: 10.1186/s13073-022-01032-y.
39. Emilsson, V., et al., *Co-regulatory networks of human serum proteins link genetics to disease.* Science, 2018. **361**(6404): p. 769-773.
40. Zimetti, F., et al., *Connection between the Altered HDL Antioxidant and Anti-Inflammatory Properties and the Risk to Develop Alzheimer's Disease: A Narrative Review.* Oxidative Medicine and Cellular Longevity, 2021. **2021**: p. 6695796.
41. Mineo, C., et al., *Endothelial and Antithrombotic Actions of HDL.* Circulation Research, 2006. **98**(11): p. 1352-1364.
42. Pi, X., L. Xie, and C. Patterson, *Emerging Roles of Vascular Endothelium in Metabolic Homeostasis.* Circulation Research, 2018. **123**(4): p. 477-494.
43. Feng, C., et al., *Effect of gene-gene and gene-environment interaction on the risk of first-ever stroke and poststroke death.* Molecular Genetics & Genomic Medicine, 2019. **7**(8): p. e846.
44. Liu, D., et al., *Association of Genetic Polymorphisms on VEGFA and VEGFR2 With Risk of Coronary Heart Disease.* Medicine, 2016. **95**(19): p. e3413.
45. Marks, E.C.A., et al., *Plasma levels of soluble VEGF receptor isoforms, circulating pterins and VEGF system SNPs as prognostic biomarkers in patients with acute coronary syndromes.* BMC Cardiovascular Disorders, 2018. **18**(1): p. 169.

46. Al Awaida, W., et al., *Association of KDR rs1870377 genotype with clopidogrel resistance in patients with post percutaneous coronary intervention.* Heliyon, 2021. **7**(2): p. e06251.
47. Yap, R.W., et al., *Association and interaction between dietary pattern and VEGF receptor-2 (VEGFR2) gene polymorphisms on blood lipids in Chinese Malaysian and Japanese adults.* Asia Pac J Clin Nutr, 2012. **21**(2): p. 302-11.
48. Nevo, O., D.K. Lee, and I. Caniggia, *Attenuation of VEGFR-2 expression by sFlt-1 and low oxygen in human placenta.* PLoS One, 2013. **8**(11): p. e81176.
49. Glubb, D.M., et al., *Novel Functional Germline Variants in the VEGF Receptor 2 Gene and Their Effect on Gene Expression and Microvessel Density in Lung Cancer.* Clinical Cancer Research, 2011. **17**(16): p. 5257-5267.
50. Wang, Y., et al., *Polymorphisms of KDRGene Are Associated With Coronary Heart Disease.* Journal of the American College of Cardiology, 2007. **50**(8): p. 760-767.
51. Masago, K., et al., *Next-generation sequencing of tyrosine kinase inhibitor-resistant non-small-cell lung cancers in patients harboring epidermal growth factor-activating mutations.* BMC Cancer, 2015. **15**(1): p. 908.
52. Suico, M.A., T. Shuto, and H. Kai, *Roles and regulations of the ETS transcription factor ELF4/MEF.* Journal of Molecular Cell Biology, 2017. **9**(3): p. 168-177.
53. Wang, K., *Pathophysiological Role of Hepatocyte Nuclear Factor 6 in Negative Regulation of Hepatic Apoptosis: A Novel Hypothesis.* Curr Mol Med, 2015. **15**(5): p. 412-7.
54. Wang, K. and A.-X. Holterman, *Pathophysiologic role of hepatocyte nuclear factor 6.* Cellular Signalling, 2012. **24**(1): p. 9-16.
55. Ahola-Olli, A.V., et al., *Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors.* American Journal of Human Genetics, 2017. **100**(1): p. 40-50.
56. Czirják, G., Z.E. Tóth, and P. Enyedi, *Characterization of the Heteromeric Potassium Channel Formed by Kv2.1 and the Retinal Subunit Kv8.2 in Xenopus Oocytes.* Journal of Neurophysiology, 2007. **98**(3): p. 1213-1222.
57. Renigunta, V., et al., *A versatile functional interaction between electrically silent KV subunits and KV7 potassium channels.* Cellular and Molecular Life Sciences, 2024. **81**(1): p. 301.
58. Guimaraes, T.A.C.D., et al., *KCNV2 retinopathy: clinical features, molecular genetics and directions for future therapy.* Ophthalmic Genetics, 2020. **41**(3): p. 208-215.
59. Chen, L., et al., *Ventricular voltage-gated ion channels: Detection, characteristics, mechanisms, and drug safety evaluation.* Clin Transl Med, 2021. **11**(10): p. e530.
60. Salami, A. and S. El Shamieh, *Association between SNPs of Circulating Vascular Endothelial Growth Factor Levels, Hypercholesterolemia and Metabolic Syndrome.* Medicina, 2019. **55**(8): p. 1-9.
61. Ghazizadeh, H., et al., *Association of rs6921438 A<G with serum vascular endothelial growth factor concentrations in patients with metabolic syndrome.* Gene, 2018. **667**: p. 70-75.
62. Ghazizadeh, H., et al., *Interaction between a genetic variant in vascular endothelial growth factor with dietary intakes in association with the main factors of metabolic syndrome.* Gene Reports, 2020. **21**: p. 100813.
63. Ghazizadeh, H., et al., *Association of a Vascular Endothelial Growth Factor genetic variant with Serum VEGF level in subjects with Metabolic Syndrome.* Gene, 2017. **598**: p. 27-31.
64. Ghazizadeh, H., et al., *Association between serum cell adhesion molecules with hs-CRP, uric acid and VEGF genetic polymorphisms in subjects with metabolic syndrome.* Mol Biol Rep, 2020. **47**(2): p. 867-875.
65. Markouli, M., D. Strepkos, and C. Piperi, *Structure, Activity and Function of the SETDB1 Protein Methyltransferase.* Life 2021. **11**(8).
66. Krauss, R.M., et al., *VLDL receptor gene therapy for reducing atherogenic lipoproteins.* Molecular Metabolism, 2023. **69**: p. 101685.
67. Gao, Y., et al., *Upregulation of hepatic VLDLR via PPARα is required for the triglyceride-lowering effect of fenofibrate.* J Lipid Res, 2014. **55**(8): p. 1622-33.
68. Petrelis, A.M., et al., *VEGF-A-related genetic variants protect against Alzheimer's disease.* Aging 2022. **14**(6): p. 2524-2536.
69. Yu, Y.H., et al., *Prostaglandin reductase-3 negatively modulates adipogenesis through regulation of PPARγ activity.* J Lipid Res, 2013. **54**(9): p. 2391-9.

70. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Research, 2002. **12**(6): p. 996-1006.
71. Fishilevich, S., et al., *GeneHancer: genome-wide integration of enhancers and target genes in GeneCards.* Database (Oxford), 2017. **2017**.
72. Takahashi, K., et al., *Involvement of extracellular vesicle long noncoding RNA (linc-VLDLR) in tumor cell responses to chemotherapy.* Mol Cancer Res, 2014. **12**(10): p. 1377-87.
73. Liu, H., et al., *Identification and functional analysis of a potential key lncRNA involved in fat loss of cancer cachexia.* J Cell Biochem, 2018. **119**(2): p. 1679-1688.
74. Kafyra, M., et al., *Associations of VEGF-A-Related Variants with Adolescent Cardiometabolic and Dietary Parameters.* Nutrients, 2023. **15**(8).
75. Xu, K., et al., *The combined effects of cardiovascular disease related SNPs on ischemic stroke.* Journal of the Neurological Sciences, 2018. **388**: p. 141-145.
76. Garfield, V. and E.L. Anderson, *A brief comparison of polygenic risk scores and Mendelian randomisation.* BMC Med Genomics, 2024. **17**(1): p. 10.
77. Yun, H., N.I. Noh, and E.Y. Lee, *Genetic risk scores used in cardiovascular disease prediction models: a systematic review.* Reviews in Cardiovascular Medicine, 2022. **23**(1): p. 1-9.
78. Martin Tobias and Maria Turley, *Health Loss in New Zealand: A report from the New Zealand Burden of Diseases, Injuries and Risk Factors Study, 2006-2016.* 2016, Ministry of Health: Wellington. p. 94.
79. Yeh, L.-C., et al., *Tatau kahukura : Māori health chart book 2015*. 3rd ed. 2015: Ministry of Health.
80. Barnes, L.A., et al., *The Prevalence of Cardiovascular Risk Factors in Different Occupational Groups in New Zealand.* Annals of Work Exposures & Health, 2020. **64**(6): p. 645-658.
81. Wang, Y., et al., *Genome-wide association study identifies 16 genomic regions associated with circulating cytokines at birth.* PLoS Genetics, 2020. **16**(11): p. e1009163.
82. Ille, A.M., H. Lamont, and M.B. Mathews, *The Central Dogma revisited: Insights from protein synthesis, CRISPR, and beyond.* Wiley Interdiscip Rev RNA, 2022. **13**(5): p. e1718.
83. Chang, H.Y. and L.S. Qi, *Reversing the Central Dogma: RNA-guided control of DNA in epigenetics and genome editing*. Molecular Cell, 2023. **83**(3): p. 442-451.

Appendices

Appendix 1.1 Research outputs

Research Output Description	Reference
Massey University finalist in the 3-minute thesis presentation. The talk focused on outlining PhD project design.	Massey University 3-minute thesis competition . **Meza Alvarado, J.C.** SNPs: The Heart of Risk Prevention. Online Submission. August 2022
Peer-reviewed literature review on VEGF-A eQTL SNPs in CVDs. Included in thesis as part of Chapter 3 – Literature Review.	**Meza Alvarado, J.C.**, Page R.A, Mallard B., Bromhead C. Palmer, B.R. (2023). VEGF-A related SNPs: a cardiovascular context. *Frontiers of Cardiovascular Medicine* 10:1190513. https://doi.org/10.3389%2Ffcvm.2023.1190513
3-minute abstract presentation focusing on VEGF-A eQTL SNPs associated with reduced survival	Queenstown Research Week Genomic Satellite 2023: **Meza-Alvarado, J.C**., Page R.A., Mallard B., Bromhead C., Pilbrow, A.P., Troughton, R.W., Frampton, C.M., Doughty, R.N., Richards, A.M., Palmer, B.R. (2023). Vascular endothelial growth factor A polymorphisms impacting patient survival. Queenstown Research Week Genomic Satellite Meeting, Rydges Hotel, Queenstown, 26-28 August 2023.
Poster presentation focusing on VEGF-A eQTL SNPs associated with reduced survival	2023 Heart Health Forum. **Meza-Alvarado, J.C.**, Page R.A., Mallard B., Bromhead C., Pilbrow, A.P., Troughton, R.W., Richards, A.M., Doughty, R.N., Palmer, B.R. Vascular endothelial growth factor-A polymorphisms impacting cardiac patient survival. Otago Museum, Dunedin, 12 October 2023
Moderated ePoster presentation focusing on findings that associated SNPs on the *VEGFA* locus (human chromosome 6) with increased risk of heart failure readmissions	European Society of Cardiology Heart Failure Congress 2024. **Meza-Alvarado, J.C.**, Page R.A., Mallard B., Bromhead C., Pilbrow, A.P., Troughton, R.W., Frampton, C.M., Doughty, R.N., Richards, A.M., Palmer, B.R. (2024). Vascular endothelial growth factor-A SNPs as predictors of hospitalisation for acute heart failure in patients with established coronary heart disease. Lisbon Congress Centre, 11-14 May 2024
Research article manuscript to be submitted for peer-review journal publication (BMC Cardiovascular Disorders) focusing on clinically relevant variants identified in the *VEGFA* locus on human chromosome 6. Presented in Thesis as Chapter 4	Meza-Alvarado, J.C., Pilbrow, A.P., Frampton, C.M., Cameron, V.A., Richards, A.M., Doughty, R.N., Page R.A, Mallard B., Bromhead C. Palmer, B.R. *VEGFA* cis-located SNPs on chromosome 6 associated with VEGF-A plasma levels and survival in a coronary disease cohort.
Drafted research article manuscript for peer-review journal publication focusing on VEGF-A eQTL SNPs of potential clinical relevance identified in the VEGFR2 gene on human chromosome 4. Presented in Thesis as Chapter 5.	Meza-Alvarado, J.C., Pilbrow, A.P., Frampton, C.M., Cameron, V.A., Richards, A.M., Doughty, R.N., Page R.A, Mallard B., Bromhead C. Palmer, B.R. Association of *VEGFR2* SNPs with cardiometabolic parameters in a post-acute coronary syndromes cohort from New Zealand.
Drafted research article manuscript for peer-review journal publication focusing on VEGF-A eQTL SNPs of potential clinical relevance located in the *VLDLR* locus on human chromosome 9. Presented in Thesis as Chapter 6.	Meza-Alvarado, J.C., Pilbrow, A.P., Frampton, C.M., Cameron, V.A., Richards, A.M., Doughty, R.N., Page R.A, Mallard B., Bromhead C. Palmer, B.R. Association of *VLDLR* region SNPs with VEGF-A levels and clinical outcome in a post-acute coronary syndromes cohort.

Appendix 1.2 Published minireview

TYPE Mini Review
PUBLISHED 23 May 2023
DOI 10.3389/fcvm.2023.1190513

OPEN ACCESS

EDITED BY
Paul Cahill,
Dublin City University, Ireland

REVIEWED BY
Sadie Slater,
University of Bristol, United Kingdom

*CORRESPONDENCE
B. R. Palmer
✉ b.palmer@massey.ac.nz

RECEIVED 29 March 2023
ACCEPTED 27 April 2023
PUBLISHED 23 May 2023

CITATION
Meza-Alvarado JC, Page RA, Mallard B, Bromhead C and Palmer BR (2023) VEGF-A related SNPs: a cardiovascular context.
Front. Cardiovasc. Med. 10:1190513.
doi: 10.3389/fcvm.2023.1190513

COPYRIGHT
© 2023 Meza-Alvarado, Page, Mallard, Bromhead and Palmer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

VEGF-A related SNPs: a cardiovascular context

J. C. Meza-Alvarado, R. A. Page, B. Mallard, C. Bromhead and B. R. Palmer*

School of Health Sciences, Massey University, Wellington, New Zealand

Cardiovascular diseases (CVDs) are the leading cause of death worldwide. Currently, cardiovascular disease risk algorithms play a role in primary prevention. However, this is complicated by a lack of powerfully predictive biomarkers that could be observed in individuals before the onset of overt symptoms. A key potential biomarker for heart disease is the vascular endothelial growth factor (VEGF-A), a molecule that plays a pivotal role in blood vessel formation. This molecule has a complex biological role in the cardiovascular system due to the processes it influences, and its production is impacted by various CVD risk factors. Research in different populations has shown single nucleotide polymorphisms (SNPs) may affect circulating VEGF-A plasma levels, with some variants associated with the development of CVDs, as well as CVD risk factors. This minireview aims to give an overview of the VEGF family, and of the SNPs reported to influence VEGF-A levels, cardiovascular disease, and other risk factors used in CVD risk assessments.

KEYWORDS
vascular endothelial growth factor, single nucleotide polymorphism, cardiovascular disease, VEGF-A eQTLs, genetic association

Introduction

Cardiovascular diseases (CVDs) are defined by the World Health Organization as a group of disorders that affect the heart and blood vessels in terms of structure or blood supply (1). Notable examples of CVDs that are a leading cause of death globally include coronary heart disease (CHD), acute coronary syndrome (ACS) and congenital heart disease (2). CHD involves inadequate coronary blood supply, which may arise from a blockage in the coronary arteries usually following progressive narrowing of the lumen of atherosclerotic blood vessels (3). Given the multifactorial nature of CVDs, there are reviews available that explore in greater detail specific diseases such as coronary artery disease (CAD) (4), CHD (5), the underlying mechanism of atherosclerosis (6, 7) and the relationship of these diseases with specific variables (8, 9).

Overall, risk factors for CVDs can be grouped as modifiable or non-modifiable. The modifiable risk factors involve lifestyle circumstances that can be behavioral (diet, physical activity, exercise, smoking, alcohol) or metabolic (circulating lipid levels and glucose levels) in nature (8). Whereas, age, genetics and ethnicity of individuals are the non-modifiable risk factors. This distinction informs the diagnosis of CVDs by determining which critical variables should be included in CVD risk assessments. Critical factors employed have included age, hypercholesterolemia, high density lipoprotein (HDL) cholesterol levels, gender, smoking, diabetes, and systolic blood pressure (10).

Genetic determinants are important non-modifiable risk factors for CVDs that have been studied intensively since the early 21st century (11–13). The influence of genetic factors on CVD development was initially explored through family history studies focused

on single gene disorders during the 1980 s (4, 14). Most CVDs are now considered to be polygenic disorders impacted by susceptibility and disease-linked genes, with major impacts from lifestyle and environmental factors (14). Susceptibility genes are associated with an increase or reduction in the risk of developing a disease. Comparatively, disease-linked genes are those whose expression is linked to a pathological phenotype (4). Both susceptibility and disease-linked genes can influence the regulation of other genes and/or factors that are directly involved in the pathobiology of different CVDs. The genetic basis for CVDs such as CAD and CHD has been reviewed in greater detail elsewhere (11, 15–17).

Considering this genetic complexity, numerous studies have focused on identifying associations between genetic variants and common cardiovascular disease traits (15, 18–21). This has been supported with the establishment of genome wide association studies (GWAS), which employ technologies that detect many gene variants simultaneously (22). The predominant variants identified through GWAS are single nucleotide polymorphisms (SNPs) (15, 22–24). SNPs can be located within a protein-coding region, where they may display a functional effect, but they can also be in non-coding and regulatory areas of the genome (e.g., introns, enhancer, etc). Moreover, SNPs can play a regulatory role by impacting gene expression and protein concentration if they are located within genetic elements such as transcription factor binding sites, splicing regions, enhancer, promoter, or silencer regions (23, 25). These are often called expression quantitative trait loci (eQTLs) and explain a proportion of the genetic variance of a particular phenotype (26). SNPs can also influence coding regions located within the same loci (cis-acting) or interact with coding regions of other chromosomes or distant loci on the same chromosome (trans-acting) (27, 28). Specifically, SNP variants can influence CVD risk through traditional risk factors, such as plasma lipid levels and blood pressure (11, 27, 28). Overall, SNP variants can have several potential effects on any given gene as summarized in Figure 1. One example covered in this review is *VEGFA*, which impacts the cardiovascular system through angiogenesis and increased endothelial cell activity.

