

COMPONENTE GENÉTICA DA DOENÇA CORONÁRIA E INTERAÇÃO COM OS FATORES DE RISCO TRADICIONAIS

ANDREIA MICAELA DE SOUSA PEREIRA

Tese para obtenção do grau de Doutor em Medicina

na Especialidade de Medicina

na NOVA Medical School | Faculdade de Ciências Médicas



**FACULDADE DE
CIÊNCIAS
MÉDICAS**

**UNIVERSIDADE
NOVA DE LISBOA**

Setembro, 2017

**COMPONENTE GENÉTICA DA DOENÇA CORONÁRIA E
INTERAÇÃO COM OS FATORES DE RISCO TRADICIONAIS**

ANDREIA MICAELA DE SOUSA PEREIRA

Orientador: Prof. Roberto Palma dos Reis

Orientadora: Dr.^a Maria Isabel Saraiva de Mendonça

**Tese para a obtenção do grau de Doutor em Medicina
na Especialidade de Medicina Interna**

Setembro, 2017

“No meio de qualquer dificuldade encontra-se a oportunidade.”

Albert Einstein

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar.

Mas o mar seria menor se lhe faltasse uma gota”.

Madre Teresa de Calcutá

“Medicine is a science of uncertainty and an art of probability.”

William Osler

A presente tese deu origem às seguintes publicações:

**1. ASSOCIAÇÃO INDEPENDENTE DA VARIANTE RS1333049, NO LOCUS 9P21, COM A DOENÇA CORONÁRIA, NUMA POPULAÇÃO PORTUGUESA.
INDEPENDENT ASSOCIATION OF THE VARIANT RS1333049 AT THE 9P21 LOCUS AND CORONARY HEART DISEASE**

Mendonça, M.I., Palma dos Reis, R., Pereira, A., Café, H., Serrão, M., Sousa, A.C., Freitas, A.I., Guerra, G., Freitas, S., Freitas, C., Ornelas, I., Brehm, A., Araújo, J.J.

• Rev Port Cardiol 2011, 30 (06):575-591

2. ASSOCIATION OF ADAMTS7 GENE POLYMORPHISM WITH CARDIOVASCULAR SURVIVAL IN CORONARY ARTERY DISEASE.

Pereira, A., Palma dos Reis R., Rodrigues R., Sousa A.C., Gomes S., Borges S., Ornelas I., Freitas A.I., Guerra G., Henriques E., Rodrigues M., Freitas S., Freitas C., Brehm A., Pereira D., Mendonça M.I.

• Physiol Genomics 2016;48 (11):810-5.

3. GENETIC RISK SCORE AND CARDIOVASCULAR MORTALITY IN A SOUTHERN EUROPEAN POPULATION WITH CORONARY ARTERY DISEASE.

Pereira, A., Mendonça M.I., Sousa A.C., Borges S., Freitas S., Henriques E., Rodrigues M., Freitas A.I., Guerra G., Ornelas I., Pereira D., Brehm A., Palma dos Reis R.

• Int J Clin Pract 2017;71 (6).

4. ADDITIONAL VALUE OF A COMBINED GENETIC RISK SCORE TO STANDARD CARDIOVASCULAR STRATIFICATION.

Pereira, A., Mendonça M.I., Sousa A.C., Borges S., Freitas S., Henriques E., Rodrigues M., Freitas A.I., Guerra G., Freitas C., Pereira D., Brehm A., Palma dos Reis R.

• Genetics and Molecular Biology. (Aceite para publicação).

5. GENETIC RISK ANALYSIS OF CORONARY ARTERY DISEASE IN A POPULATION-BASED STUDY IN PORTUGAL, USING A GENETIC RISK SCORE OF 31 VARIANTS.

Pereira, A., Mendonça M.I., Borges S., Freitas S., Henriques E., Rodrigues M., Freitas A.I., Sousa A.C., Brehm A., Palma Dos Reis R.

• Arq Bras Cardiol. 2018 Jul 2, (Epub ahead of print]

Foram apresentadas mais de três dezenas de comunicações sobre os conteúdos desta tese, das quais se destacam:

1. GENETIC RISK SCORE, FAMILY HISTORY OF CORONARY ARTERY DISEASE AND CARDIOVASCULAR RISK FACTORS.

A. Pereira, R. Palma Dos Reis, M. Neto, R. Rodrigues, J. Monteiro, E. Henriques, M. Rodrigues, A.I. Freitas, I. Ornelas, S. Borges, D. Pereira, M.I. Mendonça

Congresso: ESC 2017, Barcelona 26-31 Agosto 2017

2. From telomere length to machine learning - News in risk prediction

CORONARY ARTERY DISEASE RISK ACCORDING TO GENETIC RISK SCORE DECILES, AGE AND CARDIOVASCULAR RISK FACTORS

A. Pereira, M. Neto, R. Rodrigues, J. Monteiro, A.C. Sousa, S. Freitas, E. Henriques, S. Borges, A.I. Freitas, I. Ornelas, D. Pereira, R. Palma Dos Reis, M.I. Mendonça.

Congresso: ESC 2017, Barcelona 26-31 Agosto 2017

3. Prémio Melhor Poster moderado no Congresso Europeu de Cardiologia de 2016:
RECLASSIFICATION OF THE INTERMEDIATE GROUP CLASSIFIED ACCORDING TO HEARTSCORE TAKING IN CONSIDERATION INDIVIDUAL GENETIC PREDISPOSITION TO CORONARY ARTERY DISEASE.
A. Pereira, M. Mendonca, M. Neto, R. Rodrigues, J. Monteiro, A.C. Sousa, S. Freitas, M. Rodrigues, E. Henriques, A.I. Freitas, I. Ornelas, S. Borges, D. Pereira, R. Palma Dos Reis
Congresso: ESC 2016, Roma 27-31 Agosto 2016
Revista: Eur Heart J (2016) 37 (suppl 1): 985-1393

4. Prémio Melhor Poster moderado no Congresso Europeu de Cardiologia de 2016:
NET RECLASSIFICATION INDEX OF A GENETIC RISK SCORE AND STANDARD FRAMINGHAM RISK STRATIFICATION IN A EUROPEAN POPULATION
A. Pereira, M. Mendonça, M. Neto, R. Rodrigues, J. Monteiro, A.C. Sousa, M. Rodrigues, S. Freitas, E. Henriques, G. Guerra, I. Ornelas, C. Freitas, D. Pereira, R. Palma Dos Reis
Congresso: ESC 2016, Roma 27-31 Agosto 2016
Revista: Eur Heart J (2016) 37 (suppl 1): 599-983

5. *Advances in Science Sessions* no Congresso Europeu de Cardiologia 2016:
101. GENETIC RISK SCORE AND CARDIOVASCULAR MORTALITY IN A PORTUGUESE POPULATION
A. Pereira, M.I. Mendonça, M. Neto, R. Rodrigues, J. Monteiro, A.C. Sousa, S. Freitas, M. Rodrigues, E. Henriques, G. Guerra, I. Ornelas, S. Borges, D. Pereira, R. Palma Dos Reis
Revista: Eur Heart J (2016) 37 (suppl 1): 599-983

6. **ADDITIONAL VALUE OF A COMBINED GENETIC SCORE TO FRAMINGHAM RISK SCORE**
A. Pereira, I. Mendonça, R. Rodrigues, S. Gomes, M. Neto, A.C. Sousa, E. Henriques, G. Guerra, S. Borges, D. Pereira, R. Palma dos Reis
American Heart Association, 2015 Scientific Sessions and Resuscitation Science Symposium
Revista: Circulation. 2015; 132:A10264

7. **INCREASE THE PREDICTIVE CAPACITY OF CORONARY RISK WITH A GENETIC SCORE**
A. Pereira, I. Mendonça, R. Rodrigues, S. Gomes, M. Neto, A.C. Sousa, E. Henriques, G. Guerra, S. Borges, D. Pereira, R. Palma dos Reis
American Heart Association, 2015 Scientific Sessions and Resuscitation Science Symposium
Revista: Circulation.2015; 132:A10229

8. **WHAT IS THE BEST GENETIC SCORE TO PREDICT THE RISK FOR CORONARY DISEASE?**
A. Pereira, M.I. Mendonça, S. Gomes, R. Rodrigues, A.C. Sousa, E. Henriques, G. Guerra, C. Freitas, D. Pereira, R. Palma dos Reis.
Congresso: ESC 2015, Londres 29 Agosto a 3 Agosto de 2015
Revista: Eur Heart J (2015) 36, Issue suppl 1, 1 August 2015: 810

9. **GENETIC POLYMORPHISM AND CORONARY DISEASE – MULTIVARIATE STUDY INCLUDING TRADITIONAL RISK FACTORS**
A. Pereira, M.I. Mendonça, S. Gomes, R. Rodrigues, A.C. Sousa, S. Freitas, G. Guerra, I. Ornelas, D. Pereira, R. Palma dos Reis.
Congresso: ESC 2015, Londres 29 Agosto a 3 Agosto de 2015
Revista: Eur Heart J (2015) 36, Issue suppl 1, 1 August 2015:810

10. THE ADDED VALUE OF A COMBINED GENETIC SCORE AS CARDIOVASCULAR RISK PREDICTOR IN A PORTUGUESE POPULATION WITH INTERMEDIATE RISK ACCORDING TO THE EUROPEAN SCORE

A. Pereira, R. Palma dos Reis, R. Rodrigues, S. Gomes, A.C. Sousa, E. Henriques, G. Guerra, C. Freitas, D. Pereira, M.I. Mendonça

Congresso: ESC 2015, Londres 29 Agosto a 3 Agosto de 2015

Revista: Eur Heart J (2015) 36, Issue suppl 1, 1 August 2015:803

11. GENETIC MULTIFACTORIAL AND MULTILOCUS STUDY IN A PORTUGUESE POPULATION WITH CORONARY DISEASE

A. Pereira, R. Palma dos Reis, S. Gomes, R. Rodrigues, A.C. Sousa, E. Henriques, G. Guerra, C. Freitas, D. Pereira, M.I. Mendonça

Congresso: ESC 2015, Londres 29 Agosto a 3 Agosto de 2015

Revista: Eur Heart J (2015) 36, Issue suppl 1, 1 August 2015:804

12. MULTILOCUS GENOME-WIDE STUDY WITH VARIANTS ASSOCIATED WITH CARDIOVASCULAR RISK FACTORS AND CORONARY ARTERY DISEASE IN A SOUTHERN EUROPEAN POPULATION

A. Pereira, M.I. Mendonça, R. Palma dos Reis, S. Gomes, R. Rodrigues, A.C. Sousa, S. Freitas, G. Guerra, S. Borges, D. Pereira,

Congresso: ESC 2015, Londres 29 Agosto a 3 Agosto de 2015

Revista: Eur Heart J (2015) 36, Issue suppl 1, 1 August 2015:809

13. ADAMTS7 AS A NOVEL GENETIC LOCUS FOR CORONARY ATHEROSCLEROSIS AND CARDIOVASCULAR PROGNOSIS

A. Pereira, S. Gomes, R. Palma Reis, R. Rodrigues, M. Rodrigues, A.I. Freitas, S. Borges, S. Freitas, D. Pereira, M.I. Mendonça

Congresso: ESC 2014, Barcelona 30 Agosto a 3 Setembro de 2014

Eur Heart J (2014) 35 (suppl 1): 513-850

14. DETERMINATION OF A GENETIC RISK SCORE IN CORONARY PATIENTS from our COUNTRY

A. Pereira, M.I. Mendonça, S. Gomes, R. Rodrigues, H. Café, G. Guerra, M. Rodrigues, S. Borges, D. Pereira, R. Palma dos Reis

Congresso: ESC 2014, Barcelona 30 Agosto a 3 Setembro de 2014

Revista: Eur Heart J (2014) 35 (suppl 1): 1-172 (P394, Pág. 55)

15. GENETIC FACTORS INFLUENCING THE DEVELOPMENT OF CORONARY DISEASE IN HYPERTENSION PATIENTS

R. Palma dos Reis, M. Mendonça, A. Pereira, S. Gomes, A.C. Sousa, B. Silva, R. Rodrigues, A.I. Freitas, M. Rodrigues, D. Pereira

Congresso: ESC 2013, Amsterdão 31 Agosto a 4 Setembro de 2013

Eur Heart J (2013) 34 (suppl1): 1137-1212

16. GENETIC POLYMORPHISMS ASSOCIATED WITH CORONARY ARTERY DISEASE IN PATIENTS WITH PREVIOUS LOW CARDIOVASCULAR RISK BY TRADITIONAL RISK FACTORS

A. Pereira, R. Palma Dos Reis, B. Silva, S. Gomes, H. Cafe, S. Freitas, A.I. Freitas, A. Brehm, J.J. Araujo¹, M. Mendonça

Congresso: ESC 2012, Munique 25-29 de Agosto de 2012

European Heart Journal (2012) 33 (Abstract Supplement),124

AGRADECIMENTOS:

Agradeço ao Professor Doutor Roberto Palma dos Reis, orientador desta tese, por ter acreditado desde sempre no valor científico deste projeto, pela disponibilidade permanente e incentivo contante que proporcionou nas horas felizes e sobretudo nas horas menos felizes.

À Dr.^a Isabel Mendonça por me ter permitido integrar a Unidade de Investigação Científica e por partilhar os seus conhecimentos nesta área. Desta partilha, levo um exemplo de vida e de interesse científico genuíno com quem tenho o privilégio de partilhar o dia a dia.

À Elsa Sousa, secretária da Unidade de Investigação, Carolina Freitas, técnica de cardiopneumologia da Unidade de Investigação e à Doutora Sofia Borges também da Unidade de Investigação pelo apoio, esforço e bom ambiente de trabalho que souberam criar à sua volta.

À Sónia Freitas, Mariana Rodrigues e Eva Henriques, estatísticas do Hospital Dr. Nélio Mendonça, pela paciência exaustiva na introdução de dados, atualização da base de dados e análise estatística essenciais à concretização desta tese. À Rita Freitas pela revisão dos procedimentos estatísticos e pelos conhecimentos transmitidos nessa área.

Agradeço a colaboração do Professor Doutor António Brehm, Ana Isabel Freitas e Graça Guerra do Laboratório de Genética Humana da Universidade da Madeira, pelo trabalho exemplar desenvolvido na análise genética, na revisão de artigos e nas sugestões propostas.

Agradeço a colaboração do Dr. Ilídio Ornelas e de todo o pessoal envolvido na colheita de amostras e processamento bioquímico do Laboratório do Hospital Dr. Nélio Mendonça.

Ao Professor Doutor José Caldas do Instituto Superior Técnico e Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa, pelo apoio prestado na área da bio-informática, genética populacional e *machine learning*.

Aos meus colegas do Serviço de Cardiologia do Hospital Dr. Nélio Mendonça do Funchal por incentivarem a continuidade deste trabalho e colaborarem no envio de doentes e avaliação complementar diagnóstica dos doentes envolvidos nesta tese.

Aos internos e colegas com quem tenho o prazer de partilhar esta aventura e o entusiasmo pela investigação científica que colaboraram na avaliação dos doentes e proporcionaram dúvidas pertinentes ao longo desta caminhada.

Aos meus filhos Leonor e João que tendo nascido já durante esta jornada, aprenderam desde cedo que a descoberta move o ser humano na busca incessante da ciência.

Ao Rui, pelo incentivo inicial e constante ao longo da tese, pelos desafios “originais” lançados e pelas horas passadas em discussão profícua que deram maior relevância clínica e ética a esta tese.

Agradeço ainda à minha família e amigos que, compreendo as minhas faltas, estiveram sempre presentes dando o apoio e incentivo necessários para a realização da presente tese de dissertação.

Por fim agradeço aos Doentes, dos quais partiram as dúvidas que lançaram a razão da ser deste trabalho e para os quais espero devolver os conhecimentos apreendidos, em seu proveito.

ÍNDICE

1. Agradecimentos	1
2. Índice	3
3. Lista de Abreviaturas	5
4. Resumo	7
CAPÍTULO I- INTRODUÇÃO AO TEMA	15
I.1. Epidemiologia da doença coronária	17
I.2. Fatores de risco modificáveis	18
I.3. Doença coronária como doença multifatorial com genética complexa	33
I.4. Risco cardiovascular e risco genético	40
I.5. Justificação do estudo	47
CAPÍTULO II- OBJETIVOS GERAIS E ESPECÍFICOS	49
CAPÍTULO III- MÉTODOS	53
III.1. Amostra populacional	55
III.2. Homogeneidade e estrutura populacional	56
III.3. Critérios de inclusão e de exclusão	59
III.4. Dados colhidos e critérios de valorização das variáveis	60
III.5. Estudo genético	62
III.5.1. Variantes genéticas	62
III.5.2. Análise genética	76
III.5.3. Modelos genéticos	77
III.6. Score de risco genético (SRG)	78
III.7. Análise estatística	80
CAPÍTULO IV- RESULTADOS	91
IV.1. CARACTERÍSTICAS GERAIS DA POPULAÇÃO	95
IV.2. ESTUDO GENÉTICO e DOENÇA CORONÁRIA	98
IV.2.1. Estudo genético gene a gene	99
IV.2.2. Estudo genético multivariado	102
IV.2.3. Estudo genético multivariado ajustado	103
IV.3. INTERAÇÃO GENE-GENE e GENE-AMBIENTE	105
IV.4. SCORE DE RISCO GENÉTICO E DOENÇA CORONÁRIA	115
IV.4.1. Análise exploratória e escolha do modelo de score genético	117
IV.4.2. Estudo discriminativo do score genético	123
IV.4.3. Valor preditivo do score de risco genético	131

IV.4.4. Modelo preditor de DC por <i>machine learning</i> MDR	134
IV.4.5. Melhoria da reclassificação e discriminação com o SRG	135
IV.4.6. Comparação do SRG com novos marcadores de risco cardiovascular	137
IV.4.7. Comparação do SRG com a história familiar de doença coronária precoce	139
IV.4.8. Estudo do score no doente com doença coronária antes dos 45 anos	140
IV.5. SCORE GENÉTICO E RECLASSIFICAÇÃO do RISCO	144
CARDIOVASCULAR	
IV.5.1. Score Europeu	145
IV.5.2. Score de Framingham doença coronária	149
IV.5.3. Score de Framingham cardiovascular	152
IV.6. ESTUDO GENÉTICO PROSPETIVO DOS EVENTOS	155
CARDIOVASCULARES	
IV.6.1. Mortalidade cardiovascular	156
IV.6.2. <i>Endpoint</i> global: Enfarte + Revascularização + Mortalidade cardiovascular	162
CAPÍTULO V: DISCUSSÃO	169
V.1. Da metodologia do estudo	172
V.2. Do estudo gene a gene e doença coronária	178
V.3. Da interação gene-gene e gene- ambiente	184
V.4. Do score genético e doença coronária	188
V.5. Da reclassificação do risco cardiovascular com o SRG	199
V.6. Da análise prospetiva dos eventos cardiovasculares	203
5. FORÇAS E LIMITAÇÕES	213
6. CONCLUSÕES E PERSPETIVAS FUTURAS	215
7. BIBLIOGRAFIA	221
8. ARTIGOS PUBLICADOS	243

LISTA DE	ABREVIATURAS E ACRÓNIMOS
AA:	Antiagregação
ADN:	Ácido Desoxirribonucleico
Apo B:	Apolipoproteína B
ARN:	Ácido Ribonucleico
AUC:	Área sob a curva
AVC:	Acidente Vascular Cerebral
AHA:	<i>American Heart Association</i>
CARDIOGRAM:	<i>Coronary Artery Disease Genome-Wide Replication And Meta-analysis</i>
Cr:	Cromossoma
CI:	Cardiopatia isquémica
CV:	Cardiovascular
C Estatística:	Calibração Estatística
DC:	Doença Coronária
DC3v:	Doença Coronária de 3 vasos
DL:	Desiquilíbrio de <i>Linkage</i>
DP	Desvio-padrão
DM:	Diabetes Mellitus
DAA:	Dupla Antiagregação
EAM:	Enfarte agudo do Miocárdio
ESC:	<i>European Society of Cardiology</i>
FRT:	Fatores de risco tradicionais
G-A:	Gene-Ambiente
G-G:	Gene-Gene
GWAS:	<i>Genome Wide Association Study</i>
HapMap:	<i>International HapMap Consortium</i>
Hcy:	Homocisteína
HDL:	Lipoproteína de Alta Densidade
HF:	História familiar
H.L.:	Hosmer Lemeshow
HR:	<i>Hazard Ratio</i>
HTA	Hipertensão Arterial
IC:	Intervalo de Confiança de Cornfield
IDI:	<i>Integrated Discrimination Improvement</i>
IMC:	Índice de Massa Corporal
IDL	Lipoproteína de densidade intermédia
KM:	Kaplan-Meier
LDL:	Lipoproteína de baixa densidade
Lp(a):	Lipoproteína a
MACE:	<i>Major adverse cardiovascular events</i>

MDR:	<i>Multifactor dimensionality reduction</i>
MD4C:	<i>Malmö Diet and Cancer study</i>
MIGen:	<i>Myocardial Infarction Genetics consortium</i>
NNT:	<i>Number needed to treat</i>
NHGRI:	<i>National Human Genome Research Institute</i>
NRI:	<i>Net Reclassification index</i>
OR:	<i>Odds Ratio</i>
OMS	Organização Mundial de Saúde
PAS:	Pressão arterial sistólica
PAD:	Pressão arterial diastólica
PCA:	Análise de Componentes Principais
PCR:	Proteína C reativa
PGH:	Projeto do Genoma Humano
REGICOR	<i>Registre Gironí del Cor</i>
ROC:	<i>Receiver Operating Characteristic</i>
RR:	Risco Relativo
SNP:	<i>Single Nucleotide Polymorphism</i>
SRF:	Score de risco Framingham
SRG:	Score de risco genético
SCORE:	Systematic Coronary Risk Evaluation
WTCCC:	Welcome Trust Consortium
4C:	<i>Genetic consortium11</i>

RESUMO

Introdução: A doença das artérias coronárias (DC) representa o paradigma da doença complexa influenciada por múltiplas interações gene-gene (G-G) e gene-ambiente (G-A). A influência do componente genético na DC é ainda desconhecida.

Objetivos: Caracterizar o componente genético de risco para a DC numa população portuguesa homogénea, comparando-o com controlos saudáveis. Pretende-se analisar 33 variantes genéticas associadas à DC em diferentes eixos fisiopatológicos e estudar a sua interação com os fatores de risco tradicionais (FRT). Adicionalmente, pretende-se construir um score de risco genético (SRG) e calcular o seu valor preditivo independente e adicional aos FRT. Pretende-se também estudar a associação do componente genético com a ocorrência de eventos cardiovasculares (CV) e de mortalidade CV na população com DC.

Métodos: No estudo caso-controlo, estudámos 2888 indivíduos (idade média 53,3±8,0 anos), 1566 indivíduos com DC documentada e 1322 controlos ajustados aos casos por idade e género. Foram comparadas as suas características clínicas, do exame objetivo e bioquímico-laboratoriais. Por análise bivariada, num estudo gene a gene, foram comparados os 33 genótipos associados com a DC. Por análise multivariada, foi estabelecida a relação entre as variantes genéticas consideradas significativas e os FRT. Com recurso a diversos métodos, incluindo o MDR (*Multifactor Dimensionality Reduction*) para *data mining* dos dados, estudaram-se as interações G-G e G-A. Os melhores modelos no MDR foram selecionados com base na acuidade ponderada e consistência de validação cruzada. Foram feitas duas análises de interações G-G: das 33 variantes genéticas e das 10 com significância bivariada e foram apresentados os modelos preditores do risco para DC. As interações G-A são apresentadas por eixos fisiopatológicos na DC. Na construção do SRG, foram utilizadas 31 variantes genéticas em equilíbrio de Hardy Weinberg. Foram testados diversos modelos (aditivo e multiplicativo) na construção do SRG. Determinou-se o SRG na população total e estudou-se a sua distribuição caso-controlo, de acordo com o género, em quartis e distribuição em decis. Avaliaram-se as correlações do SRG com a idade, género, FRT e história familiar de DC precoce. Calculou-se o seu valor preditivo independente e adicional aos FRT e outros marcadores CV. É apresentado um modelo interativo para DC, obtido pelo MDR, pela combinação de FRT, SRG e marcadores CV. Estudou-se o potencial do SRG na reclassificação do risco para DC em subgrupos preditos pelo SCORE europeu, Framingham DC e Framingham CV. Calcularam-se as reclassificações (Net Reclassification Index) e Índices de Discriminação (IDI) obtidas pela introdução do SRG nos modelos de estratificação convencionais. Foram seguidos prospetivamente 1549 indivíduos com DC e avaliados os eventos adversos CV no seguimento. Numa análise gene a gene, foram determinadas as variantes genéticas associadas com maior mortalidade CV e estabelecidas as relações com os FRT. Foi feita uma análise de sobrevivência para a

mortalidade CV de acordo com o genótipo para a variante genética ADAMTS7. Foram feitas duas análises comparativas da sobrevivência livre de eventos CV nos doentes com risco genético elevado: quarto quartil versus primeiro quartil do SRG e acima e abaixo da mediana do SRG. Foram identificados os preditores independentes da mortalidade CV e de um *endpoint* combinado de EAM, necessidade de revascularização, ou mortalidade CV em modelos ajustados. O seu poder preditor foi discriminado pelo C-estatística. O potencial do SRG para a reclassificação dos doentes na predição de eventos CV foi expresso pelo NRI e pelo IDI.

Resultados: 1-Identificaram-se dez variantes genéticas associadas, de forma significativa, com a DC: LPA (OR=2,20), APOE (OR=1,26), ADIPOQ (OR=1,21), FTO (OR=1,19), PON 55 (OR=1,35), CDKN2B (OR=1,36), locus9p21 (OR=1,32), TCF721 (OR=1,54), ZC3 (OR=1,29) e PHACTR1 (OR=1,10). Destas, sete mantiveram significado no estudo multivariado (TCF21, ZC3HC1, CDKN2B, ADIPOQ, LPA, APOE, PON55, $p<0,02$) e seis (LPA, APOE, PON55, CDKN2B, TCF21 e ZC3 ($p<0,05$)) foram independentes dos FRT. Os preditores não genéticos com maior significância para a DC na nossa população foram: o tabagismo (OR=2,79), diabetes (DM) (OR=2,52) e dislipidemia (OR=3,41). A obesidade central, o sedentarismo, a homocisteína, a Lp(a) e a insuficiência renal estiveram também associados com a DC.

2-As interações G-A mais significativas por testes de associação foram a da diabetes com o TCF7L2 ($p=0,0001$) e a da dislipidemia com o ZNF259 ($p<0,0001$). As interações G-G mais significativas foram entre o ADAMTS7 GG e MTHFR677 CT ($p=8,6\times 10^{-4}$), CDKN2B AG e o PON192RR ($p=1,05\times 10^{-3}$), KIF6 GG e o ATIRCC ($p=3,29\times 10^{-4}$) e o PON192QR e o locus 9p21 GG ($p=8,57\times 10^{-4}$). Através do MDR (*data mining*), o modelo genético G-G do TCF21 e APOE foi o melhor preditor de DC com um OR de 1,51 (IC95%: 1,29-1,76; $p<0,0001$). O melhor modelo G-A foi o da variante ZC3CH1, tabagismo, HTA e DM com uma acuidade ponderada de 0,67 (OR=4,30; $p<0,0001$). No eixo da DM o modelo TAS2R50, TCF7L2 e a DM com acuidade de 0,613 (OR=2,98; $p<0,0001$). No eixo oxidativo, o melhor modelo encontrado foi MTHFD1L; PON1; MTHFR677; PON55 com 0,568 de acuidade ponderada (OR=1,73; $p<0,0001$). No eixo da dislipidemia, o modelo G-A entre o PSRC1 e a dislipidemia mostraram uma maior interação com acuidade de 0,60 (OR=3,92; $p<0,0001$). A interação G-G epistástica foi elevada entre o gene APOE e TCF21, entre o gene KIF6 e ECA e entre o par AGT e MTHFR1298. 3- Independentemente do modelo utilizado, o SRG foi sempre mais elevado nos doentes com DC do que nos controlos. A capacidade preditora máxima do SRG foi obtida com 31 variantes genéticas (AUC=0,61 (IC95%: 0,59-0,62)). O modelo multiplicativo genotípico (SRGm) apresentou a melhor capacidade preditora (AUC=0,61) com 62,3% de especificidade e 54% de sensibilidade. A mediana do SRGm na população foi de 0,36 (0,04-9,29). O SRGm com 31 variantes associou-se com risco independente para a DC (OR=1,8; IC95%:1,5-2,1; $p<0,0001$) OR=1,80; IC95%: 1,52-2,12; $p<0,0001$) mesmo ajustado para os FRT. Os indivíduos com SGRm mais elevado (4º quartil) têm cerca de 2,6 vezes maior risco

de DC comparados aos do quartil inferior (OR=2,56; IC95%: 2,07-3,17). A partir do 6º decil do SRGm existe uma associação significativa com DC. O indivíduo no 10º decil apresenta 4 vezes maior risco de DC (OR=4,04; IC95%: 2,85-5,72; $p<0,0001$) do que um indivíduo no 1º decil. O SRGm não apresentou correlações significativas com o género. O SRGm apresentou correlação com o grupo etário, a dislipidemia e a história familiar de DC precoce. O MDR identificou o modelo preditor combinado (FRT+SRG+marcador de risco CV) composto pelo SRG+dislipidemia+diabetes+tabagismo; sedentarismo; Lp(a) com um OR de 5,70 (IC95%: 4,81-6,76; $p<0,0001$) com uma acuidade ponderada de 0,70 e consistência cruzada de 10/10. O SRG foi independente dos FRT, de outros marcadores de risco CV e da história familiar de DC precoce. Este SRG proporcionou uma melhoria da capacidade preditiva em relação aos FRT (AUC (FRT) = 0,74; AUC (FRT+SRG) = 0,76; $p<0,0001$) e proporcionou uma reclassificação expressa pelo NRI de 20,9% nos indivíduos com DC e de 3,3% nos indivíduos controlos. Assim, 19,5% dos doentes com DC situaram-se numa categoria de risco mais elevada após inclusão do SGR. O comportamento do SRG foi semelhante nos doentes com <45 anos de idade. A incorporação do SRG provou incrementos positivos na curva AUC de 0,76 para 0,79 com acréscimo da sensibilidade (70%) e especificidade (77%) em relação a um modelo basal de FRT ($p=0,0004$). Por análise gene a gene, as variantes: TCF21 (OR=1,24), a ZC3HC1 (OR=1,72), o ZNF259 (OR=1,35) e o gene da ECA (OR=1,36) associaram-se com a DC neste grupo. Nestes, obteve-se um NRI de -21,6% ($p<0,0001$) e IDI de 0,6% ($p=0,017$). 4- Em relação a qualquer modelo de estratificação de risco habitual na prática clínica, o SRG demonstrou utilidade na reclassificação e melhor discriminação do risco (score Europeu NRI=30,8%, $p<0,0001$; IDI=2,0, $p<0,0001$); (score DC *Framingham* NRI=31,7%, $p<0,0001$; IDI=2,3, $p<0,0001$); (score CV *Framingham* NRI=31,7%, $p<0,0001$; IDI=2,3, $p<0,0001$). 5-Várias variantes genéticas foram associadas com eventos CV adversos no seguimento de doentes com DC. O genótipo AA do gene ADAMTS7 associou-se com uma maior mortalidade CV (HR=1,89; IC 95%: 1,00-3,67; $p=0,05$). A variável TCF21 associou-se com a ocorrência de eventos adversos (C-índice=0,519; IC95%: 0,49-0,55; $p=0,018$). Por fim, o SRG moderadamente elevado foi um preditor independente da mortalidade CV (HR=2,03; IC95%: 1,16-3,56; $p=0,014$) e de qualquer evento adverso (HR=1,35; IC95%: 1,04- 1,75; $p=0,023$) mesmo quando ajustado para potenciais co-variáveis.