Coupling our understanding of CVD pathogenesis with associations of regulatory SNPs with coronary biomarkers, there is potential for the combined use of CVD-relevant genetic risk scores (cvdGRS) in risk prevention (30). This involves using multiple SNPs identified from GWAS studies in different populations and these SNP variants can be associated with clinical outcomes or risk factors (30, 31). Overall, the goal of cvdGRS is to aid in patient risk stratification and treatment (22, 28, 31–33). The functional effects of the SNP variants may provide evidence to underpin a clinical framework for prevention, treatment, and in severe cases, genetic counselling in primary care (22, 31, 34, 35).

VEGF overview

A molecule of interest in the development and progression of CVDs is the vascular endothelial growth factor (VEGF-A), a member of the platelet-derived growth factor (PDGF)/VEGF family (36, 37). This growth factor is involved in blood vessel

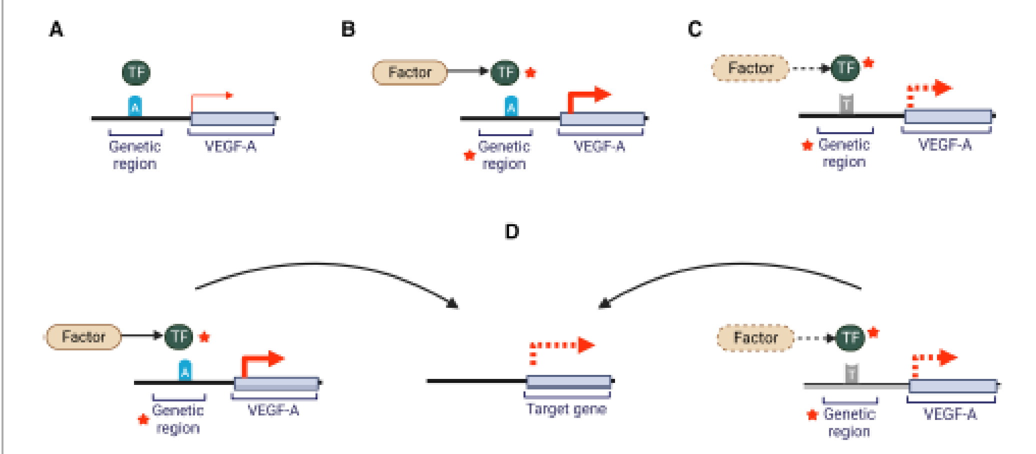


FIGURE 1
Potential effects of a SNP on VEGF-A gene expression. Gene regions where a SNP may be located include gene promoters, enhancers, introns, transcription factor binding motifs or non-coding RNAs. Dotted lines indicate hypothetical changes that could decrease or increase interactions or activity. Red arrows represent VEGF-A expression. Red stars indicate potential for an increase or decrease in activity through a variety of molecular mechanisms (A) an individual carrying the common allele "A" will show normally regulated levels of VEGF-A. (B) In the presence of a risk factor (e.g., elevated LDL-cholesterol plasma levels) and the common allele "A" there may be altered activity of genomic regions or transcription factors leading to altered VEGF-A expression and potentially impact circulating VEGF-A levels and/or activity. (C) Presence of a non-normal range risk factor (e.g., elevated plasma glucose levels) and the alternate allele "T", may impact activity of genomic regions or transcription factors leading to altered VEGF-A expression and potentially influence VEGF-A levels and/or activity (D) The changes illustrated in scenario B and C may impact expression of genes located at a distance, which may impact other pathways (e.g., lipid metabolism). Adapted from (29). Created with BioRender.com.

formation, with reported impacts on the development of CVDs, as well as potential recovery (38, 39). VEGF-A was originally referred to as a vascular permeability factor, with activity observed in tumor cells from rodents (40). In 1989 several research groups identified that this factor selectively promoted the migration of vascular endothelium and induced angiogenesis *in vivo* (41–43). Based on these findings, factors with this activity were renamed and classified as members of the VEGF family (43).

The VEGF family are glycoproteins expressed under the regulation of soluble mediators such as growth factors or cytokines (39, 44, 45). They are involved in the regulation of blood vessel formation through endothelial cell differentiation or from existing blood vessels (44, 46). Additionally, the VEGF family is involved in lymphangiogenesis, endothelial cell survival and vascular permeability regulation, amongst other functions (44, 47). However, alterations in their functionality have also been associated with the development of atherosclerosis, CHD, tumor formation, neovascularization, and other pathologies including cancer, diabetic retinopathy, preeclampsia, and endometriosis (23, 39, 47, 48).

There are five VEGF family members that directly influence the human cardiovascular system. The archetype member is VEGF-A, a potent stimulator of vasculogenesis and angiogenesis (44, 48, 49). VEGF-A production is influenced by oxygen tension, hormones (e.g., estrogen) and proinflammatory cytokines (47, 49, 50). VEGF-B induces the development of the cardiovascular system, embryonic angiogenesis and the formation of embryonic myocardium as well as participating in blood vessel survival (51). VEGF-C and VEGF-D are primarily involved in lymphangiogenesis, while the placental growth factor (PlGF) participates in both angiogenesis and wound healing (39, 47, 49).

These VEGF proteins act through one or more of three tyrosine kinase VEGF receptors (VEGFRs) found on the surface of endothelial and non-endothelial cells (44). VEGFR1 (Flt-1) and VEGFR2 (KDR) participate in angiogenesis. VEGFR2 is the primary inducer of VEGF-mediated blood vessel growth, while VEGFR3 is involved in lymphangiogenesis (47, 52, 53). Additionally, VEGFR1 has the co-receptor neuropilin-1 (NRP1), which selectively potentiates VEGFR2-mediated vascular permeability, and endothelial cell motility in vascular development (49, 54). Once activated, the signaling pathways of these receptors have the downstream effect of influencing vascular tone, blood vessel formation, endothelial cell proliferation and migration (47). VEGFR signaling is reported to also be activated in a non-VEGF-dependent manner through receptor phosphorylation due to shear stress, or recognition of alternative ligands such as lactate and low-density lipoproteins (LDLs) (36, 53, 55).

Specifically, the VEGF-A canonical pathway occurs when it binds to either VEGFR1 or VEGFR2. This promotes receptor homodimerization or heterodimerization that leads to the phosphorylation of the receptor's intracellular domains (53, 55). VEGFR1 has a soluble splice variant (sFlt-1) that acts as a decoy receptor, decreasing VEGF-A plasma concentration and limiting its binding to KDR (36, 44, 47). Also, VEGF-A activity can be potentiated when PlGF displaces it from VEGFR1 to VEGFR2 (39). These and other mechanisms surrounding the regulation of VEGF receptors have been reviewed in greater detail elsewhere (37, 55).

VEGFA gene and related SNPs in a cardiovascular context

The *VEGFA* gene has a 16.3 kb coding region located at 6p21.1 on the long arm of chromosome 6, including eight exons and seven introns (56, 57). The first five exons are constitutively present among VEGF-A isoforms, since they encode the signal sequence for protein processing and residues that bind to VEGF receptors (54, 58). Meanwhile, exons 6 and 7 contain the heparin binding domains that allow some isoforms to bind to cell surfaces and impact their activity or bioavailability depending on which are present (59, 60). Lastly, exon 8 undergoes post-translational readthrough due to a non-canonical stop codon, leading to the production of sub-exons 8a and 8b, with the latter being reported to be present in a unique isoform with anti-angiogenic activity observed in bone disorders and brain diseases (54, 61–63). So far, 16 distinct VEGF-A isoforms have been identified (47, 54). The different isoforms depend on the presence or absence of exons 6 and 7, which affect the affinity for heparin or heparan sulfate proteoglycans. For example, the most prevalent VEGFA isoform is $VEGFA_{165}$, which lacks exon 6, but has moderate heparin affinity allowing the isoform to remain bound to cell surfaces (64). Comparatively the isoform subtype $VEGFA_{121}$ lacks exon 6 and 7 so it is found only in free form (64). Despite their size difference most of the VEGF-A isoforms act as endothelial cell mitogens, upregulate the endothelial expression of adhesion molecules and present pro-angiogenic activity (36, 51, 64). Pathologies caused by increased angiogenesis include inflammatory diseases, cancers, retinopathy and atherosclerosis, while reduced angiogenesis has been observed in bone disorders and brain diseases (61). The overall *VEGFA* gene structure including SNPs with reported influence on VEGF-A expression levels (discussed below and in Supplementary Tables S1, S2) is presented in Figure 2.

Altered plasma and tissue levels of VEGF-A have been observed in various conditions including ischemic heart disease (IHD), CAD, strokes, heart failure, and myocardial infarction (38, 66–68). Due to its impact on angiogenic processes, the effect of high VEGF-A circulating levels on CVD onset varies. High VEGF-A levels are associated with various CVD risk factors including smoking, hypercholesterolemia, diabetes, hypertension, and hyperglycemia (36). Additionally, increased VEGF-A activity has been associated with inflammation, increased blood pressure and an increase in the formation of atherosclerotic lesions, leading to CHD (20, 69–71). The impact of angiogenic molecules on atherosclerosis has been reviewed elsewhere (39).

Expression of *VEGFA* can be upregulated by the hypoxia inducible factor, p53 allele polymorphisms, thyroid stimulating hormone, estrogen levels and oxygen tension (45, 47). This matches studies that show VEGF-A production is influenced by elements associated with atherosclerosis including LDL concentration, hypoxia, and interleukin activity (38, 49, 72). The increased production of VEGF-A can negatively impact human health by influencing the development of atherosclerotic plaques, by affecting vascular dilation, adhesion protein expression, monocyte migration, endothelium permeability and increased trans-endothelial lipid

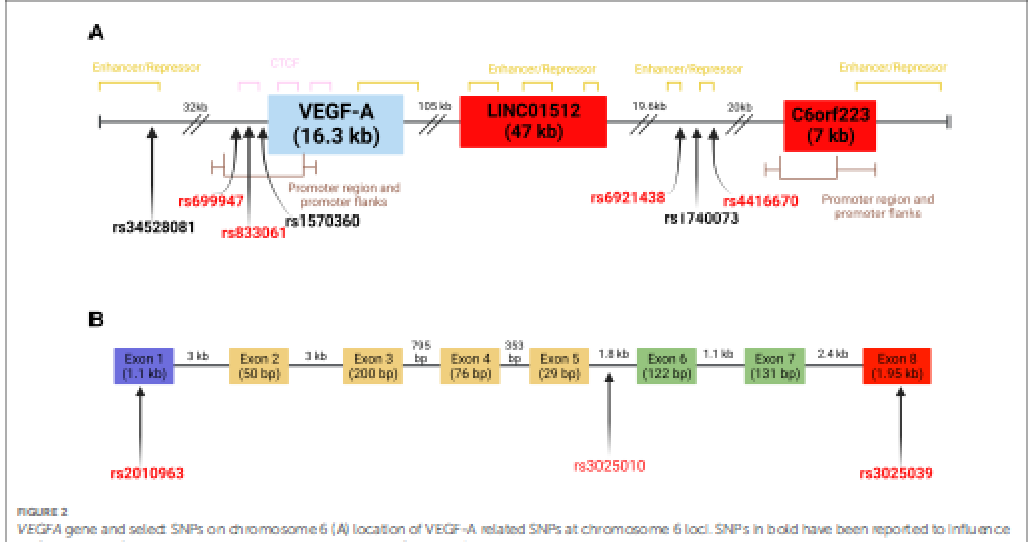


FIGURE 2
VEGFA gene and select SNPs on chromosome 6 (A) location of VEGF-A related SNPs at chromosome 6 loci. SNPs in bold have been reported to influence VEGF-A levels. SNPs in red have reported associations to CVD risk factors or biomarkers. Red boxes indicate regions that represent lncRNA. Full gene length is indicated individually. Distances between SNPs and genes are indicated above dashes. (B) Schematic representation of SNPs located within exons and introns of *VEGFA*. Individual exon and intron lengths are shown (65). The purple exon contains the signal sequence of the gene. The red exon confers anti-angiogenic capabilities when present. Green exons contain heparin-binding domains with extracellular matrix components. Yellow exons are involved in VEGF receptor binding. Location of SNPs and genomic regions determined using data from the UCSC genome browser and Ensembl databases (GRCh38-/hg38). Created with BioRender.com.

migration (38, 39). High levels of VEGF-A in plasma have been associated with increased plaque growth and subsequent lesion vulnerability that can cause intraplaque hemorrhage (73). There is evidence that proinflammatory cytokines (e.g., IL-1, IL-6, and IL-18) present during CVD onset can enhance VEGF-A production, thus exacerbating atherosclerotic lesion development (74–76).

VEGF-A is considered a highly polymorphic gene because of the 148 untranslated region (UTR), 209 exon, 779 intron, and 124 near-gene variants that have been identified (77). At least 30 SNPs within the untranslated, exon, intron and promoter regions may have the potential to influence variation in VEGF-A expression (78, 79). This genetic influence over VEGF-A circulating levels has been explored in various studies. Debette et al. (80) investigated the heritability of VEGF-A levels in healthy individuals without a cancer diagnosis. This study identified four common variants (rs6921438, rs4416670, rs6993770 and rs10738760) distributed across three chromosomes that were independently associated with circulating VEGF-A levels and explained up to 48% of the heritability of serum VEGF-A levels (80). A meta-analysis of GWAS data evaluated the association of variants with circulating VEGF-A levels (81). Choi et al. (81) found a total of ten SNPs contributed up to 52% of the variability in circulating VEGF-A levels with some SNPs associated with increased or decreased VEGF-A levels compared to median. Additional information on the study details of SNPs identified by these groups and other studies are presented in **Supplementary Tables S1, S2**. The **Supplementary Material** also includes SNPs that have been studied in relation to VEGF-A levels in healthy individuals, CVDs, or comorbidities related to the risk of CVD (e.g., diabetes, metabolic syndrome, hypertension).

Some of the SNPs that have been studied are located within exonic regions of *VEGFA* (82). One noteworthy eQTL is rs2010963 from exon 1 of *VEGFA* (Figure 2B). The CC genotype has been associated with increased VEGF-A levels in type 2 diabetes mellitus (T2DM) (83, 84). Furthermore, the rs2010963 CC genotype has been linked to risk factors including heart rate (83), blood glucose levels (77), blood pressure, cholesterol and HDL levels (83, 85). There is also evidence for this variant influencing VEGF-A levels in non-CVDs such as glioma (86) and diabetic retinopathy (48, 87).The variant rs3025039, located within exon 8 of *VEGFA*, has similar effects (Figure 2B). Dong et al. (88) observed that patients diagnosed with gestational diabetes mellitus carrying the TT genotype had higher levels of VEGFA compared to healthy pregnant women (88). Meanwhile Ruggiero et al. (89) reported that the TT genotype was associated with lower median levels of VEGFA in healthy population samples from villages in Southern Italy. Some studies showed the CT genotype of rs3025039 is associated with reduced VEGFA levels as well as reducing risk of presenting with CHD and T2DM (77, 89). The associations reported for the rs3025039 variant demonstrate its link to the cardiovascular system, but the variety of findings suggest additional studies are needed. An additional variant that has shown association to the cardiovascular system is rs3025010 located in intron 5 (Figure 2B). In a Chinese cohort diagnosed with hypertension the C allele of this variant was observed to be associated with

lower systolic and diastolic blood pressure measurements (20). Furthermore, in a Chinese case-control study, it was observed that the CC genotype of rs3025010 reduced risk of brain arteriovenous malformation (91). This evidence shows a clear link to CVD risk which could be further explored in additional ethnic groups to validate or identify other biomarker associations.

Other variants of interest can be found at the same loci, but outside the intron and exon regions of *VEGFA* (25, 80). rs69214328, is located within an enhancer region found between two long non-coding RNA genes (Figure 2A). The GWAS findings of Choi et al. (81) and Debette et al. (80) identified that the A allele of rs69214328 is associated with lower serum levels of the VEGF-A protein. Additionally, the same allele has also been reported to influence the variability of HDL and LDL in individuals of European ancestry (92). The A allele of rs6921438 appears to have eQTL activities since increased serum levels of IL-6, TNF-α and VEGF-A were observed in interaction with SNPs rs6993770 (Chr8), rs4416670 (Chr6) and rs10738760 (Chr9), respectively (93). Two additional variants (rs1740073 and rs34528081) located on chromosome 6 (Figure 2A) were identified by Choi et al. (81) to be associated with serum levels of VEGF-A (Supplementary Table S1). Furthermore, the T allele of rs34528081 was observed to be associated with increased VEGF-A serum levels in an additional GWAS study (Supplementary Table S1). Meanwhile, the T allele of rs1740073 has been reported to associate with increased VEGF-A serum in a GWAS study while analysis of IHD using 1,000 Genomes European data reported that the same allele could contribute to VEGF variance (66).

Another variant that has been studied is rs699947, which is located in the promoter region of *VEGFA* (Figure 2A). Various groups report that the AA genotype of rs699947 is associated with increased risk in cardiovascular pathologies including CAD, CHD, stroke and congenital heart diseases (Supplementary Table S1). The A allele of rs699947 has been associated with total cholesterol, LDL and apolipoprotein B (77, 83, 94). These associations have been observed across different ethnic groups, which further suggests rs699947 is a potential genetic risk marker for CVDs (89, 95, 96). For its part, rs833061 is another variant that is located within the promoter region of *VEGFA* (Figure 2A) whose CT genotype has been observed to reduce VEGF-A levels in a T2DM cohort (77). Other reports have also shown this variant is associated with hypertension and a meta-analysis of 3 cohorts implies this variant can influence congenital heart disease risk in individuals of Asian ancestry (Supplementary Table S1). A variant located further from the promoter region that presents a similar array of findings related to lipid metabolism and inflammatory molecules is rs4416670 (Figure 2A). Both its alleles have been linked to CVD risk factors and biomarkers (Supplementary Table S1). Specifically, the T allele was reported by Debette et al. (80) to increase VEGF-A serum levels while a study by Azimi-Nezhad et al. (93) reported the same allele could decrease IL-6 levels by interacting with rs6921438 (Chr6) and rs10738760 (Chr9). However, Azimi-Nezhad et al. (93) also report that the C allele of rs4416670 can increase TNF-α and IL-6 levels by interacting with the A allele of rs6921438 thus implying a link between both VEGF-A related SNPs and inflammatory molecules. Additionally, the C allele has also been observed in other studies to be associated with apolipoprotein E levels, hypertension and metabolic syndrome (92, 97). These findings demonstrate links between VEGF-A related SNPs and lipid metabolism, inflammatory biomarkers and CVD risk factors.

Some gene variants have findings of associations with molecules used in CVD risk assessment. For example, the rs1570360 variant located in the promoter region of VEGF-A (Figure 2A), was observed to contribute to an increased risk of congenital heart disease (98). Some reports showed that the GA genotype of this variant is associated with a reduced left ventricular ejection fraction and extracranial internal carotid artery (ECICA) stenosis which are both risk factors for systemic hypertension and ischemic stroke, respectively (Supplementary Table S1). However, in a Chinese study the GG genotype was observed to increase susceptibility for coronary heart disease in patients with high smoking habits and diagnosed with hypertension. As such, this variant shows consistent links to CV risk factors which, given its location, could be attributed to a potential influence on VEGF as observed in variants located within the promoter region (rs699947 and rs833061).

Similar studies have been reported for other SNPs located across the genome, often denoted as trans-acting SNPs (Supplementary Table S2). Broadly, these eQTL SNPs have been associated with increased risk of CVDs (e.g., CAD, CHD, IHD) (66, 89, 99, 100) or metabolic syndrome (81). One example rs1870377, located on chromosome 4 in exon 11 of the *VEGFR2* (*KDR*) gene (Supplementary Figure S1) can influence cardiovascular outcomes. Li et al. (72) reported that the AA genotype reduces risk of unfavorable CVD outcomes, particularly those related with disability, in an Asian ancestry cohort. Marks et al. (99) also reported that the AA genotype associated with reduced risk of heart failure readmission and the A allele associated with high levels of VEGF system components, specifically sFlt-1 and KDR (101), and increased the risk of ischemic stroke in a Korean cohort (100). The TA and TT genotypes were both associated with increased CHD prevalence in Han Chinese populations (89, 99). Location of additional SNPs influencing VEGF-A expression levels within the *VEGFR2* (*KDR*) gene is presented in Supplementary Figure S1. Additional associations observed for trans-acting SNPs are presented in Supplementary Table S2.

rs6993770 is located on chromosome 8 in intron 4 of the *ZPFM2* gene, which codes for a protein involved in heart morphogenesis and coronary vessel development. Broadly, studies on this variant have shown relationships with VEGF-A, CVD and CVD risk factors (Supplementary Table S2). In the GWAS findings of Choi et al. (81) and the Mendelian Randomization study done by Au Yeung (66), the A allele correlated with increased VEGF-A serum levels. The GWAS findings of Debette et al. (80) showed the T allele was associated with increased VEGF-A serum levels. Other studies involving individuals of European and Iranian ancestry observed the T allele was also associated with risk biomarkers of CVD, particularly fasting blood glucose, triglyceride levels, systolic blood pressure and HDL levels (92, 102). The TA genotype has been reported to increase the risk of metabolic syndrome (102), and impacts the expression of adhesion molecules (ICAM-1, E-

selectin) as well as IL-6 levels (93). Meanwhile, the TT genotype appears to contribute to metabolic syndrome risk in individuals with low iron intake (97). This spectrum of reports demonstrates the range of associations that alleles and genotypes of trans-acting SNPs, such as rs1870377 or rs6993770, may have within the cardiovascular system. Additional trans-acting SNPs (e.g., rs2071559, rs114694170, rs6993770, rs10738760, rs10761741, rs4782371 have been reported to be capable of influencing VEGF-A circulating levels (81) or soluble VEGFR levels (rs1870377) (101). Specific study details and overall findings are presented in **Supplementary Table S2**. Notably, two SNPs (rs2305948 and rs7667298) have associations with potential CVD risk, but their direct impact on VEGF system components was observed in cancer related studies (103, 104). Interestingly, trans-acting SNPs most likely involve interactions with molecules or homeostatic mechanisms that have known roles in CVD onset, including inflammatory interleukins (70, 93), triglycerides, adhesion molecules, blood cell count and blood pressure (66, 102, 105, 106). There are cases of specific variants that correlate with increased risk of presenting major adverse coronary events (rs2305948, rs7667298) (106), CHD (rs2305948, rs1870377, rs2071559, rs7667298) (89, 99), ischemic stroke (rs1870377) (100) and metabolic syndrome (rs6993770) (102). As such, some SNPs appear to be potential contributors to phenotypes (IHD, CAD, CHD) while others may increase or reduce disease risk depending on the presence or absence of risk factors (72, 89, 101).

Conclusion

Overall, the impact of VEGF-A related SNPs in various forms of heart disease has been explored in many different types of studies. The collective evidence reveals a critical subset of cis-acting SNPs mapping to the region of *VEGFA* (**Figure 2** and **Supplementary Table S1**), several trans-acting SNPs mapping in the region of the *VEGFR2* gene (**Supplementary Figure S1**) and elsewhere on the human genome (**Supplementary Table S2**), with repeatable associations with circulating levels of VEGF-A. A small group of SNPs reproducibly associate with established biomarkers and risk factors for heart disease (rs2010963, rs3025039, rs1570360, rs699947, rs6921438) or with increased susceptibility to common heart disease pathologies (rs2010963, rs3025039, rs1570360, rs699947, rs2305948, rs1870377). This minireview highlights that these SNPs can be potential markers for CVDs and may influence significant biological pathways that impact the cardiovascular system (e.g., lipid metabolism). The wide range of pathologies that VEGF-A and its related SNPs impact emphasizes the complexity of VEGF-A interactions within the cardiovascular system. Both cis- and trans-acting SNP eQTLs can affect expression levels, but there remain many unknowns around the specific mechanisms involved. There is a clear link between SNPs and VEGF-A levels as well as established cardiovascular disease biomarkers (HDL, LDL, BNP, NTproBNP). Together these have the potential to act synergistically on the development of CVDs.

The complexity of SNP influences on CVD and CVD risk factors reinforces the importance of studying them in relation to VEGF-A. Particularly considering how altered levels of VEGF-A contribute to disease onset or exacerbate an individual's health depending on the risk factors they present with. Exploring the link between CVDs, SNPs, and VEGF-A may contribute to improved cardiovascular disease risk assessment, prevention, treatment, and prognosis.

Author contributions

JCMA, BP, and RP: conceived the concept of the mini review. JM-A: canvassed the literature and drafted the manuscript. JM-A, BP, BM, CB, and RP: refined and extended the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2023.1190513/full#supplementary-material.

References

1. World Health Organization. *Cardiovascular disease (CVDs)* (2021). Available at: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (Accessed June 11, 2021).

2. Joseph P, Leong D, McKee M, Anand SS, Schwalm J-D, Teo K, et al. Reducing the global burden of cardiovascular disease, part 1. *Circ Res.* (2017) 121(6):677–94. doi: 10.1161/CIRCRESAHA.117.308903

3. Sanchis-Gomar F, Perez-Quilis C, Leischik R, Lucia A. Epidemiology of coronary heart disease and acute coronary syndrome. *Ann Transl Med.* (2016) 4(13):256. doi: 10.21037/atm.2016.06.33

4. Malakar AK, Choudhury D, Halder B, Paul P, Uddin A, Chakraborty S. A review on coronary artery disease, its risk factors, and therapeutics. *J Cell Physiol.* (2019) 234(10):16812–23. doi: 10.1002/jcp.28350

5. Khamis RY, Ammari T, Mikhail GW. Gender differences in coronary heart disease. *Heart*. (2016) 102(14):1142–9. doi: 10.1136/heartjnl-2014-306463

6. Libby P. The changing landscape of atherosclerosis. *Nature*. (2021) 592 (7855):524–33. doi: 10.1038/s41586-021-03392-8

7. Libby P, Buring JE, Badimon L, Hansson GK, Deanfield J, Bittencourt MS, et al. Atherosclerosis. *Nat Rev Dis Primers*. (2019) 5(1):56. doi: 10.1038/s41572-019-0106-z

8. Akyuz A. Exercise and coronary heart disease. *Adv Exp Med Biol*. (2020) 1228:169–79. doi: 10.1007/978-981-15-1792-1_11

9. Hartman J, Frishman WH. Inflammation and atherosclerosis. *Cardiol Rev*. (2014) 22(3):147–51. doi: 10.1097/CRD.0000000000000021

10. Damen JAAG, Hooft L, Schuit E, Debray TPA, Collins GS, Tzoulaki I, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *Br Med J*. (2016) 353:i2416. doi: 10.1136/bmj.i2416

11. Assimes TL, Roberts R. Genetics: implications for prevention and management of coronary artery disease. *J Am Coll Cardiol*. (2016) 68(25):2797–818. doi: 10.1016/j.jacc.2016.10.039

12. Roberts R. Genetics-current and future role in the prevention and management of coronary artery disease. *Curr Atheroscler Rep*. (2016) 18(12):78. doi: 10.1007/s11883-016-0628-8

13. Roberts R. A breakthrough in genetics and its relevance to prevention of coronary artery disease in LMIC. *Glob Heart*. (2017) 12(3):247–57. doi: 10.1016/j.gheart.2017.04.001

14. Roberts R, Fair J. Genetics, its role in preventing the pandemic of coronary artery disease. *Clin Cardiol*. (2021) 44(6):771–9. doi: 10.1002/clc.23627

15. McPherson R, Tybjaerg-Hansen A. Genetics of coronary artery disease. *Circ Res*. (2016) 118(4):564–78. doi: 10.1161/CIRCRESAHA.115.306566

16. 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) 375 (24):2349–58. doi: 10.1056/NEJMoa1605086

17. Roberts R. A genetic basis for coronary artery disease. *Trends Cardiovasc Med*. (2015) 25(3):171–8. doi: 10.1016/j.tcm.2014.10.008

18. LeBlanc M, Zuber V, Andreassen BK, Witoelar A, Zeng L, Bettella F, et al. Identifying novel gene variants in coronary artery disease and shared genes with several cardiovascular risk factors. *Circ Res*. (2016) 118(1):83–94. doi: 10.1161/CIRCRESAHA.115.306629

19. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. *Nature*. (2018) 558(7708):73–9. doi: 10.1038/s41586-018-0175-2

20. Zhao Z, Gong C, Gao Y, Liu X, Wu S, Zhao H, et al. Association between single nucleotide polymorphisms in cardiovascular developmental critical genes and hypertension: a propensity score matching analysis. *Int J Hypertens*. (2020) 2020:1–8. doi: 10.1155/2020/9185697

21. Earle N, Yeo Han D, Pilbrow A, Crawford J, Smith W, Shelling AN, et al. Single nucleotide polymorphisms in arrhythmia genes modify the risk of cardiac events and sudden death in long QT syndrome. *Heart Rhythm*. (2014) 11(1):76–82. doi: 10.1016/j.hrthm.2013.10.005

22. Roberts R, Chang CC, Hadley T. Genetic risk stratification: a paradigm shift in prevention of coronary artery disease. *JACC Basic Transl Sci*. (2021) 6(3):287–304. doi: 10.1016/j.jacbts.2020.09.004

23. Geiger M. *Fundamentals of Vascular Biology. Learning Materials in Biosciences*. Vienna, Austria: Springer (2019).

24. Zhang X, Johnson AD, Hendricks AE, Hwang SJ, Tanriverdi K, Ganesh SK, et al. Genetic associations with expression for genes implicated in GWAS studies for atherosclerotic cardiovascular disease and blood phenotypes. *Hum Mol Genet*. (2014) 23(3):782–95. doi: 10.1093/hmg/ddt461

25. Buroker NE. VEGFA rSNPs, transcriptional factor binding sites and human disease. *J Physiol Sci*. (2014) 64(1):73–6. doi: 10.1007/s12576-013-0293-4

26. Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. *Philos Trans R Soc Lond B Biol Sci*. (2013) 368(1620):20120362. doi: 10.1098/rstb.2012.0362

27. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. (2012) 489(7414):57–74. doi: 10.1038/nature11247

28. Roberts R, Campillo A, Schmitt M. Prediction and management of CAD risk based on genetic stratification. *Trends Cardiovasc Med*. (2020) 30(6):328–34. doi: 10.1016/j.tcm.2019.08.006

29. Wang K, Basu M, Malin J, Hannenhalli S. A transcription-centric model of SNP-age interaction. *PLoS Genet*. (2021) 17(3):e1009427. doi: 10.1371/journal.pgen.1009427

30. Yun H, Noh NI, Lee EY. Genetic risk scores used in cardiovascular disease prediction models: a systematic review. *RCM*. (2022) 23(1):1–9. doi: 10.31083/j.rcm2301008

31. O'Sullivan J, 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) 146(8):e93–e118. doi: 10.1161/CIR.0000000000001077

32. Gladding PA, Legget M, Fatkin D, Larsen P, Doughty R. Polygenic risk scores in coronary artery disease and atrial fibrillation. *Heart Lung Circ*. (2020) 29(4):634–40. doi: 10.1016/j.hlc.2019.12.004

33. Roberts R, Campillo A. Genetic stratification for primary prevention of coronary artery disease. *Curr Opin Cardiol*. (2018) 33(5):529–34. doi: 10.1097/HCO.0000000000000542

34. Knowles JW, Ashley EA. Cardiovascular disease: the rise of the genetic risk score. *PLoS Med*. (2018) 15(3):e1002546. doi: 10.1371/journal.pmed.1002546

35. Rotter JI, Lin HJ. An outbreak of polygenic scores for coronary artery disease. *J Am Coll Cardiol*. (2020) 75(22):2781–4. doi: 10.1016/j.jacc.2020.04.054

36. Smith GA, Fearnley GW, Harrison MA, Tomlinson DC, Wheatcroft SB, Ponnambalam S. Vascular endothelial growth factors: multitasking functionality in metabolism, health and disease. *J Inherit Metab Dis*. (2015) 38(4):753–63. doi: 10.1007/s10545-015-9838-4

37. Shibuya M. Vascular endothelial growth factor and its receptor system: physiological functions in angiogenesis and pathological roles in various diseases. *J Biochem*. (2013) 153(1):13–9. doi: 10.1093/jb/mvs136

38. Braile M, Marcella S, Cristinziano L, Galdiero MR, Modestino L, Ferrara AL, et al. VEGF-A in cardiomyocytes and heart diseases. *Int J Mol Sci*. (2020) 21 (15):1–18. doi: 10.3390/ijms21155294

39. Camaré C, Pucelle M, Nègre-Salvayre A, Salvayre R. Angiogenesis in the atherosclerotic plaque. *Redox Biol*. (2017) 12:18–34. doi: 10.1016/j.redox.2017.01.007

40. Senger DR, Galli SJ, Dvorak AM, Perruzzi CA, Harvey VS, Dvorak HF. Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid. *Science*. (1983) 219(4587):983–5. doi: 10.1126/science.6823562

41. Keck PJ, Hauser SD, Krivi G, Sanzo K, Warren T, Feder J, et al. Vascular permeability factor, an endothelial cell mitogen related to PDGF. *Science*. (1989) 246(4935):1309–12. doi: 10.1126/science.2479987

42. Connolly DT, Heuvelman DM, Nelson R, Olander JV, Eppley BL, Delfino JJ, et al. Tumor vascular permeability factor stimulates endothelial cell growth and angiogenesis. *J Clin Invest*. (1989) 84(5):1470–8. doi: 10.1172/JCI114322

43. Leung DW, Cachianes G, Kuang W-J, Goeddel DV, Ferrara N. Vascular endothelial growth factor is a secreted angiogenic mitogen. *Science*. (1989) 246 (4935):1306–9. doi: 10.1126/science.2479986

44. Nilklander S, Bordagaray MJ, Fernandez A, Hernandez M. Vascular endothelial growth factor: a translational view in oral non-communicable diseases. *Biomolecules*. (2021) 11(1):1–27. doi: 10.3390/biom11010085

45. Singh P, Singh M, Khinda R, Valecha S, Kumar N, Singh S, et al. Genetic scores of eNOS, ACE and VEGFA genes are predictive of endothelial dysfunction associated osteoporosis in postmenopausal women. *Int J Environ Res Public Health*. (2021) 18 (3):1–18. doi: 10.3390/ijerph18030972

46. Holmes DI, Zachary I. The vascular endothelial growth factor (VEGF) family: angiogenic factors in health and disease. *Genome Biol*. (2005) 6(2):209. doi: 10.1186/gb-2005-6-2-209

47. Shaik F, Cuthbert GA, Homer-Vanniasinkam S, Muench SP, Ponnambalam S, Harrison MA. Structural basis for vascular endothelial growth factor receptor activation and implications for disease therapy. *Biomolecules*. (2020) 10(12):1–25. doi: 10.3390/biom10121673

48. Hu L, Gong C, Chen X, Zhou H, Yan J, Hong W. Associations between vascular endothelial growth factor gene polymorphisms and different types of diabetic retinopathy susceptibility: a systematic review and meta-analysis. *J Diabetes Res*. (2021) 2021:7059139. doi: 10.1155/2021/7059139

49. Le THV, Kwon SM. Vascular endothelial growth factor biology and its potential as a therapeutic target in rheumatic diseases. *Int J Mol Sci*. (2021) 22(10):1–17. doi: 10.3390/ijms22105387