Conclusões: Nesta amostra da população portuguesa, documentaram-se diferenças genéticas significativas caso-controlo. O risco de DC aumenta consoante o número de variantes genéticas. Modelos de interação contendo variantes genéticas e marcadores de risco podem melhorar o nosso conhecimento sobre a epigenética da DC. Um SRG pode representar a melhor forma de quantificar o risco individual. Este SRG proporcionou uma melhoria da capacidade preditiva em relação a qualquer modelo de estratificação de risco habitual e poderá melhorar a estratificação do risco, mesmo após modificação dos FRT. Também na população com DC, há variantes genéticas associadas à ocorrência de eventos adversos e mortalidade CV.

ABSTRACT

Introduction: Coronary artery disease (CAD) is a complex disease influenced by multiple gene-gene (G-G) and gene-environment (G-E) interactions. So far, the relevance of the genetic component for CAD remains unclear.

Aim: In a portuguese homogeneous population, the genetic component predisposing to CAD was evaluated by comparing it with a matched healthy control population. 33 genetic variants previously associated with CAD in different fisiopathological pathways (celular, lipid, Diabetes/Obesity, Inflammation/Oxidation and Renin-Angiotensin-Aldosterone system pathways) were analysed. Their interaction with traditional risk factors (TRF) was studied. A Genetic risk score (GRS) was computed and its independent predictive power was determined. CAD patients were followed prospectively to evaluate the influence of the genetic risk component of CAD in the occurrence of major adverse CV events and CV mortality.

Methods: Case-control study of 2888 individuals with a mean age of 53.3 ± 8 years, with 1566 patients with CAD and 1322 controls, age and sex matched to patients. Baseline, clinical and biochemical characteristics were compared between cases and controls. The 33 variants genotypes were determined and gene to gene comparison between cases and controls was performed.

After bivariate analysis, multivariate analysis was used to determine which genetic variants were significantly associated with CAD and their relationship with TRF was established.

Using *Multifactor Dimensionality Reduction* (MDR) for *data mining* of data, the G-G and G-A interactions were studied. Two analyses of G-G interactions were made between the total 33 genetic variants and also between the 10 variants found to have bivariate significance, and the best risk models predictors for CAD were presented. The G-E interactions are presented by presumptive pathophysiological pathways in CAD. The best models in the MDR were selected based on the weighted accuracy and cross-validation consistency. The construction of the GRS included 31 genetic variants in Hardy Weinberg equilibrium. GRS was constructed according to several models under additive and multiplicative methods. The additive and multiplicative GRS were determined in the total population and their distribution between cases and controls, according to gender, in quartiles and distribution in decimals was studied. GRS correlations with age, gender, TRF, and family history of early CAD were assessed. Their additional predictive value of GRS was calculated taking TRF and other CV markers into equation. An interactive model for CAD, obtained by the MDR, based on the combination of TRF, GRS and CV markers is presented. We studied the GRS potential use for risk reclassification in subgroups of risk predicted by European SCORE, Framingham CAD and Framingham CV. Net reclassification indexes (NRI) obtained by the introduction of GRS in conventional stratification models were calculated. The respective Improved discrimination indices (IDI) were determined. A total of 1549 individuals were prospectively followed up and the adverse events assessed at follow-up. In a gene-to-gene

analysis, the genetic variants associated with higher CV mortality and their relationships with TRF were determined. A survival analysis for CV mortality was made according to the genotype for the genetic variant ADAMTS7. Two comparative analyzes of CV-free survival in patients with high genetic risk were performed: fourth quartile versus first quartile of GRS and above and below the median of SRG. We identified independent predictors of CV mortality and a combined endpoint of AMI, need for revascularization, or CV mortality in adjusted models. Its predictive power was discriminated by C-statistics. The potential of GRS use for patient reclassification in event prediction was expressed by NRI and IDI.

RESULTS: 1-We identified ten genetic variants that were significantly associated with CAD: LPA (OR=2.20), APOE (OR=1.26), ADIPOQ (OR=1.21), FTO (OR=1.36), PON55 (OR=1.35), CDKN2B (OR=1.36), locus9p21 (OR=1.32), TCF721 (OR=1.54), ZC3 and PHACTR1 (OR=1.1). (P <0.05) Seven genetic variants (TCF21, ZC3HC1, CDKN2B, ADIPOQ, APOE, LPA, PON55, p<0,02), were identified in the multivariate study and six (LPA, APOE, PON55, CDKN2B, TCF21, and ZC3 (p <0.05)) were found to be independent of the FRT. The genetic variants in hypertension pathway were no longer significant in multivariate model. Nongenetic independent predictors for CAD were smoking (OR=2.79), diabetes mellitus (DM) (OR=2.52) and dyslipidemia (OR=3.41). Central obesity, sedentary lifestyle, homocysteine, Lp (a) and renal failure were also associated with CAD. 2-The most significant G-E interactions were the association of diabetes mellitus with TCF7L2 (p = 0.0001) and dyslipidemia with ZNF259 (p <0.0001). The most significant G-G interactions were between ADAMTS7 GG and MTHFR677 CT (p=8.6x10⁻⁴), CDKN2B AG and PON192RR (p=1.05x10⁻³), KIF6GG and ATIRCC (p = 3.29x10⁻⁴) and PON192QR and the 9p21 GG locus (p = 8.57x10⁻⁴). With data mining MDR, the genetic model G-G of TCF21 and APOE was the best predictor of CAD with an OR of 1.51 (95% CI: 1.29-1.76; p<0.0001). The best G-E model was represented by the ZC3CH1 variant, smoking, AHT and DM with a weighted accuracy of 0.67 (OR=4.30, p<0.0001). The best model in the DM/Obesity pathway, was presented by the combination of TAS2R50, TCF7L2 and DM with accuracy of 0.613 (OR=2.98, p<0.0001). In the oxidative pathway, the best model found was MTHFD1L; PON1; MTHFR677; PON55 with 0.568 weighted accuracy (OR = 1.73, p <0.0001). In the lipidic axis, the G-E model between PSRC1 and dyslipidemia showed the highest interaction with an accuracy of 0.60 (OR=3.92, p<0.0001). Epistatic G-G interaction was elevated between the APOE and TCF21 gene, between the KIF6 and ACE gene, and between the AGT and MTHFR1298 genes. 3- Regardless of the model used, GRS was higher in patients with CAD than in controls. The maximum predictive capacity of the GRS was obtained with 31 genetic variants (AUC=0.61 (95% CI: 0.59-0.62)). The genotypic multiplicative model (GRSm) had the best predictive capacity (AUC=0.61) with 62.3% specificity and 54% sensitivity. The mean GRSm in the population was 0.36 (0.04-9.29). The GRSm with 31 variants was associated with independent risk for CAD (OR=1.80; 95% CI: 1.52-2.12; p <0.0001) even if adjusted for TRF. Individuals with higher GRSm (4th quartile) have a 2.6-fold higher risk of CAD compared

to those in the lower quartile (OR=2.56; 95% CI:2.07-3.17). There is a consistent association with CAD from the 6th decile of GRSm up. An individual in the 10th decile GRS had a 4-fold higher risk of CAD (OR= 4,05% CI: 2.85-5.72, $p<0.0001$) than an individual in the 1st decile. GRSm did not present significant correlations with gender. The GRS showed correlation with the age group, dyslipidemia and family history of early CAD. The MDR identified the combined predictor model (FRT + SRG + CV risk marker) composed of SRG + dyslipidemia + DM + smoking; sedentary lifestyle; Lp (a) with an OR of 5.70 (95% CI: 4.81-6.76, $p <0.0001$) with a weighted acuity of 0.70 and cross-consistency of 10/10. SRG was independent of FRT, other CV risk markers, and family history of early CD.

This GRS provided an improvement in the predictive capacity of TRF (AUC TRF= 0.74, AUC (TRF + GRS)= 0.76, $p <0.0001$) and provided an NRI-expressed reclassification of 20.9 % in individuals with CAD and 3.3% in controls. Thus, 19.5% of patients with CAD were in a higher risk category after inclusion of the GRS. The GRS behavior was similar in patients <45 years of age. The GRS incorporation showed positive increases in the AUC curve from 0.76 to 0.79 with increased sensitivity (70%) and specificity (77%) compared to a baseline TRF model ($p=0.0004$). Four gene variants; namely TCF21 (OR=1.24), ZC3HC1 (OR=1.72), ZNF259 (OR=1.35) and ACE variant with (OR = 1.36) were associated with CAD in this age group. In these, an NRI of -21.6% ($p <0.0001$) and IDI of 0.6% ($p=0.017$) was obtained. 4- Regarding any standard risk stratification model, SRG demonstrated utility in reclassification and better risk discrimination compared with (European Score NRI=30.8%, $p<0.0001$, IDI=2.0, $p<0.0001$); (Framingham CAD score, NRI = 31.7%, $p<0.0001$; IDI= 2.3, $p <0.0001$); (Framingham CV score, NRI=31.7%, $p<0.0001$, IDI=2.3, $p<0.0001$). 5- In CAD patients several genetic variants were associated with adverse CV events at follow-up. The AA genotype of the ADAMTS7 gene was associated with a higher CV mortality (HR=1.89, 95% CI:1.00-3.67, $p=0.05$). The TCF21 variant was associated with the occurrence of adverse events (C-index=0.519, 95% CI: 0.49-0.55, $p = 0.018$). Finally, moderately elevated SRG was an independent predictor of CV mortality (HR=2.03, 95% CI: 1.16-3.56, $p=0.014$) and any adverse events (HR=1.35, 95% CI, : 1.04-1.75, $p = 0.023$) even when adjusted for potential covariates.

Conclusions: In this population, significant genetic differences were documented between CAD patients and matched controls. The risk for CAD increases according to the number of genetic variants. Interaction models containing genetic variants and risk markers may improve our understanding of the epigenetics of CAD. A GRS may represent the best way to quantify individual risk for CAD. This GRS provided an improvement in predictive capacity over any standard risk stratification model used in clinical practice and could improve risk stratification even after modification of the TRF. Also, in CAD patients, there are genetic variants associated with the occurrence of adverse events and CV mortality.

I. INTRODUÇÃO

I.1. EPIDEMIOLOGIA DA DOENÇA CORONÁRIA

A doença das artérias coronárias, também conhecida por Doença Coronária (DC), continua a ser a causa de morte mais frequente no mundo industrializado. O estudo do *Global Burden of Disease* estimou que em 2013 cerca de 17,3 milhões de mortes no panorama mundial foram atribuíveis à doença cardiovascular (CV). Esta correspondeu a 31,5% do total de mortes e a 45% do total de mortes por doenças não transmissíveis, mais do dobro das atribuíveis às neoplasias e superior à combinação das mortes materno-infantis e nutricionais. (1)

Das doenças cardiovasculares, a DC continua a ser a causa mais frequente de mortalidade em ambos os géneros, com cerca de 19% no género masculino e de 20% no género feminino. O Acidente Vascular Cerebral (AVC) continua a ser mais representativo no género feminino (14%) e representa 9% das mortes cardiovasculares no género masculino. (1, 2) (Figura 1)

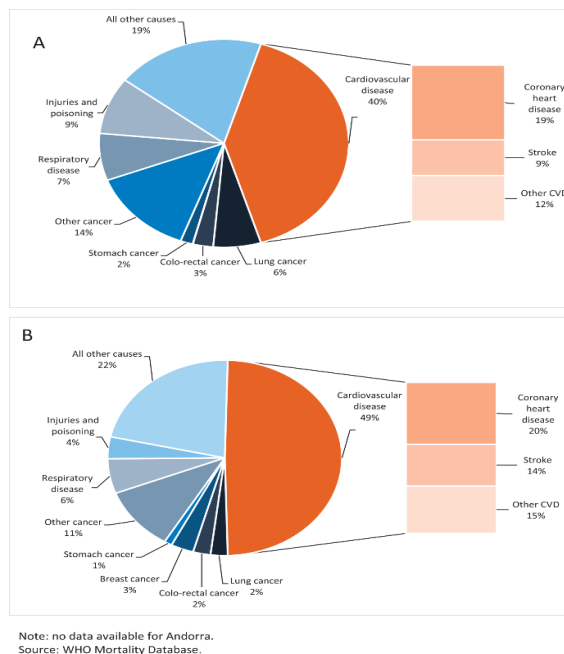


Figura 1: Causas de Mortalidade de acordo com o género. (Townsend *et al.*, 2016). (1)

Legenda: A-Principais causas de mortalidade no género masculino, B-Principais de causas de mortalidade no género feminino.

Na Europa dos 28 países, com os dados de 2013 (1), a taxa de mortalidade da DC foi de 132 indivíduos por 100 000 habitantes. Os países com mortalidade mais elevada por DC foram a Lituânia, Letónia, Eslováquia, Hungria, e República Checa — com taxas superiores a 350 mortes per 100 000 habitantes. Recentemente, países como a França, Portugal, Holanda, Espanha, Bélgica, Dinamarca, Luxemburgo, Grécia e o Liechtenstein, conseguiram reduções significativas das suas taxas de mortalidade por DC para valores abaixo

das 100 mortes per 100 000 habitantes, resultado de investimentos para disponibilizar tratamento dos Síndromes Coronários nomeadamente com a Angioplastia Primária. (1, 2)

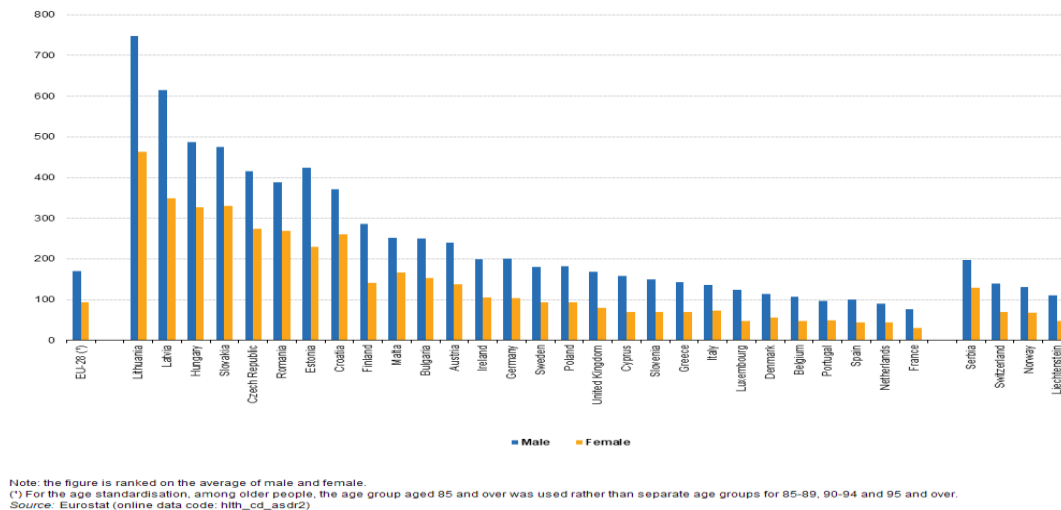


Figura 2: Taxa de Mortalidade por 100.000 habitantes por género e por nacionalidade. (Eurostat, online data code: hlth_cd_asdr2).

Legenda: As barras laranjas representam o género feminino e as barras azuis representam o género masculino.

Assim, embora globalmente a doença cardiovascular (CV) permaneça a principal causa de mortalidade, assiste-se a uma importante mudança de paradigma com a redução crescente de mortalidade associada à Doença CV nos países desenvolvidos da Europa e da América do Norte e a um aumento da incidência da Doença CV nos países em desenvolvimento nomeadamente na Índia e na China. (Figura 2) Este aumento, reflete alterações importantes nos hábitos de vida associados ao mundo industrializado e urbanizado, com maior incidência do risco de Doença CV associado ao aumento dos Fatores de Risco Tradicionais (FRT) como a pressão arterial, o tabagismo, a dislipidemia, a obesidade e diabetes mellitus (DM). (2)

1.2 FATORES DE RISCO MODIFICÁVEIS ASSOCIADOS COM DOENÇA CORONÁRIA

Diferentes Fatores de Risco podem ser avaliados como contribuidores potenciais para o aumento da incidência de Doenças CV. Incluem-se fatores ambientais, características demográficas, sociais, tipo de dieta e características fenotípicas tais como o índice de massa corporal (IMC), a hipertensão arterial (HTA), a DM, a dislipidemia e o tabagismo. Os FRT fortemente associados às doenças CV são os mais extensamente estudados.

I.2.1 HIPERTENSÃO ARTERIAL

A elevação da pressão arterial permanece um fator de risco maior para todo o espectro das doenças CV. O estudo *Global Burden of Hypertension*, com dados referentes a 8,6 milhões de participantes de 195 países, demonstrou que a incidência de pressão arterial superior a 110/115mmHg e a HTA definida como superior a 140mmHg continua a aumentar. A mortalidade associada à HTA aumentou entre 1990 e 2015 de 97,9 (IC95%: 87,5-108,1) para 106,3 (IC95%: 94,6-118,1) por cada 100 000 habitantes. (3)

A mortalidade relacionada com a HTA esteve maioritariamente associada à DC que atingiu (4,9 milhões habitantes), ao AVC hemorrágico (2 milhões de habitantes) e ao AVC isquémico (1,5 milhões de habitantes). Em 2015, mais de metade das perdas associadas à elevação da pressão arterial ocorreram em 5 países, nomeadamente na China, Índia, Federação Russa, Indonésia e Estados Unidos da América. (3)

I.2.1.1 ASSOCIAÇÃO ENTRE A HTA E A DOENÇA CORONÁRIA

Classicamente, o aumento do risco CV tem sido associado à elevação da pressão arterial sistólica (PAS) em indivíduos acima dos 60 anos e à elevação da pressão arterial diastólica (PAD) em populações mais jovens (5). Contudo, a elevação do risco CV é contínua, mesmo com valores crescentes de pressão arterial dentro dos valores considerados “normais”. (6)

Também, aos valores mais elevados da pressão arterial, associam-se muitas vezes a obesidade e a dislipidemia, que atuam em conjunto no aumento do risco CV. (4)

A evidência da importância da HTA na doença CV reflete-se na melhoria do prognóstico com o controlo tensional após introdução de terapêutica. Na meta-análise de vários estudos randomizados com fármacos antihipertensores, a redução da PAS de 3 a 10 mmHg versus placebo associou-se a uma redução de risco relativo de 20%. Por cada 10 a 12 mmHg de redução da PAS e 5 a 6 mmHg de redução de PAD estima-se 16% de redução do AVC e cerca de 13% de redução de DC. (6)

Em Portugal, segundo Espiga de Macedo, a taxa de prevalência da HTA situa-se nos 42,1%, sendo mais elevada no género masculino do que no feminino, em qualquer dos grupos etários analisados. A HTA foi mais elevada na população com mais de 64 anos de idade, com 79% dos indivíduos do género masculino e 78,7% do género feminino afetados. (7)

Apesar da evidência científica e da grande variedade de agentes farmacológicos disponíveis no panorama mundial, Portugal apresenta reduzidas taxas de adequado controlo tensional com o respetivo impacto negativo no risco CV residual.

I.2.2 DIABETES MELLITUS (DM)

A prevalência da DM tem vindo a aumentar globalmente, com especial impacto nos países em vias de desenvolvimento, onde ocorrem cerca de dois terços dos casos mundiais. Os colaboradores da meta-análise de *Risk Factor Collaboration* (7) estudaram as tendências na prevalência da diabetes avaliada como glicémia em jejum superior a 7mmol/dl ou 126 mg/dl ou terapêutica hipoglicemiante. Analisando dados desde 1980 até 2014 em cerca de 200 países, demonstrou-se que desde 1980 houve um aumento da prevalência da diabetes ajustada à idade de 4,3% para 9% no género masculino, e de 5% para 7,9% no género feminino. O número de adultos diabéticos aumentou de 108 milhões em 1980 para 422 milhões em 2014. Este incremento deve-se ao aumento da incidência da DM em 28,5% dos casos e ao aumento e envelhecimento da população em 39,7% dos casos.

Na Europa Ocidental a incidência da DM tem-se mantido estável, devendo-se o aumento da sua prevalência, ao envelhecimento populacional. No género feminino, a menor prevalência de diabetes ajustada à idade foi encontrada na Europa Ocidental com menos de 5% de prevalência. Na maioria dos restantes países analisados a glicémia média aumentou ou permaneceu inalterada. As regiões que demonstraram um maior aumento na prevalência da diabetes foram as regiões dos arquipélagos das ilhas da Polinésia e da Micronésia onde a diabetes atingiu cerca de 25% da população, seguidas das Regiões do Médio Oriente e Norte de África (8). O aumento da glicémia e da prevalência da Diabetes associou-se à tendência, já reportada previamente, do incremento do IMC em quase todas as regiões, demonstrando a já consistente associação entre excesso ponderal, intolerância aos hidratos de carbono e resistência à insulina.

I.2.2.1 ASSOCIAÇÃO ENTRE A DIABETES E A DOENÇA CORONÁRIA

A DM é bem conhecida por aumentar o risco CV. Segundo dados epidemiológicos clássicos a DM duplica o risco de doença CV no género masculino e quadruplica o risco no género feminino. (9) Por isso, a doença CV representa a principal causa de mortalidade na população diabética.

Haffner *et al.* (10) após um seguimento de 7 anos de 3 grupos de doentes (diabéticos, não diabéticos e com DC) concluíram que a incidência de Enfarte Agudo do Miocárdio (EAM) nos doentes diabéticos foi de 2,7%, semelhante à incidência de EAM no grupo de doentes com DC (2,9%/ano) e cerca de seis vezes mais do que a incidência no 3º braço do estudo, doentes não diabéticos e sem DC (0,5%/ano). Os autores concluíram que a mortalidade por DC em doentes diabéticos sem antecedentes de EAM era sobreponível à mortalidade dos doentes não diabéticos com antecedentes de EAM com *hazard ratio* (HR) de 1,40 (IC 95%: 0,7-2,6). Este

facto reforça o grau de intervenção proposto para o doente diabético na maioria das orientações/*guidelines* de prevenção cardiovascular. (9) Numa meta-análise de 37 estudos prospetivos com 447 064 participantes o risco de mortalidade por DC no diabético foi 3,5 vezes superior à mortalidade nos não diabéticos. (11) Este excesso de risco é mais acentuado no género feminino, com risco relativo de evento fatal por DC 1,46 vezes mais elevado do que no género masculino, parcialmente explicado pela presença de outros fatores de risco CV e *clustering* de variáveis. (11-13).

Assim, assume-se que o risco de desenvolver DM é largamente modificável e que está muitas vezes associado à tríade metabólica que compreende um perfil lipídico adverso, a insulinoresistência e o excesso ponderal. O aumento da prevalência desta tríade no mundo urbanizado reflete a perda de hábitos de vida saudáveis, menor atividade física e alterações da dieta. Dada a importante correlação da DM com a DC e a sua associação a outros fatores de risco independentes para a DC, é importante reforçar e implementar novas estratégias para reduzir a incidência e melhorar o controlo dos doentes já diagnosticados com DM.

I.2.3 DISLIPIDEMIA

O termo dislipidemia reflete alterações dos níveis dos lípidos plasmáticos, que estão intimamente relacionados com a formação de placas ateroscleróticas vasculares. A nível mundial, têm-se verificado tendências opostas nos níveis médios de colesterol plasmático, com reduções discretas na América do Norte, Europa, Austrália e Ásia central, enquanto se observa uma tendência crescente dos níveis médios de colesterol no sudoeste e este asiático e nalgumas regiões do Pacífico. (14) Estas variações parecem estar associadas às mudanças de hábitos de vida, conseqüentes à urbanização e a disponibilidade e acesso à terapêutica farmacológica. De qualquer modo, os níveis médios ajustados mais elevados ainda se medem na Austrália, Ásia Central, América do Norte e Europa Ocidental sendo mais baixos na África Subsariana. Segundo dados de 2008, a nível mundial, o nível médio de colesterol ajustado para a idade foi de 4,64 mmol/l (179,4 mg/dl) no género masculino e de 4,76 mmol/l (184 mg/dl) no género feminino. (14)

I.2.3.1 ASSOCIAÇÃO DA DISLIPIDEMIA COM A DOENÇA CORONÁRIA

As principais partículas de colesterol associadas à aterosclerose são as lipoproteínas de baixa densidade (LDL), as lipoproteínas de densidade intermédia (IDL), as lipoproteínas de muito baixa densidade (VLDL) e a lipoproteína(a) (Lp(a)). A apolipoproteína B (ApoB), o principal componente das LDL, IDL e VLDL, facilita a produção de VLDL e medeia as relações entre as lipoproteínas aterogénicas e a parede vascular. A

apolipoproteína A é o principal componente das lipoproteínas de alta densidade (HDL). A lipoproteína (a)-Lp(a) é uma lipoproteína com estrutura básica semelhante à LDL-colesterol, com a apolipoproteína(a) ligada à ApoB 100. A apolipoproteína (a) apresenta semelhança estrutural com a molécula do plasminogénio, facilitando assim, o aparecimento de fenómenos trombogénicos. A variabilidade da apolipoproteína(a) é altamente dependente do tipo e número de cópias do domínio *Kringle*, uma parte fundamental para a semelhança da sua estrutura ao plasminogénio. Quanto menor o número de repetições do *Kringle* tipo dois, ou seja, sequências de apolipoproteína (a) menores, maiores as concentrações da Lp(a) e maior a atividade pró-aterotrombogénica. Assim, há uma relação inversa entre o peso molecular das isoformas de apolipoproteína(a) e a concentração plasmática da Lp(a). Níveis de Lp(a) acima de 25 mg/dl estão associados a aumento do risco de doença aterosclerótica e EAM, independente de outros fatores de risco. (15)

Numa meta-análise de 36 estudos prospetivos com 126 634 participantes, a incidência de DC no tercil superior e inferior dos valores basais de Lp(a) foram respetivamente de 5,6/1000 pessoas-ano e de 4,4/1000 pessoas-ano. (16) O risco para DC após ajuste para covariáveis de potencial confusão como os outros parâmetros lipídicos, revelou um risco independente de 1,13 (1,09-1,18). (16)

Na meta-análise *Prospective Studies Collaboration* envolvendo cerca de 900 000 adultos europeus e norte americanos confirmou-se existir uma relação gradual entre o colesterol total e a mortalidade por DC sem limite inferior bem definido. (17) A redução de 1 mmol/l (38,7 mg/dl) nos valores do colesterol total associa-se a uma redução do risco relativo da mortalidade por DC de 50% na faixa etária dos 40-49 anos de idade, de 33% na faixa etária dos 50-69 anos de idade e de 17% na faixa etária dos 70-89 anos de idade.

Outra meta-análise com 170 000 participantes de 26 estudos randomizados com estatina versus placebo constatou-se que a redução de 1 mmol/L (38,7 mg/dl) associa-se à redução do risco relativo de 10% da mortalidade de qualquer causa (RR=0,90; IC 95%: 0,87-0,93) maioritariamente, à custa da redução de 19% para a mortalidade associada a DC (RR=0,80; IC 95%: 0,74-0,87). (18) Efeitos similares na redução do risco CV foram observados com a redução dos valores de Apolipoproteína B em doentes medicados com estatinas.

Consistente com os estudos observacionais, onde já se tinha defendido *the lower the better* para os lípidos plasmáticos, no estudo IMPROVE-IT demonstrou-se um acréscimo de redução de eventos (EAM, AVC e revascularização a 30 dias) com reduções mais intensas dos lípidos plasmáticos (<70mg/dl), evidenciando uma relação estreita entre a redução dos lípidos e a redução da carga aterosclerótica global. (19)

I.2.4 OBESIDADE

A nível mundial, estima-se que 1,5 bilhões de indivíduos tenham excesso ponderal ou obesidade. Globalmente o Índice de Massa Corporal (IMC), introduzido por Quetelet, tem vindo consistentemente a aumentar em todas as regiões do globo. Dados da meta-análise referente a 19 milhões de adultos (9,9 do género masculino e 9,3 do género feminino) de 186 países revelaram aumentos consistentes da prevalência da obesidade desde 1975 a 2014. (20) Globalmente, o IMC ajustado para a idade aumentou no género masculino de 21,7 kg/m² em 1975 para 24,2 kg/m² em 2014; e no género feminino de 22,1 kg/m² em 1975 para 24,4 kg/m² em 2014. (20) Os valores mais elevados do IMC ocorreram nas regiões da Oceânia, da Polinésia e da Micronésia, com um IMC médio de 29,2 kg/m² no género masculino e de 32,2 kg/m² no género feminino. (20) No panorama global, de 1975 para 2014, a prevalência da obesidade no género masculino aumentou de 3,2% para 10,8% e no género feminino passou de 6,4% para 14,9%. Na população mundial, em 2014, eram muito obesos (IMC≥35 kg/m²) 2,3% (2,0-2,7) dos indivíduos do género masculino e 5% (4,4-5,6) do género feminino. A prevalência da obesidade mórbida (IMC>40) foi de 0,64% (0,46-0,86) no género masculino e de 1,6% (1,3-1,9) no género feminino. (20)

I.2.4.1 ASSOCIAÇÃO ENTRE A OBESIDADE E DOENÇA CORONÁRIA

A Organização mundial de Saúde (OMS) definiu os valores de excesso ponderal para um IMC ≥ 25 kg/m² e de obesidade para um IMC ≥ 30 kg/m². O tipo de obesidade mais frequentemente associada com DC é caracterizada pela acumulação de gordura visceral em localização central entre a anca e a cintura. Denominada obesidade visceral, relaciona-se com a deposição de gordura “castanha”, de composição mais aterogénica e frequentemente associada à Síndrome Metabólica, à insulinoresistência e à HTA. Assim, em alternativa ao IMC, surgiram dois índices antropométricos da gordura visceral como discriminadores de risco cardio-metabólico utilizados na avaliação CV: a cintura-anca e a cintura-altura. A circunferência da cintura > 90 cm no género masculino e > 84cm no género feminino, a razão cintura-anca > 0,9 no género masculino e > 0,80 no género feminino ou uma razão cintura/altura superior a 0,5, mesmo após ajuste para o IMC, têm sido apontadas como fatores de risco associados à DC. (9, 21,22)

Para alguns autores, uma vez que a cintura-altura aumentou o poder discriminativo avaliado pela *Area under the curve* (AUC) dado pelo IMC (23), existe evidência científica de que, as medidas da gordura visceral são melhores discriminadores do risco cardio-metabólico, quer no género feminino, quer no género masculino. Independentemente do método utilizado, o excesso ponderal e a obesidade condicionam um aumento do risco para o desenvolvimento da DC.