50. Gu H, Chen W, Yin J, Chen S, Zhang J, Gong J. Methionine sulfoxide reductase a rs10903323 G/A polymorphism is associated with increased risk of coronary artery disease in a Chinese population. *Clin Biochem*. (2013) 46(16-17):1668–72. doi: 10.1016/j.clinbiochem.2013.07.011

51. Zhou Y, Zhu X, Cui H, Shi J, Yuan G, Shi S, et al. The role of the VEGF family in coronary heart disease. *Front Cardiovasc Med*. (2021) 8:1–16. doi: 10.3389/fcvm.2021.738325

52. Oszajca K, Szemraj J, Wyrzykowski D, Chrzanowska B, Salamon A, Przewratil P. Single-nucleotide polymorphisms of VEGF-A and VEGFR-2 genes and risk of infantile hemangioma. *Int J Dermatol*. (2018) 57(10):1201–7. doi: 10.1111/ijd.14127

53. Apte RS, Chen DS, Ferrara N. VEGF in signaling and disease: beyond discovery and development. *Cell*. (2019) 176(6):1248–64. doi: 10.1016/j.cell.2019.01.021

54. Peach CJ, Mignone VW, Arruda MA, Alcobia DC, Hill SJ, Kilpatrick LE, et al. Molecular pharmacology of VEGF-A isoforms: binding and signalling at VEGFR2. *Int J Mol Sci*. (2018) 19(4):1264. doi: 10.3390/ijms19041264

55. Simons M, Gordon E, Claesson-Welsh L. Mechanisms and regulation of endothelial VEGF receptor signalling. *Nat Rev Mol Cell Biol*. (2016) 17(10):611–25. doi: 10.1038/nrm.2016.87

56. Pagès G, Pouysségur J. Transcriptional regulation of the vascular endothelial growth factor gene–a concert of activating factors. *Cardiovasc Res.* (2005) 65(3):564–73. doi: 10.1016/j.cardiores.2004.09.032

57. Tischer E, Mitchell R, Hartman T, Silva M, Gospodarowicz D, Fiddes JC, et al. The human gene for vascular endothelial growth factor. Multiple protein forms are encoded through alternative exon splicing. *J Biol Chem.* (1991) 266(18):11947–54. doi: 10.1016/S0021-9258(18)99049-6

58. Keyt BA, Berleau LT, Nguyen HV, Chen H, Heinsohn H, Vandlen R, et al. The carboxyl-terminal domain (111–165) of vascular endothelial growth factor is critical for its mitogenic potency. *J Biol Chem.* (1996) 271(13):7788–95. doi: 10.1074/jbc.271.13.7788

59. Krilleke D, DeErkenez A, Schubert W, Giri I, Robinson GS, Ng Y-S, et al. Molecular mapping and functional characterization of the VEGF164 heparin-binding domain. *J Biol Chem.* (2007) 282(38):28045–56. doi: 10.1074/jbc.M700319200

60. Stevens M, Oltean S. Modulation of receptor tyrosine kinase activity through alternative splicing of ligands and receptors in the VEGF-A/VEGFR axis. *Cells.* (2019) 8(4):288. doi: 10.3390/cells8040288

61. Dehghanian F, Hojati Z, Kay M. New insights into VEGF-A alternative splicing: key regulatory switching in the pathological process. *Avicenna J Med Biotechnol.* (2014) 6(4):192–9. PMID: 25414781

62. Eswarappa SM, Potdar AA, Koch WJ, Fan Y, Vasu K, Lindner D, et al. Programmed translational readthrough generates antiangiogenic VEGF-Ax. *Cell.* (2014) 157(7):1605–18. doi: 10.1016/j.cell.2014.04.033

63. Bridgett S, Dellett M, Simpson DA. RNA-sequencing data supports the existence of novel VEGFA splicing events but not of VEGFAxxxb isoforms. *Sci Rep.* (2017) 7(1):58. doi: 10.1038/s41598-017-00100-3

64. Al Kawas H, Saaid I, Jank P, Westhoff CC, Denkert C, Pross T, et al. How VEGF-A and its splice variants affect breast cancer development–clinical implications. *Cell Oncol.* (2022) 45(2):227–39. doi: 10.1007/s13402-022-00665-w

65. Arcondéguy T, Lacazette E, Millevoi S, Prats H, Touriol C. VEGF-A mRNA processing, stability and translation: a paradigm for intricate regulation of gene expression at the post-transcriptional level. *Nucleic Acids Res.* (2013) 41(17):7997–8010. doi: 10.1093/nar/gkt539

66. Au Yeung SL, Lam H, Schooling CM. Vascular endothelial growth factor and ischemic heart disease risk: a Mendelian randomization study. *J Am Heart Assoc.* (2017) 6(8):1–8. doi: 10.1161/JAHA.117.005619

67. Zhang Y, Ma T, Hu H, Wang J, Zhou S. Serum vascular endothelial growth factor as a biomarker for prognosis of minor ischemic stroke. *Clin Neurol Neurosurg.* (2020) 196:106060. doi: 10.1016/j.clineuro.2020.106060

68. Rodríguez-Rodríguez L, García-Bermúdez M, Martín J, Fernández-Gutiérrez B, González-Juanatey C, Vazquez-Rodriguez TR, et al. Vascular endothelial growth factor a and cardiovascular disease in rheumatoid arthritis patients. *Tissue Antigens.* (2011) 77(4):291–7. doi: 10.1111/j.1399-0039.2010.01625.x

69. Heinonen SE, Kivelä AM, Huusko J, Dijkstra MH, Gurzeler E, Mäkinen PI, et al. The effects of VEGF-A on atherosclerosis, lipoprotein profile, and lipoprotein lipase in hyperlipidaemic mouse models. *Cardiovasc Res.* (2013) 99(4):716–23. doi: 10.1093/cvr/cvt148

70. Ruotsalainen SE, Partanen JJ, Cichonska A, Lin J, Benner C, Surakka I, et al. An expanded analysis framework for multivariate GWAS connects inflammatory biomarkers to functional variants and disease. *Eur J Hum Genet.* (2021) 29(2):309–24. doi: 10.1038/s41431-020-00730-8

71. Zhao X, Meng L, Jiang J, Wu X. Vascular endothelial growth factor gene polymorphisms and coronary heart disease: a systematic review and meta-analysis. *Growth Factors.* (2018) 36(3-4):153–63. doi: 10.1080/08977194.2018.1477141

72. Li Z, Wang M, Gu J, Zhao L, Guo Y, Zhang Z, et al. Missense variants in hypoxia-induced VEGFA/VEGFR2 signaling predict the outcome of large artery atherosclerotic stroke. *Cell Mol Neurobiol.* (2021) 41(6):1217–25. doi: 10.1007/s10571-020-00890-7

73. Michel J-B, Virmani R, Arbustini E, Pasterkamp G. Intraplaque haemorrhages as the trigger of plaque vulnerability. *Eur Heart J.* (2011) 32(16):1977–85. doi: 10.1093/eurheartj/ehr054

74. Mitrokhin V, Shim A, Aksyonov A, Zotov A, Konev A, Ovchinnikov R, et al. Circulating interleukin-18: association with IL-8, IL-10 and VEGF serum levels in patients with and without heart rhythm disorders. *Int J Cardiol.* (2016) 215:105–9. doi: 10.1016/j.ijcard.2016.04.002

75. Tzeng HE, Tsai CH, Chang ZL, Su CM, Wang SW, Hwang WL, et al. Interleukin-6 induces vascular endothelial growth factor expression and promotes angiogenesis through apoptosis signal-regulating kinase 1 in human osteosarcoma. *Biochem Pharmacol.* (2013) 85(4):531–40. doi: 10.1016/j.bcp.2012.11.021

76. Fahey E, Doyle SL. IL-1 family cytokine regulation of vascular permeability and angiogenesis. *Front Immunol.* (2019) 10:1–15. doi: 10.3389/fimmu.2019.01426

77. Sellami N, Lamine LB, Turki A, Sarray S, Jailani M, Al-Ansari AK, et al. Association of VEGFA variants with altered VEGF secretion and type 2 diabetes: a case-control study. *Cytokine.* (2018) 106:29–34. doi: 10.1016/j.cyto.2018.03.003

78. Stevens A, Soden J, Brenchley PE, Ralph S, Ray DW. Haplotype analysis of the polymorphic human vascular endothelial growth factor gene promoter. *Cancer Res.* (2003) 63(4):812–6. PMID: 12591731

79. Kim Y-J, Chung WC, Jun K-H, Chin H-M. Genetic polymorphisms of vascular endothelial growth factor (VEGF) associated with gastric cancer recurrence after curative resection with adjuvant chemotherapy. *BMC Cancer.* (2019) 19(1):483. doi: 10.1186/s12885-019-5702-5

80. Debette S, Visvikis-Siest S, Chen MH, Ndiaye NC, Song C, Destefano A, et al. Identification of cis- and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels. *Circ Res.* (2011) 109(5):554–63. doi: 10.1161/CIRCRESAHA.111.243790

81. Choi SH, Ruggiero D, Sorice R, Song C, Nutile T, Vernon Smith A, et al. Six novel loci associated with circulating VEGF levels identified by a meta-analysis of genome-wide association studies. *PLoS Genet.* (2016) 12(2):e1005874. doi: 10.1371/journal.pgen.1005874

82. Cui QT, Li Y, Duan CH, Zhang W, Guo XL. Further evidence for the contribution of the vascular endothelial growth factor gene in coronary artery disease susceptibility. *Gene.* (2013) 521(2):217–21. doi: 10.1016/j.gene.2013.03.091

83. Skrypnik D, Mostowska A, Jagodziński PP, Bogdański P. Association of rs699947 (−2578 C/A) and rs2010963 (−634 G/C) single nucleotide polymorphisms of the VEGF gene, VEGF-A and leptin serum level, and cardiovascular risk in patients with excess body mass: a case-control study. *J Clin Med.* (2020) 9(2):1–16. doi: 10.3390/jcm9020469

84. Merlo S, Starcevic JN, Mankoc S, Letonja MS, Vujkovac AC, Zorc M, et al. Vascular endothelial growth factor gene polymorphism (rs2010963) and its receptor, kinase insert domain-containing receptor gene polymorphism (rs2071559), and markers of carotid atherosclerosis in patients with type 2 diabetes mellitus. *J Diabetes Res.* (2016) 2016:1–6. doi: 10.1155/2016/1482194.

85. Totomoch-Serra A, de Lourdes Muñoz M, Burgueño J, Revilla-Monsalve MC, Diaz-Badillo A, Muñoz MdL. Association of common polymorphisms in the and genes with type 2 diabetes-related traits in Mexicans. *Arch Med Sci.* (2018) 14(6):1361–73. doi: 10.5114/aoms.2018.74757

86. Linhares P, Viana-Pereira M, Ferreira M, Amorim J, Nabiço R, Pinto F, et al. Genetic variants of vascular endothelial growth factor predict risk and survival of gliomas. *Tumour Biol.* (2018) 40(3):1010428318766273. doi: 10.1177/1010428318766273

87. Petrovič MG, Korošec P, Košnik M, Osredkar J, Hawlina M, Peterlin B, et al. Local and genetic determinants of vascular endothelial growth factor expression in advanced proliferative diabetic retinopathy. *Mol Vision.* (2008) 14:1382–7. PMID: 18682813

88. Dong P-P. Association of vascular endothelial growth factor expression and polymorphisms with the risk of gestational diabetes mellitus. *J Clin Lab Anal.* (2019) 33(2):e22686. doi: 10.1002/jcla.22686

89. Ruggiero D, Dalmasso C, Nutile T, Sorice R, Dionisi L, Aversano M, et al. Genetics of VEGF serum variation in human isolated populations of cilento:importance of VEGF polymorphisms. *PLoS One.* (2011) 6(2):e16982. doi: 10.1371/journal.pone.0016982

90. Liu D, Song J, Ji X, Liu Z, Cong M, Hu B. Association of genetic polymorphisms on VEGFA and VEGFR2 with risk of coronary heart disease. *Medicine.* (2016) 95(19):e3413. doi: 10.1097/MD.0000000000003413

91. Chen H, Gu Y, Wu W, Chen D, Li P, Fan W, et al. Polymorphisms of the vascular endothelial growth factor a gene and susceptibility to sporadic brain arteriovenous malformation in a Chinese population. *J Clin Neurosci.* (2011) 18(4):549–53. doi: 10.1016/j.jocn.2010.08.025

92. Stathopoulou MG, Bonnefond A, Ndiaye NC, Azimi-Nezhad M, El Shamieh S, Saleh A, et al. A common variant highly associated with plasma VEGFA levels also contributes to the variation of both LDL-C and HDL-C. *J Lipid Res.* (2013) 54(2):535–41. doi: 10.1194/jlr.P030551

93. Azimi-Nezhad M, Stathopoulou MG, Bonnefond A, Rancier M, Saleh A, Lamont J, et al. Associations of vascular endothelial growth factor (VEGF) with adhesion and inflammation molecules in a healthy population. *Cytokine.* (2013) 61(2):602–7. doi: 10.1016/j.cyto.2012.10.024

94. Yadav BK, Yadav R, Chang H, Choi K, Kim JT, Park MS, et al. Genetic polymorphisms rs699947, rs1570360, and rs3025039 on the VEGF gene are correlated with extracranial internal carotid artery stenosis and ischemic stroke. *Ann Clin Lab Sci.* (2017) 47(2):144–55. PMID: 28442515

95. Ma W-Q, Wang Y, Han X-Q, Zhu Y, Liu N-F. Association of genetic polymorphisms in vascular endothelial growth factor with susceptibility to coronary artery disease: a meta-analysis. *BMC Med Genet.* (2018) 19(1):1–12. doi: 10.1186/s12881-018-0628-3

96. Palmer BR, Paterson MA, Frampton CM, Pilbrow AP, Skelton L, Pemberton CJ, et al. Vascular endothelial growth factor-A promoter polymorphisms, circulating VEGF-A and survival in acute coronary syndromes. *PLoS One.* (2021) 16(7):1–5. doi: 10.1371/journal.pone.0254206

97. Hoseini Z, Ghayour-Mobarhan M, Avan A, Nematy M, Safarian M, Ferns GA, et al. VEGF gene polymorphism interactions with dietary trace elements intake in determining the risk of metabolic syndrome. *J Cell Biochem.* (2019) 120(2):1398–406. doi: 10.1002/jcb.27171

98. Wang W, Xu A, Xu H. The roles of vascular endothelial growth factor gene polymorphisms in congenital heart disease: a meta-analysis. *Growth Factors.* (2018) 36(5-6):232–8. doi: 10.1080/08977194.2018.1513506

99. Li L, Pan Y, Dai L, Liu B, Zhang D. Association of genetic polymorphisms on vascular endothelial growth factor and its receptor genes with susceptibility to coronary heart disease. *Med Sci Monit.* (2016) 22:31–40. doi: 10.12659/MSM.895163

100. Oh S-H, Min K-T, Jeon Y-J, Kim M-H, Moon J-S, Kim H-S, et al. Association between kinase insert domain-containing receptor gene polymorphism and haplotypes and ischemic stroke. *J Neurol Sci.* (2011) 308(1):62–6. doi: 10.1016/j.jns.2011.06.012

101. Marks ECA, Wilkinson TM, Frampton CM, Skelton L, Pilbrow AP, Yandle TG, et al. Plasma levels of soluble VEGF receptor isoforms, circulating pterins and VEGF system SNPs as prognostic biomarkers in patients with acute coronary syndromes. *BMC Cardiovasc Disord.* (2018) 18(1):169. doi: 10.1186/s12872-018-0894-1

102. Azimi-Nezhad M, Mirhafez SR, Stathopoulou MG, Murray H, Ndiaye NC, Bahrami A, et al. The relationship between vascular endothelial growth factor cis- and trans-acting genetic variants and metabolic syndrome. *Am J Med Sci.* (2018) 355(6):559–65. doi: 10.1016/j.amjms.2018.03.009

103. Góra-Tybor J, Szemraj J, Robak T, Jamroziak K. Clinical relevance of vascular endothelial growth factor type A (VEGFA) and VEGF receptor type 2 (VEGFR2) gene polymorphism in chronic lymphocytic leukemia. *Blood Cells Mol Dis.* (2015) 54 (2):139–43. doi: 10.1016/j.bcmd.2014.11.022

104. Mafu T, September A, Shamley D. KDR inferred haplotype is associated with upper limb dysfunction in breast cancer survivors of mixed ancestry. *Cancer Manag Res.* (2019) 11:3829–45. doi: 10.2147/CMAR.S191969

105. Salami A, El Shamieh S. Association between SNPs of circulating vascular endothelial growth factor levels, hypercholesterolemia and metabolic syndrome. *Medicina.* (2019) 55(8):1–9. doi: 10.3390/medicina55080464

106. Zhang LJ, Zhang YQ, Han X, Zhang ZT, Zhang ZQ. Association of VEGFR-2 gene polymorphisms with clopidogrel resistance in patients with coronary heart disease. *Am J Ther.* (2016) 23(6):e1663–70. doi: 10.1097/MJT.0000000000000231

Appendix 4.1 Imputed genotype comparison for five clinically relevant *VEGFA* locus SNPs in the CDCS cohort

		rs7767396 Genotypes					
Variable	**n**	**AA**	**n**	**AG**	**n**	**GG**	**p**
Anthropometric Variables							
Age (years)	589	65.9 ± 0.517	922	67.1 ± 0.401	402	67.3 ± 0.601	0.112
Activity (Sed, Mild, Mod, Act)	588	20.1%, 13.9%, 13.4%, 52.6%	919	20%, 12%, 14.9%, 53.1%	402	26.4%, 11.2%, 13.9%, 48.5%	0.149
Male Gender	589	436 (74.0%)	922	647 (70.2%)	402	290 (72.1%)	0.264
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	589	86.1%, 7.1%, 6.1%, 0.7%	922	90.9%, 5.9%, 3%, 0.2%	402	95%, 2.7%, 2%, 0.2%	**0.0002**
Height (m)	580	1.69 ± 0.004	910	1.69 ± 0.003	399	1.70 ± 0.004	0.534
Weight (kg)	582	81.01 ± 0.731	910	78.9 ± 0.544	399	79.0 ± 0.823	**0.047**
BMI (kg/m^2)	580	27.9 ± 0.233	907	27.3 ± 0.157	398	27.1 ± 0.223	**0.006**
Systolic BP (mmHg)	576	127.3 ± 0.903	906	129.1 ± 0.719	390	131.1 ± 1.13	**0.026**
Diastolic BP (mmHg)	576	74.7 ± 0.503	906	74.6 ± 0.389	390	75.1 ± 0.587	0.85
Beta blocker use	589	506 (85.9%)	922	799 (86.7%)	402	346 (86.1%)	0.907
Previous MI	584	181 (31%)	917	268 (29.2%)	399	124 (31.1%)	0.694
Tobacco use (Current, Ex Smoker, Non-smoker)	589	6.5%, 58.4%, 35.1%	922	6%, 56.9%, 37.1%	402	7.2%, 56%, 36.8%	0.852
Analytes							
Cholesterol (mmol/L)	455	4.98 ± 0.058	722	4.86 ± 0.042	313	4.85 ±0.068	0.152
Creatinine (mmol/L)	573	99.7 ± 2.19	892	99.2 ± 1.91	392	100.5 ± 1.86	0.913
Urate (mmol/L)	350	0.376 ± 0.005	562	0.366 ± 0.004	235	0.386 ± 0.007	**0.04**
Troponin I (ng/L) $	569	49.02 (12.6 - 85.4)	880	39.4 (18.8 – 59.9)	378	42.9 (19.9 – 65.8)	0.294
ANP (pg/mL) $	588	40.6 (38.3 – 42.9)	914	45.5 (43.4 – 47.5)	399	45.2 (41.8 – 48.6)	**0.015**
NT-ANP (pmol/L) $	588	1.25 (1.17 – 1.33)	914	1.41 (1.33 – 1.49)	399	1.38 (1.26 – 1.50)	**0.013**
BNP (pmol/L) $	588	23.5 (20.9 – 25.9)	914	28.6 (26.3 – 30.9)	399	27.4 (23.9 – 30.9)	**0.003**
NTproBNP (pg/ml) $	588	117.9 (105.5 – 130.5)	914	144.3 (131.6 – 157.1)	399	146 (124.9 – 167.1)	**0.002**
CNP (pmol/L) $	581	0.638 (0.603 – 0.673)	902	0.672 (0.642 – 0.703)	390	0.681 (0.63 – 0.731)	0.546
NT-CNP (pmol/L) $	580	23.1 (21.3 – 24.9)	899	22.8 (21.1 – 24.6)	391	23.5 (21.1 – 25.8)	0.326
Aldosterone (pmol/L) $	579	171.6 (147.1 – 196.1)	885	162.9 (155.7-170.1)	389	166.4 (154.9 – 177.8)	0.656
Endothelin(pmol/L) $	588	2.65 (2.57 – 2.73)	914	2.69 (2.62-2.75)	399	2.69 (2.59 – 2.80)	0.745
Adrenomedullin (pg/ml) $	567	8.08 (7.74 – 8.44)	882	8.81 (8.46 – 9.17)	391	8.66 (8.11 – 9.21)	**0.009**
VEGF-A (pg/mL) $	166	50.1 (44.2 – 55.8)	252	45.6 (40.9 – 50.2)	107	38.4 (33.6 – 43.1)	**0.003**

Variable	n	rs6921438 genotypes GG	n	GA	n	AA	p
Anthropometric Variables							
Age (years)	536	66.1 ± 0.539	946	66.9 ± 0.398	431	67.1 ± 0.583	0.359
Activity (Sed, Mild, Mod, Act)	535	20.6%, 13.5%, 13.6%, 52.3%	943	19.6%, 12.2%, 14.8%, 53.4%	431	26.2%, 11.8%, 13.7%, 48.3%	0.177
Male Gender	536	399 (74.4%)	946	664 (70.2%)	431	310 (71.9%)	0.217
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	536	87.9%, 5.6%, 6%, 0.6%	946	90.2%, 6.4%, 3.1%, 0.3%	431	93.5%, 3.7%, 2.6%, 0.2%	**0.017**
Height (m)	529	1.70 ± 0.004	932	1.69 ± 0.003	428	1.70 ± 0.004	0.480
Weight (kg)	531	80.9 ± 0.765	932	78.9 ± 0.537	428	79.2 ± 0.805	0.093
BMI (kg/m^2)	529	27.9 ± 0.241	929	27.3 ± 0.155	419	27.2 ± 0.228	**0.035**
Systolic BP (mmHg)	524	126.6 ± 0.931	929	129.2 ± 0.710	419	131.5 ± 1.10	**0.002**
Diastolic BP (mmHg)	524	74.1 ± 0.525	929	74.9 ± 0.385	419	75.1 ± 0.568	0.338
Beta blocker use	536	461 (86.0%)	946	819 (86.6%)	431	371 (86.1%)	0.943
Previous MI	531	166 (31.3%)	941	276 (29.3%)	428	131 (30.6%)	0.721
Tobacco use (Current, Ex Smoker, Non-smoker)	536	6.7%, 57.8%, 35.4%	946	5.8%, 57.7%, 36.5%	431	7.2%, 55.2%, 37.6%	0.792
Analytes							
Cholesterol (mmol/L)	415	4.96 ± 0.059	740	4.87 ± 0.041	335	4.86 ± 0.066	0.365
Creatinine (mmol/L)	523	98.7 ± 2.20	916	99.2 ± 1.87	418	101.5 ± 2.09	0.682
Urate (mmol/L)	321	0.378 ± 0.005	579	0.365 ± 0.004	247	0.386 ± 0.007	**0.018**
Troponin I (ng/L) [$]	518	50.9 (11.1 – 90.9)	906	36.4 (17.1 – 55.7)	403	48.2 (23.8 – 72.6)	0.118
ANP (pg/mL) [$]	535	40.8 (38.3 – 43.2)	939	45.2 (43.2 – 47.2)	427	44.9 (41.7 – 48.3)	**0.028**
NT-ANP (pmol/L) [$]	535	1.26 (1.17-1.35)	939	1.39 (1.32 – 1.48)	426	1.39 (1.27 – 1.50)	**0.045**
BNP (pmol/L) [$]	535	23.1 (20.5 – 25.7)	939	28.5 (26.2 – 30.9)	427	27.6 (24.2 – 30.9)	**0.006**
NTproBNP (pg/ml) [$]	535	118.4 (104.9 – 131.8)	939	142.5 (130 – 155)	427	146.1 (128 – 145)	**0.010**
CNP (pmol/L) [$]	528	0.643 (0.606 – 0.680)	929	0.668 (0.638 – 0.699)	416	0.679 (0.631 – 0.726)	0.734
NT-CNP (pmol/L) [$]	527	22.4 (21.2 – 23.7)	926	22.8 (21.1 – 24.5)	417	24.4 (21.4 – 27.3)	0.247
Aldosterone (pmol/L) [$]	527	171.7 (144.9 – 198.4)	910	164.1 (156.7 – 171.4)	416	164.6 (154.3 – 175)	0.632
Endothelin (pmol/L) [$]	535	2.65 (2.57 – 2.74)	939	2.67 (2.61 – 2.74)	427	2.71 (2.61 – 2.81)	0.665
Adrenomedullin (pg/ml) [$]	515	8.20 (7.83 – 8.57)	907	8.68 (8.36 – 9.00)	418	8.72 (8.11 – 9.33)	0.109
VEGF-A (pg/mL) [$]	152	51.3 (45.1 – 57.5)	261	45.4 (40.9 – 49.9)	112	37.9 (42.5 - 48.5)	**0.0009**

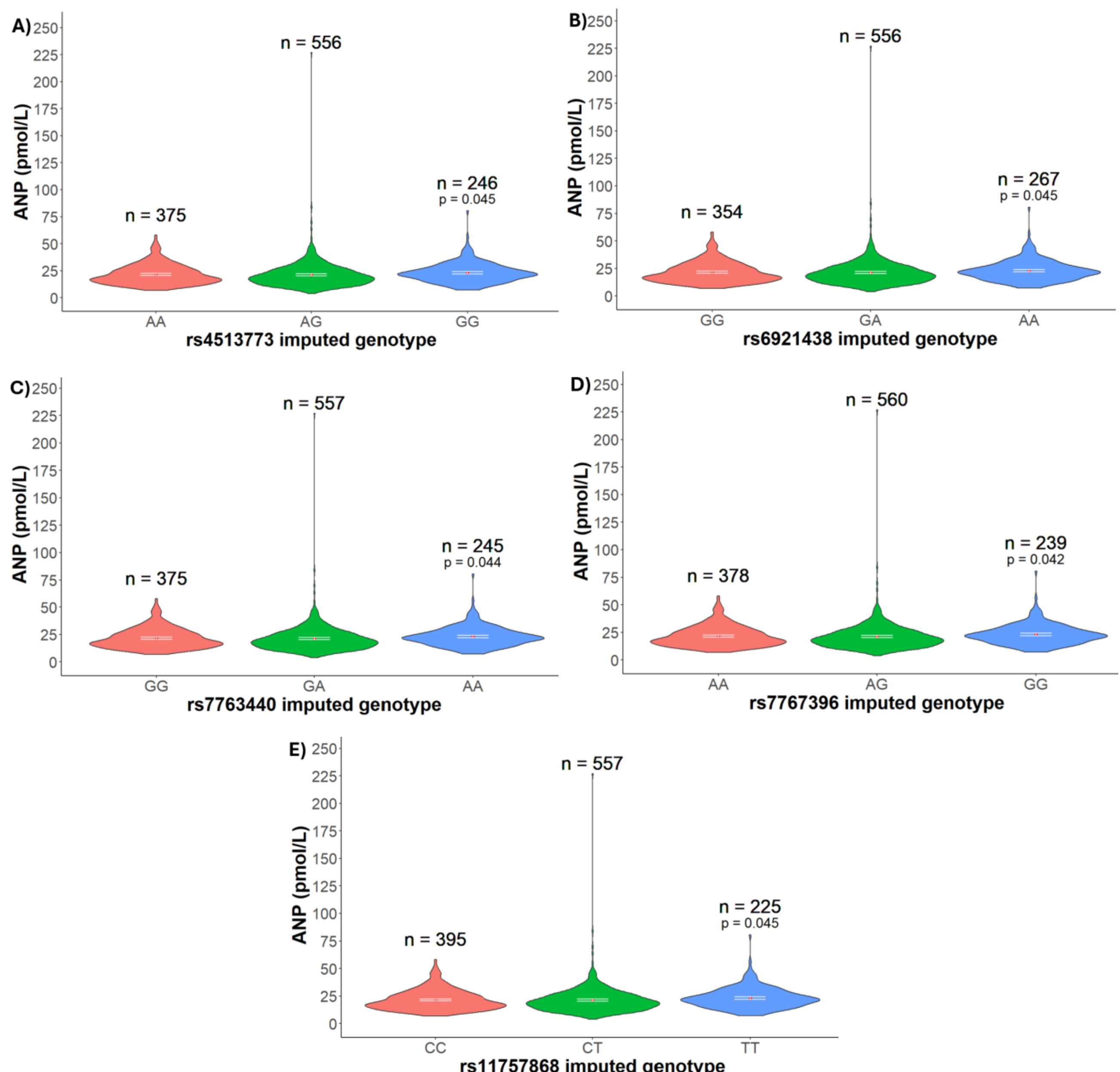
Variable	n	AA	n	AG	n	GG	p
				rs4513773 Genotypes			
Anthropometric Variables							
Age (years)	587	66.0 ± 0.513	927	67.0 ± 0.402	399	67.3 ± 0.605	0.193
Activity (Sed, Mild, Mod, Act)	586	20.3%, 14%, 13.5%, 52.2%	924	19.9%, 11.9%, 14.9%, 53.3%	399	26.3%, 11.3%, 13.8%, 48.6%	0.155
Male Gender	587	435 (74.1%)	927	651 (70.2%)	399	287 (71.9%)	0.262
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	587	86.2%, 7%, 6.1%, 0.7%	927	90.8%, 5.9%, 3.1%, 0.2%	399	95%, 2.8%, 2%, 0.2%	**0.0002**
Height (m)	578	1.69 ± 0.004	915	1.69 ± 0.003	396	1.70 ± 0.005	0.506
Weight (kg)	580	81.1 ± 0.732	915	79.9 ± 0.542	396	78.9 ± 0.829	**0.038**
BMI (kg/m^2)	578	27.9 ± 0.234	912	27.3 ± 0.156	395	27.1 ± 0.224	**0.006**
Systolic BP (mmHg)	574	127.2 ± 0.908	911	129.2 ± 0.715	387	131.2 ± 1.14	**0.021**
Diastolic BP (mmHg)	574	74.6 ± 0.504	911	74.6 ± 0.387	387	75.1 ± 0.59	0.803
Beta blocker use	587	504 (85.9%)	927	802 (86.5%)	399	345 (86.5%)	0.932
Previous MI	582	181 (31.1%)	922	269 (29.2%)	396	123 (31.1%)	0.663
Tobacco use (Current, Ex Smoker, Non-smoker)	587	6.3%, 58.8%, 34.9%	927	6%, 56.7%, 37.2%	399	7.3%, 55.9%, 36.8%	0.806
Analytes							
Cholesterol (mmol/L)	453	4.98 ± 0.058	727	4.86 ± 0.041	310	4.85 ± 0.068	0.148
Creatinine (mmol/L)	571	99.8 ± 2.20	897	99.1 ± 1.90	389	100.5 ± 1.88	0.905
Urate (mmol/L)	349	0.376 ± 0.005	565	0.366 ± 0.004	233	0.386 ± 0.007	**0.028**
Troponin I (ng/L) $	567	49.2 (12.7 – 85.7)	885	39.2 (18.7 – 59.7)	375	43.2 (11.8 – 20.1)	0.323
ANP (pg/mL) $	586	40.9 (38.5 – 43.3)	919	45.2 (43.2 – 47.3)	396	45.3 (41.8 – 48.7)	**0.034**
NT-ANP (pmol/L) $	586	1.26 (1.18 – 1.35)	919	1.40 (1.32 – 1.48)	395	1.38 (1.26 – 1.51)	**0.030**
BNP (pmol/L) $	586	23.7 (21.2 – 26.2)	919	28.4 (26.1 – 30.8)	396	27.5 (23.9 – 31.0)	**0.009**
NTproBNP (pg/ml) $	586	119.5 (106.8- 132.1)	919	143.2 (130.5 – 155.8)	396	146.5 (125.3 – 167.7)	**0.007**
CNP (pmol/L) $	579	0.639 (0.605 – 0.675)	907	0.672 (0.642 – 0.703)	387	0.679 (0.629 – 0.729)	0.534
NT-CNP (pmol/L) $	578	23.1 (21.3 – 24.9)	904	22.8 (21.1 – 24.6)	388	23.5 (21.1 – 25.8)	0.345
Aldosterone (pmol/L) $	577	171.5 (146.9 – 196.1)	890	163.7 (156.3 – 171.2)	386	164.7 (157.6 – 175.1)	0.702
Endothelin (pmol/L) $	586	2.66 (2.58 – 2.74)	919	2.68 (2.61- 2.75)	396	2.69 (2.59 – 2.80)	0.877
Adrenomedullin (pg/ml) $	565	8.11 (7.76 – 8.46)	887	8.79 (8.44 – 9.15)	388	8.65 (8.09 – 9.21)	**0.016**
VEGF-A (pg/mL) $	165	49.9 (44.1 – 55.7)	255	45.7 (41.1 – 50.2)	105	38.3 (33.5 – 43.1)	**0.004**