Uma meta-análise de 2011, com cerca de 200 000 participantes de 17 países, indica que o maior risco de DC foi associado ao incremento de 4,5 kg/m², do desvio padrão acima do IMC de 20 kg/m², de 12,6 cm do desvio padrão da cintura e de 0,083 do desvio-padrão da cintura-altura. Após ajuste para as variáveis de idade, género e tabagismo, o aumento de risco de DC foi estimado em: 1,23 por cada aumento do desvio padrão do IMC; 1,27 por cada incremento do desvio padrão da circunferência da cintura; e de 1,25 por cada incremento do desvio padrão da cintura-anca. (24)

I.2.5 TABAGISMO

A epidemia mundial do tabagismo atingiu proporções pandémicas com cerca de 1,3 biliões de fumadores e cerca de 6 milhões de mortes anuais associadas ao consumo de tabaco, sendo a segunda causa mais importante de mortalidade CV, depois da HTA. (25)

De acordo com o *Global Burden Disease 2015 Tobacco Collaborators*, a prevalência mundial de consumo regular de tabaco foi de 25% no género masculino e de 5,4% para o género feminino. Em 2015, estimou-se que 11,5% da mortalidade mundial possa ser atribuída ao tabagismo, atingindo 6,4 milhões de indivíduos. 52,2% desta mortalidade ocorreu apenas em 4 países: China, Índia, Estados Unidos da América e Federação Russa. (26) A China apresenta o maior número de consumidores de tabaco com 288 milhões do género masculino e 13 milhões do género feminino, seguida pela Índia com 197 milhões do género masculino e 78 milhões do género feminino. A prevalência mais elevada foi observada no Bangladesh (43%) e a mais reduzida no Panamá e na Nigéria (6%). A prevalência continua a ser mais elevada no género masculino do que no género feminino, com cerca de 12 países apresentando prevalências superiores a 40% no género masculino: Indonésia 67%, Federação Russa 61%, Bangladesh 58%, China 53%, Grécia 51%, Ucrânia 50%, Filipinas 49%, Turquia 48%, Índia 48%, Vietname 48%, Tailândia 46% e Malásia 44%. O Egito apresenta a maior diferença associada ao género na relação de consumidores, com uma relação de 61,1 maior para o género masculino. Seis países apresentaram prevalências para o género feminino superiores a 20%: Bangladesh 29%, Grécia 26%, Polónia 24%, Rússia com 22% e a Índia e Uruguai ambos com 20%. (27)

Dos fumadores de tabaco, cerca de 50% são considerados fumadores pesados por consumirem mais de 20 cigarros/dia. À exceção da Índia, o predomínio de fumadores pesados é no género masculino. No género masculino, o número médio de cigarros fumados por dia variou de 6 na Índia para 21 na Grécia. No género feminino variou de 7 nas Filipinas e na Índia para 17 na Grécia. (27)

I.2.5.1 TABAGISMO COMO FATOR DE RISCO PARA DOENÇA CORONÁRIA

Em vários estudos observacionais e prospectivos constata-se que os fumadores têm, em média, um risco relativo de DC de 1,6 a 3 vezes superior ao risco de um não fumador. O risco associado de DC aumenta quer com a exposição recorrente e a quantidade de tabaco por unidade de tempo, quer com a quantidade de tempo em que ocorre o consumo. É um fator de risco particularmente importante para indivíduos jovens (menos de 50 anos), quer do género masculino, quer do género feminino, e em mulheres que utilizam contraceptivos orais.

O risco para DC associado ao tabagismo é independente de outros fatores de risco tradicionais, mas também reflete o seu efeito no aumento da prevalência da HTA, redução da tolerância ao esforço, efeito pró-trombótico, redução do HDL e disfunção endotelial, efeitos bem conhecidos associados ao tabaco. (28, 29)

O risco para a DC associado ao tabaco é independente da sua forma de utilização, mantendo-se o risco no tabaco de mascar, enrolar e os chamados cigarros eletrónicos.

Estudos observacionais demonstram a redução do risco de DC com a cessação tabágica. Mesmo após o diagnóstico de doença CV, a cessação tabágica é importante, uma vez que está acompanhada da redução de cerca de 50% do risco potencial para enfarte, morte súbita e mortalidade de qualquer causa. (30) Contudo, é referida uma grande variabilidade na velocidade do decréscimo do risco de DC. Para que o risco de DC associado ao tabaco fique equiparado ao do não fumador, a maioria dos estudos refere, que são necessários pelo menos 5 anos de cessação tabágica. (31, 32)

I.2.6 OUTROS FATORES DE RISCO PARA DOENÇA CORONÁRIA

A identificação dos fatores de risco CV convencionais é essencial para otimizar as estratégias de prevenção primária e secundária com vista à redução do risco CV.

Assim, constata-se que existe risco CV residual associado a outros fatores de risco que, entretanto, foram identificados. Constituem um grupo heterogéneo, constituído quer por influências ambientais, tais como o consumo de álcool e o nível de atividade física, quer por marcadores doseados no plasma como a proteína C-reativa, a homocisteína e o fibrinogénio.

Também marcadores da idade vascular, como a velocidade da onda de pulso têm sido apontados como potenciais fatores de risco cardiovascular independentes dos FRT.

I.2.6.1 NÍVEL DE ATIVIDADE FÍSICA/SEDENTARISMO

De acordo com dados da OMS o sedentarismo é ainda uma das dez causas mais significativas de mortalidade global. Quando a atividade física (de moderada intensidade) é inferior a 150 minutos por semana, o risco de mortalidade por qualquer causa aumenta de 20 a 30%. (33)

O nível de atividade física tem vindo a reduzir-se na maioria dos países industrializados como consequência da urbanização. Assim, o sedentarismo constitui uma alteração do estilo de vida, inerente à vida no meio citadino, onde a grande parte dos seus habitantes transitou de atividades maioritariamente agrícolas e industriais, para as atividades relacionadas com a prestação de serviços. Associado ao sedentarismo, há a mudança da dieta associada ao meio urbano, composta por uma elevada quantidade de comida processada de elevado componente calórico e de gorduras saturadas. Assim, o sedentarismo não pode ser dissociado do excesso ponderal e da síndrome metabólica.

De acordo com dados de 142 países coletados pelo grupo *Lancet Physical Activity Series*, 31,1% dos adultos não são fisicamente ativos, 33,9% do género feminino e 27,9% do género masculino. Os valores mais baixos de sedentarismo situam-se no Sudoeste Asiático (17%) e os mais elevados no continente Americano e no Este do Mediterrâneo com 43% de adultos inativos. (34) Nesta avaliação, fisicamente ativo significou: atividade física mais do que ligeira de pelo menos 20 minutos, 3 vezes por semana, ou 600 equivalentes metabólicos por semana. Sedentarismo foi quantificado como estando sentado 4 ou mais horas por dia. O sedentarismo aumentou com a idade, foi mais elevado na mulher do que no homem, e mais elevado nos países com maior rendimento *per capita*. Nesta análise, 41,5% dos adultos a nível mundial preencheram os critérios para sedentarismo. Este é mais elevado na Europa (64%) e na América (55%). Nos adolescentes norte americanos cerca de 80,3% pratica menos do que 60 minutos de atividade física moderada e vigorosa /dia, sendo que os rapazes são mais ativos que as raparigas. (34)

Foi proposto que a definição de atividade física seja feita de acordo com o gasto energético despendido na atividade. Assim, é considerada sedentária, qualquer atividade que envolva um gasto energético $\leq 1,5$ equivalente metabólico. (35)

I.2.6.1.1 SEDENTARISMO COMO FATOR DE RISCO PARA DOENÇA CORONÁRIA

A redução da atividade física tem uma relação inversa com o risco de DC. De igual modo, níveis moderados de atividade física estão associados com a redução significativa na incidência de DC. Várias meta-análises recentes avaliaram prospetivamente a associação do comportamento sedentário com os eventos CV. (36-38)

Numa meta-análise para avaliação prospetiva da incidência de DC de acordo com níveis de atividade física, os participantes com mais de 150 minutos/semana de atividade física moderada ou intensa apresentaram uma redução do risco relativo para a incidência de DC de 0,86. Dados do Grupo *European Prospective Investigation Into Cancer and Nutrition* demonstraram que, cada hora adicional de visualização de televisão, esteve associada a um HR de 1,06 para doença CV incidente fatal e não fatal e de 1,08 para DC após ajuste para múltiplas co-variáveis. Nesta análise, o IMC demonstrou ter um papel menos significativo neste aumento do risco para DC. (39)

Existe alguma evidência de que a obesidade e o sedentarismo possam ser parcialmente geneticamente herdados. Vários genes foram testados neste contexto dando resultados ainda pouco esclarecedores. A relação entre o componente genético ligado à obesidade e risco para DC será abordada adiante nesta tese.

I.2.6.2 PROTEÍNA C-REATIVA

A proteína C-Reativa (PCR) é um marcador da inflamação sistémica e um mediador da aterotrombose. A inflamação está presente em todos os estádios da aterosclerose e é particularmente crítica na estabilidade da placa aterosclerótica. A PCR, após ligação aos LDL circulantes, promove a sua captação pelos macrófagos. Esta etapa é tida como crucial na aterosclerose.

Por um lado, a concentração plasmática da PCR sobe consideravelmente após qualquer surto infeccioso ou lesão de órgão, por outro lado fora dos eventos agudos, a sua concentração plasmática é estável e está dentro dos limites da normalidade. Recentemente, a disponibilidade de ensaios de alta sensibilidade, para detetar valores baixos e subnormais, dentro dos limites da normalidade da PCR, permitiu a sua avaliação como fator de risco para eventos aterotrombóticos.

Da meta-análise do *Emerging Risk Factors Collaboration* (40), com dados de 54 estudos prospetivos com 160 309 indivíduos, a PCR demonstrou o aumento do risco relativo de 1,63 para DC por cada aumento do desvio-padrão (DP) do logaritmo da PCR quando ajustado para a idade e género. Após ajuste para os FRT, e apesar da conhecida relação linear com outros fatores de risco CV nomeadamente o fibrinogénio, o tabagismo e os lípidos, a PCR manteve-se como fator de risco independente (RR=1,37) para DC. Também foi encontrada uma relação positiva entre a PCR e o risco para a mortalidade vascular (RR=1,55) e não vascular (RR=1,54). (40)

I.2.6.3 VELOCIDADE DE ONDA DE PULSO

A pressão de pulso determinada pela rigidez arterial e calculada pela diferença entre a PAS e a PAD está associada ao aumento do risco CV. (5)

A rigidez arterial está associada a alterações adversas no fluxo coronário, que resultam numa onda de pressão arterial refletida precoce. Esta onda precoce condiciona elevação da PAS e redução da PAD, e consequente aumento da pressão de pulso central. A elevação da PAS condiciona o aumento da pós carga, e a redução da PAD reduz a perfusão coronária, cujo efeito pode propiciar eventos CV. (41)

Existem várias metodologias que tentam quantificar a rigidez arterial. Estas incluem cálculos simples como: a pressão de pulso (PAS-PAD), o índice de rigidez arterial ambulatória que avalia as mudanças relativas da PAD numa curva de regressão da PAS nas 24h, e o índice de aumentação (magnitude do aumento da curva de pressão sistólica na aorta em relação à curva de pressão refletida). De todos os cálculos, a velocidade da onda de pulso carotídeo-femoral (VOP) é considerado o método não invasivo *gold standard* para a avaliação da rigidez arterial. (42,43) Um valor de VOP >10 m/s está associado ao aumento do risco cardiovascular. (42)

I.2.6.3.1 VOP E RISCO CARDIOVASCULAR

Múltiplos estudos epidemiológicos correlacionaram vários marcadores da rigidez arterial como: a VOP, o índice de aumentação, a pressão de pulso e o risco de eventos CV fatais e não fatais. De igual modo, em vários estudos com doentes controlados adequadamente com fármacos anti-hipertensores, o risco CV residual apresentou correlação com os marcadores da rigidez arterial.

Numa meta-análise, que incluiu 15 877 indivíduos seguidos em média durante 7,7 anos, observou-se uma relação linear entre a VOP e eventos CV. (44) O risco relativo de mortalidade CV e mortalidade de qualquer causa foi de 2,26 no tercil mais elevado, 2,02 no tercil intermédio e 1,9 no tercil inferior. Um aumento de 1m/s de VOP correspondeu a um aumento ajustado do risco relativo para eventos CV de 14%, de mortalidade CV de 15% e de mortalidade de qualquer causa de 15%. (44)

I.2.6.4 CONSUMO DE ÁLCOOL

De acordo com o *Global Status Report on Alcohol and Health 2014* reportado pela OMS (45), em 2010, o consumo médio anual de álcool foi de 6,2 litros por pessoa com mais de 15 anos, o que se traduz em cerca de 13,5 gramas de álcool por dia. Cerca de um quarto do consumo refere-se a álcool não controlado, correspondendo na sua maioria à produção caseira de bebidas alcoólicas. Em 2010, 61,7% da população mundial era não consumidora ou ex-consumidora com mais de 12 meses de abstinência.

O consumo de álcool continua a ser mais prevalente no género masculino (50%) do que no género feminino (32%) e nestes, é responsável por 7,6% da mortalidade total versus 4 % da mortalidade no género feminino. Como regra geral, o consumo de álcool é tanto maior quanto a riqueza económica ou rendimento *per capita*

do país, sendo mais elevado no continente europeu e América do Norte. Dentro do continente europeu a Europa de Leste (Rússia, Ucrânia, Polónia, Hungria e ex-Jugoslávia) e Portugal apresentam os consumos mais elevados, superiores 12,5 L por ano e por pessoa. Também, consoante aumenta o rendimento disponível aumenta a percentagem de consumidores episódicos (consumidor social do fim de semana) e reduz-se o número de abstinente de longa duração. Neste campo, Portugal, França, Reino Unido e países nórdicos apresentaram mais de 30% de consumidores episódicos esporádicos, definido como pelo menos um consumo superior a 60 gr por episódio no último mês.

I.2.6.4.1 CONSUMO DE ÁLCOOL COMO FATOR DE RISCO PARA DOENÇA CORONÁRIA

O consumo de álcool apresenta uma relação complexa com o risco CV global, quando consumido em níveis ligeiros. O álcool confere um efeito protetor para eventos coronários nomeadamente EAM, mas para graus mais elevados de consumo, este confere um aumento significativo do risco de AVC. (46)

A relação, entre o consumo regular de álcool e o aparecimento de aterosclerose, obedece a uma curva em forma de U. Os consumidores ligeiros de álcool (<50 gr/dia) têm um risco mais baixo que os extremos da curva que correspondem aos consumidores de elevado grau (>100gr/dia) dum lado e aos abstinente do outro. (46) Ao consumo de álcool menos que uma vez por semana (consumidor ocasional) não se associou qualquer aumento de risco para aterosclerose. (47)

O efeito protetor do consumo ligeiro de álcool (<100 gr/semana) parece ser mediado pela inibição parcial dos efeitos deletérios das LDL, embora não haja relação com os valores medidos de LDL no plasma. Outros mecanismos plausíveis para os efeitos cardioprotetores incluem ainda a elevação das HDL, a redução da pressão arterial, redução da trombogénese e da agregação plaquetária. (48)

Do mesmo modo, num estudo multicêntrico envolvendo 52 países, 12 195 casos e 15 583 controlos, níveis ligeiros de consumo de álcool foram considerados protetores para a ocorrência de EAM no seguimento, com um OR associado de 0,87. (49)

I.2.6.5 HOMOCISTEÍNA

A hiperhomocisteinémia (HHcy) é uma condição médica definida como a elevação sérica dos níveis de homocisteína (Hcy) superior a 15 µmol/L. Para valores entre 16-30 µmol/L a HHcy é considerada ligeira, entre 31-100 µmol/L é considerada moderada e para valores superiores a 100 µmol/L a HHcy é considerada severa. (50)

Os principais determinantes da concentração da Hcy plasmática são: fatores genéticos, nutricionais, hormonais, uso de drogas e estado de oxidação-redução. O gênero masculino apresenta valores plasmáticos cerca de 21% mais elevados que o gênero feminino. A idade também está associada à elevação dos níveis plasmáticos de homocisteína. A diminuição dos estrógenos na menopausa, aumenta os níveis de Hcy que se atenua com a utilização da terapêutica hormonal de substituição. Várias drogas, nomeadamente o metotrexato, anticonvulsivantes e a teofilina aumentam os valores de Hcy. As duas principais mutações associadas à elevação da Hcy plasmática ocorrem em duas distintas vias de metabolização da Hcy, a transulfuração e a transmetilação. Na presença de estados pró-oxidativos a via da desmetilação é inibida e favorecida a transulfuração da Hcy para aumentar a produção de cistationina e glutatião. (Figura 3)

A mutação no gene da cistationina beta sintetase (C β S) (enzima catabolizadora de 70% da Hcy na presença de dieta rica em proteínas contendo metionina) ocorre no cromossoma 21 (21q22.3). Estão-lhe associadas pelo menos 30 mutações pontuais que condicionam menor atividade catabolizadora da Hcy por via da transulfuração. A mutação C β S ocorre em 1% da população geral e em 30-40% dos doentes com doença vascular precoce. (52)

A mutação no gene da metileno-tetraidrofolato redutase (MTHFR) ocorre no cromossoma 1p36.3. Apresenta pelo menos 10 mutações, sendo que a mais significativa condiciona uma forma termolábil (C677T) com perda de mais de 60% da atividade enzimática da enzima catabolizadora da Hcy em metionina. Cerca de 8% da população em geral é homozigótica para a variante termolábil e nestes, os níveis plasmáticos da Hcy estão dependentes da concentração do ácido fólico. Assim, regra geral, as hiperhomocisteinémias graves só ocorrem com deficiência de folatos plasmáticos <15,4 nmol/l. (53)

A HHcy severa ocorre, na maioria das vezes, em homozigotos para a deficiência de C β S e em apenas 5%-10% para deficiência da MTHFR, enquanto que HHcy moderada ocorre em heterozigotos para a mutação da C β S e de MTHFR na presença de deficiência de folato e, ou com o uso de drogas. (54)

A acumulação da Hcy e dos seus derivados oxidativos parece propiciar progressão da aterosclerose por via da lipoperoxidação lipídica com lesão da parede vascular e subsequente proliferação de células musculares lisas. Por outro lado, existe evidência crescente que a ativação plaquetária e disfunção endotelial estão na etiopatogénese de muitos eventos pro-trombóticos ligados à Hcy.

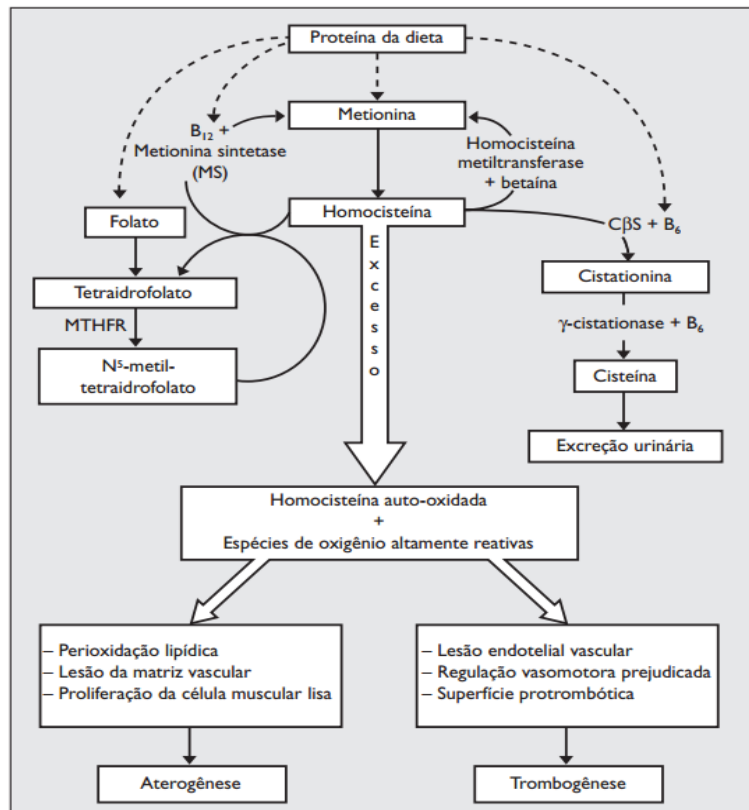


Figura 3: Vias de metabolização habitual da homocisteína e relação com a doença aterotrombótica. (Adaptado de: Neves *et al.*, 2004). (51) Legenda: *MTHFR*- metileno-tetraidrofolato redutase; *CBS* - cistationina β -sintetase

1.2.6.5.1 HOMOCISTEÍNA E RISCO CARDIOVASCULAR

Tem sido observada uma associação entre os níveis de Hcy plasmática e o risco CV. Contudo, a concentração plasmática de Hcy é influenciada por um grande número de fatores genéticos, nutricionais e de estilo de vida, que limitam a sua utilização como fator de risco CV. A ausência de informação nalguns estudos sobre a influência dos níveis de Hcy e os níveis de ácido fólico nos resultados têm sido apontados como dificuldades à interpretação desses dados. De outro modo, numa revisão recente da *Cochrane* a suplementação da dieta com vista à redução da Hcy com ácido fólico e, ou vitamina B6, B9 ou B12 não se traduziu na redução de risco de eventos CV. (55)

Várias meta-análises (56-58) evidenciam a associação entre o genótipo TT para a MTHFR e risco para DC e EAM em diferentes grupos populacionais. A meta-análise mais recente que envolveu 6 912 indivíduos demonstrou que o genótipo TT associou-se com DC prematura (OR=1,19); no subgrupo com Hcy >15 $\mu\text{mol/L}$ a associação do genótipo TT com DC prematura foi ainda mais elevada (OR=2,50). (58)

I.2.6.6 HISTÓRIA FAMILIAR DE DOENÇA CORONÁRIA

I.2.6.6.1 ESTUDOS GENÉTICOS EM FAMÍLIAS

Para além dos FRT existe um claro padrão de hereditariedade na DC, uma vez que, as fratrias de pais com doença têm risco superior de incidência da mesma, quando comparados com a população em geral. (59) Estudos realizados em pais e em populações de gémeos indicam que a história familiar de DC é um fator de risco independente para o aparecimento da mesma. (60-62) Este *Clustering* familiar salienta a importância do componente genético no risco CV, embora também os fatores ambientais e comportamentais exibam muitas vezes um padrão familiar. Assim, nos estudos genéticos familiares clássicos estimou-se que 40-60% da DC está associada a fatores hereditários. (62)

I.2.6.6.2 HISTÓRIA FAMILIAR DE DC PRECOCE E RISCO PARA DC

Vários estudos demonstraram que a história familiar de doença coronária precoce (HF de DC precoce) é um fator de risco independente para eventos CV. Embora sejam aceites outras definições, é usual definir a HF de DC precoce como a presença de DC em familiar do género masculino, antes dos 55 anos (pai ou irmão) e em familiar do género feminino antes dos 65 anos (mãe ou irmã). (60)

No estudo caso controlo internacional europeu INTERHEART demonstrou-se que: o risco de EAM aumenta cerca de 1,67 vezes na presença de história paterna de EAM depois dos 50 anos; aumenta 2,36 se existe antecedentes de EAM paterno antes dos 50 anos; e aumenta 6,56 na presença de EAM precoce em ambos os progenitores. (63)

Numa análise do estudo Framingham, a HF de DC precoce foi associada com o aumento do risco, a 8 anos, de duas vezes para o género masculino e de 1,7 vezes para o género feminino. (60) O risco acrescido, conferido pela HF de DC precoce, mantém-se elevado ao longo da vida.

Assim, numa meta-análise de 49 255 indivíduos do género masculino seguidos prospetivamente, a história de DC precoce (definida como <50 anos) aumentou em 50% o risco para DC e mortalidade CV, respetivamente de 8,9% para 13,7% e de 14,1% para 21%. (64)

I.3 GENÉTICA COMPLEXA DA DOENÇA CORONÁRIA

I.3.1 ORIGEM DOS CONCEITOS DE GENÉTICA E LEIS DE MENDEL

Gregor Mendel, no século XIX, ao observar a penetrância de características fenotípicas em linhagens da ervilha *Pisum sativum* lançou os dois fundamentos da genética. A primeira lei da segregação ou da pureza dos gametas, segundo a qual cada característica herdada é determinada por dois alelos. Cada ser tem dois alelos, mas transmite apenas um alelo ao seu descendente, pois estes separam-se na formação dos gametas. A segunda lei ou da segregação independente dos genes, em que, os alelos de um gene se separam, independentemente dos alelos do outro gene, na formação dos gametas e combinam-se ao acaso na descendência.

Este padrão de hereditariedade foi posteriormente observado em muitas doenças, nascendo assim o conceito de doença Mendeliana. Até à data, as bases genéticas de mais de 2000 doenças Mendelianas estão descritas, tais como: a hipercolesterolemia familiar (65), a deficiência familiar de Apolipoproteína B-100 (66) e o Síndrome de Brugada. (67) Regra geral, este tipo de doenças tem uma baixa incidência, afetando um número reduzido de indivíduos e, por isso, não têm grande impacto na saúde pública da população.

Posteriormente, Bateson, Saunders e Punnet observaram precocemente exceções à lei de Mendel, referindo que alguns fenótipos eram observados mais frequentemente do que a proporção determinada pela lei de Mendel. Com Morgan, surgiu o pressuposto de que alguns alelos devem estar acoplados ou em *linkage*, sendo herdados de modo preferencial, ou seja, em desequilíbrio de *linkage*. Este princípio primordial continua a guiar os estudos na era da genética moderna.

Para além da observação de certas características que eram herdadas em conjunto, Morgan e os seus colegas também definiram o primeiro mapa genético de um organismo (*Drosophila melanogaster*) e deduziram que estes traços herdados em bloco estavam proximamente localizados nos cromossomas. (68) Por fim, surgiu com estes autores a ideia que algumas características fenotípicas daquela espécie eram determinadas pela ação combinada de dois ou mais genes, definindo-se assim o tipo de hereditariedade observada nas chamadas DOENÇAS COMPLEXAS.

A DC representa um paradigma da doença poligénica complexa, influenciada por múltiplas combinações de interações gene-gene (G-G) e gene-ambiente (G-A). (69, 70) Muitas investigações têm sido conduzidas no sentido de identificar os fatores de risco adquiridos e hereditários para DC. A complexa combinação de vários fatores de risco torna a identificação do componente genético das doenças complexas muito mais difícil do que para as doenças Mendelianas clássicas.

I.3.2 ADN COMO VEÍCULO DA INFORMAÇÃO GENÉTICA

Em 1944, após anos de estudo duma bactéria nossa contemporânea, o pneumococo, foram publicadas no *Journal of Experimental Medicine* as conclusões de Oswald Avery e da sua equipa que identificou um ácido nucleico, então denominado ADN, como base bioquímica da transmissão de características genéticas. (71) Nessa altura, algumas dúvidas permaneceram sobre como seria contida a informação genética no ADN, julgando-se que o efeito genético do ADN estaria sob o controlo direto das proteínas. Surgiu, então, o objetivo crucial de identificar e caracterizar a molécula de ADN de modo a determinar como seria armazenado e transmitido o património genético.

Com a descrição da estrutura em dupla hélix do ADN por Watson and Crick em 1953 (72), generalizou-se o conceito de *um gene uma enzima*. Esta corrente estava centrada na ideia que os genes contidos no ADN codificavam proteínas, que por sua vez tinham ações moleculares conhecidas.

O “desvendar do genoma” propriamente dito, com a descoberta das características estruturais do ADN só viria a ser, tecnicamente possível, na década de 50 com a identificação por Ochoa e Kornberg da enzima PNPase (polinucleotideo-fosforilase) e da sua capacidade de sintetizar ácido ribonucleico (ARN) *in vitro*. A enzimologia do ADN desenvolveu-se mais rapidamente com Arthur Kornberg (galardoado Nobel) e a sua equipa, que estudaram em profundidade a enzima ARN polimerase como responsável do processo de transcrição do ADN. O conhecimento da função desta enzima permitiu o desenvolvimento de variadas técnicas laboratoriais, muitas das quais utilizadas hoje em dia para a identificação de variantes genéticas como a *Polimerase Chain Reaction* (PCR) descrita por Mullis em 1986. (73)

I.3.3 O GENOMA DESCODIFICADO

Na década de 90, iniciou-se o projeto ambicioso denominado Projeto do Genoma Humano (PGH) que, durante mais de uma década, perseguiu o objetivo de sequenciar a totalidade dos pares-bases de nucleótidos do ADN e mapear todos os genes do genoma humano, quer do ponto de vista físico (mapa da localização dos genes nos cromossomas) como funcional (proteínas codificadas pelos genes e respetiva função biológica). Finalmente, em junho de 2000, Craig Venter e Francis Collins anunciaram, na Casa Branca, o primeiro rascunho do genoma humano que ficaria concluído com as respetivas anotações em 2003. (74)

Os genomas variam grandemente em tamanho, sendo que o mais pequeno para um microrganismo vivo (bactéria) contém cerca de 600,000 pares-bases de ADN, enquanto que as células humanas têm cerca de 3

bilhões de pares-bases distribuídas por 23 cromossomas. Cada cromossoma, moléculas fisicamente separadas que diferem em tamanho desde 50 a 250 milhões pares-bases, contem muitos genes que são as unidades básicas funcionais e físicas da hereditariedade.

Em 2003, era dado início a outro projeto fundamental na área da Genética, o denominado *Projeto Internacional HapMap*, que permitiu aprofundar o conhecimento sobre a variação genética humana, dando melhores estimativas sobre as frequências alélicas. (75) Uma vez que os cromossomas existem aos pares, cada ser humano tem dois alelos para cada gene, um em cada par de cromossomas, herdado de cada progenitor. Os dois alelos herdados dos pais podem ser iguais (homozigotia), ou diferentes (heterozigotia).

Do projeto *HapMap* retiraram-se novas variantes com significado clínico e definiram-se novos padrões de relação entre as variantes como por exemplo o desequilíbrio de *linkage* (DL). Estas variantes em *linkage* são herdadas em bloco como resultado da recombinação cromossômica presente ao longo da história da nossa espécie. Ao cruzar diferentes padrões de DL ao longo do genoma humano, o projeto *HapMap* conseguiu chegar a um número relativamente reduzido de *Single Nucleotide Polymorphisms* (SNP) chamados de *Tags* que funcionam como marcadores, tornando mais rápida a genotipagem da maioria das variantes genéticas. (76)

I.3.4 TIPOS DE VARIANTES GENÉTICAS

As variantes genéticas classificam-se consoante a quantidade de material genético que se encontra alterado que podem ir desde variações no número de cromossomas (numéricas), variações na estrutura dos cromossomas (estruturais) e subcromossômicas afetando pequenas quantidades de informação genética mutada de indivíduo para indivíduo. Quanto maior a quantidade de informação genética alterada, maior é o impacto no fenótipo do indivíduo e menos frequente é essa variante genética na população em geral afetando, por isso, um menor número de pessoas. No outro extremo, as variantes genéticas mais comuns representadas apenas pela troca de um elemento do par base (SNP) são as mais frequentes na população sem, *per si*, condicionarem alterações no fenótipo. (Figura 4)

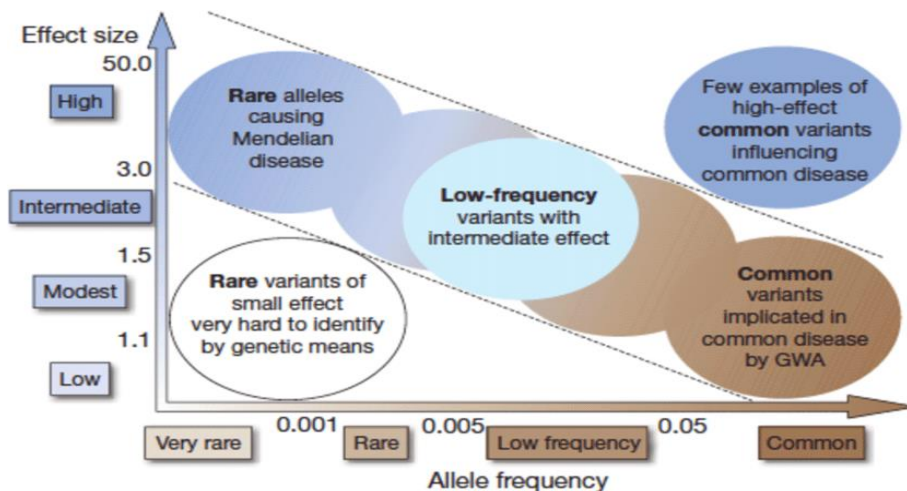


Figura 4: Tipo de variantes genéticas e risco associado. (Manolio *et al.*, 2009). (77) Relação inversa entre a frequência das variantes genéticas na população e o risco de doença conferido pelas mesmas. Alelos raros conferem risco mais elevado e associam-se a doenças do tipo Mendeliano, alelos comuns muito frequentes na população conferem baixo risco individual e associam-se a doenças comuns como a DC.

1.3.5 GENES CANDIDATOS E GWAS: LIMITAÇÕES

Com a melhoria do conhecimento sobre os padrões de hereditariedade dos genes em diferentes populações, tentou usar-se essa informação para localizar genes associados com doenças comuns como a aterosclerose, a HTA, e a DM.

A fase inicial do estudo genético denominada de *genes candidatos* partia da fisiopatologia da doença e tentava encontrar o gene candidato em alvos teoricamente relacionados com os eixos fisiopatológicos da doença. Dos mais consistentes encontrados ao nível destes estudos salientam-se os genes relacionados com o metabolismo lipídico, com o eixo renina-angiotensina-aldosterona, com o *stress* oxidativo, com a hemostase e trombogénese, com a inflamação e com a disfunção endotelial. Assim, um grande número de variantes genéticas foi estudado nestas duas últimas décadas.

Para além de falhas em estudos de replicação foi apontada também outra limitação metodológica aos genes candidatos. Uma vez que a seleção dos *loci* a serem estudados era feita com base no seu conhecimento fisiopatológico, vários *loci* com mecanismo desconhecido, foram à partida potencialmente perdidos da análise de associação real com a doença.

Apesar dos resultados controversos relacionados com diferenças entre populações e o tamanho das amostras, estes estudos têm permitido seguir muitas pistas para explorar os mecanismos subjacentes à

etiopatogênese da aterosclerose e melhorar os nossos conhecimentos sobre a difícil interligação de várias vias que culminam na manifestação da DC.

Com a possibilidade de genotipagem em grande escala, imputação e sequenciação de segunda geração tornou-se possível testar a associação de centenas de milhares de variantes genéticas com a doença aterosclerótica nos denominados estudos de associação genômica de larga escala ou *Genome-wide association studies* (GWAS). A maioria destes estudos foram consórcios multicêntricos com extensa partilha de dados, vindo a aumentar de forma significativa o número de variantes genéticas associadas com múltiplas patologias. Estes consórcios foram essenciais para a integração da informação genética em meta-análises que confirmaram muitas associações já identificadas previamente em populações mais pequenas. Além da integração da informação genética, permitiram assinalar associações falsas positivas e, conferiram maior poder estatístico para a identificação de novas variantes associadas com a DC. As duas meta-análises *CARDIoGRAM* (76) e *C4D* (78) são consideradas como as mais robustas e têm sido suportadas por vários resultados consistentes.

O consórcio *CARDIoGRAM* em 2011 realizou uma meta-análise de 14 estudos GWAS que incluiu 22,233 casos e 64,762 controlos de origem europeia. (76) Foram replicados os achados encontrados em 60,738 amostras adicionais. No global, foram incluídos dados de 143 677 indivíduos de origem europeia. Neste estudo, foram confirmados 10 SNPs previamente conhecidos, cujas frequências dos alelos de risco foram relativamente comuns (entre 15-87%), com exceção do polimorfismo, pouco frequente, da Lipoproteína (a). Foram identificados 13 novos SNPs com associação para DC e frequências do alelo de risco entre 13 e 91% da população. Nesta meta-análise, os OR para DC variaram entre 1,06 e 1,20 e apenas 3 destes novos *loci* estavam associados com pelo menos um fator de risco tradicional. (76)

No mesmo ano, o consórcio genético *C4D* realizou uma meta-análise de quatro estudos GWAS com 15 420 casos e 15 062 controlos e com replicação dos achados em 21,408 casos e 19,185 controlos, atingindo um total de 71,075 indivíduos de origem europeia e sul asiáticos. (78) Onze variantes associadas com a DC na população inicial e 5 novos *loci* identificados na população global (inicial e de replicação) atingiram o nível de significância imposto nos GWAS (5×10^{-8}). Este nível de significância permite reduzir a possibilidade dos achados serem apenas encontrados por acaso, que ocorre quando se fazem análises com elevado número de amostras (erro de amostragem múltipla). Neste estudo, as frequências dos alelos de risco variaram entre 32 e 80% e associaram-se a risco modesto de DC com OR entre 1,05-1,10 por alelo de risco. (78)

Com o advento dos estudos GWAS, muitas variantes genéticas têm sido associadas com a DC. Destas, a variante genética identificada no cromossoma 9p21.3 tem sido a mais consistente. (79)

Nesta associação, que se estende ao longo de uma região superior a 100 kilobases (Kb), na qual foram estudados vários polimorfismos foi identificado o rs 1333049 que apresentou a associação mais forte com a

DC. A zona de interesse é atravessada, quase completamente, por um gene recentemente identificado, não codificante, denominado ANRIL. Os genes mais próximos que codificam proteínas são os CDKN2A, CDKN2B e MTAP). O CDKN2A codifica a proteína p16INK4a e o CDKN2B codifica a p15INK4B, sendo a P14 ARF codificada por exões alternativos do gene CDKN2A e do gene MTAP, situado também nesta zona. Destas variantes genéticas, o CDKN2A e o CDKN2B, são considerados supressores tumorais e codificam proteínas inibidoras do ciclo celular, as quinases ciclina-dependentes (CDKs). Estas desempenham um papel importante na regulação do ciclo celular, podendo as suas mutações estar implicadas na patogénese da arterosclerose e de certos tumores malignos como o melanoma. (80,81)

Em 2013, o número previamente estabelecido de 34 SNPs consistentemente associados com DC com base nas meta-análises anteriores, subiu para 49 SNPs com a primeira meta-análise conjunta (*CARDIoGRAM E C4D*) de 63,746 casos e de 130,681 controlos. Nesta meta-análise, 15 novos *loci* foram associados com a DC e identificadas 104 novas variantes, com possível associação com DC. (82)

Estas variantes genéticas, encontradas nos GWAS, colocam vários desafios. (83) Por um lado, trazem múltiplas possibilidades de conhecer novas vias implicadas na aterosclerose, por outro, têm sido associados a múltiplos erros que condicionam a sua utilidade imediata. (84)

Uma vez que se baseiam em puras associações matemáticas entre milhões de SNPs e o fenótipo de interesse, incorrem no erro das análises múltiplas, ou seja, à medida que o número de análises de associação aumenta, aumenta também o número de falsas associações (*False Discovery Rate*) encontradas apenas por acaso. Os métodos propostos para minorar esta limitação encontram-se em desenvolvimento constante, sendo já clássica a utilização da correção de Bonferroni com valor de P corrigido para 5×10^{-8} (corrigindo para 1000000 testes independentes). (85)

Um sumário de outras limitações técnicas envolve problemas com a estratificação populacional e a utilização de técnicas de imputação para permitir deduzir os *missing data*. Estes consórcios utilizam amostras de populações de zonas geográficas diferentes que podem ter diferentes frequências alélicas, devido à seleção natural, de genótipos mais favoráveis, o denominado desvio genético. Contudo, foram desenvolvidos métodos para minorar as diferenças genéticas entre populações que serão abordados mais adiante nesta tese.

As críticas teóricas aos GWAS salientam a ausência de orientação nestes estudos para a fisiopatologia da aterosclerose e da realidade populacional. A utilização apenas de SNPs, deixando de fora da análise outras formas de material genético como as *copy variant number*, é outra importante consideração.

Desde cedo se percebeu, ser de interesse primordial, demonstrar em populações independentes a associação dos SNPs situados nos *loci* de risco, de forma a se validarem os resultados dos estudos GWAS. Uma vez que se tratam de porções de ADN sem função fisiopatológica conhecida, estas variantes podem vir

a impulsionar novas linhas de investigação, quer de replicação, quer funcionais.

Assim reveste-se de especial importância a replicação destas variantes em populações homogêneas e reais e a realização de estudos funcionais, ou seja, dedicados ao escrutínio da função associada à variante genética.

Uma vez que os GWAS estudam o risco ao nível teórico e não populacional, na era pós GWAS, é de muita relevância a identificação dos mecanismos envolvidos na trama genética desta doença complexa. O estudo das interações destas variantes genéticas com os fatores de risco clássico, e G-G pode revelar-se importante. (86) É possível que alguns destes genes tenham um impacto significativo no risco para DC se as interações G-G e G-A forem estudadas.

Tendo consciência que a tecnologia evoluiu rapidamente e que permite analisar porções cada vez mais extensas do genoma, em 2015, foram publicadas informações tendo por base o estudo integral do genoma de cerca de 1000 indivíduos de diferentes populações (The 1000 Genomes Project Consortium), fornecendo os mapas de haplótipos mais minuciosos conhecidos até hoje. (87)

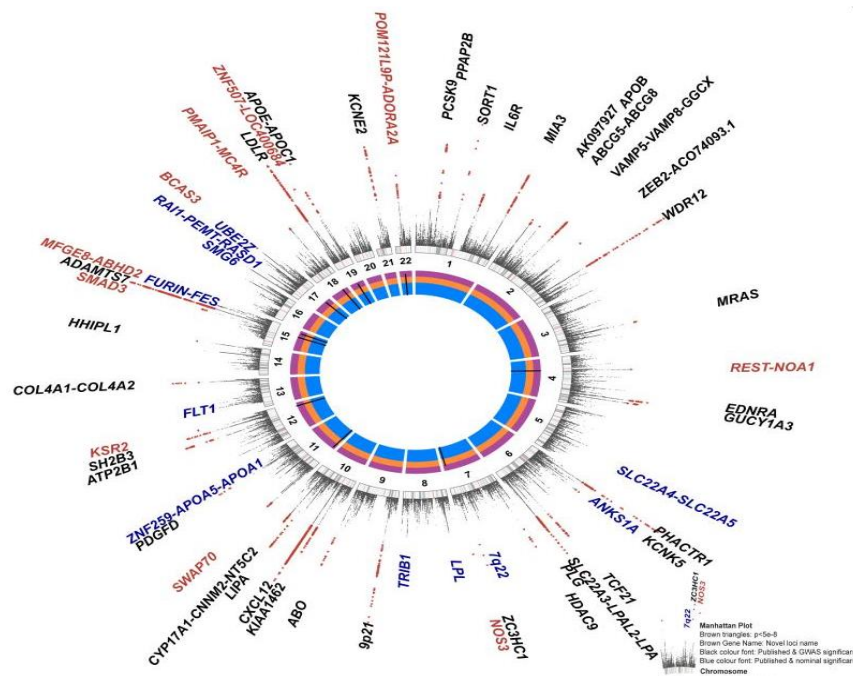


Figura 5: Diagrama dos cromossomas e localização atualizada de variantes genéticas com associação à DC em estudos GWAS. (Nikpay *et al.*, 2015). (87). Legenda: A preto estão as variantes genéticas de estudos GWAS iniciais que se confirmam como associadas a DC ($p < 5 \times 10^{-8}$). A vermelho estão as novas variantes genéticas identificadas em GWAS a partir de 2015 ($p < 5 \times 10^{-8}$). A azul estão variantes genéticas identificadas previamente que nesta meta-análise tiveram significado marginal ($p < 0,05$).

Neste caso, e após a análise de um número cada vez maior de indivíduos, pelas técnicas avançadas de imputação, estão a surgir dados referentes a 185,000 casos e controlos, interrogando cerca de 6,7 milhões de variantes comuns (frequência do alelo *minor* > 0,05) assim como 2,7 milhões de variantes de baixa

frequência ($0,005 < \text{frequência do alelo minor} < 0,05$) associadas com DC. Para além de confirmarem o risco já estabelecido para a maioria das variantes identificadas previamente, 10 novos *loci* foram já identificados e estabelecida a sua relação com os processos biológicos da parede vascular. (87) (Figura 5)

No global, o genoma humano apresenta mais de 19,000 genes que, no seu todo correspondem apenas a 1,5% do genoma humano. Em média, cada um desses genes codifica cerca de 3 proteínas.

As regiões não codificantes, outrora consideradas “lixo” genético, são hoje estudadas com recurso à inovação tecnológica denominada de *Next Generation Sequencing*. Estes estudos têm demonstrado que essas regiões incluem funções integrativas sinalizadoras e reguladoras do ADN génico ou codificante.

A DC representa um paradigma da doença de componente poligénico complexo, influenciada por múltiplas combinações de interações G-G e G-A. (88,89) A carga genética hereditária da DC estima-se rondar os 40-50%, mas só uma pequena parte dos mecanismos genéticos intrínsecos é conhecida.

Embora o número de variantes genéticas associadas à DC ainda não se encontre fechado, e seja provável que continue a aumentar, as variantes identificadas explicam uma proporção próxima a 10% do risco de DC (82). Para explicar este fenómeno, que tem sido denominado de “mistério da hereditariedade escondida”, várias hipóteses têm sido avançadas, com duas linhas de pensamento. Uma em que os cálculos da proporção atribuível ao componente genético foi subestimada e outra, mais recente e com mais seguidores, que estima que fatores genéticos e não genéticos ainda não conhecidos e a sua interação podem explicar o componente ainda por descobrir.

O caminho atual tem sido no sentido de continuar a identificar o maior número de variantes, identificar variantes ainda mais raras que escapem às redes de genotipagem atuais e por fim estudar as interações gene-ambiente. As interações específicas com cada fator de risco começam agora a ser estudadas para cada variante genética (G-G e G-A) e estas podem vir a tornar-se uma importante ferramenta no combate à aterosclerose.

I.4 RISCO GENÉTICO E RISCO CARDIOVASCULAR GLOBAL

I.4.1 MODELOS PREDITORES DE RISCO CV: MODELO DE FRAMINGHAM E MODELO SCORE

Um dos principais objetivos dos estudos genéticos reside no desenvolvimento de testes de ADN, a utilizar na prática clínica para a avaliação do risco genético CV. Esta necessidade advém da ainda incompleta estratificação pelos modelos habituais que deixam uma parte significativa do risco individual por explicar e raramente incorporam a HF de DC precoce.

Por outro lado, o conhecimento científico sobre a contribuição do componente genético para o risco de DC tem vindo a aumentar.

A incorporação da informação genética na estratificação do risco CV deve ser provida de significado clínico, ou seja, deve adicionar conhecimento à estratificação de risco realizada atualmente com os fatores de risco clássicos.

Tradicionalmente têm-se utilizado scores para estratificação do risco de eventos cardio e cerebrovasculares num tempo pré-definido, geralmente a 10 anos. Estes modelos de risco, utilizados na prática clínica atualmente, utilizam combinações de variáveis como a idade, o género, a HTA, a DM, o tabagismo e o colesterol total ou o colesterol HDL.

Embora existam outros, os dois modelos de maior utilidade na prática clínica para avaliação do risco CV são o score de risco a 10 anos de Framingham e o Euro-Score (SCORE) para estimativa do risco de eventos CV fatais a 10 anos, utilizado preferencialmente no continente Europeu.

1.4.1.1 SCORE DE RISCO FRAMINGHAM (SRF)

O primeiro modelo preditor de risco CV conhecido foi desenvolvido em Framingham (90) uma cidade do estado de Massachusetts com uma amostra total de 14,428 homens e mulheres, composta por 3 gerações dos residentes locais. Na coorte inicial foram recrutados apenas indivíduos com mais de 50 anos e sem DC manifesta e foram estudados os fatores associados com o aparecimento de DC. Para além da idade e do género masculino viriam a ser comprovados como risco para DC vários fatores de risco modificáveis como a pressão arterial, o tabagismo, os lípidos e a obesidade, ou seja, os atuais FRT. Por análise multivariada, estes fatores modificáveis foram associados ao risco de DC para eventos fatais – Mortalidade CV – e não fatais – EAM não fatal, angina de peito, insuficiência cardíaca, AVC isquémico e hemorrágico, acidente isquémico transitório e doença arterial periférica – tão cedo como após 6 anos de seguimento. (90) A segunda fase do recrutamento (1971) corresponde à geração descendente da coorte inicial e foram incluídos indivíduos mais jovens. Foi comprovado o risco CV aumentado em indivíduos com história familiar de DC em pai ou irmãos, antes dos 55 anos e em mãe ou irmãs com DC, antes dos 65 anos de idade.

O risco estimado de doença CV a 10 anos inclui risco de diversos eventos como a DC, a doença cerebrovascular, a doença vascular periférica, e a insuficiência cardíaca e, é baseado na ponderação das variáveis de cada doente, pelos coeficientes de regressão (B) da análise de sobrevivência de Cox pelo método de proporção de Hazard para cada fator de risco.

Na apresentação inicial, as variáveis idade, pressão arterial, colesterol e HDL foram introduzidas como variáveis contínuas. Numa fase posterior foram categorizadas, dando origem a tabelas mais facilmente

utilizáveis na prática clínica. Cada categoria corresponde a um determinado valor de pontos que somados dão a estimativa do risco de CV a 10 anos de acordo com tabelas para o género feminino e masculino. Em 1998 é publicado um modelo preditor de risco para DC simplificado, dirigido para a prática clínica baseada em variáveis categóricas em vez de variáveis contínuas. (91)

Há várias críticas ao modelo de score de risco de Framingham. O SRF tem melhor *performance* em populações semelhantes à coorte inicial, o que não é o caso da maioria das populações europeias. Na população norte americana existem diferenças significativas entre o padrão dietético e os níveis de colesterol e tabagismo da década de 90 para a atual. Por fim, a incorporação de *end points* intermédios como: a angina, o Acidente Isquémico Transitório e a Insuficiência Cardíaca, por vezes, tornaram a avaliação do risco para a DC menos precisa.

D' Agostino *et al.* (92) avaliaram o SRF em sete coortes correspondentes a sete etnias diferentes incluindo hispânicos, etnia branca, etnia negra, nativos americanos e americanos japoneses. Nestes, o SRF demonstrou boa capacidade discriminativa com C-estatística variando entre 0,63-0,79 no género masculino e 0,66-0,83 no género feminino. (92) Vários estudos em outras subpopulações incluindo italianas, mediterrâneas e ainda em chineses americanos concluíram que o SRF sobrestima o verdadeiro risco para a DC. (93)

I.4.1.2 EUROSCORE (SCORE)

Em 2003 foi criado um modelo de risco CV com base numa população europeia, tendo em conta a diferença da prevalência da DC no território europeu. O SCORE europeu estima o risco de evento CV fatal a 10 anos, utilizando as seguintes variáveis: idade, género, colesterol total ou a razão colesterol/ HDL e a PAS, que foram categorizadas em tabelas de risco (HEARTSCORE), ajustadas de acordo com o elevado ou baixo risco CV global no país, onde o doente reside. Para cada participante são calculadas quatro probabilidades de acordo com a possibilidade de DC ou não, para a idade real e repetidas estas análises para a idade real com mais 10 anos. A diferença de sobrevida esperada corresponde ao risco de DC fatal a 10 anos de acordo com a fórmula em suplemento. (94)

A computação do EuroSCORE envolveu dados de 12 estudos em diferentes populações europeias incluindo 205,178 participantes dos quais 88,080 mulheres e 117,098 homens com idades compreendidas entre os 45 e os 64 anos de idade. (94)

Uma importante modificação na execução do SCORE foi que a sociedade europeia de cardiologia (ESC) considerou para o cálculo do risco CV fatal a 10 anos apenas a análise de *hard endpoints*, ou seja, eventos fatais, tornando a sua análise mais clara. O SCORE europeu obteve os coeficientes *B* para a ponderação dos

fatores de risco, após análise de sobrevivência de Regressão de Weibull e não de COX *proportional hazards* como exposto para o SRF. (94)

O SCORE demonstrou, no global, boa capacidade discriminatória com C-estatística entre 0,71 e 0,84 para a análise do evento de morte por DC (5 652 eventos) e morte por doença CV fatal (7934 eventos). (94)

I.4.1.3 LIMITAÇÕES DOS MODELOS PREDITORES DE RISCO PARA DC

Os scores de risco convencionais têm-se demonstrado úteis, mas com limitações significativas. Por um lado, é nos 75% da população que é classificada como tendo risco baixo ou intermédio, que ocorrem 60% dos eventos CV. Por outro lado, em subgrupos particulares, como os jovens e o género feminino, a aplicabilidade das habituais tabelas de risco CV parece subvalorizar o risco.

No global estes algoritmos têm um valor preditivo para falsos negativos a 10 anos de cerca de 5%, sugerindo que ainda devem existir outros fatores significativos que não são contemplados nos scores de risco atual (95). Têm havido várias tentativas de adicionar mais fatores de risco clássico aos *scores*, mas com pouco aumento no seu poder preditivo. Ou porque têm relação com fatores já incluídos, ou porque são de baixa prevalência.

A *performance* dos modelos de risco é em geral avaliada segundo três critérios: calibração, discriminação e reclassificação, sendo estes independentes uns dos outros. Um modelo bem calibrado estima corretamente o risco de um grupo de indivíduos e é geralmente expresso pela *area under de curve (AUC)* da *Receiver operating curve (ROC)* e respetivo C-estatística discutidos na metodologia desta tese. A má calibração leva à reduzida acuidade do modelo que poderá ser global ou apenas em certas subcategorias.

Um modelo com boa discriminação hierarquiza corretamente o risco na ordem adequada, comparando-o com o risco global da população. Assim, um modelo pode discriminar bem, mas estar pouco calibrado se nesse modelo uma determinada variável estiver homogeneamente distribuída, mas incorretamente excluída pelo modelo.

A reclassificação representa a percentagem de indivíduos reclassificados corretamente noutra categoria de risco, após a incorporação de um novo modelo, geralmente expressa pelo *net reclassification index (NRI)*.

I.4.2 A INCORPORAÇÃO DO COMPONENTE GENÉTICO NA ESTRATIFICAÇÃO DO RISCO CV

Face a estas limitações, existe necessidade de melhorar a avaliação de risco CV com recurso a novos discriminadores de risco, sendo o risco genético para DC uma forma particularmente atrativa.

A introdução de genótipos, referentes a genes alvo envolvidos na DC nos scores de risco, pode aumentar o

seu poder preditivo e aumentar a sensibilidade. O risco para DC para cada genótipo isolado é muito modesto (OR de 1,05-1,4) pelo que a utilização conjunta de vários genótipos significativos pode melhorar o seu poder preditivo.

Teoricamente, se falarmos para uma população de 100,000 indivíduos com uma incidência da DC a 10 anos, de pelo menos 10%, para elevarmos a curva AUC para 0,90 dos habituais 0,6 necessitaríamos de 100 variantes genéticas com OR de 1,5 e frequências genotípicas de 50%. (96) Do ponto de vista realístico, com as variantes já consagradas como risco para DC com nível de significância 5×10^{-8} e OR de 1,13-1,40 conseguiu-se elevar para 0,63 com apenas 20 variantes e para 0,69 com 50 variantes. No máximo, em populações reais, conseguiu-se demonstrar incrementos positivos até pelo menos 100 variantes com AUC correspondentes de 0,76. (96)

Com a identificação de múltiplos polimorfismos associados com DC pelos estudos GWAS, tornou-se fundamental avaliar o seu impacto cumulativo no risco CV. Vários estudos tentaram, então, avaliar o impacto cumulativo associado a estas múltiplas variantes através do desenvolvimento de scores de risco genético.

1.4.3 ESTUDOS DE INCORPORAÇÃO DO SCORE DE RISCO GENÉTICO NO RISCO CV GLOBAL

O score de risco genético (SRG) tenta sumarizar o risco genético individual que advém dum número significativo de variantes genéticas que vão sendo identificadas.

Ripatti *et al.* (97) avaliaram o peso cumulativo de 13 SNPs associados com DC num estudo prospetivo caso controlo numa população europeia norte-europeia. Neste estudo, o SRG foi construído de acordo com o modelo aditivo, sendo ponderado, pelo efeito de cada alelo descrito previamente na literatura. O quintil mais elevado do SRG foi associado com um risco de eventos de 1,66 (IC95%: 1,35 – 2,04) maior do que o quintil mais baixo após ajuste para a idade, género e FRT. Comparado com os FRT o risco foi menor do que o risco conferido pela elevação dos lípidos e similar em magnitude ao risco conferido pela HTA. Apesar de ser fator de risco independente para DC, o SRG não melhorou de forma significativa a capacidade preditiva dos fatores de risco modificáveis baseado no C- estatística ou no NRI. (97)

Em 2012, Thanassoulis *et al.* (98) estudaram numa subpopulação de 3014 participantes do estudo de Framingham, com seguimento médio de 11 anos, a associação de um score genético alélico com 13 variantes, quer com maior incidência de DC quer com mais cálcio coronário. O SRG associou-se com risco para a doença CV com o HR de 1,07 por cada aumento de alelos de risco, mesmo após ajuste para os FRT e a história familiar. A adição do SRG à idade, género e FRT traduziu-se numa elevação modesta da curva do C-estatística (0,819 para 0,822) e com um NRI 0,17. Curiosamente, após a adição de outras 16 variantes

associadas a risco de DC, o score com 29 SNPs não traduziu uma capacidade preditiva adicional maior do que o SRG com 13 variantes. Finalmente, a adição dos 89 SNPs associados a fatores de risco modificáveis reduziu o impacto do SRG para um HR de 1,01. (98)

Em 2013, Ganna *et al.* (99) investigaram prospectivamente 6 coortes suecas com 10 612 participantes sem DC à data de inclusão no estudo. Nesta população foram testados 2 scores: um SRG dirigido às 46 variantes consagradas nos estudos GWAS e um SRG alargado contendo 395 polimorfismos associados com fenótipos intermediários. Ambos os SRG, dirigido e alargado foram associados com o risco para aparecimento de DC no quarto quartil do SRG associado a um HR de 1,54 quando comparado com o quartil inferior, e melhorou a reclassificação com um NRI de 4,2% e de 4,9%. A melhoria da capacidade discriminativa pelo C index foi de novo modesta. (C-estatística de 0,004). (99)

Em 2015, no estudo de Roterdão, (100) foi avaliado o poder aditivo de 152 polimorfismos na estimativa do risco a 10 anos de DC. Foram comparados 3 SRG ponderados com diferente número de SNPs, 49 SNPs GWAS, 103 SNPs associados com a DC mas sem nível GWAS (ou seja com $p < 0,05$ e não $< 5 \times 10^{-8}$) e um score conjunto com 152 SNPs. Os scores foram associados com a DC independentemente dos FRT e da HF. Por cada aumento do desvio-padrão do SRG, aumentou em 13% o risco de eventos coronários. Todos os scores aumentaram a curva do C- estatística 0,006% para além do género e idade, 0,003 para além dos FRT e 0,003 para além dos FRT e história familiar. Nenhum dos scores aumentou de forma significativa as medidas de reclassificação para eventos coronários, sendo, no entanto, significativa a discriminação e reclassificação para a DC prevalente com um aumento de 0,009 na curva C- estatística. (100)

Finalmente em 2016, Morris *et al.* (101) publicaram dados referentes a sete estudos prospetivos de 11,851 indivíduos com um seguimento médio de 10 anos. A curva AUC foi de 0,635 para a estratificação de risco tradicional que reduziu após a adição do SRG para 0,623. Para um risco CV a 10 anos de 10% o NRI foi de 0,25% após a adição do SRG. Outros estudos em populações caucasianas demonstraram resultados similares.

Se com a adição de mais SNPs associados com fatores de risco para a DC, era expectável uma maior força da associação do SRG com a DC, a inclusão de polimorfismos associados a fatores de risco ou fenótipos intermediários resultou num efeito de diluição, em algumas coortes, por serem absorvidos por outros polimorfismos mais robustos.

No global, em múltiplos estudos em populações caucasianas um grande número de polimorfismos genéticos estiveram associados com o aumento do risco para a DC. Consistente com os modestos efeitos

individuais de cada alelo, os SRG estiveram associados de forma independente e moderada, mostrando a capacidade discriminativa dos seus quartis e quintis superiores versus inferiores à volta dos 30%, atingindo o máximo de 66% apenas num estudo. A melhoria da capacidade preditiva do risco CV, para além dos FRT e da HF de DC precoce, parece, no global, moderada.

I.5 JUSTIFICAÇÃO DO ESTUDO

O racional para este estudo, parte do desconhecimento sobre o componente genético comum ou poligénico da DC na população portuguesa. Embora a DC seja maioritariamente influenciada por FRT e comportamentos, é fundamental aprofundar os conhecimentos sobre a influência da genética na DC e da sua interação com os FRT e fatores comportamentais.

Há vários estudos internacionais, por vezes com resultados contraditórios, sobre polimorfismos genéticos associados a DC. Não se conhecem estudos semelhantes na população portuguesa.

A maioria dos estudos, até agora publicados, focou genes associados com a DC, excluindo genes associados aos FRT. Pensamos que um estudo multilocus, com variantes genéticas associadas aos FRT, ligados aos eixos fisiopatológicos da aterosclerose coronária e da doença coronária *per si* baseadas no GWAS, representará uma mais-valia. Este modelo genético multifatorial vai ao encontro da perceção clínica da DC, como entidade resultante da ação de múltiplos fatores.

Assim, o estudo de genes localizados em eixos fisiopatológicos conhecidos, com mecanismos óbvios, tem maior possibilidade de ter impacto na prática clínica. Conhecendo o seu mecanismo de ação, o estudo destas variantes pode levar a mudanças na terapêutica ou ao aparecimento de novos fármacos.

Por outro lado, os genes emanados dos estudos de associação genómica ligados à DC carecem de replicação nas populações reais e, ainda não foram validados na população portuguesa. O seu valor estatístico é inequívoco, mas o seu valor clínico, embora promissor, é desconhecido.

Uma vez que os polimorfismos encontrados se apresentam com um risco individual baixo, será pertinente proceder à construção de um score de risco genético, que sume o risco do indivíduo tendo em conta as variantes genéticas de que se é portador. Faz pouco sentido, em termos clínicos, falar isoladamente de um polimorfismo protetor ou de risco, sem ter em conta o risco associado ao restante perfil genético.

A informação genética, a ser utilizável na prática clínica, deve adicionar conhecimento à estratificação de risco cardiovascular usual baseada nos FRT. Os estudos demonstram resultados controversos quanto à introdução dos SRG na prática clínica. Assim, justifica-se determinar se estes polimorfismos associados com a DC, isolados ou associados num score, podem melhorar a nossa capacidade preditiva, nos modelos de estratificação de risco cardiovascular tradicionais na população.

Nos doentes com DC, apesar da extensa farmacoterapia disponível para o controlo dos FRT, a ocorrência de eventos cardiovasculares é elevada. Nestes doentes, o componente genético da DC poderá ter implicações na evolução da placa aterosclerótica e, conseqüentemente, na ocorrência de eventos. Deste modo, justifica-se estudar a influência dos polimorfismos genéticos nas complicações cardiovasculares dos doentes com DC.

Sabendo-se que os genes atuarão em conjunto e que genes e ambiente terão interações no aparecimento da DC, justifica-se avaliar a interação das variantes genéticas de risco entre si, com os fatores de risco cardiovascular estabelecidos, e o seu impacto na história natural da aterosclerose coronária.

Conhecendo as interações gene-ambiente mais significativas e criando um score de risco baseado no perfil genético particular do doente e que adicione poder preditivo aos FRT, podemos apontar para novos caminhos na prevenção primária ou secundária, pela modificação dos estilos de vida e pela farmacoterapia adequada, desenhada para o indivíduo em risco.

Deste modo, justifica-se uma investigação que pretenda caracterizar e aprofundar conhecimentos sobre os mecanismos genéticos subjacentes à aterosclerose coronária em áreas ainda pouco conhecidas, e que poderá possibilitar o surgimento de novas estratégias diagnósticas, de Medicina preventiva e de terapêutica, que poderão contribuir para o melhor controlo individual desta patologia de distribuição universal.

II. OBJETIVOS

II.1 OBJETIVOS GERAIS

Com o presente trabalho pretende-se avaliar a influência duma componente genética comum no aparecimento e na evolução da DC numa população portuguesa.