Variable	rs7763440 Genotypes n	GG	n	GA	n	AA	p
Anthropometric Variables							
Age (years)	591	65.9 ± 0.516	925	67.1 ± 0.400	397	67.3 ± 0.606	0.120
Activity (Sed, Mild, Mod, Act)	590	20.2%, 14%, 13.4%, 52.4%	922	20.1%, 11.8%, 15%, 53.1%	397	26.2%, 11.3%, 13.9%, 48.6%	0.163
Male Gender	591	438 (74.1%)	925	649 (70.2%)	397	286 (72.0%)	0.247
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	591	86.1%, 7.1%, 6.1%, 0.7%	925	90.9%, 5.8%, 3.1%, 0.2%	397	95%, 2.8%, 2%, 0.3%	**0.0003**
Height (m)	582	1.69 ± 0.004	913	1.69 ± 0.003	394	1.70 ± 0.004	0.521
Weight (kg)	584	81.0 ± 0.729	913	78.9 ± 0.542	394	79.0 ± 0.833	**0.048**
BMI (kg/m^2)	582	27.9 ± 0.233	910	27.3 ± 0.156	393	27.1 ± 0.225	**0.007**
Systolic BP (mmHg)	578	127.2 ± 0.902	909	129.2 ± 0.716	385	131.2 ± 1.14	**0.019**
Diastolic BP (mmHg)	578	74.6 ± 0.503	909	74.6 ± 0.387	385	75.1 ± 0.592	0.820
Beta blocker use	591	508 (86.0%)	925	800 (86.5%)	397	343 (86.4%)	0.956
Previous MI	586	182 (31.1%)	920	268 (29.1%)	394	123 (31.2%)	0.639
Tobacco use (Current, Ex Smoker, Non-smoker)	591	6.4%, 58.4%, 35.2%	925	5.9%, 56.9%, 37.2%	397	7.3%, 56.2%, 36.5%	0.838
Analytes							
Cholesterol (mmol/L)	457	4.98 ± 0.057	724	4.86 ± 0.041	309	4.85 ± 0.069	0.175
Creatinine (mmol/L)	575	99.7 ± 2.19	895	99.2 ± 1.90	387	100.6 ± 1.88	0.905
Urate (mmol/L)	351	0.376 ± 0.005	564	0.366 ± 0.004	232	0.387 ± 0.007	**0.022**
Troponin I (ng/L) [$]	571	48.9 (12.6 – 85.1)	883	39.3 (18.8 – 59.8)	373	43.3 (27.4 – 58.9)	0.275
ANP (pg/mL)	590	40.8 (38.4 – 43.1)	917	45.3 (43.2 – 47.4)	394	45.3 (41.8 – 48.8)	**0.020**
NT-ANP (pmol/L) [$]	590	1.26 (1.18 – 1.34)	917	1.40 (1.32 – 1.49)	393	1.39 (1.26 – 1.51)	**0.017**
BNP (pmol/L) [$]	590	23.59 (21.11 – 26.1)	917	28.5 (26.2 – 30.8)	394	27.5 (23.9 -31.1)	**0.005**
NTproBNP (pg/ml) [$]	590	118.9 (106.3 – 131.5)	917	143.5 (130.8 – 156.2)	394	146.8 (125.5 – 168.1)	**0.004**
CNP (pmol/L) [$]	583	0.639 (0.605 – 0.674)	905	0.672 (0.641- 0.703)	385	0.681 (0.630 – 0.731)	0.568
NT-CNP (pmol/L) [$]	582	23.1 (21.3 – 24.8)	902	22.8 (21.1 – 24.6)	386	23.5 (21.2 – 25.9)	0.325
Aldosterone (pmol/L) [$]	581	171.4 (146.9 – 195.7)	888	163.7 (156.2 – 171.1)	384	165.1 (154.2 – 175.9)	0.668
Endothelin (pmol/L) [$]	590	2.65 (2.57 – 2.73)	917	2.68 (2.62 – 2.75)	394	2.69 (2.59 – 2.80)	0.788
Adrenomedullin (pg/ml) [$]	569	8.09 (7.75 – 8.45)	884	8.82 (8.46 – 9.17)	387	8.63 (8.07 -9.19)	**0.011**
VEGF-A (pg/mL) [$]	166	50.0 (44.2 – 55.8)	255	45.4 (40.9 – 49.9)	104	38.5 (33.6 – 43.4)	**0.004**

Variable	n	CC	n	CT	n	TT	p
			rs11757868 genotypes				
Anthropometric Variables							
Age (years)	629	65.7 ± 0.494	909	67.2 ± 0.408	375	67.4 ± 0.617	**0.032**
Activity (Sed, Mild, Mod, Act)	628	19.6%, 13.9%, 14.3%, 62.2%	906	20.5%, 11.7%, 14.3%, 53.5%	375	26.4%, 11.7%, 13.9%, 48%	0.185
Male Gender	629	471 (74.9%)	909	634 (69.7%)	375	268 (71.5%)	0.088
Ethnicity (European, Māori & Pasifika, Asian, MELAA)	629	85.7%, 7.6%, 6%, 0.7%	909	91.1%, 5.8%, 2.9%, 0.2%	375	96%, 1.6%, 2.1%, 0.3%	**8 x10^{-6}**
Height (m)	620	1.70 ± 0.004	897	1.69 ± 0.003	372	1.70 ± 0.005	0.614
Weight (kg)	622	81.1 ± 0.704	897	78.9 ± 0.557	372	78.8 ± 0.817	**0.031**
BMI (kg/m^2)	620	27.9 ± 0.224	894	27.3 ± 0.160	371	27.1 ± 0.220	**0.011**
Systolic BP (mmHg)	616	127.6 ± 0.863	891	129.2 ± 0.733	364	130.9 ± 1.16	**0.072**
Diastolic BP (mmHg)	616	74.8 ± 0.484	891	74.6 ± 0.395	364	75.0 ± 0.598	0.816
Beta blocker use	629	540 (85.9%)	909	786 (86.5%)	375	325 (86.7%)	0.918
Previous MI	624	194 (31.1%)	904	265 (29.3%)	372	114 (30.6%)	0.739
Tobacco use (Current, Ex Smoker, Non-smoker)	629	6.3%, 57.9%, 35.8%	909	6.5%, 57.6%, 35.9%	375	6.2%, 54.9%, 38.9%	0.866
Analytes							
Cholesterol (mmol/L)	490	4.96 ± 0.055	707	4.85 ± 0.042	293	4.88 ± 0.070	0.292
Creatinine (mmol/L)	610	100.5 ± 2.33	880	98.7 ± 1.80	367	100.3 ± 1.89	0.779
Urate (mmol/L)	373	0.376 ± 0.005	552	0.365 ± 0.357	222	0.388 ± 0.007	**0.015**
Troponin I (ng/L) $	607	50.9 (15.9 – 85.8)	868	36.7 (16.5 – 56.9)	352	45.6 (21.0 – 70.3)	0.351
ANP (pg/mL) $	627	41.1 (38.8 – 43.4)	900	45.1 (43.1 – 47.2)	374	45.7 (42.2 – 49.3)	**0.034**
NT-ANP (pmol/L) $	627	1.30 (1.20 – 1.40)	900	1.38 (1.31 – 1.45)	373	1.39 (1.26 – 1.52)	0.065
BNP (pmol/L) $	627	24.1 (21.6 – 26.6)	900	28.2 (25.9 - 30.6)	374	27.7 (25.9 – 30.5)	**0.010**
NTproBNP (pg/ml) $	627	122.9 (109.3 – 136.6)	900	141.6 (129.5 – 153.7)	374	147.2 (124.9 – 169.5)	**0.006**
CNP (pmol/L) $	619	0.648 (0.613 – 0.683)	889	0.666 (0.636 – 0.697)	365	0.682 (0.629 – 0.735)	0.757
NT-CNP (pmol/L) $	618	23.5 (21.5 – 25.6)	886	22.5 (20.9 – 24.2)	366	23.4 (20.9 – 25.9)	0.340
Aldosterone (pmol/L) $	614	169.7 (146.6 – 192.8)	876	164.3 (156.9 – 171.6)	363	165.6 (153.6 – 177.6)	0.575
Endothelin (pmol/L) $	627	2.65 (2.57 – 2.72)	900	2.68 (2.61 – 2.75)	374	2.72 (2.61 – 2.83)	0.617
Adrenomedullin (pg/ml) $	603	8.09 (7.75 – 8.42)	869	8.84 (8.48 – 9.20)	368	8.65 (8.08 – 9.23)	**0.006**
VEGF-A (pg/mL) $	173	50.9 (44.9 – 56.8)	253	44.9 (40.6 – 49.4)	99	37.5 (32.9 – 42.1)	**0.001**

$Log10 transformed p-values are reported. Mean ± standard error or Mean (95% CI range) or incidence (%) are reported. Significantly associated variables and their p-values are shown in **bold.** Abbreviations: Act: active (≥30 minutes on ≥3 days/week), ANP: atrial natriuretic peptide, BP: blood pressure, BMI: body mass index, BNP: B-type natriuretic peptide, CNP: C-type natriuretic peptide, MELAA: Middle Eastern/Latin American/African, MI: Myocardial infarction, Mod: moderate (≥30 minutes on 2 days/week), NT-ANP Amino terminal atrial natriuretic peptide, NT-CNP: Amino terminal C-type natriuretic peptide NTproBNP = amino-terminal pro-B type natriuretic peptide, Sed: Sedentary, VEGF-A: Vascular endothelial growth factor A. Sed, Mild, Mod, Act

Appendix 4.2 Variants associated with ANP levels in HVOL cohort



Violin plots of ANP levels in the HVOL cohort for variants A) rs4513773, B) rs6921438, C) rs7763440, D) rs7767396 and E) rs11757868. Red represents the homozygous reference genotype, green heterozygous genotype, and blue homozygous for the minor allele genotype. Red dot indicates mean ANP levels. Lines represent 95% CI SE error bars. p-value when compared to the reference genotype group of the respective variant

Appendix 4.3 Genotype frequencies for rs6921438 and rs7767396

SNP ID	Cohort	Genotype	Imputed genotype Frequencies n (%)	Manual genotype Frequencies n (%)
rs692148	CDCS	GG	536 (28%)	529 (26.1%)
		GA	946 (49.5%)	983 (48.5%)
		AA	431 (22.5%)	514 (25.4%)
		Total	1913	2026
	HVOL	GG	71 (28.4%)	53 (23%)
		GA	123 (49.2%)	122 (53%)
		AA	56 (22.4%)	55 (24%)
		Total	250	230
rs7767396	CDCS	AA	589 (30.8%)	575 (28.4%)
		AG	922 (48.2%)	983 (48.6%)
		GG	402 (21%)	464 (23%)
		Total	1913	2022
	HVOL	AA	77 (30.8%)	54 (23.8%)
		AG	121 (48.4%)	122 (53.7%)
		GG	52 (20.8%)	51 (22.5%)
		Total	250	227

Abbreviations: CDCS: Coronary Disease Cohort Study HVOL: Canterbury Healthy Volunteers Study

Appendix 4.4. Multivariate HF readmissions model on the CDCS cohort using manual genotypes for rs6921438 and rs7767396 (n = 1147, 240 (20.9%) events).

Predictor	Coeff.	SE	Wald	P-value	HR	95% CI for HR	
						Lower	Upper
Gender	0.292	0.153	3.62	0.057	1.33	0.991	1.81
Ethnicity			2.23	0.525			
European v Pasifika	0.658	0.483	1.85	0.173	1.93	0.75	4.97
European v Asian	0.652	1.01	0.415	0.519	1.91	0.264	13.9
European v MELAA	-6.501	124.4	0.003	0.958	0.002	1.9×10^{-109}	1.1×10^{103}
***Physical Activity (scale 1–4) $$**	-0.374	0.057	43.2	***4.8×10^{-11}**	0.688	0.615	0.769
***Previous MI**	0.609	0.137	19.6	***9.2×10^{-6}**	1.83	1.41	2.41
***Age**	0.050	0.009	32.1	***1.3×10^{-8}**	1.05	1.03	1.06
Body mass index	0.009	0.017	0.262	0.609	1.009	0.976	1.04
***Urate**	1.91	0.68	7.902	***0.005**	6.76	1.78	25.6
Creatinine	0.001	0.002	0.672	0.412	1.001	0.998	1.005
***Beta blocker**	-0.529	0.176	9.06	***0.003**	0.589	0.418	0.832
***Log10 NTproBNP$**	1.73	0.206	71.4	***2.9×10^{-17}**	5.68	3.79	8.51
rs6921438 genotype			2.19	0.333			
GG v GA	-0.433	0.529	0.671	0.413	0.648	0.230	1.82
GG v AA	0.086	0.638	0.018	0.892	1.09	0.312	3.81
rs7767396 genotype			3.05	0.217			
GG v AA	0.24	0.635	0.143	0.706	1.27	0.366	4.41
GG v AG	0.662	0.413	2.57	0.109	1.93	0.863	4.35
***Time to sampling**	0.02	0.005	15.6	***7.7×10^{-5}**	1.02	1.01	1.03

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or VEGF-A level.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Appendix 4.5. Multivariate MACE readmissions model on the CDCS cohort using manual genotypes for rs6921438 and rs7767396 (n = 1132, 503 (44.4%) events)

Predictor	Coeff.	SE	Wald	P-value	HR	95% CI for HR	
						Lower	Upper
Gender	0.145	0.104	1.94	0.164	1.15	0.943	1.41
Ethnicity			2.69	0.440			
European v Pasifika	-0.36	0.391	0.847	0.358	0.698	0.324	1.502
European v Asian	0.277	0.587	0.224	0.636	1.32	0.418	4.16
European v MELAA	1.27	1.009	1.58	0.208	3.56	0.493	25.7
***Physical Activity (scale 1–4)** [$$]	-0.322	0.039	68.2	***1.4 x10^{-16}**	0.725	0.671	0.782
***Previous MI**	0.364	0.096	14.5	***1.3 x10^{-4}**	1.44	1.19	1.73
***Age**	0.025	0.005	21.6	***3.1 x10^{-6}**	1.02	1.01	1.03
***Body mass index**	-0.03	0.012	6.49	***0.011**	0.97	0.948	0.993
Urate	0.55	0.478	1.32	0.25	1.73	0.679	4.41
***Creatinine**	0.003	0.001	11.1	***0.001**	1.003	1.001	1.005
Beta blocker	-0.184	0.130	2.005	0.157	0.832	0.644	1.07
***Log10 NTproBNP** [$]	0.904	0.134	45.1	***1.8 x10^{-11}**	2.46	1.89	3.21
rs6921438 genotype			4.08	0.129			
GG v GA	-0.041	0.310	0.018	0.894	0.959	0.523	1.76
GG v AA	0.523	0.378	1.92	0.166	1.68	0.805	3.54
rs7767396 genotype			3.74	0.154			
GG v AA	0.601	0.377	2.54	0.111	1.82	0.872	3.81
GG v AG	0.537	0.290	3.41	0.065	1.71	0.968	3.02
***Time to sampling**	0.019	0.003	29.2	***6.4 x 10^{-8}**	1.01	1.01	1.026

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or VEGF-A level.

$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.

Significant variables and their p-values are in bold and "*."

Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Appendix 4.6. Multivariate NSTEMI readmissions model on the CDCS cohort using manual genotypes for rs6921438 and rs7767396 (n = 1126, 287 (25.5%) events)

Predictor	Coeff.	SE	Wald	P-value	HR	95% CI for HR	
						Lower	Upper
Gender	0.154	0.139	1.22	0.269	1.16	0.888	1.53
Ethnicity			5.45	0.142			
European v Pasifika	-0.791	0.595	1.76	0.184	0.454	0.141	1.45
European v Asian	-9.69	116.5	0.007	0.934	0.0001	3.9×10^{-104}	9.5×10^{94}
European v MELAA	1.92	1.01	3.57	0.059	6.83	0.932	50.1
***Physical Activity (scale 1–4) $$**	-0.337	0.051	43.1	***5.1×10^{-11}**	0.714	0.646	0.789
***Previous MI**	0.577	0.126	21.1	***4.5×10^{-6}**	1.78	1.39	2.27
***Age**	0.017	0.007	5.71	***0.017**	1.02	1.003	1.03
Body mass index	-0.03	0.016	3.64	0.056	0.970	0.941	1.001
Urate	0.552	0.635	0.754	0.385	1.73	0.500	6.03
***Creatinine**	0.003	0.001	5.41	***0.02**	1.003	1.000	1.005
Beta blocker	-0.008	0.182	0.002	0.966	0.992	0.695	1.41
***Log10 NTproBNP$**	1.02	0.180	32.2	***1.3×10^{-8}**	2.78	1.95	3.96
***rs6921438 genotype**			7.13	***0.028**			
GG v GA	-1.26	0.679	3.45	0.063	0.283	0.075	1.07
GG v AA	-0.450	0.732	0.377	0.539	0.638	0.152	2.68
***rs7767396 genotype**			6.11	***0.047**			
GG v AA	-0.470	0.731	0.413	0.521	0.625	0.149	2.62
GG v AG	0.740	0.388	3.64	0.056	2.09	0.980	4.48
***Time to sampling**	0.029	0.005	41.3	***1.2×10^{-10}**	1.03	1.02	1.03

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or VEGF-A level.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Appendix 4.7. Multivariate unstable angina readmissions model on the CDCS cohort using manual genotypes for rs6921438 and rs7767396 (n = 1126, 231 (20.5%) events)

Predictor	Coeff.	SE	Wald	P-value	HR	95% CI for HR	
						Lower	Upper
Gender	-0.046	0.156	0.088	0.767	0.955	0.704	1.29
Ethnicity			3.06	0.382			
European v Pasifika	-0.394	0.471	0.700	0.403	0.674	0.268	1.69
European v Asian	-0.177	1.01	0.031	0.861	0.838	0.115	6.09
European v MELAA	1.53	1.02	2.25	0.133	4.65	0.625	34.6
***Physical Activity (scale 1–4) [$$]**	-0.367	0.058	40.1	***2.4 x10^{-10}**	0.693	0.619	0.77
***Previous MI**	0.938	0.141	44.4	***2.6 x10^{-11}**	2.55	1.93	3.36
Age	0.014	0.008	2.99	0.084	1.01	0.998	1.02
Body mass index	-0.005	0.017	0.084	0.772	0.995	0.963	1.02
Urate	0.611	0.730	0.700	0.403	1.84	0.440	7.71
Creatinine	0.002	0.002	1.01	0.314	1.002	0.998	1.006
Beta blocker	-0.181	0.193	0.879	0.349	0.834	0.571	1.21
Log10 NTproBNP[$]	0.285	0.199	2.06	0.151	1.33	0.901	1.96
rs6921438 genotype			5.63	0.060			
GG v GA	-0.078	0.631	0.015	0.902	0.925	0.268	3.18
***GG v AA**	-0.877	0.395	4.92	***0.027**	0.416	0.192	0.903
rs7767396 genotype			4.19	0.123			
GG v AA	0.178	0.631	0.080	0.778	1.19	0.347	4.11
GG v AG	0.798	0.410	3.79	0.052	2.22	0.995	4.96
***Time to sampling**	0.028	0.005	30.2	***3.8 x10^{-8}**	1.02	1.01	1.03

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or VEGF-A level.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Appendix 5.1 Imputed data for 13 VEGFR2 SNPs