Pretende-se, assim, caracterizar o perfil genético de risco para doença coronária numa amostra homogénea de doentes e caracterizada angiograficamente, comparando-a com controlos saudáveis. Pretende-se analisar os polimorfismos genéticos mais frequentemente associados com a doença coronária em genes alvo já anteriormente associados à fisiopatologia da aterosclerose coronária e simultaneamente analisar sequências de ADN produzidas recentemente nos estudos de associação genómica.

Sendo o risco genético individual cumulativo, como demonstrado em vários estudos de SRG, pretende-se sumarizar o risco individual com recurso à construção de um score de risco genético. Colocamos a hipótese que scores mais elevados demonstrarão risco crescente para a DC.

II.2 OBJETIVOS ESPECÍFICOS

O presente trabalho tem os seguintes objetivos específicos:

1- Estudo genético:

Avaliação das variantes genéticas que possam influenciar o aparecimento da DC numa população portuguesa. Conhecendo a importância dos FRT, pretende-se determinar se as variantes genéticas se mantêm como preditores independentes da DC após o controlo com os FRT.

2- Estudo das interações gene-gene (G-G) e gene-ambiente (G-A) na DC:

Estudar as interações G-G e G-A por eixos fisiopatológicos (eixo celular, eixo da dislipidemia, eixo da oxidação, eixo da diabetes/obesidade e eixo da HTA) na DC.

3- Seleção dum score de risco genético:

Construir um SRG que sumarie o risco genético individual para a DC. Avaliar a sua distribuição, calibração, interações e correlações. Avaliar a força da associação do SRG com a DC e estabelecer a sua relação com fatores de risco modificáveis e com a história familiar de DC precoce. Avaliar a associação e poder preditor do SRG nos doentes com DC antes dos 45 anos.

4- Avaliar o poder reclassificativo e discriminativo do SRG:

Avaliar a capacidade preditiva e discriminativa adicional do SRG nos métodos de estratificação cardiovascular usuais (scores de Framingham e Europeu).

Determinar a capacidade do SRG na reclassificação do risco CV avaliado pelos scores convencionais.

5- Avaliar prospetivamente a importância do SRG na ocorrência de eventos cardiovasculares:

Estabelecer quais as variantes genéticas associadas aos eventos adversos nos doentes com DC, nomeadamente a mortalidade cardiovascular, o enfarte agudo do miocárdio e a necessidade de revascularização.

III. MÉTODOS

III.MÉTODOS

Começamos por fazer a revisão bibliográfica sobre o tema e calcular a dimensão da amostra que seria necessária para cumprirmos os objetivos.

Foram definidos os critérios de inclusão e exclusão, de emparelhamento e definidas as técnicas laboratoriais, de genotipagem e a metodologia para a execução do score.

Foram definidos os parâmetros a avaliar e os critérios de valorização das variáveis.

A revisão bibliográfica foi feita inicialmente pela consulta da *PUBMED* com recurso à pesquisa sistemática cruzada das palavras *Coronary e atherosclerosis* com *gene, genotype, genetic score, candidate, GWAS, Framingham, SCORE, cardiovascular mortality, major adverse cardiovascular events (MACE), gene-gene interaction, genotype-phenotype*.

Nos raros artigos em língua que não nos era acessível ou referentes a trabalhos apresentados em congressos em que não existia o texto completo, limitamo-nos à consulta do resumo.

Os instrumentos do estudo consistiram num conjunto de formulários, realizados para o efeito, onde foram recolhidos os diferentes dados de acordo com os objetivos do trabalho. Estes dados foram posteriormente transferidos para uma folha de Excel e para um ficheiro SPSS.

III.1 AMOSTRA POPULACIONAL

III.1.1 DIMENSÃO DA AMOSTRA POPULACIONAL

Estimaram-se as amostras populacionais de acordo com os diferentes objetivos deste estudo:

- Encontrar diferenças significativas quanto às características clínico-laboratoriais e FRT entre a população com DC e a população controlo.
- Encontrar diferenças significativas caso-controlo quanto às variantes genéticas analisadas (gene a gene).
- Encontrar diferenças significativas caso-controlo quanto ao SRG.

A prevalência da DC referida na literatura varia entre 50% para doentes com síndrome metabólico e 7% para populações sem fatores de risco tradicionais. Para Portugal, foi estimada em 11% de acordo com dados estudos epidemiológicos do *European cardiovascular disease status* referentes a 2015. (102)

Assim, para detetar diferenças significativas quanto à DC entre os grupos serão necessários cerca de 800 indivíduos para uma precisão de 5% do valor real e com uma potência de 80% para detetar diferenças entre

$$n = \frac{\sigma^2 \times p \times q \times N}{e^2 (N - 1) + \sigma^2 \times p \times q}$$

os grupos. (Campbell *et al*, 1995). (103)

n = Tamanho da amostra, σ = Nível de confiança escolhido, expresso em números de desvio padrão (95%) = 1.96

p = Percentagem com a qual o fenómeno se verifica (valor encontrado na população geral); q = Percentagem complementar (100-p), N = Tamanho da população (Indivíduos com mais de 35 anos e menos de 65) N=113,528 (104), e = Erro máximo permitido (5%)

Para avaliar diferenças, quanto às variantes genéticas, utilizamos como referência o gene no locus9p21. Os OR reportado na literatura de 1,25 (alélico) e 1,5 (genotípico) foram utilizados como referência pelo que, para uma precisão de 5% e uma potência de 80% seriam necessários 900 indivíduos por grupo para detetar diferenças significativas entre os grupos.

Para avaliação de diferenças, quanto ao SRG, a literatura refere OR entre 1,2 e 1,8 pelo que utilizamos como referência um valor médio de 1,35 para cálculo da amostra. Assim, prevê-se uma amostra de 1305 indivíduos por grupo, para uma potência de 80% e precisão de 5% do valor real.

Os cálculos foram realizados pelo *SAMPsize EpiTools epidemiological calculator*.

Assim, com base nesta informação os nossos objetivos de inclusão foram de 2600 indivíduos no total.

III.2 GENÉTICA POPULACIONAL: ESTRUTURA E HOMOGENEIDADE POPULACIONAL

O estudo de variantes genéticas deve assegurar que a amostra populacional utilizada seja homogénea, sem indivíduos muito díspares dos demais (*outliers*) e que a população escolhida tenha uma representação equilibrada dos diferentes genótipos.

Populações de zonas geográficas diferentes podem apresentar diferenças nas frequências alélicas devido ao fenómeno de seleção natural dos genótipos mais favoráveis em cada população, o denominado efeito de derivação genética (*Genetic Drift*). (105)

Quando se utilizam populações de zonas geográficas diferentes, ou contendo várias etnias, as diferenças encontradas entre as variantes genéticas podem ser atribuídas apenas às diferenças de frequências alélicas entre as populações. Este efeito denomina-se estratificação populacional. (85,106)

Existem vários métodos propostos para minimizar o impacto da estratificação populacional na análise genética. Os mais utilizados nos estudos de genética populacional são o *Structure* e o *Principal Component*

Analysis (PCA). (107,108)

Com recurso a um modelo de mistura denominado *Structure*, analisamos a estrutura populacional. Este método, associa observações a populações, e caracteriza cada população de acordo com os genótipos das observações mapeadas para essa população. (Figura 6) Este método, automaticamente, calcula o número de populações através de um critério de informação, *Bayesian information criterion* e é utilizado para testar a hipótese de existirem populações latentes "escondidas" no conjunto de casos e controlos. A existência de tais populações, diferentes das descritas para o sul da Europa, pode levar à descoberta de falsos positivos nas associações testadas. (109) O número encontrado de populações mais provável nesta amostra foi de apenas uma, pelo que se deduziu não existirem populações latentes nesta amostra. Concluiu-se a representatividade do sul da Europa, semelhante do ponto de vista genético, às populações de Espanha, Grécia e Itália.

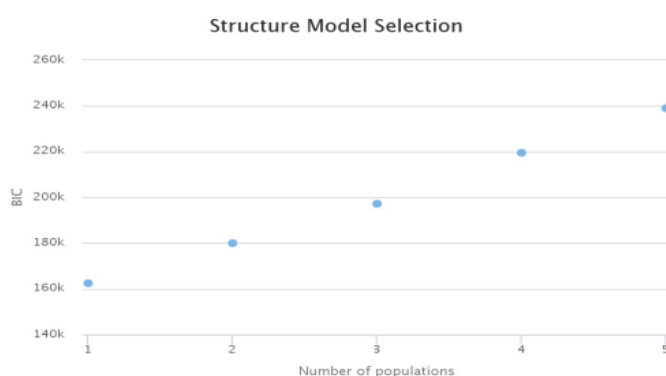


Figura 6: Análise da estrutura populacional pelo Método Structure para avaliação de populações latentes.

Legenda: O *Bayesian Information Criterion* atribui uma pontuação a um número de populações. A pontuação mais baixa corresponde ao número de populações mais prováveis na amostra.

Estes resultados, estão de acordo com Gonçalves *et al.* (110) que analisou os haplótipos ligados ao Y e DNA mitocondrial, numa amostra de 553 indivíduos portugueses continentais, do Arquipélago da Madeira e dos Açores. Neste estudo, a análise do DNA mitocondrial, revelou um componente africano sobreponível ao do Sul da Europa, fruto da influência do património genético de escravas africanas e não documentou presença significativa de componentes subsarianos na análise de haplótipos ligados ao Y. (110)

Com o método de análise de Componentes Principais do *Principal Component Analysis (PCA)* estudamos a presença de *outliers*, indivíduos cujos dados genéticos divergiam demasiado dos demais, possivelmente correspondendo a dados erróneos. (Figura 7) Este método, usa alelos relacionados para quantificar as

relações genéticas entre indivíduos por intermédio de matrizes que depois derivam num conjunto de variáveis principais ou associadas (o PCA) que refletem estas relações. (107,111)

Estes componentes principais, derivados do PCA, podem ser usados para ajustar a variabilidade étnica numa população e agrupar indivíduos em grupos ancestrais do modo a identificar 'outliers'. Com o método PCA constatou-se que apenas ~5% dos doentes da amostra excedem o limiar de significância sendo que nenhum doente aparenta ser demasiado diferente dos demais. Concluiu-se que, não se tendo detetado 'outliers' significativos, não foi necessário excluir doentes para a análise.

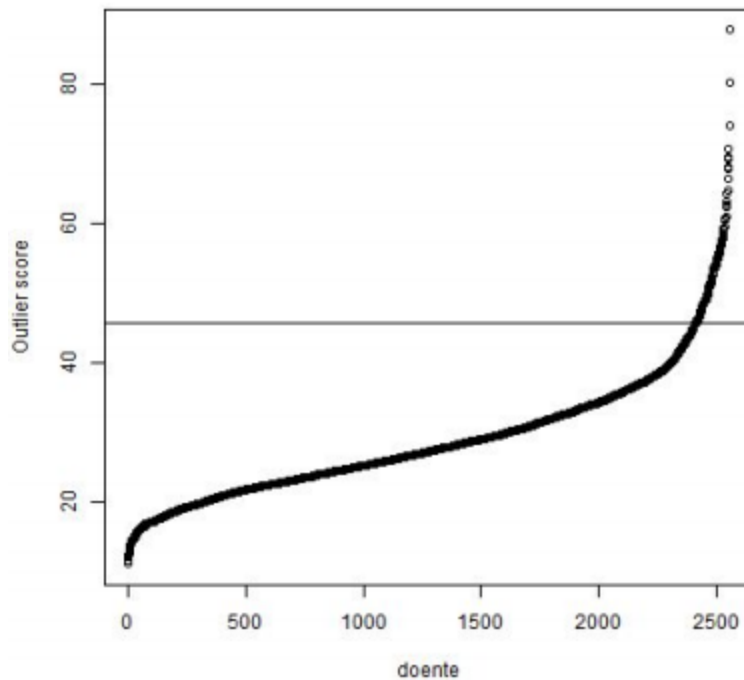


Figura 7: Análise de componentes principais (PCA) para avaliação da homogeneidade populacional.

III.3 CRITÉRIOS DE INCLUSÃO E EXCLUSÃO

Todos os participantes foram selecionados do Genemacor (*Genes in Madeira and Coronary Disease*), um estudo observacional, transversal, de caso-controlo com inclusão prospectiva de indivíduos desde 2006 até à data. É considerada uma população geneticamente isolada e homogénea do Sul da Europa, como referido previamente.

Para esta análise, incluímos um total de 2888 participantes que assinaram o termo de consentimento informado e foram informados acerca da colheita de sangue para análise de ADN e de dados clínicos relevantes.

Este estudo foi aprovado pela Comissão de Ética do SESARAM (Serviço de Saúde da Região Autónoma da Madeira) e foi conduzido de acordo com as diretrizes éticas internacionais da Declaração de Helsínquia.

Todos os participantes deste estudo (i) têm idades compreendidas entre os 30 e os 65 anos; (ii) são oriundos e residentes na Ilha da Madeira, há pelo menos 3 gerações; (iii) a presença de doença coronária foi avaliada angiograficamente nos casos; (iv) e a ausência de evidência de doença coronária nos controlos saudáveis.

A determinação da presença e da gravidade de DC foi feita pela equipa do laboratório de Hemodinâmica, que incluiu um Cardiologista de intervenção dedicado.

A DC foi definida como angiograficamente significativa pela presença de uma ou mais lesões coronárias de estenose $\geq 70\%$ de diâmetro em pelo menos uma artéria coronária epicárdica principal ou o seu ramo primário, ou de 50% de estenose do tronco comum. (112) Os doentes sem DC ou com DC não obstrutiva e valvulopatia severa não foram considerados para este estudo.

A população controlo foi recrutada através de cadernos eleitorais de acordo com a idade e género dos doentes com a DC. Após o primeiro contato telefónico os controlos vieram à Unidade de Investigação para a avaliação clínico laboratorial completa. Registou-se um grau de rejeição de apenas 2%, motivado por doença do próprio ou familiar, ou devido à situação laboral que impossibilitou a avaliação inicial. Estes indivíduos foram, mais tarde, substituídos por outros do mesmo grupo etário e género.

O grupo dos controlos incluiu indivíduos sem antecedentes ou história sugestiva de DC, significativamente semelhantes aos casos em termos de idades e género. Neste grupo, a presença de DC é critério de exclusão do estudo, pelo que foram questionados acerca dos sintomas sugestivos de cardiopatia isquémica, fatores de risco cardiovasculares e submetidos a um eletrocardiograma e, quando necessário, a outros exames complementares de diagnóstico. Após a admissão dos indivíduos, foi registada a história clínica de cada um, com referência a dados antropométricos, FRT e dados demográficos, incluindo idade, local de nascimento e residência.

III.4. DADOS COLHIDOS E CRITÉRIOS DE VALORIZAÇÃO DAS VARIÁVEIS:

III.4.1. ANÁLISE DE PARÂMETROS ANTROPOMÉTRICOS

Foram determinadas as seguintes medidas antropométricas e sinais vitais: altura, peso, perímetro abdominal, relação cintura-anca, frequência cardíaca, pressão arterial e velocidade de onda de pulso.

A altura foi medida ao centímetro mais próximo e o peso ao quilograma mais próximo. Calculou-se posteriormente o IMC a partir da fórmula “Peso (Kg) / Altura² (m)”. Um IMC normal foi definido com valores $< 25 \text{ kg/m}^2$, o excesso de peso com valores entre $25 - 29,9 \text{ kg/m}^2$ e obesidade com valores $\geq 30 \text{ kg/m}^2$. (113)

O perímetro abdominal foi medido a uma distância média entre a grelha costal inferior e a crista ilíaca lateral em centímetros. A obesidade central foi definida como cintura superior a 94 cm no sexo masculino e 80cm no sexo feminino. (114)

A relação cintura/anca foi calculada, dividindo o valor do perímetro abdominal (cm) pelo valor do perímetro da anca (cm). Para as mulheres, o valor da relação cintura/anca deve ser inferior a 0,80 cm e para os homens, o valor deve ser inferior a 0,90 cm. (115)

A pressão arterial foi medida, após 10 minutos de repouso, no braço direito, usando um esfigmomanómetro (Welch Allyn) (fases I a V). Registou-se a média de três leituras feitas com dois minutos de intervalo (116).

A velocidade de onda de pulso (VOP) foi determinada por Tonometria através de um dispositivo automático, o Complior®, que obtém simultaneamente duas ondas de pulso registadas com dois sensores, um colocado na artéria carótida comum direita e o outro sobre a artéria femoral direita, possibilitando o cálculo automático da VOP como previamente descrito. (117)

III.4.2. FATORES DE RISCO CARDIOVASCULAR TRADICIONAIS

Os doentes foram classificados como tendo HTA se referirem este antecedente, cumprirem medicação antihipertensora há pelo menos 3 meses ou apresentarem uma PAS ≥ 140 mmHg e/ou uma PAD ≥ 90 mmHg, após a média de três medições. (116)

A DM tipo 2 (DMT2) foi definida pela presença de uma glicémia em jejum ≥ 126 mg/dL ou hemoglobina glicosilada (HBA1c) $\geq 6,5\%$ ou em doentes sob terapêutica antidiabética específica. (118)

A dislipidemia foi definida como LDL >150 mg/dl, HDL <40 mg/dL para os homens e <45 mg/dL para as mulheres, e triglicéridos >150 mg/dl. (119)

Considerou-se que os indivíduos tinham história familiar de DC, se o pai ou irmão tivessem sido diagnosticados com DC antes dos 55 anos de idade ou a mãe ou irmã antes dos 65 anos de idade. (120)

O sedentarismo foi considerado quando os indivíduos praticavam menos de 40 minutos por semana de atividade física moderada.

Considerou-se fumador um indivíduo com esse hábito à entrada do estudo ou com menos de 5 anos de cessação tabágica. (121)

Quantificou-se o consumo de álcool em gramas por semana, sendo considerado significativo se fosse superior a 60 gr/semana no género masculino e 40gr/semana no género feminino. O abuso de álcool foi quantificado em mais do que 300gr/semana que corresponde a exceder duas bebidas diárias.

III.4.3. VARIÁVEIS BIOQUÍMICO-LABORATORIAIS

Todos os indivíduos efetuaram um estudo analítico de rotina na admissão (hemograma completo, estudo da coagulação, função renal e hepática, ionograma sérico). As colheitas foram realizadas após um período de jejum noturno de, pelo menos, 12 horas.

Efetou-se também a determinação do perfil lipídico (colesterol total, colesterol LDL, colesterol HDL, triglicéridos), glicémia em jejum, e determinação da hemoglobina glicosilada (HbA1c) no caso dos diabéticos.

Os valores de colesterol total, colesterol LDL, HDL, triglicéridos e glucose foram depois quantificados no soro, com um auto analisador "AU 5400" (Beckman Coulter) através de uma técnica enzimática automática. Dada a conhecida interferência da lipoproteína(a) na determinação dos valores de LDL pelos métodos de precipitação, foi usada uma técnica enzimática direta para o HDL e LDL. (122)

Os marcadores bioquímicos de risco: Apo B, e a PCR foram quantificados por um aparelho automático "AU 5400" (Beckman Coulter) pela técnica de Imunoturbidimetria.

Para a determinação enzimática da lipoproteína (a) foram utilizados anticorpos monoclonais extra *kringle* e turbidimetria para minorar o efeito da dispersão de pesos moleculares. (123,124)

Para a medição de homocisteína, a amostra foi mantida a 2-8°C, pelo menos durante 30 minutos (no máximo 60 minutos) e, após a retração do coágulo, o soro foi dividido em alíquotas e mantido a -20 ° C até à medição pelo método cinético (AU 5400, Beckman Coulter).

Em relação ao fibrinogénio a colheita também foi feita com o doente em jejum, para um tubo com citrato de sódio, utilizando-se um analisador automático "ACL TOP 700".

III.5 ESTUDO GENÉTICO

III.5.1 SELEÇÃO DAS VARIANTES GENÉTICAS

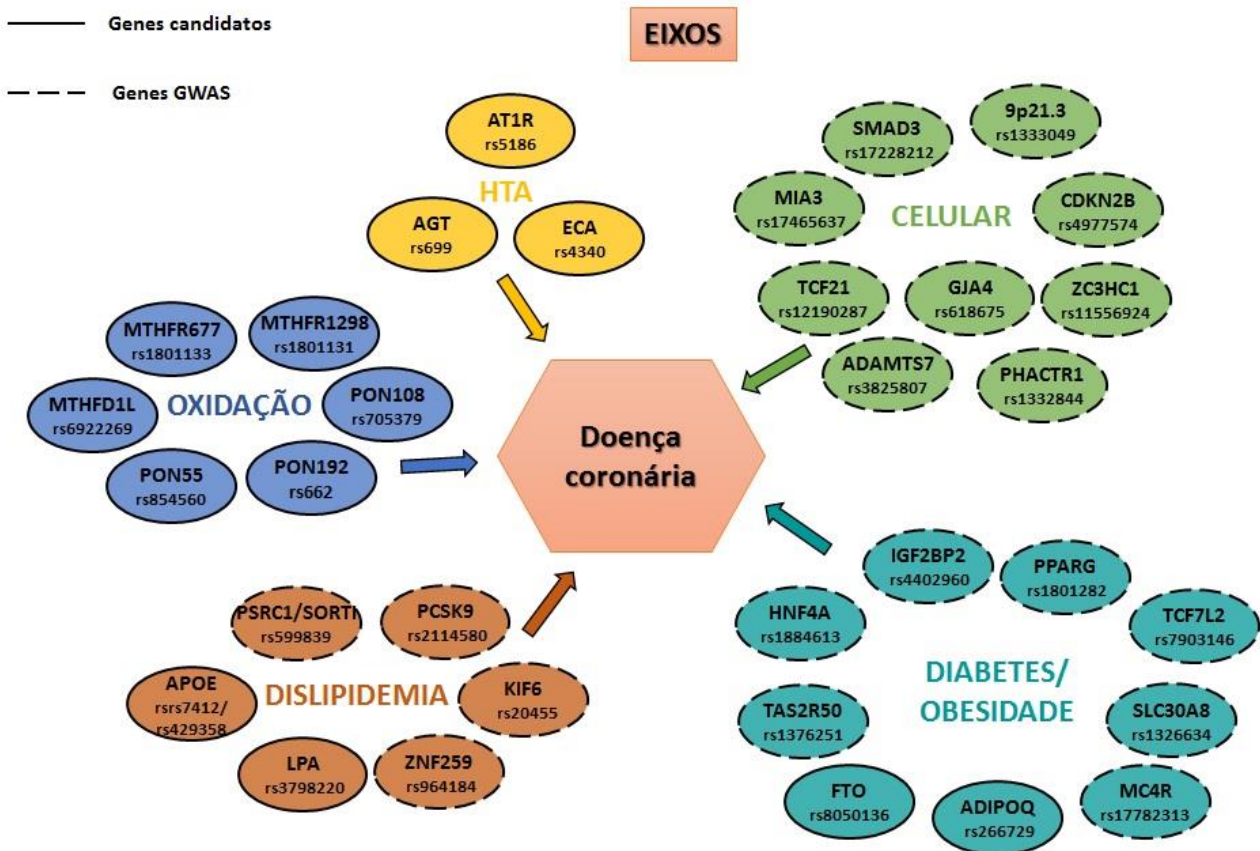


Figura 8: Variantes genéticas incluídas e localização presumível em eixos fisiopatológicos para a DC. A tracejado estão representados variantes genéticas dos estudos GWAS e a traço contínuo estão representados genes candidatos.

Para o estudo do componente genético da DC selecionaram-se 13 variantes genéticas dos estudos de genes candidatos e 20 variantes consideradas promissoras nos estudos GWAS publicados até 2013. (fig 8) Os genes candidatos atuam em vias biologicamente plausíveis, como o eixo da HTA no sistema Renina-Angiotensina-Aldosterona (SRAA), eixo da oxidação, eixo da dislipidemia e eixo da diabetes/obesidade.

As variantes GWAS foram entretanto sendo estudadas, e algumas delas têm já mecanismo fisiopatológico conhecido, enquanto outras, permanecem com mecanismo ainda por clarificar. Dado o elevado número de variantes emanadas dos GWAS, limitámos a inclusão de variantes que tivessem sido testadas em populações caucasianas e cujo risco relativo fosse $\geq 1,1$ e simultaneamente a frequência do alelo menor (FAM) fosse suficiente para ser encontrado de forma representativa na nossa amostra.

III.5.1.1 VARIANTES GENÉTICAS ASSOCIADAS COM O METABOLISMO LIPÍDICO

III.5.1.1.1 *Apolipoprotein E (APOE)*

O gene APOE, localizado no cromossoma 19, codifica uma proteína conhecida como apolipoproteína E, que desempenha um papel fundamental no metabolismo do colesterol e triglicéridos ao se ligar a receptores hepáticos, estimulando a captação de quilomicrons e lipoproteínas de muito baixo peso molecular (VLDL) da corrente sanguínea.

A elevação dos níveis de colesterol está na origem do risco CV acrescido, aterosclerose, DC e EAM. (125) São conhecidos pelo menos três alelos da variante genética rs429358 APOE: Apoε2, Apoε3 e Apoε4, sendo o ε3 o mais comum, presente em mais de metade da população mundial.

Reportou-se uma relação praticamente linear entre os diferentes genótipos com os níveis de LDLc e risco coronário, afirmando que em oposição ao genótipo comum ε3/ε3, os portadores do alelo ε2 teriam 20% menos risco de DC e os portadores do alelo ε4, um risco mais elevado. (126)

III.5.1.1.2 *Lipoprotein (a) (LPA)*

O gene LPA, localizado no cromossoma 6, codifica uma proteína com atividade de serina proteinase conhecida como lipoproteína a Lp(a), cujas concentrações plasmáticas têm forte componente hereditária e dependem da expressão do referido gene LPA, com a evidência científica a mostrar resultados limitados com terapêutica antilipidêmica. (127). Níveis elevados de Lp(a) estão tradicionalmente associados a aterosclerose, DC e risco CV global independentemente de outros FRT. (128-130)

A ApoA contém domínios semelhantes aos do plasminogénio. A acumulação na parede do vaso de Lp(a), impedindo a ligação do plasminogénio à superfície endotelial, reduzindo a produção de plasmina (aumentando o risco pró-trombótico) e promovendo a proliferação de células musculares lisas, mecanismos propostos para justificar a aterosclerose e o risco de coágulos em placas instáveis. (131)

A prevalência e relação da variante rs3798220 LPA com o tamanho da apolipoproteína(a) e com os níveis de Lp(a) parecem ser altamente variáveis e etnia-dependentes. Contudo, o Locus LPA revelou, na maioria dos GWAS, ter uma associação genética com DC potente, sendo já intensamente estudada como alvo terapêutico.

III.5.1.1.3 *Kinesin Like Protein 6 (KIF6)*

A variante KIF6 codifica uma proteína da família das kinesinas (*Kinesin-like 6*). Esta proteína é um dos vários componentes moleculares mediadores do transporte de organelos, complexos proteicos e mRNA.

Uma variante genética rs20455 do gene KIF6 revelou em vários estudos uma forte associação com a DC.

(132,133) Esta variante está presente em, pelo menos, 30% dos caucasianos, afro-americanos e asiáticos. Nestes condiciona aumento do risco CV, embora com boa resposta à terapêutica hipolipemiante instituída, observando-se uma redução de eventos CV. (134,135)

Os mecanismos moleculares e fisiopatológicos precisos, pelo qual este gene contribui para um aumento do risco CV, não são ainda conhecidos.

III.5.1.1.4 Zinc Finger Protein 259 (ZNF259)

Também conhecida como ZPR1, a proteína codificada pela variante ZNF259, localizada no cromossoma 11, tem um papel regulador na proliferação celular e transdução de sinal, necessárias à homeostasia celular.(136)

A região promotora do ZNF259 encontra-se ligada ao PPAR γ 1 e 2 (*proxisome proliferator-activated receptor gamma*), com conhecida função na insulinoresistência e obesidade por intermédio do HNF4a (*hepatocyte nuclear factor 4 alpha*). A variante ZNF259 rs964184 é capaz de ativar uma variedade de genes, envolvidos no metabolismo da glucose, ácidos gordos, metabolismo de colesterol no fígado, rins, intestino e pâncreas. O polimorfismo, nesta variante genética, associou-se a níveis mais elevados de colesterol total (CT), TG, LDL-C e níveis mais baixos de HDL-C. (137) Esta variante pertence a um *cluster* lipídico (APOA4-APOA5-ZNF259-BUD139) que em populações europeias e asiáticas demonstrou estar associado com a DC. (76,82)

III.5.1.1.5 Sortilin 1 (PSRC1/SORT1)

O *cluster* genético CELSR2-PSRC1-SORT1 mostrou uma importante associação com o risco CV e DC pelo seu papel no metabolismo lipídico e nos níveis circulantes de colesterol. Entre os vários polimorfismos conhecidos nesta região (1p13.3) descritos na literatura, o SNP rs599839, constitui mais uma das variantes identificadas nos GWAS como relevante na regulação dos níveis de LDL-C. (138-140)

O gene SORT1 codifica a proteína sortilina, que atua como um receptor/ligando da apolipoproteína B100, promovendo a formação e libertação de lipoproteínas-ApoB100 hepáticas, que regulam os níveis de LDL-C. Sabe-se, ser recetor de múltiplos ligandos, incluindo a lipoproteína lípase (LPL). (141) O nível de expressão/tradução deste gene é inversamente proporcional aos níveis circulantes de ApoB e de LDL-C. Foi proposto, como mecanismo molecular subjacente, que a sortilina sinalizaria a apoB100 e o LDL-C para a degradação lisossomal. (142) O mecanismo concreto pelo qual a sortilina se relaciona com o metabolismo lipídico é ainda controverso.

O alelo G, na variante rs599839, parece ter um papel protetor CV ao estar associado a níveis mais baixo de LDL-C e mais elevados de HDL. No entanto, a evidência científica parece conflituosa no que toca à real contribuição desta variante ou *cluster* genético para a DC, parecendo haver diferenças étnicas a considerar.

III.5.1.1.6 *Proprotein convertase subtilisin/kexin type 9 (PCSK9)*

Localizada no cromossoma 1, o gene PCSK9 tem um reconhecido papel na homeostasia lipídica e tem sido alvo de enorme interesse científico pelo potencial efeito terapêutico hipolipemiante resultante da inibição da enzima que codifica. (143)

Este gene, reconhecido em 2003, em estudos de *linkage*, é expresso em diferentes tecidos e células, incluindo o endotélio vascular, músculo liso vascular e macrófagos com reconhecido efeito proaterosclerótico. Este gene codifica uma enzima conhecida como *proprotein convertase subtilisin/kexin type 9 (PCSK9)*, capaz de se ligar aos receptores de LDL (LDLr) hepáticos, promovendo a sua destruição lisossomal e impedindo a sua reciclagem. (144,145) Assim, com a menor expressão de receptores de LDL hepáticos à superfície celular, menor a capacidade de captar LDLs circulantes que contribui para a hipercolesterolemia/dislipidemia. Certas mutações, relacionadas com o ganho de função da sua atividade proteolítica, estão na gênese da rara hipercolesterolemia familiar monogénica, tal como mutações com perda de função estão na base de menores níveis de LDL-c e menor risco cardiovascular. (146)

Secretada inicialmente como um zimogénio inactivo, a PCSK9 sofre ativação por clivagem no retículo endoplasmático e mais tarde interage com a sortilina, mencionada previamente e codificada pelo PSRC1. De facto, os níveis plasmáticos de ambas parecem correlacionar-se entre si e com o ciclo de síntese de colesterol. A variante comum rs2114580 é uma das variantes consistentes na associação com DC.