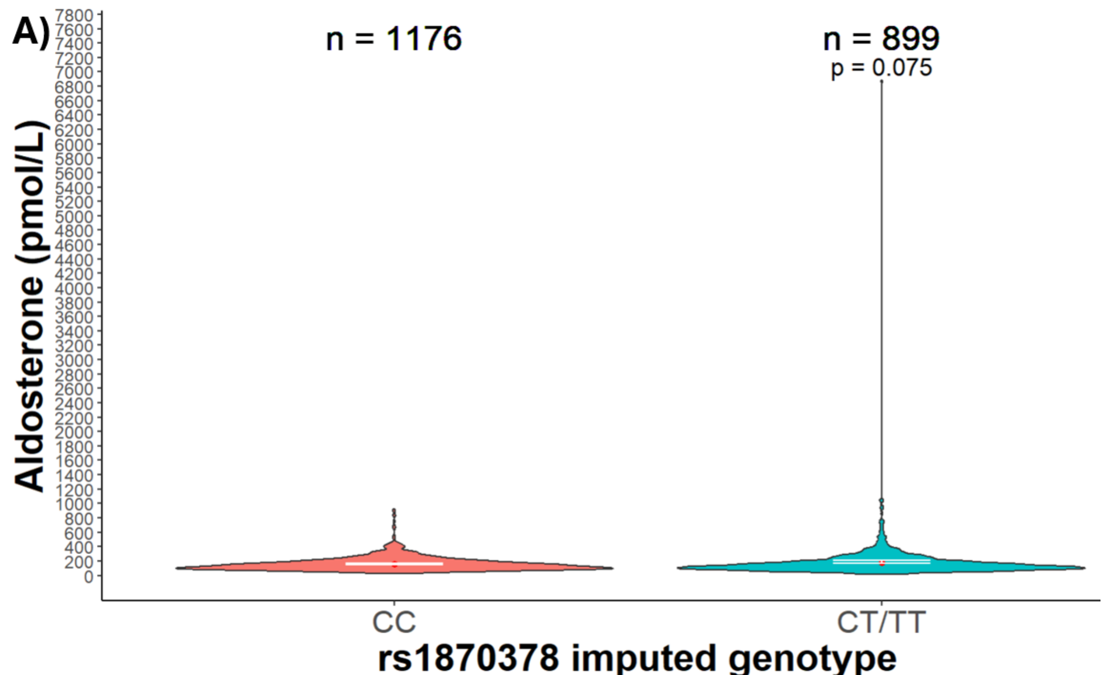
rsID	A*	a*	Cohort	Total	AA n (%)	Aa n (%)	aa n (%)	MAF	$^{\$}R^2$
rs1870379	A	C	CDCS	1935	1216 (62.8%)	658 (34.0%)	61 (3.2%)	0.238	0.667
			HVOL	1177	744 (63.2%)	386 (32.8%)	47 (4.0%)	0.236	0.672
rs2305948	C	T	CDCS	1935	1705 (88.1%)	222 (11.5%)	8 (0.4%)	0.097	0.523
			HVOL	1177	1037 (88.1%)	138 (11.7%)	2 (0.2%)	0.095	0.519
rs1870377	T	A	CDCS	1935	1216 (62.8%)	657 (34.0%)	62 (3.2%)	0.238	0.678
			HVOL	1177	760 (64.6%)	371 (31.5%)	46 (3.9%)	0.233	0.682
rs10016064	G	C	CDCS	1935	1182 (61.1%)	685 (35.4%)	68 (3.5%)	0.248	0.676
			HVOL	1177	731 (61.7%)	393 (33.4%)	53 (4.5%)	0.247	0.68
rs13136007	C	A	CDCS	1935	1204 (62.2%)	668 (34.5)	63 (3.3%)	0.240	0.676
			HVOL	1177	745 (63.3%)	383 (32.5%)	49 (4.2%)	0.239	0.681
rs7677779	C	T	CDCS	1935	1217 (62.9%)	655 (33.8%)	63 (3.3%)	0.236	0.675
			HVOL	1177	754 (64.1%)	377 (32.0%)	46 (3.9%)	0.233	0.679
rs1870378	C	T	CDCS	1935	1216 (62.8%)	658 (34.0%)	61 (3.2%)	0.238	0.667
			HVOL	1177	744 (63.2%)	386 (32.8%)	47 (4.0%)	0.236	0.672
rs7667298	T	C	CDCS	1935	387 (20%)	981 (50.7%)	567 (29.3%)	0.488	0.602
			HVOL	1177	245 (20.8%)	551 (46.8%)	381 (32.4%)	0.489	0.634
rs9994560	A	G	CDCS	1935	559 (28.9%)	991 (51.2%)	385 (19.9%)	0.485	0.626
			HVOL	1177	355 (30.2%)	559 (47.5%)	263 (22.3%)	0.495	0.666
rs55713360	A	G	CDCS	1935	706 (36.5%)	950 (49.1%)	279 (14.4%)	0.429	0.615
			HVOL	1177	447 (38%)	538 (45.7%)	192 (16.3%)	0.431	0.655
rs28695311	G	C	CDCS	1935	559 (28.9%)	991 (51.2%)	385 (19.9%)	0.486	0.625
			HVOL	1177	355 (30.2%)	562 (47.7%)	260 (22.1%)	0.495	0.666
rs2071559	A	G	CDCS	1935	558 (28.8%)	993 (51.3%)	384 (19.8%)	0.485	0.625
			HVOL	1177	355 (30.2%)	561 (47.7%)	261 (22.1%)	0.495	0.666
rs1380069	G	C	CDCS	1935	406 (21%)	989 (51.1%)	540 (27.9%)	0.498	0.602
			HVOL	1177	254 (21.6%)	547 (46.5%)	376 (31.9%)	0.497	0.639

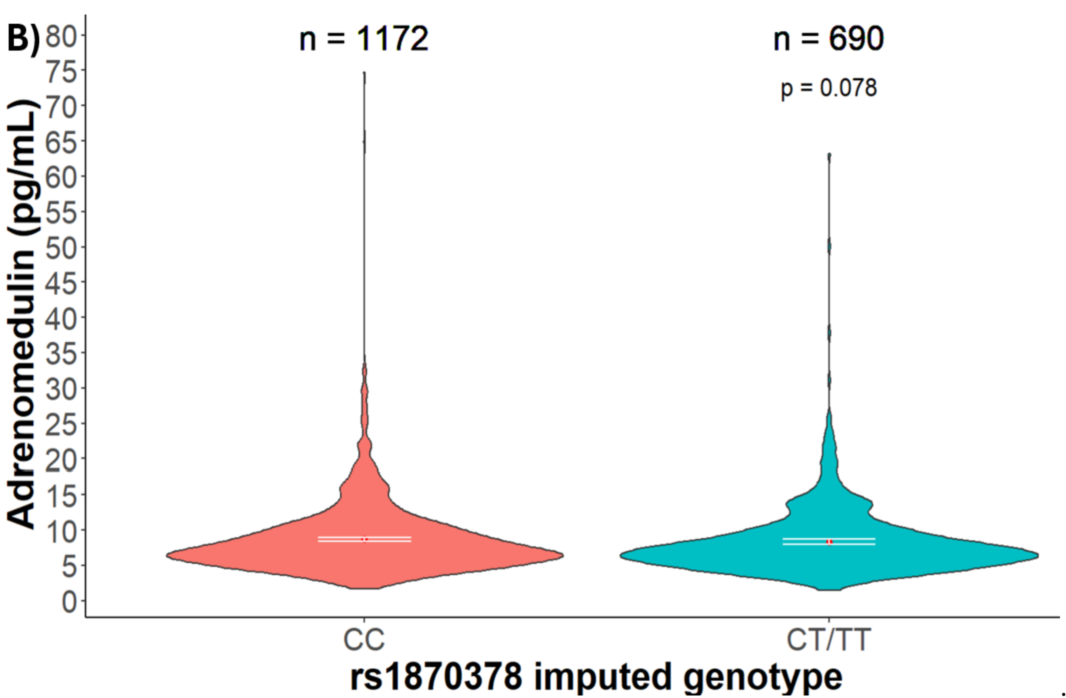
*A= reference allele a = minor allele

$Imputation parameter on the estimated correlation between imputed genotypes and unobserved genotypes.

Abbreviations: CDCS Coronary Disease Cohort Study, HVOL: Canterbury Healthy Volunteers Study, MAF: Minor allelic Frequency

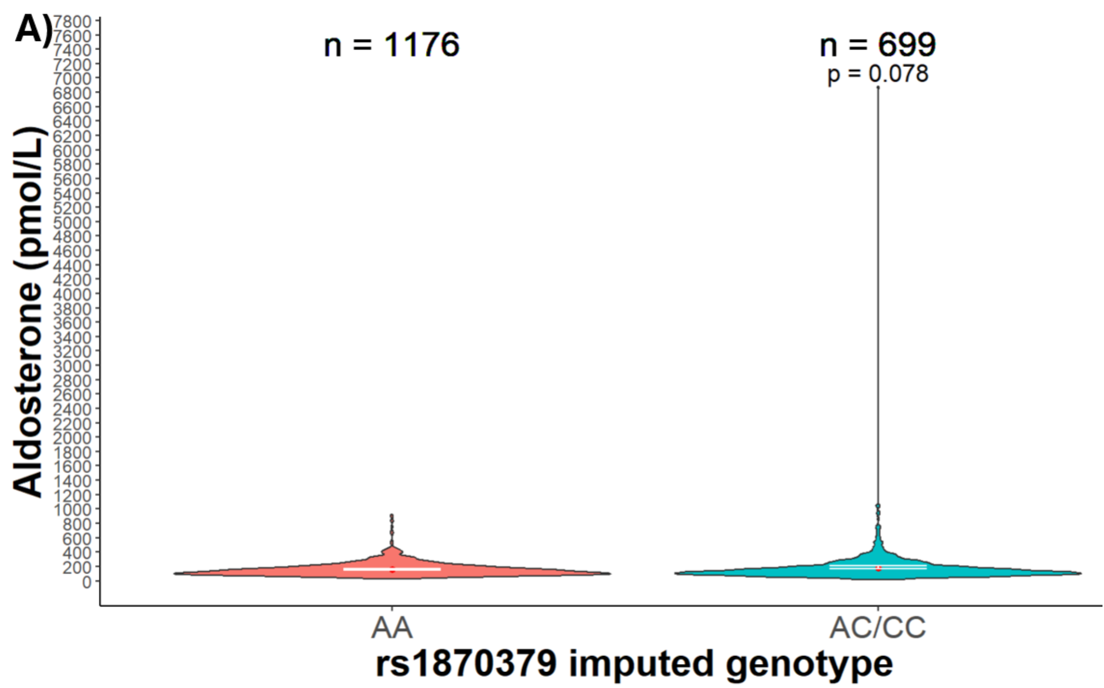
Appendix 5.2 rs1870378 imputed genotype associations within the CDCS cohort

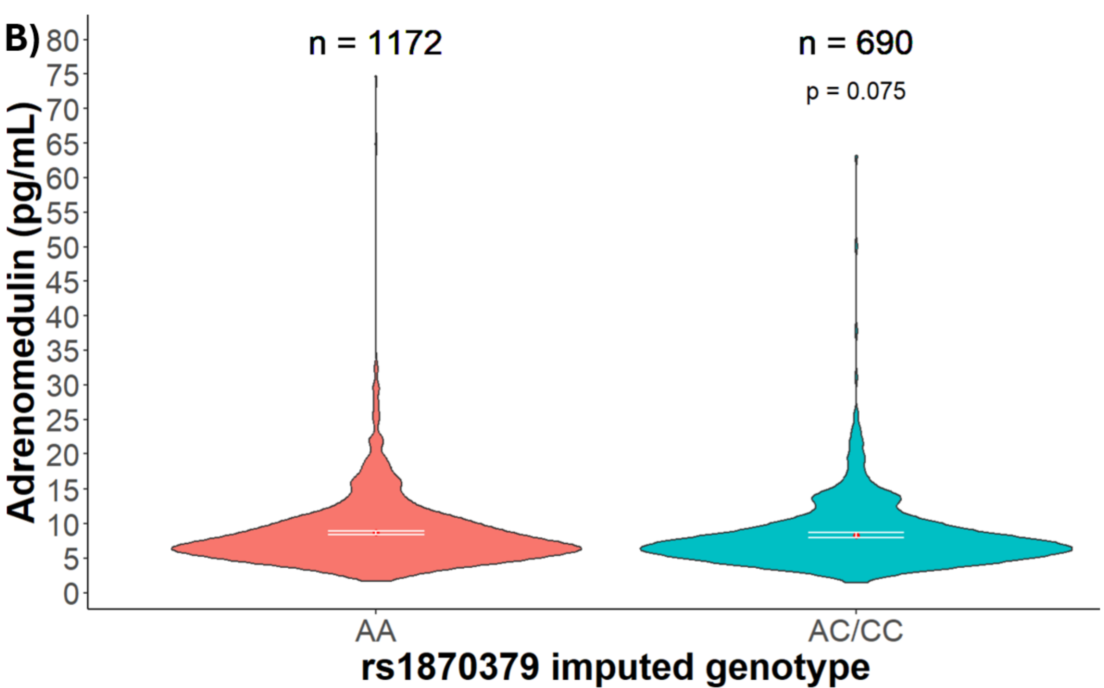




A) aldosterone and **B)** adrenomedullin levels. Red dot indicates mean ANP levels. Lines represent 95% CI SE error bars.

Appendix 5.3 rs1870379 imputed genotype associations within the CDCS cohort



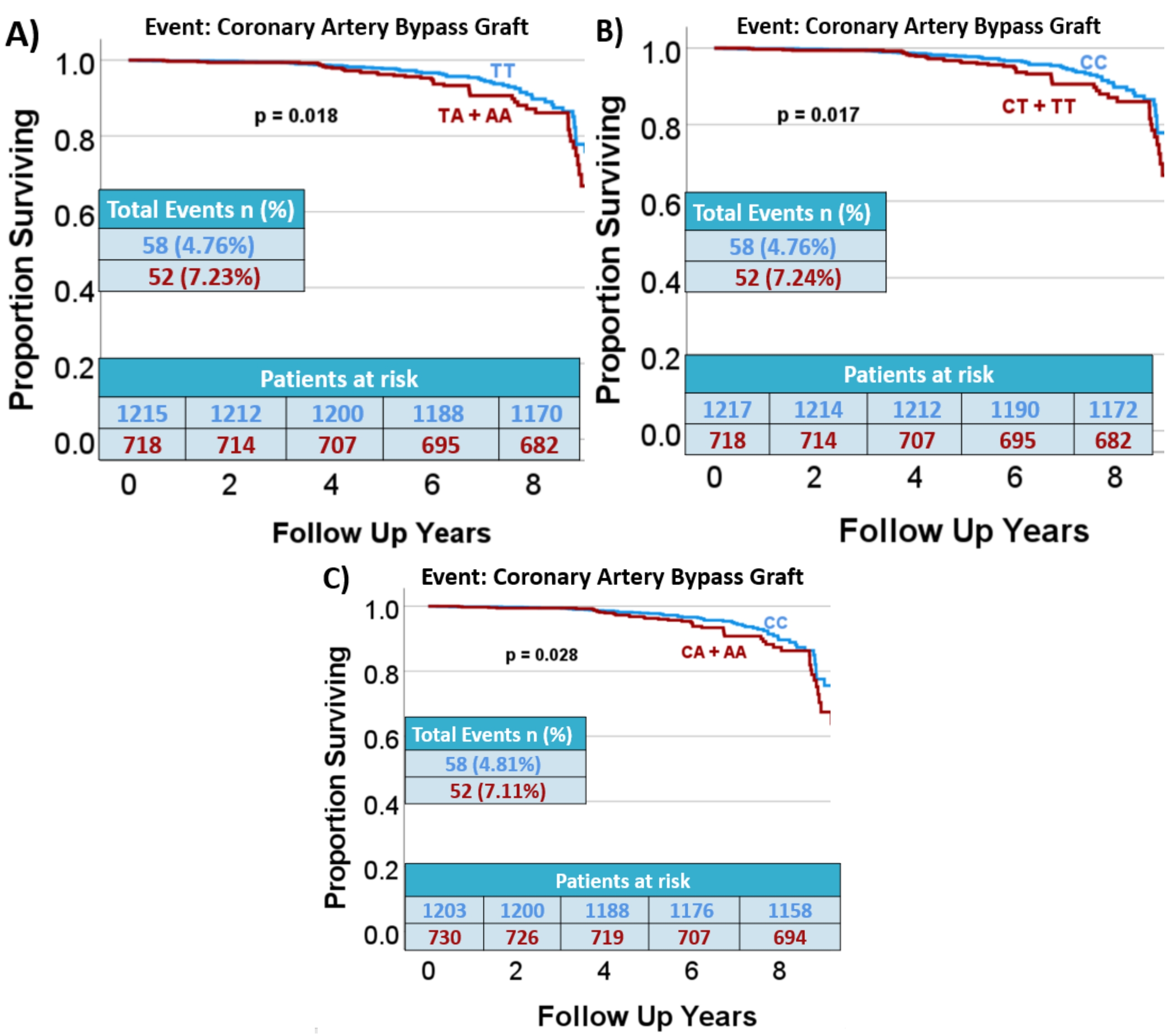


A) aldosterone and **B)** adrenomedullin levels. Red dot indicates mean ANP levels. Lines represent 95% CI SE error bars.

Appendix 5.4. Variables with p<0.1 from one-way CDCS ANOVA analysis using rs2305948 imputed genotype group comparison.

Variable	SS	df	Mean Square	F	p value
B-type natriuretic peptide (pmol/L)	0.583	1	0.583	3.722	0.054
Creatinine (mmol/L)	8710.68	1	8710.68	3.248	0.072

Appendix 5.5. Kaplan-Meier survival plots using imputed genotype data from rs1870377, rs7677779 and rs13136007 for CABG events within the CDCS cohort.



A) rs1870377 **B)** rs7677779 and **C)** rs13136007. Genotypes are colour-coded blue for homozygous reference and red for homozygous for the minor allele genotypes. Patients at risk reported for every 2-year interval.

Appendix 5.6. Multivariate all cause death model on the CDCS cohort using rs1870377 imputed genotype data (n = 1842, 436 (23.7%) deaths)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.196	0.109	3.19	0.074	1.22	0.981	1.51
Ethnicity			2.44	0.486			
European v Māori/Pasifika	0.288	0.306	0.888	0.345	1.34	0.732	2.43
European v Asian	0.473	0.364	1.69	0.193	1.61	0.786	3.28
European v MELAA	-7.29	89.6	0.006	0.935	7×10^{-4}	0	1.3×10^{73}
***Physical Activity (scale 1–4) $$**	-0.275	0.039	47.7	**$*4.9 \times 10^{-12}$**	0.759	0.702	0.821
***Previous MI**	0.51	0.099	26.1	**$*3.2 \times 10^{-7}$**	1.67	1.37	2.03
***Age**	0.063	0.006	109.7	**$*1.1 \times 10^{-25}$**	1.07	1.05	1.08
***Log10 NTproBNP$**	1.32	0.137	93.9	**$*3.4 \times 10^{-22}$**	3.77	2.88	4.94
***Beta blocker**	-0.43	0.128	11.3	**$*8 \times 10^{-4}$**	0.649	0.505	0.835
***Creatinine**	0.003	0.0007	20.1	**$*7.4 \times 10^{-6}$**	1.003	1.001	1.004
***Hypertension**	0.266	0.103	6.55	***0.011**	1.30	1.06	1.59
rs1870377 genotype (TT v TA/AA)	0.178	0.099	3.19	0.074	1.19	0.982	1.45

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B-type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Appendix 5.7. Multivariate all cause death model on the CDCS cohort using rs7677779 imputed genotype data (n = 1842, 436 (23.7%) deaths)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.196	0.11	3.19	0.074	1.22	0.981	1.51
Ethnicity			2.42	0.489			
European v Māori/Pasifika	0.287	0.307	0.877	0.349	1.33	0.731	2.43
European v Asian	0.473	0.364	1.69	0.194	1.61	0.786	3.28
European v MELAA	-7.29	89.6	0.007	0.935	0.001	0	1.1×10^{73}
***Physical Activity (scale 1–4)** [$$]	-0.275	0.04	47.6	***5.1×10^{-12}**	0.759	0.702	0.821
***Previous MI**	0.511	0.1	26.2	***3.1×10^{-7}**	1.67	1.37	2.03
***Age**	0.063	0.006	109.9	***1×10^{-25}**	1.07	1.053	1.08
***Log10 NTproBNP** [$]	1.33	0.137	93.7	***3.7×10^{-22}**	3.77	2.88	4.93
***Beta blocker**	-0.429	0.128	11.3	***0.001**	0.651	0.507	0.836
***Creatinine**	0.003	0.001	20.2	***6.9×10^{-6}**	1.003	1.002	1.004
***Hypertension**	0.266	0.104	6.54	***0.011**	1.30	1.06	1.59
rs7677779 genotype (CC v CT/TT)	0.17	0.1	2.88	0.09	1.19	0.974	1.44

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."** Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B-type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Appendix 5.8 Multivariate all cause death model on the CDCS cohort using rs13136007 imputed genotype data (n = 1842, 436 (23.7%) deaths)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.196	0.11	3.18	0.075	1.22	0.981	1.51
Ethnicity			2.44	0.486			
European v Māori/Pasifika	0.290	0.307	0.896	0.344	1.34	0.733	2.44
European v Asian	0.473	0.364	1.69	0.194	1.605	0.786	3.28
European v MELAA	-7.29	89.4	0.007	0.935	0.001	0.000	8.6×10^{72}
***Physical Activity (scale 1–4) $$**	-0.275	0.04	47.6	***5.3×10^{-12}**	0.759	0.702	0.821
***Previous MI**	0.512	0.1	26.3	***2.9×10^{-7}**	1.67	1.37	2.03
***Age**	0.063	0.006	109.8	***1.1×10^{-25}**	1.07	1.05	1.08
***Log10 NTproBNP$**	1.33	0.137	93.4	***4.3×10^{-22}**	3.76	2.88	4.92
***Beta blocker**	-0.429	0.128	11.3	***7.9×10^{-4}**	0.651	0.506	0.837
***Creatinine**	0.003	0.001	20.1	***7.6×10^{-6}**	1.003	1.002	1.004
***Hypertension**	0.266	0.104	6.52	***0.011**	1.30	1.06	1.60
rs13136007 genotype (CC v CA/AA)	0.155	0.1	2.42	0.119	1.17	0.961	1.42

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP or Adrenomedullin level.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B-type natriuretic peptide, SE = standard error, VEGF-A = Vascular endothelial growth factor A

Appendix 6.1. Imputed frequencies for *VLDLR* locus SNPs in the CDCS cohort

rsID	A$	A$	AA n (%)	Aa n (%)	aa n (%)	MAF	*predicted R^2
rs7043199	T	A	1244 (64.3%)	602 (31.1%)	89 (4.6%)	0.334	0.736
rs7030781	A	T	765 (39.5%)	914 (47.2%)	256 (13.3%)	0.538	0.781
rs10738760	A	G	574 (29.7%)	974 (50.3%)	387 (20%)	0.603	0.795
rs2375981	C	G	636 (32.9%)	957 (49.5%)	342 (17.7%)	0.583	0.783

*Imputation parameter on the estimated correlation between imputed genotypes and unobserved genotypes.
$A= reference allele a = minor allele
r^2=1 occurs when a physical and imputed genotype match perfectly

Appendix 6.2. Imputed frequencies for *VLDLR* locus SNPs in the HVOL cohort

rsID	A$^{\$}$	A$^{\$}$	AA n (%)	Aa n (%)	aa n (%)	MAF	*predicted R^2
rs7043199	T	A	763 (64.8%)	368 (31.3%)	46 (3.9%)	0.332	0.739
rs7030781	A	T	460 (39.1%)	555 (47.2%)	162 (13.8%)	0.541	0.795
rs10738760	A	G	343 (29.1%)	598 (50.8%)	236 (20.1%)	0.608	0.807
rs2375981	C	G	378 (32.1%)	591 (50.2%)	208 (17.7%)	0.590	0.795

*Imputation parameter on the estimated correlation between imputed genotypes and unobserved genotypes.
$A= reference allele a = minor allele
r^2=1 occurs when a physical and imputed genotype match perfectly

Appendix 6.3. Multivariate all cause death model on the CDCS cohort using rs2375981 imputed genotype data (n = 1702, 386 (22.7%) events)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.046	0.120	0.149	0.699	1.047	0.828	1.32
Ethnicity			2.77	0.427			
European v Pasifika	0.334	0.311	1.15	0.283	1.39	0.759	2.56
European v Asian	0.490	0.365	1.804	0.179	1.63	0.798	3.33
European v MELAA	-7.57	101.8	0.006	0.941	0.001	9.8×10^{-91}	2.6×10^{83}
***Physical Activity (scale 1–4) $$**	-0.260	0.042	37.4	***9.6×10^{-10}**	0.771	0.710	0.838
***Previous MI**	0.501	0.107	22.0	***2.7×10^{-6}**	1.65	1.33	2.03
***Age**	0.062	0.006	95.8	***1.2×10^{-22}**	1.06	1.05	1.07
***Log10 NTproBNP$**	1.304	0.168	60.5	***7.2×10^{-15}**	3.68	2.65	5.11
***Beta blocker**	-0.416	0.134	9.61	***0.0019**	0.660	0.507	0.858
***Creatinine**	0.003	0.001	23.5	***1.2×10^{-6}**	1.003	1.002	1.005
rs2375981 genotype			4.86	0.088			
***CC v CG**	-0.247	0.113	4.81	***0.028**	0.781	0.626	0.974
CC v GG	-0.170	0.156	1.18	0.276	0.844	0.622	1.14
LVEF	-0.006	0.004	2.06	0.151	0.994	0.986	1.002

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*"**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF= left ventricular ejection fraction, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error,

Appendix 6.4. Multivariate NSTEMI readmissions model on the CDCS cohort using rs2375981 imputed genotype data (n = 1697, 382 (22.5%) events)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.056	0.123	0.205	0.651	1.05	0.830	1.34
Ethnicity			4.43	0.219			
European v Pasifika	-0.050	0.284	0.031	0.859	0.951	0.546	1.65
European v Asian	0.324	0.313	1.07	0.300	1.38	0.749	2.55
European v MELAA	1.341	0.725	3.41	0.064	3.82	0.923	15.8
***Physical Activity (scale 1–4) $$**	-0.268	0.042	39.9	***2.6 x10^{-10}**	0.765	0.704	0.831
***Previous MI**	0.651	0.107	36.7	***1.3 x10^{-9}**	1.91	1.55	2.36
***Age**	0.026	0.006	20.5	***5.8 x10^{-6}**	1.02	1.02	1.04
***Log10 NTproBNP$**	0.874	0.162	29.02	***7.1 x10^{-8}**	2.39	1.74	3.29
Beta blocker	-0.107	0.150	0.507	0.477	0.899	0.67	1.21
***Creatinine**	0.003	0.001	27.6	***1.4 x10^{-7}**	1.003	1.002	1.005
rs2375981 genotype			4.76	0.092			
***CC v CG**	-0.248	0.113	4.76	***0.029**	0.781	0.625	0.975
CC v GG	-0.129	0.155	0.689	0.407	0.879	0.648	1.19
LVEF	-0.005	0.004	1.23	0.267	0.995	0.987	1.004

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF= left ventricular ejection fraction, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error,

Appendix 6.5. Multivariate NSTEMI readmissions model on the CDCS cohort using rs10738760 imputed genotype data (n = 1697, 382 (22.5%) events)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.057	0.123	0.215	0.642	1.06	0.831	1.34
Ethnicity			4.57	0.205			
European v Pasifika	-0.033	0.283	0.013	0.906	0.967	0.555	1.68
European v Asian	0.316	0.313	1.02	0.311	1.37	0.743	2.53
European v MELAA	1.38	0.724	3.65	0.056	3.99	0.965	16.5
***Physical Activity (scale 1–4) $$**	-0.267	0.042	39.9	***2.5 x10^{-10}**	0.765	0.704	0.831
***Previous MI**	0.652	0.107	36.9	***1.2 x10^{-9}**	1.92	1.55	2.36
***Age**	0.025	0.006	20.5	***5.8 x10^{-6}**	1.03	1.01	1.03
***Log10 NTproBNP$**	0.874	0.162	29.1	***6.9 x10^{-8}**	2.39	1.74	3.29
Beta blocker	-0.107	0.150	0.513	0.474	0.898	0.669	1.21
***Creatinine**	0.004	0.0007	27.5	***1.5 x10^{-7}**	1.004	1.002	1.004
***rs10738760 genotype**			6.62	***0.037**			
***AA v AG**	-0.291	0.115	6.39	***0.011**	0.747	0.596	0.936
AA v GG	-0.228	0.152	2.23	0.135	0.796	0.590	1.073
LVEF	-0.005	0.004	1.27	0.258	0.995	0.986	1.003

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold**.
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF= left ventricular ejection fraction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error

Appendix 6.6. Multivariate cardiovascular death model on the CDCS cohort using rs10738760 imputed genotype data (n = 1701, 223 (13.1%) events)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.128	0.161	0.63	0.427	1.13	0.828	1.56
Ethnicity			6.48	0.090			
European v Pasifika	0.571	0.382	2.23	0.135	1.77	0.838	3.74
***European v Asian**	0.873	0.397	4.84	***0.028**	2.39	1.10	5.21
European v MELAA	-7.27	123	0.003	0.953	0.001	8 $x10^{-109}$	6 x 10^{101}
***Physical Activity (scale 1–4)** $^{\$\$}$	-0.259	0.056	21.3	***3.7 x 10^{-6}**	0.772	0.692	0.86
***Previous MI**	0.741	0.142	27.1	***1.8 $x10^{-7}$**	2.09	1.59	2.77
***Age**	0.052	0.008	37.7	***7.9 $x10^{-10}$**	1.05	1.04	1.07
***Log10 NTproBNP** $^{\$}$	1.78	0.229	60.6	***6.7 $x10^{-15}$**	5.95	3.80	9.33
Beta blocker	-0.171	0.193	0.783	0.376	0.843	0.577	1.23
***Creatinine**	0.002	0.001	5.02	***0.025**	1.002	1.000	1.004
***rs10738760 genotype**			6.55	***0.038**			
***GG v AA**	0.405	0.204	3.94	***0.047**	1.49	1.005	2.23
GG v GA	0.066	0.199	0.110	0.741	1.06	0.724	1.57
LVEF	-0.006	0.005	1.28	0.257	0.994	0.984	1.004

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF= left ventricular ejection fraction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error

Appendix 6.7. Multivariate cardiovascular death model on the CDCS cohort using rs2375981 imputed genotype data (n = 1667, 211 (12.7%) events)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.130	0.166	0.616	0.433	1.13	0.823	1.57
Ethnicity			6.19	0.103			
European v Pasifika	0.456	0.409	1.24	0.264	1.57	0.708	3.51
European v Asian	0.926	0.399	5.39	0.020	2.52	1.15	5.51
European v MELAA	-7.51	132.4	0.003	0.955	0.001	9.1 x10^{-117}	3.3 x 10^{109}
***Physical Activity (scale 1–4) $$**	-0.231	0.058	15.8	***6.9 x10^{-5}**	0.794	0.708	0.889
***Previous MI**	0.765	0.147	27.1	***1.9 x10^{-7}**	2.15	1.61	2.86
***Age**	0.050	0.009	29.6	***5.1 x10^{-8}**	1.051	1.03	1.07
***Log10 NTproBNP$**	1.87	0.239	61.2	***4.9 x10^{-15}**	6.48	4.06	10.3
Beta blocker	-0.217	0.195	1.24	0.265	0.805	0.550	1.17
***Creatinine**	0.002	0.001	4.61	***0.032**	1.002	1.000	1.004
rs2375981 genotype			5.71	0.058			
***CC v CG**	-0.348	0.153	5.21	***0.022**	0.706	0.523	0.952
CC v GG	-0.328	0.213	2.36	0.124	0.720	0.474	1.09
Systolic blood pressure	0.005	0.004	1.41	0.235	1.005	0.997	1.01
Diastolic blood pressure	-0.004	0.007	0.26	0.609	0.996	0.982	1.01
LVEF	-0.008	0.006	1.71	0.191	0.992	0.981	1.004

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF= left ventricular ejection fraction, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error

Appendix 6.8. Multivariate NSTEMI readmissions model on the CDCS cohort using rs10738760 imputed genotype data (n = 1663, 372 (22.4%) events)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.093	0.125	0.551	0.458	1.09	0.858	1.40
Ethnicity			4.27	0.233			
European v Pasifika	-0.124	0.295	0.178	0.673	0.883	0.495	1.57
European v Asian	0.322	0.313	1.05	0.304	1.38	0.747	2.55
European v MELAA	1.27	0.733	3.04	0.081	3.59	0.854	15.1
***Physical Activity**	-0.263	0.043	37.1	***1.1 $x10^{-9}$**	0.769	0.707	0.837
***Previous MI**	0.675	0.109	38.1	***6.8 $x10^{-10}$**	1.96	1.58	2.43
***Age**	0.025	0.006	16.5	***4.6 $x10^{-5}$**	1.02	1.01	1.03
***Log10 NTproBNP**	0.911	0.166	30.2	***3.8 $x10^{-8}$**	2.48	1.79	3.43
Beta blocker	-0.123	0.151	0.668	0.413	0.884	0.657	1.18
***Creatinine**	0.003	0.001	27.4	***1.6 $x10^{-7}$**	1.004	1.002	1.005
rs10738760 genotype			5.53	0.063			
***AA v AG**	-0.273	0.117	5.43	***0.020**	0.761	0.606	0.958
AA v GG	-0.198	0.155	1.63	0.201	0.821	0.606	1.11
Systolic blood pressure	0.003	0.003	1.19	0.275	1.003	0.997	1.01
Diastolic blood pressure	0.003	0.006	0.199	0.656	1.003	0.991	1.01
LVEF	-0.005	0.005	1.31	0.253	0.995	0.986	1.004

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazards ratio, LVEF= left ventricular ejection fraction, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error

Appendix 6.9. Multivariate cardiovascular death model on the CDCS cohort using rs7030781 imputed genotype data (n = 1701, 223 (13.1%) events)

Predictor	Coeff.	SE	Wald	P - value	HR	95% CI for HR	
						Lower	Upper
Gender	0.109	0.161	0.458	0.499	1.11	0.813	1.53
Ethnicity			6.97	0.073			
European v Pasifika	0.587	0.384	2.34	0.126	1.79	0.848	3.81
***European v Asian**	0.913	0.397	5.28	***0.022**	2.49	1.144	5.42
European v MELAA	-7.317	121.7	0.004	0.952	0.001	1.6×10^{-107}	2.6×10^{100}
***Physical Activity (scale 1–4)** [$$]	-0.256	0.056	21.1	***4.4×10^{-6}**	0.774	0.694	0.864
***Previous MI**	0.744	0.142	27.3	***1.6×10^{-7}**	2.11	1.59	2.78
***Age**	0.052	0.008	37.9	***7.2×10^{-10}**	1.05	1.036	1.07
***Log10 NTproBNP** [$]	1.76	0.230	58.8	***1.6×10^{-14}**	5.82	3.71	9.12
Beta blocker	-0.175	0.193	0.827	0.363	0.839	0.575	1.22
***Creatinine**	0.002	0.001	5.29	***0.021**	1.002	1.000	1.004
rs7030781 genotype			5.48	0.064			
AA v AT	-0.259	0.144	3.23	0.072	0.772	0.582	1.02
***AA v TT**	-0.452	0.231	3.84	***0.050**	0.636	0.405	1.000
LVEF	-0.007	0.005	1.634	0.201	0.993	0.983	1.004

$Hazard Ratio represents the change in risk for every 10-fold increase in NTproBNP.
$$Score of 1 = sedentary, 2 = <30 minutes activity on >2 days/week, 3 = ≥30 minutes on 2 days/week, 4 = ≥30 minutes on ≥3 days/week.
Significant variables and their p-values are in **bold and "*."**
Abbreviations: CI = confidence interval, Coeff = Coefficient, HR = hazard ratio, LVEF= left ventricular ejection fraction, MI = Myocardial Infarction, NTproBNP = amino-terminal pro-B type natriuretic peptide, SE = standard error,