III.5.1.2 VARIANTES GENÉTICAS ASSOCIADAS COM O SISTEMA RENINA-ANGIOTENSINA-ALDOSTERONA

III.6.1.2.1 *Angiotensin-converting enzyme (ACE)*

O gene ACE, localizado no braço longo do cromossoma 1 (17q23.3), é responsável por codificar a conhecida ECA (enzima conversora de angiotensina), uma metaloproteinase de zinco capaz de clivar a angiotensina I em angiotensina II. O seu papel no sistema renina-angiotensina-aldosterona e consequentemente no balanço da regulação de sódio, água e pressão arterial é bem conhecido e estudado.

Os genótipos DD, ID e II da variante rs4340 estão associados com elevados, intermédios e baixos níveis de ECA circulante. (147) Esta variante genética associou-se com risco de DC, EAM e HTA em vários estudos, havendo no entanto resultados inconsistentes. (148-150) Variações étnicas parecem explicar as diferenças encontradas.

III.5.1.2.2 *Angiotensinogen (AGT)*

O gene AGT está localizado no cromossoma 1 e codifica o angiotensinogénio, um polipéptido produzido primariamente no fígado, que será clivado em angiotensina I por ação da renina. Sendo conhecido o seu papel no eixo renina-angiotensina-aldosterona e nos mecanismos da hipertensão arterial. Polimorfismos no gene AGT parecem correlacionar-se com níveis plasmáticos de angiotensinogénio diferentes, bem como variações na pressão arterial.

A relação da variante rs699 com o risco para DC é controversa e ainda não foi totalmente esclarecida. (151,152) Em variados estudos demonstrou-se o aumento de risco para DC, EAM e HTA e AVC. Com efeito, numa meta-análise de 2012, com base no global dos estudos prévios, documentou-se um risco para DC de acordo com o modelo recessivo (OR=1,37; IC95%: 0,98-1,91). (153) Numa meta-análise de 2013, documentou-se o risco aumentado para EAM com um OR de 1,04 (IC95%: 0,92-1,17), tendo-se, no entanto, encontrado disparidades étnicas. (154)

No global, a relação desta variante com o risco de DC parece independente dos valores da pressão arterial. Com efeito, num estudo, esteve associada com o score de cálcio em indivíduos com história familiar de DC, e foi independente dos valores de pressão arterial. (155)

III.5.1.2.3 *Angiotensin II type 1 receptor (AT1R)*

O agonismo dos recetores de angiotensina I (AT1R) desencadeia uma resposta hipertensiva à custa da vasoconstrição e da expansão volémica associadas. A nível vascular, a estimulação de recetores AT1 induz ainda a hipertrofia do músculo liso vascular, aumento da produção de matriz extracelular e inflamação local que poderão contribuir para a progressão da aterogénese e risco de DC. O polimorfismo rs5186 descrito para este gene demonstrou uma associação com maior risco CV, independentemente da pressão arterial. No entanto, a controvérsia persiste com outros estudos a não mostrarem o mesmo. Na meta-análise de 2014 e com base nos estudos previamente publicados, a variante AT1R esteve associada ao risco de EAM com um OR de 1,12 (IC95%: 1,01-1,25). (156) O alelo C da variante AT1R aumentou o risco, enquanto que o alelo A foi protetor, sobretudo nos estudos com populações caucasianas. Também nesta meta-análise foram referidas disparidades entre os diferentes grupos populacionais. (156) Apesar das heterogeneidades, numa meta-análise de 2012, o risco para DC do alelo C, em populações asiáticas foi também confirmado, tendo-se demonstrado um OR de 1,59 (IC95%: 1,22-2,06). (157)

III.5.1.3 VARIANTES GENÉTICAS ASSOCIADAS COM A DIABETES/OBESIDADE

III.5.1.3.1 *Insulin-like Growth factor 2 mRNA binding Protein (IGF2BP2)*

A variante localizada no cromossoma 3q27 codifica uma proteína de ligação que regula a tradução do IGF2 por ligação ao seu mRNA. Este IGF2BP2 pertence a uma família de proteínas de ligação oncofetais, implicadas no crescimento e desenvolvimento embrionário e na organogénese do tecido adiposo e pancreático. A ligação ao fator de crescimento tipo insulina tem sido apontada como regulador do metabolismo glicémico e a suscetibilidade à DM. Esta variante rs4402960 foi associada à disfunção de células pancreáticas β , à redução da libertação da insulina basal e ao aumento da gordura visceral, conduzindo à insulinoresistência. (158,159) Para além de DM e obesidade, esta variante genética foi também associada à HTA, e à elevação dos triglicéridos. (160)

III.5.1.3.2 *Solute Carrier Family 30 Member 8 (SLC30A8)*

Esta variante genética, também conhecido por ZNT8, está localizada no cromossoma 8q24. Esta codifica uma proteína co-transportadora do Zinco que se expressa sobretudo nos ilhéus de Langerhans e interfere com a maturação e acumulação de insulina nas células B pancreáticas. (161) Também se expressa em menor grau no gordura visceral e no tecido celular subcutâneo dos indivíduos obesos.

O alelo C do rs13266634, localizado no gene SLC30A8 (ZNT8), foi associado com a redução da libertação de insulina, sendo este o mecanismo proposto para a sua associação com a DM. Este mecanismo ainda não foi validado em modelo animal.

Esta variante genética, nos GWAS, esteve associada a 14% do aumento de risco para DM. (162,163) Em populações europeias, foi posteriormente replicada e foram demonstradas significativas interações sinérgicas com outras variantes genéticas expressas nas células pancreáticas. (164)

III.5.1.3.3 *Adiponectin, C1Q And Collagen Domain Containing (ADIPOQ)*

O gene ADIPOQ, localizado no cromossoma 3, codifica a adiponectina, uma proteína similar ao colagénio X e VIII e ao Fator 1C do complemento. Expressa-se, apenas, no tecido adiposo. A variante rs266729 esteve associada à redução dos níveis de adiponectina.

A adiponectina está envolvida no metabolismo da gordura e na sensibilidade à insulina com efeitos diretos anti diabéticos, anti aterogénicos e anti-inflamatórios. Após a ativação no fígado e no músculo-esquelético, aumenta a combustão dos ácidos gordos e da glucose. Também antagoniza os efeitos do fator de necrose tumoral alfa. Pode, por isso, ter um papel no crescimento celular, angiogénese e *remodeling* tecidular por ligar-se a vários fatores de crescimento. Variantes genéticas do ADIPOQ foram associadas com DC, obesidade, DM, hiperuricémia e com esteatose hepática não alcoólica. Apesar de existirem resultados

inconsistentes em diferentes populações, a meta análise mais recente, publicada em 2017, confirma a associação de variantes da adiponetina com o risco de DC. (165)

III.5.1.3.4 *Peroxisome proliferator-activated receptor gamma, coactivator 1 alpha (PPAR gama)*

O gene PPAR gama está localizado no cromossoma 3 e codifica um recetor nuclear da família dos peroxissomas proliferadores ativados. Tem um papel essencial na reprogramação metabólica em resposta à disponibilidade de dieta, aumentando a expressão de múltiplos genes envolvidos no metabolismo glicídico, lipídico e na diferenciação dos adipócitos. Assim, na presença de ácidos gordos, os ligandos do PPAR gama promovem a acumulação dos mesmos nas células adiposas e regulam a expressão de hormonas secretadas pelos adipócitos, relacionadas com o metabolismo glicídico. De igual modo, é responsável pela ativação da via β oxidativa dos ácidos gordos, aumentando a transcrição da acyl-CoA oxidase podendo, por isso, conduzir à hiperlipidémia. É também o local de ação dos fibratos e das glitazonas. (166)

Duas variantes genéticas do PPAR gama, associadas com o risco de DM, foram já replicadas em várias populações mundiais. No global, a variante rs1801282 produziu um OR de 1,14 (IC95%: 1,08-1,20) no risco de DM (167).

De igual modo, a associação da variante C161T do PPAR gama com EAM produziu um OR de 1,43 (IC95%: 1,09-1,87). (168)

De igual modo, variantes no gene PPAR gama foram determinantes no índice de massa corporal, numa maior espessura da íntima média carotídea, e na HTA. A associação com a aterosclerose parece existir de forma independente dos FRT.

III.5.1.3.5 *Transcription Factor 7-Like 2 (TCF7L2)*

A variante genética Fator de Transcrição 7 está localizada no cromossoma 10 e é expresso nas células pancreáticas β envolvidas no controlo da secreção da insulina. É um elemento essencial da via de sinalização *Wingless*. Esta via de sinalização está envolvida em diversos processos biológicos. Em particular, esta via demonstrou ter um importante papel na regulação do *remodeling* vascular, através da regulação da proliferação de células musculares lisas e crescimento endotelial.

Em populações caucasianas, a presença da variante genética rs7903146 TCF7L2 aumenta o risco de DM em 50%. (159). Embora, ainda não totalmente esclarecido, este efeito parece ser exarcebado pela insulinoresistência mediada pelos ácidos gordos. Cada alelo T aumenta o risco para DM tipo II em 1,45. (169) Esta variante, mesmo em não diabéticos, esteve associado com o risco significativo para DC com um OR de 1,90 a 2,30; maior mortalidade CV e maior evidência de eventos adversos. (170) Em estudos humanos, a redução da secreção de insulina, mediada pelo TCF7L2, durante a fase aguda do enfarte associou-se à disfunção endotelial, maior carga trombótica coronária e maior mortalidade a curto prazo.

III.5.1.3.6 *Melanocortin 4 Receptor (MC4R)*

A variante genética Receptor da Melanocortina 4 está localizado no cromossoma 18 e codifica uma proteína ligada à membrana que é membro da família dos recetores da melanocortina. Esta proteína interage com hormonas adrenocorticotrópicas por intermédio de proteína G, tendo por essa via um papel central na homeostasia e crescimento somático.

Em estudos europeus, cerca de 4 a 5,8% dos adultos obesos apresentam mutação no gene MC4R. (171,172) Polimorfismos desta variante genética como o rs17782313 representam a causa genética mais comum para a obesidade. Em humanos, polimorfismos desta variante genética e do seu promotor estão consistentemente associados com o aumento do índice de massa corporal e obesidade. (173)

III.5.1.3.7 *Hepatocyte nuclear factor 4 (HNF4)*

A variante do Fator Hepatocitário Nuclear 4 alfa codifica um fator de transcrição que é responsável pela regulação da transcrição de genes nas células Beta pancreáticas, para além de regular a transcrição de vários genes hepatocitários. Entre vários genes, esta variante é capaz de ativar a transcrição do gene da insulina. (174) Variantes genéticas do factor hepatocitário nuclear 4 (rs1884613) foram associadas com a hiperglicémia, a obesidade, a intolerância à glicose e a insulinoresistência. (174) Esta variante genética associou-se com um risco para DM de 1,33 (IC95%: 1,06-1,65). (174)

Atualmente, estão em estudo interações gene-gene, entre esta variante e a família das sirtuínas, e as variantes do gene PPAR gama, na modulação hepática dos ácidos gordos mitocondriais e peroxissomais, no metabolismo dos ácidos biliares e das lipoproteínas. (175,176) Esta interação parece fundamental na homeostasia do metabolismo lipídico e glicídico.

III.5.1.3.8 *Fat mass and obesity-associated protein (FTO)*

Esta variante, localizada no cromossoma 16, é uma proteína nuclear da superfamília de oxigenases dependentes do 2-oxoglutarato, cujo mecanismo fisiopatológico não se encontra ainda muito esclarecido. Contribui para a regulação da massa corporal e da acumulação de gordura, controlando o metabolismo e gasto de energia.

A variante genética rs 8050136 contribui para a regulação da termogénese e o controlo da diferenciação dos adipócitos em células castanhas (gastadoras de energia) ou brancas (acumuladoras de energia). (169)

Estudos humanos indicam uma forte associação com o IMC, o risco de obesidade e com a DM. (177) Estudos animais demonstraram o aumento da expressão do gene ao nível cerebral, incluindo nos núcleos hipotalâmicos ligados à regulação da ingestão de alimentos. Assim em ratinhos laboratoriais a maior expressão do gene FTO levou a aumento da ingestão e adiposidade sem aumento do metabolismo e gasto energético. (169)

III.5.1.3.9 *Taste 2 Receptor Member 50 (TAS2R 50)*

Esta variante rs 1376251, localizada no cromossoma 12, pertence a uma superfamília de recetores TAS2R. São expressos na superfície das células gustativas e medeiam a percepção do paladar por uma via de segundo mensageiro acoplada à proteína G. Também, sinalizam a composição bioquímica do conteúdo gástrico. Teoricamente, influenciam as escolhas e a procura de alimentos, influenciando a saciedade.

Em estudos humanos iniciais, associou-se com a DC e com maior incidência de eventos cardiovasculares como o EAM e/ou AVC com HR de 1,76. (178,179) Igualmente esteve associada a uma menor resposta às estatinas com reduções menores dos valores de LDL. No entanto, vários estudos posteriores, não conseguiram reproduzir estas associações. (180)

III.5.1.4 VARIANTES GENÉTICAS ASSOCIADAS COM O STRESS OXIDATIVO

III.5.1.4.1 *Metiltetrahidrofolato reductase: (MTHFR677 e MTHGR1298)*

Localizado no cromossoma 1, o gene MTHFR codifica uma proteína redutase (metiltetrahidrofolato reductase) com um papel chave no ciclo do ácido fólico. Esta promove a redução do 5,10-metiltetrahidrofolato a 5-metiltetrahidrofolato (5-MTHF), reação importante na transformação da Hcy em metionina.

Mutações ou polimorfismos neste gene (onde as mais comuns e mais bem estudadas são a 677 C/T e 1298 A/C) estão associadas a uma menor atividade enzimática de MTHF redutase e a uma maior acumulação de Hcy. Como mecanismos por detrás do efeito pró-aterosclerótico da HHcy foram propostos: a interferência com o fibrinogénio e com outros fatores da coagulação e a disfunção estrutural arterial, desde o colagénio à elastina e aos proteoglicanos. Embora as variantes genéticas do MTHF condicionem o aumento da Hcy plasmática, a sua relação com a doença aterosclerótica é ainda controversa. As interações com outros genes, níveis de ácido fólico e a dieta, têm sido apontados como potenciais moduladores desta variante genética. Na meta-análise mais recente o risco para DC prematura, associada à variante MTHF677, foi de 1,13 (IC95%: 1,01-1,27) quando comparada com o genótipo protetor. (181)

III.5.1.4.2 *Methylenetetrahydrofolate dehydrogenase 1-like (MTHFD1L)*

Como descrito previamente, as mutações nos genes reguladores do metabolismo do ácido fólico, têm sido associadas a um maior risco CV. Localizado no cromossoma 6, a variante genética MTHFD1L, não é exceção. Esta codifica uma proteína com atividade de sintetase, envolvida no metabolismo do folato a nível mitocondrial, ao promover a síntese de Tetrahidrofolato mitocondrial, importante na síntese de purinas e regeneração de metionina a partir de Hcy. Um polimorfismo neste gene (rs6922269) foi associado à DC,

além de outras doenças de foro neurológico e oncológico. (182)

III.5.1.4.3 Paroxonase 1 (PON108, PON192 e PON55)

O cromossoma 7 aloja um *cluster* genético com atividade paraoxonase. O gene PON1 codifica uma proteína (paraoxonase 1) com atividade de esterase, responsável pela hidrólise de paroxona (um componente dos organofosforados dos pesticidas) e de outros substratos tóxicos.

Hoje, é claro que esta enzima desempenha um papel importante na proteção e destoxificação do nosso organismo, perante a exposição a estes compostos (organofosforados e gases tóxicos).

A paraoxonase é, maioritariamente, sintetizada no fígado, circula ligada às HDL e previne ou retarda a oxidação das LDL. A oxidação lipídica tem também um papel fundamental na génese e progressão de diversos estados inflamatórios como é exemplo a aterosclerose. Dada a sua capacidade de remover lipoproteínas oxidadas potencialmente lesivas, foi igualmente proposto que a PON1 tivesse papel anti-aterosclerótico. Uma variante PON hiperfuncionante pode estar associada com a redução do risco CV.

Variantes genéticas da PON (rs854560, rs662 e rs705379) são responsáveis por 25% dos valores séricos da PON. A restante variabilidade dos valores séricos é atribuída a possíveis interações gene-ambiente com estatinas, ácidos gordos, vitaminas antioxidantes e fenómenos de metilação. (183,184)

III.5.1.5 VARIANTES GENÉTICAS POTENCIALMENTE ASSOCIADAS COM O CICLO CELULAR, MIGRAÇÃO DE CÉLULAS E/OU INFLAMAÇÃO/MECANISMO DESCONHECIDO

III.5.1.5.1 Locus 9p21 (9p21.3)

A relação entre o *locus* 9p21 e o risco CV está bem estabelecida na literatura, sendo independente de FRT como a HTA, dislipidemia ou hiperglicemia. (185) Apesar de tradicionalmente associado à aterosclerose, as consequências moleculares da variação genética deste cromossoma e os mecanismos precisos por detrás dos diversos eventos cardiovasculares, a que se associa, permanecem por esclarecer, especulando-se ainda sobre o seu potencial papel na trombogénese e estabilidade da placa coronária.

A associação do gene 9p21.3 com a DC e o EAM, permanece uma das mais fortes em vários estudos como o *Wellcome Trust Case Control Consortium (WTCCC)*, salientando a existência de uma variante genética principal de risco neste *locus*. (186-188) O alelo de risco (rs1333049) ocorre em 75% da população humana e está associado a um aumento de 25% no risco relativo para a DC. Embora a relação com a DC já se encontre bem estabelecida, a relação com EAM é ainda controversa. (189)

III.5.1.5.2 *Cyclin-dependent kinase inhibitor (CDKN2A/2B)*

Próximo da região 9p21.3 localizam-se as sequências de genes que codificam dois inibidores da cinase ciclina-dependentes 2A/B (CDKN2A/CDKN2B), que regulam o ciclo celular e que poderão estar na fisiopatologia da aterosclerose pelo seu papel na inibição da TGF- β .

Também localizado no *cluster* genético CDKN2B/CDKN2A, está o gene ANRIL, cujo transcrito é uma molécula de ARN funcional que ao interagir com os complexos PCR1 e PCR, contribui para o silenciamento epigenético de outros genes neste cluster. Esta região (rs4977574) é tida na literatura como sendo de significativa suscetibilidade genética para a doença CV.

O papel do ANRIL na fisiopatologia da DC foi ainda demonstrado em diversos estudos de expressão. A sua região promotora é crítica para a transcrição de CDKN2A e CDKN2B, cuja deleção seria responsável pela hiperproliferação celular vascular e diminuição da senescência das células vasculares, mecanismos na gênese da DC. (190)

III.5.1.5.3 *SMAD family member 3 (SMAD3)*

Localizado no cromossoma 15, a variante genética SMAD3 codifica um conjunto de proteínas que participam na normal sinalização intracelular da família TGF- β . A sua expressão ou regulação inadequada poderá estar na gênese de diversas patologias cardiovasculares como a insuficiência cardíaca (por hipertrofia e fibrose miocárdica), a aterogênese, o EAM e a hipertensão pulmonar.

A “superfamília” das proteínas da via TGF- β são proteínas reguladoras do crescimento celular, da proliferação, da apoptose, da diferenciação e da reparação tecidual, contribuindo para a homeostasia celular. Estes processos estão na base da formação, progressão e instabilidade da placa aterosclerótica e do seu carácter dinâmico ao longo do tempo. (183)

Várias evidências apontam para um papel duplo da SMAD3 no coração, um benéfico e cardioprotetor (ao limitar a resposta hipertrófica à sobrecarga hemodinâmica), e um prejudicial (por fibrose, diminuição da *compliance* ventricular, disfunção sistólica e risco aumentado de arritmias). Esta variante genética é importante para a integridade da parede vascular, estando por esse motivo associada, também, a síndromes aórticas e anomalias do desenvolvimento embrionário.

O papel das SMADs e da sua via de sinalização (via *Transforming growth factor- β (TGF β)*) tem sido alvo de interesse científico, dado o seu potencial terapêutico CV e anti-neoplásico. Em estudos europeus, a variante rs 17228212 SMAD foi associada a um risco para a DC com um OR de 1,19 (IC95%: 1,09-1,3). (183)

III.5.1.5.4 *Melanoma Inhibitory Activity Family, Member 3 (MIA3)*

O gene MIA3 localizado no cromossoma 1 é necessário para a normal exportação do colagénio VII do retículo endoplasmático para o aparelho de Golgi.

Várias variantes deste gene têm revelado maior risco CV, sendo este tradicionalmente associado à DC em populações ancestrais europeias. Nestes, e tendo em conta os vários estudos, esta variante genética (rs17465637) esteve associada com um risco para a DC de 1,33 (IC95%: 1,20-1,48). (97,138,182)

Em ratinhos *Knockout* para este gene, a secreção inapropriada de colagénio I, II, III, IV, VII e IX associou-se a uma maior extensão de EAM, quando este era induzido. (191) Mudanças na composição do colagénio em humanos, poderá resultar num compromisso da integridade da matriz extracelular vascular (promovendo *remodelling* arterial) e facilitando a migração de monócitos/macrófagos através do fibrinogénio e células endoteliais, iniciando e perpetuando o processo aterosclerótico. (192)

III.5.1.5.5 *Zinc Finger C3HC-Type Containing 1 (ZC3HC1)*

Identificado como um dos principais locus de particular relevo na DC nos estudos genéticos, a variante rs11556924 ZC3HC1 localizado no cromossoma 7, codifica a proteína NIPA *Nuclear Interacting Partner of Anaplastic Lymphoma Kinase*, que é um componente de um complexo regulador da divisão celular. (76) Nestes estudos, em populações europeias, a variante genética rs11556924 do ZC3HC1 esteve associada a um risco para a DC de 1,09 (IC95%: 1,07- 1,12). (76) Este complexo sinaliza a ciclina B1 para degradação, evitando desta forma a transição para mitose na fase G2/M do ciclo celular. (193)

Variantes neste gene afetam a expressão da KLHDC10 *Kelch domain-containing protein 10*, sugerindo este último como o potencial gene relacionado com a DC, dado o papel desta proteína na apoptose e morte celular induzida por *stress* oxidativo, mecanismos inerentes à formação e progressão da placa aterosclerótica.

III.6.1.5.6 *Transcription Factor 21 (TCF21)*

O gene TCF21 localizado no cromossoma 6 parece associar-se a uma maior taxa de formação de complexo ARN-ARN e maior ligação ao miR-224, nas células musculares lisas de artérias coronárias humanas, levando à expressão desregulada do gene TCF21 perante doença. Estudos humanos mostraram, já, níveis de expressão aumentada de TCF21 e miR-224 nas lesões ateroscleróticas. O seu mecanismo, relacionado com a regulação de células progenitoras de células musculares lisas, encontra-se ainda em estudo. (194) Em estudos europeus, a variante genética rs12190287 teve um risco consistente para a DC de 1,09, sendo aparentemente independente dos FRT e ambientais. (82)

III.5.1.5.7 A disintegrin and metalloproteinase with thrombospondin motifs 7 (ADAMTS7)

As proteínas codificadas por este gene inserem-se na família das “metzincinas”, uma família de metaloproteinases de zinco capazes de processar todos os componentes da matriz extracelular. Estas são importantes reguladoras do *remodelling* tecidual, particularmente a nível vascular com a progressão da aterosclerose. (195) As alterações da matriz, induzidas por estas “metzincinas”, induzem mudanças no comportamento das células endoteliais e musculares lisas da parede vascular, promovendo a formação de neoíntima. Esta variante foi recentemente localizada em placas ateroscleróticas vulneráveis, em modelo humano. (196)

Em 3 GWAS de grande dimensão foi estabelecida uma relação entre variantes do gene ADAMTS7 (rs3825807) e a aterosclerose, a doença arterial periférica e o risco CV. (76,78,197) Nestes estudos europeus, o risco conferido por esta variante no aparecimento de DC foi de 1,08 (IC95%: 1,06-1,10).

III.5.1.5.8 Gap Junction Protein, alpha 4 (GJA4)

Localizado no cromossoma 1, este gene codifica um membro da família das conexinas. A proteína codificada (conexina 37) é um componente importante nas *gaps junctions* intercelulares, permitindo a comunicação direta entre as células adjacentes e a difusão de componentes de baixo peso molecular (metabolitos, iões, água). Diversos estudos de expressão localizaram a conexina 37 nas células do endotélio vascular, nos monócitos e nos macrófagos, todos intervenientes no processo aterosclerótico.

Diversos estudos genéticos mostraram uma relação entre a variante rs618675 no gene da conexina 37 (Cx37) e a aterosclerose, a DC e a EAM. Em média, em estudos europeus, o risco para a DC foi de 1,32 (IC95%: 1,01-1,72). (178)

III.5.1.5.9 Phosphatase and Actin regulator 1 (PHACTR1)

Localizado no cromossoma 6, o gene PHACTR1 codifica uma proteína da família de proteínas moduladoras da fosfatase (PP1) e actina, participando na reorganização do citoesqueleto, tubulogénese e consequentemente, na manutenção e integridade das células endoteliais. (198)

A variante rs1332844, neste gene, esteve associada a uma menor expressão da proteína nas artérias coronárias humanas. Esta variação foi associada, não só com a DC, mas também com a ocorrência de EAM, disseção aórtica e maior mortalidade cardiovascular a 5 anos. O risco desta variante genética para a DC foi de 1,13 (IC95%: 1,09-1,17). (138)

Assim, de acordo com o seu potencial mecanismo de ação na aterosclerose coronária foram selecionadas 33 variantes genéticas, que estão resumidas na tabela 1.

Variante genética						Literatura	
SNP ID	Gene próximo	Alelo Risco	Cr	Posição	Mecanismo de ação	OR	REF
rs1333049	9p21.3	G/C	9	22125504	Celular	1,33	182
rs4977574	CDKN2B	A/G	9	22098575	Celular	1,29	76
rs618675	GJA4	T/C	1	34922761	Celular	1,73 [*]	178
rs17228212	SMAD3	T/C	15	65245693	Celular	1,21 [*]	182
rs17465637	MIA3	A/C	1	222650187	Celular	1,20	182
rs12190287	TCF21	G/C	6	134256218	Celular	1,08	76
rs3825807	ADAMTS7	A/G	15	76876166	Celular	1,08	76
rs11556924	ZC3HC1	T/C	7	130023656	Celular	1,09	76
rs1332844	PHACTR1	T/C	6	12927312	Celular	1,13	78
rs2114580	PCSK9	G/A	1	55167236	Lípidos	1,49 [*]	178
rs20455	KIF6	C/T	6	39357302	Lípidos	1,50	294
rs7412/ rs429358 ¹	APOE ¹	E4	19	44908822/ 44908684	Lípidos	1,06	126
rs964184	ZNF259	C/G	11	116778201	Lípidos	1,13	76
rs599839	PSRC1/ SORT1	G/A	1	109279544	Lípidos	1,29	182
rs3798220	LPA	T/C	6	160540105	Lípidos	1,51	76
rs1801133	MTHFR 677	C/T	1	11796321	Oxidação	1,19 [*]	181
rs1801131	MTHFR 1298	A/C	1	11794419	Oxidação	1,71	181
rs705379	PON -108	C/T	7	96324583	Oxidação	1,23	224
rs662	PON192	A/G	7	95308134	Oxidação	1,10	224
rs854560	PON55	T/A	7	95316772	Oxidação	1,44	224
rs6922269	MTHFD1L	G/A	6	150931849	Oxidação	1,23	182
rs5186	AT1R	A/C	3	148742201	SRAA	1,65 [*]	156
rs699	AGT	T/C	1	230710048	SRAA	1,87 [*]	153
rs4340	ACE	I/D	17	61565892	SRAA	1,34	147
rs4402960	IGF2BP2	G/T	3	185793899	Diab/Obes	1,14 ²	167
rs1326634	SLC30A8	C/T	8	117172544	Diab/Obes	1,12 ²	167
rs266729	ADIPOQ	C/G	3	186841685	Diab/Obes	1,09	165
rs7903146	TCF7L2	C/T	10	112998590	Diab/Obes	1,29	169
rs17782313	MC4R	T/C	18	60183864	Diab/Obes	1,18 ²	173
rs1801282	PPARG	C/G	3	12351626	Diab/Obes	1,14 ²	167
rs1884613	HNF4A	C/G	20	44351775	Diab/Obes	1,11 ²	174
rs8050136	FTO	C/A	16	53782363	Diab/Obes	1,17 ²	167
rs1376251	TAS2R 50	G/A	12	11030119	Diab/Obes	1,71 [*]	178

Legenda: A Tabela 1 mostra as 33 variantes genéticas associadas com a DC analisadas na nossa população de 2888 indivíduos. O potencial mecanismo de ação é baseado no que se sabe acerca da função do gene próximo. Inclui "Celular" (genes associados ao ciclo celular, migração e inflamação); "Oxidação" (genes envolvidos no stress oxidativo); genes associados com os FRT como o metabolismo lipídico; o SRAA e a DM/Obesidade. 'SNP – Single Nucleotide Polymorphism'; Cr – Cromossoma; OR – Odds Ratio; SRAA – Sistema Renina-Angiotensina-Aldosterona; Diab/Obes – Diabetes/Obesidade; ¹Resultante de um Haplótipo; ²OR relativamente ao risco de diabetes/obesidade

III.5.2 ANÁLISE GENÉTICA

A análise genética deste estudo foi efetuada no Laboratório de Genética Humana da Universidade da Madeira, sob a supervisão do Prof. Doutor António Brehm.

O ADN genómico de todos os indivíduos foi extraído a partir de 80 µl de sangue periférico, usando o método de fenol-clorofórmio, com precipitação por ação do etanol.

A análise das variantes genéticas em estudo foi efetuada em placas de 96 poços (Micro Amp), utilizando sondas oligonucleotídicas marcadas com fluorescência específica para cada um dos alelos, num ensaio que combina a técnica convencional de PCR e a técnica TaqMan (Applied Biosystems). As sondas e os “primers” usados para cada polimorfismo foram pré-estabelecidos pelo fornecedor (TaqMan SNP Genotyping Assays, Applied Biosystems) (Tabelas 2 e 3).

Tabela 2: Sequências de genotipagem e os respetivos “primers”/sondas pré-estabelecidos.

SNP ID	Gene	Sequência de Genotipagem
rs4340	ACE	GGTGAGCTAAGGGCTGGAGCTCAAG[C/G]CATTCAAACCCCTACCAGATCTGAC
rs699	AGT	CAGGGTGCTGTCCACACTGGCTCCC[A/G]TCAGGGAGCAGCCAGTCTCCATCC
rs5186	ATR1	TGCAGCACTTCACTACCAATGAGC[A/C]TTAGCTACTTTTCAGAATTGAAGGA
rs662	PON 192	TAAACCCAAATACATCTCCCAGGAT[C/T]GTAAGTAGGGGTCAAGAAAATAGTG
rs705379	PON -108	GGGGCAGCGCGATTGGCCGCCCC[A/G]CCCCCTCCCGCCGGTCCGGCAGCTA
rs854560	PON 55	GCCAGTCCATTAGGCAGTATCTCCA[A/T]GTCTTCAGAGCCAGTTTCTGCCAGA
rs1333049	9p21.3	CATACTAACCATATGATCAACAGTT[C/G]AAAAGCAGCCACTCGCAGAGGTAAG
rs7412	APOE	CCGCGATGCCGATGACCTGCAGAAG[C/T]GCCTGGCAGTGACCAGGCCGGGGC
rs429358	APOE	GCTGGGCGCGGACATGGAGGACGTG[C/T]GCGGCCGCTGTGTCAGTACCDCGG
rs1801133	MTHFR 677	GAAAAGCTGCGTGATGATGAAATCG[G/A]CTCCCGCAGACACCTTCTCCTTCAA
rs1801131	MTHFR 1298	AAGAACGAAGACTTCAAAGACACTT[G/T]CTTCACTGGTCACTCTCCCCCA
rs7903146	TCF7L2	TAGAGAGCTAAGCACTTTTATGATA[C/T]TATATAATTAATTGCCGTATGAGG
rs1884613	HNF4A	ATTGGCTTGTGGACATCCGCCCCC[C/G]ACCCCACTGCATGGTACTGGAGCC
rs8050136	FTO	CATGCCAGTTGCCACTGTGGCAAT[A/C]AATATCTGAGCCTGTGGTTTTGTTG
rs618675	GJA4	TTCCGGAGCTCAGCAGATTTCAACA[C/T]TTTACAACCTGCTTCAACCTGGTT
rs2114580	PCSK9	GTTTCTCTATTCTGTCCAAITC[A/G]TACATGACACAGCTGTGGGAAACC
rs1376251	TAS2R50	GAGCTGCATCTTCTGAGATGTTA[C/T]ACAGAGAACAGATTAGCATCAGAAA
rs20455	KIF6	CTTTTGTAGAGAGGAGTTGGGACC[A/G]TTCATGCTGGGAGTCAGATGCTGG
rs4402960	IGF2BP2	AGTAAGGTAGGATGGACAGTAGATT[G/T]AAGATACTGATTGTGTTTGCAAACA
rs13266634	SLC30A8	TGCTTCTTTATCAACAGCAGCCAGC[C/T]GGGACAGCCAAGTGGTTCGGAGAGA
rs266729	ADIPOQ	TTGCAAGAACCGCTCAGATCTGCG[C/G]CTTCAAAAACAAAACATGAGCGTGC
rs17782313	MC4R	GTTTAAAGCAGGAGAGATTGTATCC[C/T]GATGGAAATGACAAGAAAAGCTTCA
rs1801282	PPARG	AACTCTGGGAGATTCTCCTATTGAC[C/G]CAGAAAGCGATTCTTCACTGATAC
rs4977574	CDKN2B	GGGTACATCAAATGCATTCTATAGC[A/G]CAGGATGTTCCAGTCACTTAACAA
rs3825807	ADAMTS7	CGCTGTCCCAACGCTCCCGTCGAG[A/G]CTCCAGCTCTGGGTACACTGGAGGC
rs3798220	LPA	GCTCCAAGAACAGCCTAGACACTTC[C/T]ATTTCTGAACATGAGATTCCGAGGT
rs964184	ZNF259	TCACCATCTGATGACTGTTTTCT[C/G]ATCTGTTTATTGTCATTTTCCCA
rs17228212	SMAD3	ACACTGTCTTTGCCGTCATTGAACT[C/T]GCAACCTAACTGCTGAGTGAGGACA
rs6922269	MTHFD1L	CAGTTTTTACTGTAAGTCCAATAA[A/G]TAATACTCATCTTTAAAAAGACATC
rs17465637	MIA3	ACCATAATAGTTATGCTGAGAAGTT[A/C]TTTTTTGTCATAGTCAAGATAACA
rs1332844	PHACTR1	ACTTGCAATATGATTTCTCTCA[C/T]GTTATATGTGTGCTGGTGGTCAACC
rs599839	PSRC1	AAGAGAAAAGAAATAGGAGCAGGATC[A/G]ACTTCCAGATATACAGAGAATATAA
rs11556924	ZC3HC1	GGGGCTAGCAGCCTCTGGCTCAGGA[C/T]GGTCAACAGGACTGGAAGAGTCCCA
rs12190287	TCF21	CTCCAAGGGCTGAGAACCTCGGTGA[C/G]TTCATCCACCTGTCTATTGACAT

Os oligonucleótidos foram sintetizados e os marcadores fluorogênicos FAM e VIC acoplados às extremidades 5' das sondas, de modo a alcançar a discriminação alélica. A reação de polimerização em dois passos consistiu em 40 ciclos de desnaturação a 92°C durante 15 segundos e de “primer annealing” e extensão a 60°C durante 1 minuto e foi efetuada num aparelho 7300 Real-Time PCR System (Applied Biosystems).

Os genótipos foram determinados, recorrendo ao Software 7300 System SDS (Applied Biosystems), sem qualquer conhecimento prévio dos dados clínicos individuais.

A qualidade da genotipagem foi verificada através da inclusão de um controlo negativo em cada placa de 96 poços, de duplicados cegos para aproximadamente 20% de todas as amostras e confirmação dos genótipos por sequenciação *Sanger*[®] em 10% de amostras.

Tabela 3: Lista de “primers” usados para a reação de sequenciação do fragmento amplificado correspondente a cada gene.

SNP ID	Gene	Primers (Forward / Reverse)
rs4340	ACE	GTCACITTTATGTGGTTTCGCCAAT / GCAGCCTGGTTGATGAGTTC
rs699	AGT	CTCCAAGGCCTGACTGGCTGAT / TCTCTATCTGGGAGCCTTGGGA
rs5186	AGTR1	TGGATGTATTGATTCAACTAGGCATCA / GGCTAGGGAGATTGCATTTCTGTCA
rs662	PON1	TCACCCCTGAAAAATTAAGACATT / AAGGATTGTATCGGCAGGACTA
rs705379	PON1	TAACAGCCTGGACCCAACTTTCTG / AGTGCTCACCAAAGCTTGACT
rs854560	PON1	AGCACAAGTGAAATGCATACACAA / CATGGCTTTTGTACGTTTTGTGAA
rs7412	APOE	CTCTTGGGTCTCTCTGGCTCAT / CTGCCATCTCCTCCATCCG
rs429358	APOE	CTCTTGGGTCTCTCTGGCTCAT / CTGCCATCTCCTCCATCCG
rs1801133	MTHFR	TCAGCGAACTCAGCACTCCAC / CACCTCCCCTGAGAGTCATCTCTG
rs1801131	MTHFR	TTCTCCCTTTGCCATGTCCACAG / CCTGCAGACCTTCTTGCAAATA
rs1376251	TAS2R50	ACAGATGCACACATAGATACACT / ACTCTGTCCCTGATATCTTTCTGA
rs20455	KIF6	CTACCCACTGTCCCACCACCA / TTCAAGCTAACACTGTGAAACTCCTT
rs13266634	SLC30A8	AGACAAAGTACTTGAAGTTGGAGTCA / GCTGTACTTTCGGCTCCACTCA
rs1801282	PPARG	CAAGCCCAGTCTTTCTGTGTTTA / GAAATGGGAGCCATGCACAGAGA
rs3825807	ADAMTS7	CTAGTGGCATGACACCCATATGG / CCAGAGAGAGCTCACTGAGGAGAG
rs3798220	LPA	ATAGACAGTTCCTCCCACTGC / GATGAAACAGATTGAGAACAGACCTT
rs17228212	SMAD3	CCTGGGTTAGTTTACTGGGCTGAGA / CTCACGCAGGACCGGATCAA
rs11556924	ZC3HC1	GTTTGTCCCTGTATCACACGGTTGTT / CGTGTACTTCCATTATCTCCCTTT
rs12190287	TCF21	GGTTATGACCGCAGTGTTTACAAGA / GGAGAGAACACATGAATTTGAGGAA

III.5.3. MODELOS GENÉTICOS

A penetrância da doença associada a um determinado genótipo é o risco da doença em indivíduos portadores desse genótipo. A maneira como herdamos cada genótipo obedece a diferentes modelos, padrões genéticos para a penetrância da doença. Assim, para cada variante genética, foram testados quatro modelos genéticos: multiplicativo, aditivo, recessivo e dominante, com recurso a tabelas de contingência e regressão logística, conforme exposto na Tabela 4.

Considerando um marcador genético constituído por um único locus com dois alelos A e B, os possíveis genótipos são: A/A, A/B e B/B. O risco para a DC é estimado pela diferença, observada nos genótipos ou alelos entre casos e controlos.

Tabela 4: Tabela de contingência para a análise de casos-controlos, por diferentes modelos genéticos.

	AA	AB	BB
Casos	a	b	c
Controlos	d	e	f
Modelo dominante: o alelo B confere risco (AA versus AB+BB)			
	AA	AB + BB	
Casos	a	b + c	
Controlos	d	c + f	
Modelo recessivo: 2 cópias do alelo B são requeridas para conferir risco (AA+AB versus BB)			
	AA + AB	BB	
Casos	a + b	c	
Controlos	d + e	f	
Modelo multiplicativo: o risco aumenta γ -vezes para o genótipo AB e γ^2 -vezes para o BB			
	A	B	
Casos	2a + b	b + 2c	
Controlos	2d + e	e + 2f	
Modelo aditivo: o risco aumenta γ -vezes para o genótipo AB e 2γ -vezes para BB			
	AA	AB	BB
Casos	A	b	c
Controlos	D	e	f

Legenda: a, b, c, d, e, f são as frequências dos genótipos observados nos casos e controlos. γ é um parâmetro de penetrância genética.

III.6. SCORE DE RISCO GENÉTICO

Com a construção de um SRG pretende-se sumarizar o risco genético individual, baseado em múltiplas variantes genéticas analisadas.

As variantes genéticas, a serem incluídas na construção do SRG, devem respeitar o equilíbrio de *Hardy-Weinberg* e não se encontrar em desequilíbrio de *linkage* com outra variante genética. Pelo exposto, as variantes genéticas da LPA e do CDKN2B não foram incorporadas no cálculo do SRG.

O número de variantes genéticas, a incluir no SRG, foi determinado pelo melhor poder preditor expresso pela AUC, testada com todas as variantes genéticas, com apenas as variantes genéticas significativas na análise bivariada e com apenas as variantes genéticas significativas na análise de regressão logística multivariada.

A construção do SRG pode ou não ser ponderada, consoante se pondera o risco de cada alelo ou genótipo pelo seu peso individual. (199,200)

No modelo não ponderado, também denominado de modelo aditivo, é calculada apenas a soma do número de alelos de risco. Este representa o modelo básico de construção do SRG. O número de cópias do alelo de risco que um indivíduo pode ter, varia entre zero e dois. O score calculado varia entre zero e o dobro do número de *loci* incorporados na análise. Este modelo assume que, o peso relativo de cada alelo é igual, e que o efeito no risco é proporcional à soma das partes, sendo esta a sua principal limitação. Na nossa amostra o número máximo de alelos de risco, possuídos por um indivíduo, seria o dobro do número de variantes incluídas.

O modelo ponderado ou multiplicativo tem sido o mais testado nos últimos anos, mas ainda sem consenso científico, quanto à forma de ponderação. Na literatura, estão descritas diversas formas; ou pelo respetivo efeito publicado de cada alelo numa população semelhante, ou pelo coeficiente individual dado pela regressão logística utilizando cada variante como co variável. Outras formas de ponderação descritas são: a utilização do efeito obtido pelo melhor modelo genético descrito acima ou pelo logaritmo do OR obtido na própria população. No global, a ponderação nos SRG parece apresentar a limitação de extrapolar o risco nos subgrupos mais elevados.

Dada a ausência de uma forma definitiva de cálculo do SRG, realizámos a modelagem do mesmo, na nossa população, testando todos os modelos possíveis. Assim foram testados seis modelos de SRG:

1. SRG aditivo que corresponde à soma do nº de alelos de risco. (\sum Nº ALELOS DE RISCO)
2. SRG multiplicativo genótipo (SRG MG) que corresponde ao produto dos OR dos genótipos pressupondo a codominância. (\prod OR GENÓTIPOS)
3. SRG multiplicativo melhor modelo genético (SRG MM) que corresponde a soma do produto dos OR do melhor modelo genético pelo número de alelos em cada variante. (\sum OR MELHOR MODELO GENÉTICO X Nº ALELOS)
4. SRG multiplicativo B (SRG B) que corresponde a soma do produto do coeficiente B da regressão logística pelo número de alelos. (\sum BETA LOGISTICA X Nº ALELOS). Obteve-se o Beta pela análise de regressão logística simples, para cada variante genética.
5. SRG multiplicativo Log (SRG log L) que corresponde a soma do produto do logaritmo dos Odds ratio da literatura pelo número de alelos de risco em cada variante. (\sum LOG (OR LITERATURA) X Nº ALELOS em cada variante genética.
6. SRG multiplicativo Log (SRG log P) que corresponde à soma do produto do logaritmo do OR do melhor modelo genético da nossa amostra, pelo número de alelos de risco em cada variante. (\sum LOG (OR MODELO) X Nº ALELOS)

Após a modelagem do SRG, testaram-se as medidas de calibração e discriminação dos modelos de SRG com uma análise de sensibilidade e especificidade para cada modelo de SRG, de forma a determinar qual o modelo final a utilizar em análises subsequentes.

Para validação dos cálculos do SRG selecionámos uma amostra aleatória de 20% de indivíduos, que utilizamos como amostra para validação. Nesta amostra aleatória foi recalculada a calibração, a sensibilidade e a especificidade do SRG, de acordo com o modelo escolhido e comparado com os valores previamente obtidos, de forma a verificar disparidade ou sobreposição de resultados.

III.7 ANÁLISE ESTATÍSTICA

III.7.1 ANÁLISE UNIVARIADA E BIVARIADA

Na análise estatística, as variáveis foram expressas em percentagem (%) ou média \pm desvio-padrão, quando a distribuição foi normal, avaliada pelo teste de *Kolmogorov-Smirnov*. Assim, as características gerais da população foram comparadas entre os casos e controlos através do teste paramétrico do Qui-quadrado para as variáveis categóricas, e para as variáveis contínuas, através do teste paramétrico t de *Student*, se comparando dois grupos e o teste *Anova One-way*, se comparando diferenças entre mais de dois grupos.

Nas distribuições assimétricas, os dados foram apresentados através da mediana e amplitude interquartil (percentil 25-percentil 75). Mesmo com a violação dos pressupostos de homogeneidade ou normalidade, os testes paramétricos t Student e Anova One Way são geralmente robustos sobretudo em amostras grandes como é o caso da população em estudo. No entanto, para as variáveis expressas pela mediana, como por exemplo as variáveis bioquímicas e o álcool, que tiveram distribuição assimétrica, recorreremos ao teste não-paramétrico de *Mann-Whitney* para comparação de duas amostras ou de *Kruskal Wallis* para comparação de mais de duas amostras.

A fim de testarmos a existência ou não de uma relação linear entre duas variáveis contínuas, fizemos a correlação linear de *Pearson* (paramétrica), se ambas as variáveis tivessem distribuição normal, caso contrário foi feita a regressão linear simples.

O coeficiente de correlação, ρ , varia entre -1 e 1. Se $\rho > 0$, existe uma relação linear direta entre as variáveis em estudo. Se $\rho < 0$, existe relação linear inversa entre as variáveis em estudo. Se $\rho = 0$, existe independência linear entre as variáveis em estudo.

III.7.2 ODDS RATIO (OR)

As frequências genóticas foram determinadas a partir das contagens observadas e comparadas com a análise Qui-quadrado (análise bivariada). As frequências alélicas foram calculadas a partir das frequências dos genótipos e comparadas entre os casos e os controlos, também utilizando o teste Qui-quadrado.

Foram calculados os odds ratios (OR) e respetivos intervalos de confiança de *Cornfield* 95% para os FRT, para cada variante genética e para o SRG em geral. Calculámos ainda os OR para diferentes valores limite do SRG, nomeadamente em quartis ou em decis.

O OR é uma medida de associação entre o fator de risco e a doença, medindo quantas vezes o risco de um evento é maior numa ou noutra situação. Quando OR é igual a 1 significa que existe uma distribuição equivalente do fator nos dois grupos. É equivalente a dizer que não existe associação ou que as categorias são independentes. Para valores de OR superiores a 1, a frequência de exposição é maior entre os casos, indicando um risco aumentado de doença com a exposição. Assim, quanto mais forte for a associação entre a exposição e a doença, maior será o OR. Se o OR for inferior a 1, significa que a frequência de exposição é mais baixa entre os casos, indicando que a exposição é um fator protetor em relação à doença.

A associação entre o alelo e o risco de ocorrência de DC denomina-se OR alélico. No caso de avaliarmos o risco conferido por um determinado genótipo versus os outros dois genótipos possíveis, denominamos de OR genotípico.

III.7.3 EQUILÍBRIO DE HARDY-WEINBERG

Para sabermos se a nossa amostra está em equilíbrio de *Hardy-Weinberg*, é necessário aplicarmos um teste estatístico que nos informará, se as frequências genóticas observadas são estatisticamente diferentes, ou não das esperadas. Assim, o teste de equilíbrio de *Hardy-Weinberg* foi realizado a todas as variantes genéticas estudadas com recurso ao teste do Qui-quadrado de *Pearson*.

Sendo o nível de significância de 5% para um grau de liberdade = 3,84, se o valor obtido for inferior a 3,84, não rejeitamos a hipótese nula (H_0), logo a população estará em equilíbrio de *Hardy-Weinberg*.

Para a avaliação do equilíbrio de *Hardy-Weinberg* utilizou-se a correção para as análises múltiplas pela Correção de Bonferroni ($p < 0,05 / \#genes$ (33)=0,0015). As variantes genéticas que não demonstrem equilíbrio de *Hardy-Weinberg* não foram utilizadas para o cálculo do SRG.

III.7.4 ANÁLISE MULTIVARIADA

III.7.4.1 MODELO DE REGRESSÃO LOGÍSTICA

O modelo de regressão logística tem sido amplamente utilizado para a identificação de preditores independentes, bem como para estimar a força de associação com a variável resposta expressa pelo OR.

O objetivo da análise de regressão logística é encontrarmos o melhor ajustamento, de forma a ser o mais parcimonioso, quer do ponto de vista estatístico, quer do ponto de vista clínico.

Quando se pretendeu avaliar as relações de potenciais fatores com variáveis categóricas (dicotómicas), utilizou-se a regressão logística simples.

Quando se pretendeu avaliar as relações de múltiplas variáveis dependentes e/ou independentes recorreu-se, então, aos métodos estatísticos de regressão logística multivariável. Assim, a análise multivariada permitiu explorar a influência ou importância das variantes genéticas ou do SRG, na presença dos FRT.

As variáveis candidatas à análise multivariada foram selecionadas pela análise bivariada, utilizando o método proposto por Hosmer & Lemeshow, 2013 (201), em que qualquer variável que apresentar um valor- $p < 0,15$ na análise univariada deverá ser candidata ao modelo multivariado, assim como as variáveis com importância biológica, mesmo que não significativas.

Nos *softwares* utilizados existem métodos automáticos que podem ser utilizados, na decisão de inserir e remover variáveis. Seguidamente, descrevem-se os métodos implementados no *SPSS inc*, 2007:

- 1– *Enter*: é um procedimento para a seleção de variáveis, em que todas elas em bloco entram no processo uma única vez;
- 2– *Forward*: Método de selecção Stepwise, este procedimento inicia-se com um modelo que não contenha variáveis explicativas. A ideia do método é adicionar uma variável de cada vez, seleccionando em primeiro lugar aquela que apresentar um valor de correlação mais elevado, em módulo, com a variável resposta, e assim consequentemente, até que o processo para, quando o aumento do coeficiente de determinação, devido à inclusão de uma nova variável explicativa no modelo, não é mais importante.
- 3– *Forward (condicional)*: baseado na significância da estatística de pontuação e testes de remoção com base na probabilidade de uma estatística de razão de verosimilhança, com base em estimativas de parâmetros condicionais;
- 4– *Forward (Likelihood Ratio)*: baseado na significância da estatística de pontuação e testes de remoção com base na probabilidade de uma estatística de razão de verosimilhança, baseada na máxima verosimilhança parcial das estimativas;
- 5– *Forward (Wald)*: Método de selecção Stepwise baseado na significância da estatística de pontuação

e testes de remoção com base na probabilidade da estatística de *Wald*;

- 6– *Backward*: Enquanto o método *Forward* começa sem nenhuma variável no modelo e adiciona variáveis a cada passo, o método *Backward* faz o oposto. Este incorpora, inicialmente, todas as variáveis e ao longo do processo cada uma pode ou não ser eliminada. A primeira variável a ser removida é aquela que apresenta um menor coeficiente de correlação parcial com a variável resposta.

O método preferencial que utilizaremos para a seleção das variáveis é o método passo a passo *Stepwise*, no qual as variáveis são selecionadas por inclusão e/ou exclusão no modelo de forma sequencial, com base em algoritmos estatísticos, que verificam a importância dessas mesmas variáveis.

Este método *Stepwise* é um dos métodos mais utilizados e consiste na combinação dos dois métodos descritos (*Forward* e *Backward*). Este inicia com uma variável (a que apresentar maior correlação com a variável resposta), e a cada passo do *Forward*, depois de incluir uma variável, aplica o *Backward* para ver se será descartada alguma variável. Continua-se o processo até este não incluir ou excluir nenhuma variável.

Este método considera uma regra de decisão com um valor-p de entrada e um valor-p de saída. Os limites máximos e mínimo, que iremos fixar para inclusão e exclusão das variáveis no modelo *Stepwise*, são os sugeridos por *Hosmer & Lemeshow*, 2013 (201) isto é 0,15 como valor-p de entrada e 0,20 como valor p de saída.

A significância dos modelos obtidos foi analisada pelo teste da razão das verossimilhanças.

A qualidade do ajuste do modelo logístico foi avaliada pela determinação da estatística do qui-quadrado e do valor-p do teste de *Hosmer & Lemeshow* com o teste de qualidade do ajustamento (*goodness of fit*). (202)

III.7.4.1.1 ANÁLISE DOS RESÍDUOS E VALIDAÇÃO DOS PRESSUSPOSTOS DO MODELO

Quando existem preditores quantitativos relativamente frequentes, em alguns contextos clínicos como é o caso da DC, são necessários cuidados adicionais na etapa de análise do modelo para minimizar potenciais vieses decorrentes de observações influentes. Estas estão usualmente associadas a variáveis contínuas com valores extremos nos preditores comuns.

Adotaram-se procedimentos adicionais, denominados de análise de resíduos necessários para o ajuste do modelo, visando minimizar vieses nas estimativas das associações. Os métodos usados para a análise de resíduos mais extensamente descritos na literatura são atribuídos a *Hosmer & Lemeshow*, 2013 (201) e estão disponíveis em diversos procedimentos computacionais como o *R* e o *SPSS*.

Na nossa análise, procedeu-se à validação dos pressupostos em vários passos: análise gráfica dos resíduos e diagnóstico de casos influentes.

Na regressão logística a validação do modelo é realizada através da diferença entre o valor observado e o valor ajustado. Esta diferença foi medida com recurso à análise dos resíduos de *Pearson* e/ou resíduos da função desvio para detetar *outliers* e aos resíduos *DfBetas* e à distância de *Cook* para detetar observações influentes.

A existência de multicolinearidade foi analisada com recurso à avaliação da matriz de correlações e expressa pela tolerância ou pelo seu recíproco o fator de inflação de variância, *Variance Inflation Factor* (VIF). Um valor de tolerância menor que 0,2 sugere que pelo menos 80% da variância de uma variável independente é explicada por outras variáveis independentes. Um VIF superior a dez foi considerado evidência de colinearidade. Um valor de VIF superior a 5 foi considerado como possível colinearidade.

III.7.4.2 ANÁLISE DE SOBREVIVÊNCIA

Realizamos uma análise de sobrevivência na população com DC, seguidos prospetivamente para avaliação de eventos CV adversos.

Neste grupo foi registado o tempo decorrido desde o instante inicial, a entrada no estudo, até à ocorrência do evento. Este tempo designa-se por tempo de sobrevida. Os dados censurados ocorreram quando para alguns indivíduos, não foi observado nenhum evento de interesse, durante o período em que estão em observação.

Através de uma análise não paramétrica, recorremos às curvas de *Kaplan-Meier* (KM) que traduzem a probabilidade de sobrevivência em cada um dos tempos em estudo. Para comparar as curvas de KM de dois ou mais grupos de indivíduos com características diferentes, recorreu-se ao teste de *Log-rank*, sendo que se considerou existir diferenças entre as probabilidades de sobrevivência para valores p inferiores a 0,05, de acordo com *Hosmer & Lemeshow*, 2008. (202)

Através de técnicas semi-paramétricas, ajustaram-se modelos de riscos proporcionais de *Cox* para avaliar o efeito de várias variáveis na sobrevida. A medição do efeito das variáveis explicativas foi efetuada, estimando a razão de riscos, *hazard ratio* (HR) e os respetivos intervalos de confiança a 95% (IC 95%). Tendo em conta que este tipo de modelos assume que a razão entre o risco de ocorrência do evento para quaisquer dois indivíduos é constante no tempo, validou-se o pressuposto de riscos proporcionais, através do teste de *Harrell*. Assim para valores p superiores a 0,05, assumimos a validação do pressuposto.

Avaliou-se a qualidade de ajustamento e calibração do modelo através do índice de prognóstico do modelo (preditor linear do modelo de *Cox*), sendo que se comparou as funções de sobrevivência empírica e as funções de sobrevivência de KM para grupos de alto, médio e baixo índice de prognóstico. (202)

III.7.5 MEDIDAS DE DISCRIMINAÇÃO E RECLASSIFICAÇÃO NUMÉRICA DOS MODELOS

III.7.5.1 CAPACIDADE DISCRIMINATIVA PARA DOENÇA CORONÁRIA: CURVAS ROC

A acuidade preditiva dos modelos finais da análise multivariada foi verificada mediante análise da capacidade discriminativa do modelo pela determinação da área sob as curvas, *Area Under Curve* (AUC) das curvas *Receiver Operating Characteristics* (ROC).

As curvas (ROC) descrevem a capacidade discriminativa de um teste diagnóstico para um determinado número de valores *cut off point*. Isto permite pôr em evidência os valores para os quais existe maior otimização da sensibilidade em função da especificidade. O ponto, numa curva ROC, onde isto acontece é aquele que se encontra mais próximo do canto superior esquerdo do diagrama. Por outro lado, as curvas ROC permitem quantificar a exatidão de um teste diagnóstico, já que, esta é proporcional à área sob a curva ROC, isto é, é tanto maior quanto mais a curva se aproxima do canto superior esquerdo do diagrama. Sabendo isto, a curva será útil também na comparação de testes diagnósticos, tendo um teste uma exatidão tanto maior, quanto maior for a área sob a curva ROC. A calibração do modelo foi avaliada pela determinação da estatística do qui-quadrado e do valor-p do teste de *Hosmer & Lemeshow* com o teste de qualidade do ajustamento (*goodness of fit*). (202)

Uma vez que o *software* SPSS não realiza a comparação de AUC das curvas ROC, recorreu-se ao método não paramétrico teste de *DeLong* (203), usando o programa MedCalc versão 14.10.2.0.

III.7.5.2 NET RECLASSIFICATION INDEX (NRI) E INTEGRATED DISCRIMINATORY IMPROVEMENT (IDI)

As medidas de melhoria de reclassificação global, *Net reclassification improvement* (NRI) e melhoria integrada discriminatória, *Integrated discriminatory improvement* (IDI) representam medidas adicionais para avaliar as mudanças, na previsão de risco associadas a um novo marcador ou teste, quando comparado a uma medida padrão. Em geral, eles são mais sensíveis a mudanças na capacidade discriminatória de um novo preditor, quando comparado com a C-estatística.

O NRI mede o número de indivíduos, cujo risco predito de desenvolver a doença se altera com um novo marcador de risco ou teste. Ao contrário da C-estatística, ele mede diretamente as mudanças na previsão de risco para cada participante, e depois então calcula a percentagem de indivíduos que mostra uma mudança adequada no risco. O NRI é calculado usando a seguinte fórmula (204):

$$\text{NRI} = [(\# \text{ eventos sobem categoria} / \# \text{ de eventos}) - (\# \text{ eventos descem de categoria} / \# \text{ de eventos})] + [(\# \text{ não eventos descem de categoria} / \# \text{ de não eventos}) - (\# \text{ não eventos sobem de categoria} / \# \text{ de não eventos})]$$

Os métodos estatísticos do NRI podem ser utilizados de forma categórica (categórica) ou livre de categorias (contínuo). O NRI categórico utiliza categorias de risco predefinidas para a população e caracteriza as mudanças no risco como mudanças de uma categoria para a outra. Diferentes categorias devem representar diferenças clinicamente significativas no risco.

Por exemplo, usando a previsão do risco de *Framingham*, os doentes foram classificados quanto ao risco de eventos CV a 10 anos em grupos de baixo risco (<10%), moderado (10-20%) ou alto (> 20%); com diferentes alvos terapêuticos. Utilizando estes limiares de risco, calculamos o NRI categórico para determinar a percentagem de indivíduos que demonstram uma alteração adequada na categoria de risco (por exemplo, de <10% para 10-20%) com base no novo marcador de risco, ou seja o SRG. A principal limitação deste método é depender do número de categorias e dos limites impostos a cada uma delas.

Quando as categorias de risco não podem ser estabelecidas numa população, ou quando a prevalência da doença é desconhecida na população estudada, pode ser calculado um NRI categórico livre (contínuo), que representa a percentagem de indivíduos que demonstram qualquer alteração no risco previsto na direção adequada. (204)

Apesar do NRI contínuo avaliar a proporção de participantes nos quais o SRG melhora a predição de risco basal, este não testa o grau de melhoria com impacto clínico, pois mesmo pequenas melhorias na capacidade preditiva (dentro da mesma categoria de risco) resultam em mudanças no índice. Esta é a principal limitação deste método estatístico.

O IDI é uma medida da melhoria média na previsão de risco que ocorre com um novo marcador de risco ou teste. O teste mede a melhoria média na sensibilidade, além da melhoria na '1-especificidade', comparando modelos sem (modelo antigo) e com (modelo novo) o novo preditor de interesse. Portanto, o teste caracteriza a melhoria média na sensibilidade de um novo preditor sem especificidade. (205) O IDI pode ser representado da seguinte forma (206):

$$\text{IDI} = (\text{IS modelo novo} - \text{IS modelo antigo}) - (\text{IP modelo novo} - \text{IP modelo antigo})$$

Assim, o IS representa a sensibilidade integral (ou média) e IP representa a "1-especificidade" integral (ou média) de cada modelo. O IDI também pode ser estimado usando as probabilidades preditas de eventos entre modelos (206):

$$\text{IDI} = (\text{probabilidade de um evento no modelo novo} - \text{probabilidade de um evento no modelo antigo}) + (\text{probabilidade de um não evento no modelo antigo} - \text{probabilidade de um não evento no modelo novo})$$

No nosso estudo calculámos o IDI para determinar a melhoria média na sensibilidade preditiva do modelo associado ao SRG.

Ambas as medidas (NRI e IDI) foram obtidas através do pacote 'PredictABEL' disponível no *software* estatístico R (versão 3.2.0).

III.7.5.3 CAPACIDADE DISCRIMINATIVA NA SOBREVIDA– MÉTODO DE C- ESTATÍSTICA

O método de concordância-Estatística (C-estatística) é uma medida discriminativa global dos modelos de regressão logística, avaliando a capacidade discriminatória do modelo distinguir entre os indivíduos que têm um evento versus os que não têm evento na análise de sobrevivência, em todos os possíveis pares de indivíduos que tiveram ou não um evento.

A C-estatística representa a proporção de pares em que a probabilidade de ocorrência de um evento predita pelo modelo foi mais elevada no indivíduo que teve o evento. Se todos os casos tiverem uma predição de risco superior a todos os controlos, a capacidade discriminatória do modelo logístico é perfeita ou C-estatística de 1. Pelo contrário, quando a C- estatística é 0,5 o modelo logístico é incapaz de distinguir entre os que tiveram o evento e os que não tiveram o evento. Entre 0,5 e 0,7 considera-se a capacidade discriminativa limitada; entre 0,7 e 0,8 considera-se a moderada capacidade discriminativa; e acima de 0,8 considera-se uma boa capacidade discriminativa. Quando novas variáveis são adicionadas ao modelo, a mudança da C-estatística representa a capacidade discriminatória adicional veiculada pelas novas variáveis. Nesta tese, a C-estatística foi utilizada para avaliar a capacidade discriminativa do SRG em relação aos FRT. O SRG foi adicionado ao modelo logístico contendo o tabagismo, a dislipidemia, a DM e a HTA. Um terceiro modelo, testado, envolveu adicionar ao modelo outras variáveis (PCR, Clearance Creatinina, IMC, VOP), ao modelo tradicional de estratificação de risco e calcular a variação da C-estatística entre os modelos. Utilizou-se o teste de Homer-Lemeshow para avaliar a robustez dos modelos (*Goodness of fit*) e assegurar que se mantém a calibração do mesmo, apesar da introdução do SRG.

Embora seja uma boa medida discriminativa, a C-estatística apresenta a limitação de apenas se referir ao aumento da proporção de indivíduos corretamente classificados. Um aumento da C-estatística significa uma maior proporção de indivíduos corretamente classificados com tendo a doença, mas não avalia a mudança individual do risco. Assim, indivíduos já classificados como tendo o evento, não apresentarão diferença na C-estatística, mesmo que o seu risco individual seja modificado pela introdução da ou das novas variáveis introduzidas no modelo logístico.

A análise da C-estatística foi feita com o pacote 'survcomp' do *software* R (versão 3.2.0). (207,208)

III.7.6 INTERAÇÕES GENE GENE (G-G) E GENE-AMBIENTE (G-A)

Os métodos estatísticos disponíveis para avaliação das interações G-G e G-A podem ser classificados de acordo com três estratégias (209): modelos de filtragem ou sinergismos, de associação/regressão logística e de mineração de dados, *Data mining*.

Neste estudo aplicamos os modelos de associação, regressão logística e de *Data Mining*, uma vez que, os modelos de filtragem ou sinergismo são cada vez menos utilizados na análise G-G e G-A.

III.7.6.1 MODELOS DE REGRESSÃO LOGÍSTICA

Representa o método habitual para a análise das interações G-G E G-A. Pretende identificar combinações de genótipos ou gene-ambiente, entre duas variáveis que melhorem um modelo para prever a DC. Um modelo é calculado para cada variante isolada e para a sua interação com outra variante ou fator. Verificaram-se as diferenças significativas ($p < 0,05$) entre modelos, através do teste *likelihood ratio* (verosimilhança).

III.7.6.2 MODELOS COM BASE EM DATA MINING POR MACHINE LEARNING

Os modelos baseados em *data mining*, um subgrupo da Inteligência Artificial, pressupõem deixar os dados falar e permitem à análise computacional identificar e selecionar padrões sem assumir qualquer premissa sobre os dados. Existem vários exemplos que incluem métodos de reconhecimento automático de padrões, árvores de decisão, métodos de redução multidimensional e métodos de particionamento, *partitioning*. Destes, o mais popular corresponde a uma aplicação dos métodos de redução multidimensional.

Assim, recorreremos ao método *Multifactorial dimensionality reduction* (MDR) baseado em *data mining*. Este método deteta e caracteriza interações não lineares entre variáveis categóricas ou discretas que são preditivas de um resultado discreto (caso-controlo). O *software* MDR combina a seleção e construção de atributos e classificação com a validação cruzada para fornecer uma abordagem poderosa para modelar interações. (210,211)

O *software* MDR 3.0.2 foi usado, no contexto deste estudo, de modo a avaliar as interações G-G e G-A com a DC e a interação do SRG com os FRT. Após a deteção das interações não lineares, entre as melhores variáveis preditoras num conjunto de treino que corresponde a 1/10 da amostra, o *software* recalcula o valor preditor em outros nove conjuntos (9/10 da amostra) denominados de teste. A validação cruzada corresponde ao número de testes, onde se verifica o encontrado no conjunto de treino, podendo ser a máxima de 10/10 e aceitável a partir de 7/10. Quando a validação cruzada é inferior a 50%, o modelo de treino é considerado sobredimensionado (*overfitting*). A acuidade ponderada corresponde a uma estimativa do poder preditor do modelo.

O melhor padrão de interação é baseado no princípio, tanto da consistência de validação cruzada, como da acuidade ponderada do teste que são maximizados para avaliar a interação do(s) polimorfismo(s) genético(s) com outras variáveis (genéticas ou ambientais) em relação à DC.

O *software* MDR expõe os resultados em modelos interativos, denominados diagramas de entropia ou dendogramas. A força das interações é expressa por uma escala colorida em que a vermelho estão as interações sinérgicas elevadas (epistasis); a laranja as interações sinérgicas de grau moderado, (epistasis); a amarelo estão as interações que são independentes; a verde estão as interações ligeiramente redundantes e a azul a ausência de qualquer interação ou redundância. As variantes genéticas hipostáticas são as que sofrem epistasis (interação) por outra variante genética dominante ou epistática.

III.7.7 SOFTWARE UTILIZADO

O *software* utilizado foi o *Statistical Package for the Social Sciences Software* (SPSS) versão 19.0 (IBM, Armonk, NY, USA) e o programa estatístico R, *R Project for Statistical Computing*, (versão 3.2.0) com utilização de pacotes específicos para análise do NRI e IDI *PredictABEL* e C estatística *survcomp*.

Tal como referido, foram ainda utilizados, o programa *MedCalc* versão 14.10.2.0 para comparação de AUC das curvas ROC. O programa de *machine learning* MDR versão 3.0.2 foi utilizado para avaliação das interações G-G e G-A.

IV. RESULTADOS

O IV capítulo dos RESULTADOS será apresentados em tabelas e figuras, pela seguinte ordem:

A- POPULAÇÃO:

As tabelas 5 a 7 mostram as características clínicas, do exame objetivo e parâmetros bioquímico-laboratoriais expostas por caso-controlo.

B- ESTUDO GENÉTICO:

Descrição dos resultados obtidos no estudo das variantes genéticas. As tabelas numeradas 8 a 12 mostram as diferenças encontradas na análise gene a gene (Bivariada) e na análise multivariada (Regressão logística) antes e após controlo pelos FRT.

C- INTERAÇÃO G-G E G-A:

As tabelas 13 a 16 exemplificam as interações G-G e G-A obtidas pelos estudos de associação e por análise de regressão logística. As tabelas 17 a 20 representam as interações G-A por eixos fisiopatológicos na DC (eixo celular, eixo dislipidemia, eixo DM/obesidade, eixo da oxidação e eixo da HTA). As tabelas 21 a 24 representam modelos de *machine learning* (MDR) para avaliar a interação G-G das 10 variantes genéticas significativas na análise gene a gene e uma análise de interação G-G para todas as 33 variantes genéticas analisadas.

D- SCORE DE RISCO GENÉTICO:

As tabelas numeradas de 25 a 30 mostram a análise exploratória conducente à escolha do modelo de SRG. As tabelas 28 e 29 mostram o número de variantes genéticas a incluir de acordo com o poder preditor estimado do SRG. A figura 20 apresenta o poder preditor do SRG numa amostra aleatória de 20%.

As tabelas 30 a 37 mostram o estudo discriminativo do SRG (aditivo e multiplicativo) caso-controlo em quartis, decis e por género. As tabelas 33 a 35 mostram as correlações e interações do SGR encontradas, com variáveis de interesse (idade, género, FRT e História familiar de DC precoce).

As tabelas 38 e 40 mostram o valor preditivo do SRG após controlo pelos FRT. Nas tabelas 42 e 43 são apresentados os melhores modelos preditores combinados (FRT + SRG + marcador de risco CV) para a DC com recurso ao MDR. São apresentados os resultados de acuidade ponderada e de consistência de validação cruzada para os melhores modelos na tabela 43. O poder discriminativo do SRG (AUC), potencial na reclassificação por categorias de risco (NRI) e índice de Discriminação (IDI) são apresentados nas tabelas 44 a 46.

Nas tabelas 47 e 48 encontramos o estudo comparativo do desempenho do SGR com os novos marcadores de risco CV avaliados por análise de regressão logística multivariada. Nas tabelas 49 e 50 é expressa a comparação do poder preditor do SRG com os FRT e com a História Familiar de DC precoce.

As tabelas 51 a 58 mostram o estudo do SRG no doente jovem (<45 anos de idade) com DC comparado com os controlos. A tabela 53 mostra quais as variantes genéticas associadas com DC com significância no estudo gene a gene. Nas tabelas 56 a 58 é apresentado o poder preditor (AUC), reclassificativo (NRI) e discriminativo (IDI) do SRG nesta subpopulação.

E- PODER PREDITOR DO SRG na RECLASSIFICAÇÃO do risco CV:

Nas tabelas 59 a 67 constam os resultados obtidos com a inclusão do SRG nos subgrupos de risco classificados pelo score Europeu, Framingham DC e Framingham CV. Apresentam-se resultados da reclassificação (NRI) e discriminação (IDI) do risco CV com a introdução do SRG.

F- ESTUDO PROSPETIVO:

As tabelas numeradas de 68 a 70 mostram as características basais, clínicas, bioquímicas e do exame objetivo nos doentes com ocorrência de eventos versus doentes sem eventos CV no seguimento.

Na tabela 71 apresentam-se as diferenças gene a gene entre o grupo com mortalidade CV e o grupo livre do evento no seguimento. Na figura 30 estão expressas as curvas de KM para a mortalidade CV, de acordo com o genótipo da variante genética do ADAMTS7. Nas Figuras 31 a 33 mostram-se as análises de sobrevivência de KM de acordo com a mediana do SRG e de acordo com quartis do SRG. Nas tabelas numeradas de 72 a 79 apresentam-se os preditores multivariados da mortalidade CV e do evento global. O seu poder preditor foi discriminado pela C-estatística exposto na figura 34. Na tabela 80 é apresentado o potencial do SRG para a reclassificação dos doentes na predição de eventos, expresso pelo NRI e pelo IDI.

IV.1 CARACTERIZAÇÃO DA POPULAÇÃO

Na tabela 5 apresenta-se a caracterização da população em estudo, ajustada por género e idade. Avaliaram-se 2888 indivíduos, com predomínio do género masculino (77,8%).

Encontrou-se maior prevalência dos principais fatores de risco CV na população com DC documentada. Com efeito, na população com DC a prevalência dos fatores de risco convencionais variou entre 34% para a DM, 46% para o tabagismo, 71% para a HTA e 90% para a dislipidemia.

Os restantes fatores de risco associados com a DC também foram mais prevalentes no grupo de doentes, contudo para a população controlo encontrou-se menor prevalência dos FRT.

A população com DC, à data da observação, apresentou menor frequência cardíaca basal, valores mais elevados de PAS e mais baixos da PAD e valores mais elevados dos índices de rigidez arterial. Na população com DC, 35,7% apresentou-se no tercil mais elevado da VOP versus 37,5% dos controlos no tercil inferior.

O risco CV de Framingham foi estimado com recurso às três funções de previsão de risco publicadas por D'Agostino *et al*, tendo por base a população Framingham: risco CV a dez anos (Framingham CV), risco de DC a dez anos (Framingham DC) e risco de morte por DC ou EAM (Framingham *Hard DC*).

O risco CV estimado a 10 anos, quer pelo score Europeu (SE) quer pelo score de Framingham, situou a maioria dos casos e controlos no baixo e intermédio risco CV. No entanto, foi sempre mais alto na população com DC do que na população controlo.

Nesta população a média do SRGa foi de $24 \pm 3,6$, sendo mais elevada nos casos do que na população controlo ($24,3 \pm 3,7$ versus $23,5 \pm 3,5$; $p < 0,0001$). A média do SRGm foi de $0,58 \pm 0,65$, mais elevada nos casos do que na população controlo ($0,67 \pm 0,73$ versus $0,48 \pm 0,53$; $p < 0,0001$). Nesta amostra a utilização da mediana revelou igual distribuição. Com efeito, a mediana do SRGm foi de 0,41 (0,25 – 0,67), mais elevada nos casos do que na população controlo (0,47 (0,28 – 0,76) versus 0,36 (0,22 – 0,57); $p < 0,0001$). A mediana do SRGa foi de 24 (21 – 26), também mais elevada nos indivíduos com DC do que no grupo controlo (24 (22 – 27) versus 23 (21 – 26); $p < 0,0001$).

Tabela 5: Características gerais da população (n=2888):

Variáveis	Casos (n=1566)	Controlos (n=1322)	Valor p
Idade, anos	53,3 ± 8	52,7 ± 7,8	0,053
Gênero masculino, n (%)	1238 (79,1)	1010 (76,4)	0,087
HF DC precoce, n (%)	373 (23,8)	167 (12,6)	<0,0001
Fatores de risco modificáveis			
Dislipidemia [†] , n (%)	1414 (90,3)	930 (70,3)	<0,0001
Colesterol total, mg/dl	180 (154 – 213)	205 (181 – 234)	<0,0001*
LDL, mg/dl	104,6(82,8 – 128,7)	127,2 (104,7 – 152,3)	<0,0001*
HDL, mg/dl	41 (35 – 49)	48 (41 – 57)	<0,0001*
Triglicéridos, mg/dl	141 (102 – 210)	121 (89 – 174)	<0,0001*
APO B, mg/dl	93,9 (75,5 – 113,3)	92,5 (43 – 115,8)	<0,0001*
LP(a), mg/dl	20,4 (9,2 – 62)	12,8 (8,8 – 29,3)	<0,0001*
LP(a)>25, n (%)	714 (45,6)	374 (28,3)	<0,0001
Diabetes, n (%)	533 (34)	175 (13,2)	<0,0001
Glicémia	106 (96 – 129)	99 (91 – 109)	<0,0001*
Tabagismo*, n (%)	730 (46,6)	309 (23,4)	<0,0001
HTA, n (%)	1114 (71,1)	700 (53)	<0,0001
PAS, mmHg	137,9 ± 20,8	136,2 ± 18,1	0,024
PAD, mmHg	82,6 ± 11,8	83,9 ± 11,1	0,002
VOP, m/s	8,6 ± 1,9	8,3 ± 1,7	<0,0001
1º Tercil (≤7,5)	6,8 ± 0,6	6,8 ± 0,5	0,770
2º Tercil (>7,5-8,8)	8,1 ± 0,4	8,1 ± 0,4	0,430
3º Tercil (>8,8)	10,6 ± 1,6	10,3 ± 1,7	0,005
VOP>10 m/s, n (%)	296 (18,9)	169 (12,8)	<0,0001
Frequência cardíaca, bpm	68,8 ± 12,5	72,3 ± 11,5	<0,0001
Exercício físico*, n (%)	573 (36,6)	761 (57,6)	<0,0001
Álcool, gr/dia	24,7 ± 49,7	18,2 ± 28,2	<0,0001
Álcool >50gr/semana [‡] , n (%)	653 (42)	586 (44,3)	0,213
Álcool >300gr/semana [§] , n (%)	263 (16,9)	181 (13,7)	0,017
Score de Risco Genético			
SRG aditivo (média)	24,3 ± 3,7	23,5 ± 3,5	<0,0001
SRG aditivo (mediana)	24 (22 - 27)	23 (21 - 26)	<0,0001*
SRG multiplicativo (média)	0,67 ± 0,73	0,48 ± 0,53	<0,0001
SRG multiplicativo (mediana)	0,47 (0,28 – 0,76)	0,36 (0,22 – 0,57)	<0,0001*
Score de risco:			
score Framingham CV**, (%)	24,1 ± 17,4	17,1 ± 13,5	<0,0001
Baixo risco, <10, n (%)	333 (21,3)	519 (39,3)	
10-20, n (%)	481 (30,7)	368 (27,8)	<0,0001
Alto risco, >20, n (%)	752 (48)	435 (32,9)	
score Framingham DC, (%)	14,5 ± 9,8	11,7 ± 8,1	<0,0001
score Framingham Hard DC, (%)	11,6 ± 8,4	9,3 ± 7,8	<0,0001
SCORE Europeu	2,52 ± 2,99	2,15 ± 2,49	<0,0001
Risco Baixo e Intermédio, <5, n (%)	1371 (87,5)	1198 (90,6)	
Alto Risco, 5-10, n (%)	162 (10,3)	104 (7,9)	0,031
Muito Alto Risco, >10, n (%)	33 (2,1)	20 (1,5)	
SCORE Europeu® (morte + eventos)	7,83 ± 9	6,68 ± 7,46	<0,0001

[†]LDL>150 ou HDL<40 ou Triglicéridos>150; ^{*}Mais de 40min/semana; ^{*}Fumadores ou <5 anos de cessação; [‡]Mais de 50gr/semana; [§]Mais de 300gr/semana; [§]IMC>30; HDL – High density lipoprotein; LDL – Low density lipoprotein; HTA – Hipertensão arterial; PAS – Pressão arterial sistólica; PAD – Pressão arterial diastólica; VOP – Velocidade de onda de pulso; Bpm – batimentos por minuto; SRG – Score de risco genético; [®]Score europeu multiplicado por 4 para o homem e por 3 para a mulher. Valores contínuos expressos pela média±desvio padrão ou pela mediana (Percentil 25 – Percentil 75). *Valores obtidos através do teste não paramétrico Mann-Whitney; **Porcentagem média de risco a 10 anos; estatisticamente significativo para p<0,05.

Na tabela 6 destacamos que os indivíduos com DC apresentam valores séricos de leucócitos e de fibrinogénio mais elevados e valores séricos mais baixos de hemoglobina (medianas). A Hcy, considerada quer como variável contínua ou categórica, apresentou medianas mais elevadas no grupo com DC. A PCR apresentou medianas semelhantes entre os dois grupos, sendo que valores elevados dessa variável foram mais frequentes no grupo dos doentes, com significado estatístico.

Verifica-se que, do ponto de vista bioquímico – laboratorial, o grupo com DC apresenta perfil aterogénico e inflamatório mais agressivo.

Tabela 6: Características bioquímico-laboratoriais da população estudada

Variáveis	Casos (n=1566)	Controlos (n=1322)	Valor p
Hemoglobina, g/dl	14,6 (13,8 – 15,4)	14,7 (14 – 15,4)	0,001*
Plaquetas, 10 ³ /µl	215 (183 – 254)	220 (190 – 252)	0,171*
Leucócitos, 10 ³ /µl	7,1 (6 – 8,3)	6,6 (5,6 – 7,8)	<0,0001*
Fibrinogénio, mg/dl	387 (337 – 444)	361 (315 – 409)	<0,0001*
Homocisteína, µmol/L	12,2 (10 – 14,9)	11,4 (9,7 – 13,6)	<0,0001*
>10, n (%)	1166 (74,5)	929 (70,3)	0,012
>30, n (%)	25 (1,6)	13 (1)	0,150
Creatinina, mg/dl	1 (0,87 – 1,1)	0,99 (0,89 – 1,1)	0,831*
Clearance Creatinina ml/min**	94,1 (76,9 – 113)	92,7 (78,9 – 108,7)	0,496*
<60 ml/min, n (%)	123 (7,9)	48 (3,6)	<0,0001
PCR, mg/L	2,7 (1,6 – 6,1)	2,7 (1,2 – 5,3)	<0,0001*
>3, n (%)	648 (41,4)	496 (37,5)	0,035

PCR – Proteína C-reativa; Variáveis contínuas apresentadas pela mediana (Percentil 25 – Percentil 75);
*Valores obtidos através do teste não paramétrico Mann-Whitney; Estatisticamente significativo para $p < 0,05$. ** Cockcroft-Gault

A tabela 7 apresenta os índices antropométricos da população. O IMC foi mais elevado na população com DC, sendo também mais significativo o número de doentes com excesso ponderal ou obesidade. Os restantes índices antropométricos de avaliação da massa corporal da obesidade central foram também mais elevados na população com DC, sobretudo no género masculino.

De salientar que, em mais de 90% dos casos e dos controlos, o perímetro abdominal encontrava-se elevado em ambos os géneros, indicando elevados índices de obesidade central.

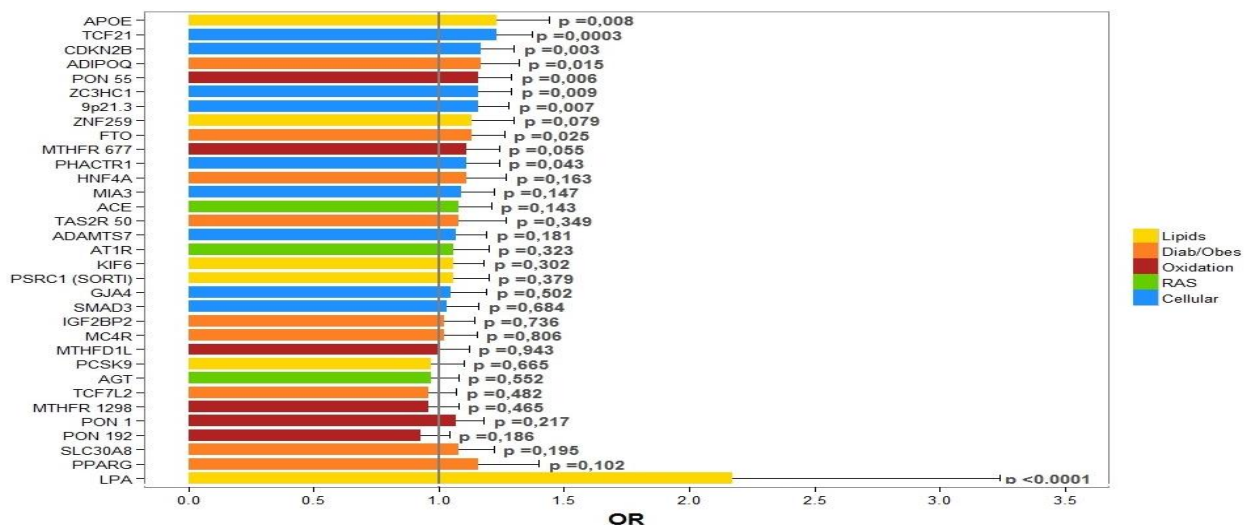
Tabela 7: Índices antropométricos da população

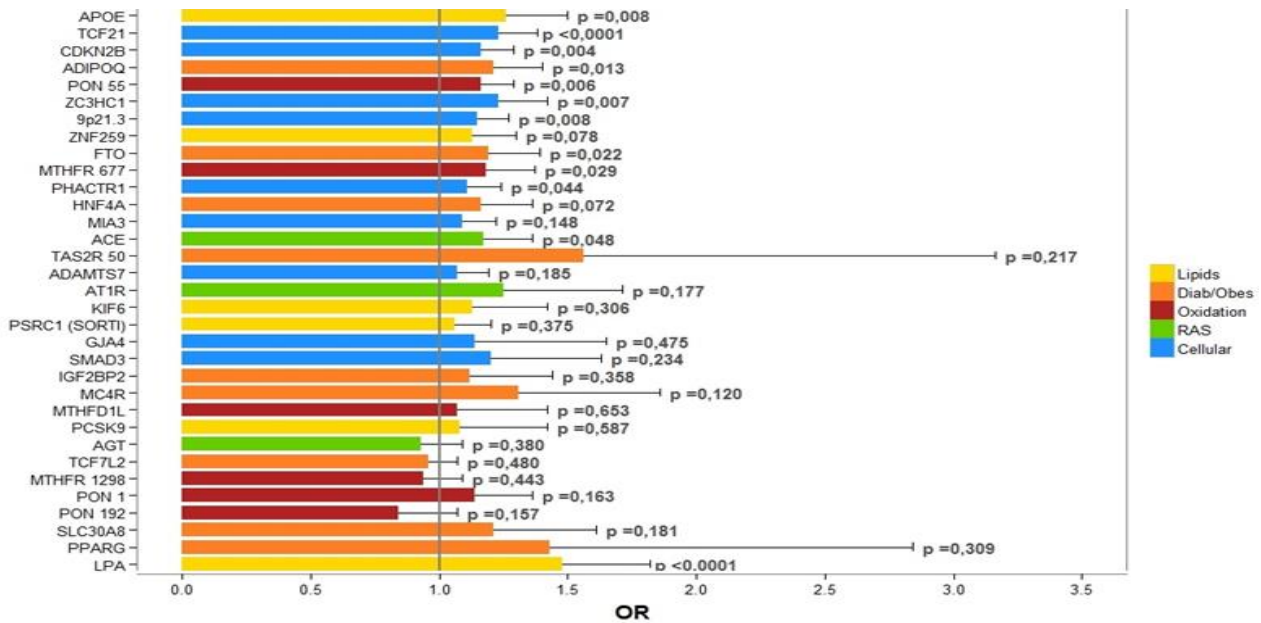
Variáveis	Casos (n=1566)	Controlos (n=1322)	Valor p
PAS, mmHg	137,9 ± 20,8	136,2 ± 18,1	0,024
PAD, mmHg	82,6 ± 11,8	83,9 ± 11,1	0,002
VOP, m/s ²	8,6 ± 1,9	8,3 ± 1,7	<0,0001
Índice de Quetelet (IMC) (kg/m ²)	28,6 ± 4,2	28,1 ± 4,5	0,007
>25, n (%)	1268 (81)	1008 (76,2)	0,002
>30 ⁵ , n (%)	524 (33,5)	386 (29,2)	0,014
Perímetro Abd.,cm (obesidade central)	101,4 ± 10,7	98,4 ± 11	<0,0001
Homem> 94, n (%)	1205 (97,3)	962 (95,2)	0,008
Mulher> 80, n (%)	325 (99,1)	307 (98,4)	0,495
Cintura/Anca	0,98 ± 0,07	0,96 ± 0,06	<0,0001
Homem> 0,9, n (%)	1194 (96,4)	932 (92,3)	<0,0001
Mulher> 0,8, n (%)	321 (97,9)	297 (95,2)	0,064
Cintura/Altura	0,61 ± 0,06	0,59 ± 0,07	<0,0001
Mulher	0,62 ± 0,08	0,59 ± 0,07	<0,0001
Homem	0,60 ± 0,06	0,59 ± 0,06	<0,0001
>0,5, n (%)	1506 (96,2)	1220 (92,3)	<0,0001

PAS – Pressão arterial sistólica; PAD – Pressão arterial diastólica; VOP – Velocidade de onda de pulso; Índice Quetelet =IMC – Índice de massa corporal; Abd- Abdominal. Valores contínuos expressos pela média±desvio padrão; Estatisticamente significativo para p<0,05.

IV.2 ESTUDO GENÉTICO e DC

Das trinta e três variantes genéticas incluídas neste estudo, associadas com a DC ou FRT, muitas não tinham mecanismo fisiopatológico atribuído na data das publicações iniciais. Atualmente, ainda em estudo, podem ser-lhes atribuídos os seguintes eixos fisiopatológicos: seis variantes genéticas do eixo lipídico (APOE, LPA, ZNF259, KIF6, SORTI, PSCK9); nove variantes do eixo Diabetes/Obesidade e insulinoresistência (ADIPOQ, FTO, HNF4A, IGF2BP2, MC4R, TCF7L2, SLC30A8, PPAR G, TAS2R50); três variantes do eixo da HTA (ECA, ATIR e AGT); seis variantes associadas ao estado pró-oxidativo (MTHFD1L, MTHFR 677, MTHFR 1298, PON108, PON55, PON 192) e nove variantes cujo mecanismo fisiopatológico para a DC não se encontra totalmente esclarecido.





Figuras 9 e 10: Presumível mecanismo de ação das 33 variantes genéticas organizada por eixos fisiopatológicos na DC. OR alélico (em cima) e genotípico (em baixo) para DC para cada variante genética e nível de significância na análise gene a gene.

Maioritariamente estarão envolvidas no ciclo celular, na transcrição de genes, na diferenciação de células musculares ou como *binders* de cotransporte (SMAD3, GJA4, ADAMTS7, MIA3, 9p21, CDKN2B, PHACTR1, ZC3HC1 e TCF21). (Figura 9 e Figura 10)

Foram encontradas dez variantes genéticas associadas com a DC com significado estatístico na análise gene a gene (Tabela 8); nomeadamente as variantes LPA (OR=2,20), APOE (OR=1,26), ADIPOQ (OR=1,21), FTO (OR= 1,19), PON 55 (OR=1,35), CDKN2B (OR=1,36), *locus* 9p21 (OR=1,32), TCF721 (OR=1,54), ZC3HC1 (OR=1,29) e PHACTR1 (OR=1,11). Estas variantes significativas foram avaliadas de acordo com os diferentes modelos genéticos possíveis. Em cinco variantes genéticas a força de associação com DC manteve-se em todos os modelos genéticos. Nas restantes variantes genéticas a associação com DC manteve-se em pelo menos dois ou três modelos genéticos.

A variante genética 9p21, a mais representativa dos GWAS, confirmou-se neste estudo estar associada com a DC com um OR genotípico de 1,32. 32,2% dos indivíduos com DC e 28,7% dos indivíduos do grupo controlo possuem esta variante em homocigotia (CC) e 47% dos casos e controlos em heterocigotia (CG).

As variantes genéticas, encontradas como estatisticamente significativas na nossa amostra, abrangem curiosamente diferentes eixos na aterosclerose nomeadamente, o lipídico (LPA, APOE), o eixo da DM/obesidade (ADIPOQ, FTO) e o eixo oxidativo (PON 55), demonstrando que na nossa amostra, também a nível genético as diferentes variantes concorrem para a história natural da doença em diferentes passos.

As variantes TCF721, ZC3HC1, PHACTR1 cujo mecanismo fisiopatológico ainda não se encontra esclarecido, também foram associadas com a DC.

Tabela 8: Análise bivariada gene a gene de acordo com diferentes modelos de hereditariedade genética

Genes		Casos (n=1566)	Controlos (n=1322)	Modelos				
				Dominante	Recessivo	Aditivo	Multiplicativo	Co-dominante
LPA	TT	1480 (94,5%)	1288 (97,4%)	2,20	2,20	1,48	2,17	2,20
	CT	86 (5,5%)	34 (2,6%)	(1,47 - 3,30)	(1,47 - 3,30)	(1,21 - 1,82)	(1,45 - 3,24)	(1,47 - 3,30)
	CC	0 (0,0%)	0 (0,0%)	p<0,0001	p<0,0001	p<0,0001	p<0,0001	p<0,0001
TCF21	GG	140 (8,9%)	160 (12,1%)	1,40	1,27	1,23	1,23	1,54
	CG	680 (43,4%)	609 (46,1%)	(1,10 - 1,78)	(1,09 - 1,47)	(1,10 - 1,38)	(1,10 - 1,37)	(1,20 - 1,98)
	CC	746 (47,6%)	553 (41,8%)	p=0,006	p=0,002	p<0,0001	p=0,0003	p=0,001
CDKN2B	AA	275 (17,6%)	281 (21,3%)	1,27	1,20	1,16	1,17	1,36
	AG	711 (45,4%)	605 (45,8%)	(1,05 - 1,53)	(1,03 - 1,39)	(1,05 - 1,29)	(1,06 - 1,30)	(1,11 - 1,67)
	GG	580 (37,0%)	436 (33,0%)	p=0,012	p=0,023	p=0,004	p=0,003	p=0,004
PON55	LL	527 (33,7%)	500 (37,8%)	1,20	1,25	1,16	1,16	1,35
	LM	761 (48,6%)	627 (47,4%)	(1,03 - 1,40)	(1,02 - 1,52)	(1,04 - 1,29)	(1,04 - 1,29)	(1,09 - 1,69)
	MM	278 (17,8%)	195 (14,8%)	p=0,020	p=0,030	p=0,006	p=0,006	p=0,007
Locus9p21	GG	323 (20,6%)	321 (24,3%)	1,23	1,18	1,15	1,16	1,32
	CG	738 (47,1%)	621 (47,0%)	(1,04 - 1,47)	(1,01 - 1,38)	(1,04 - 1,27)	(1,04 - 1,28)	(1,08 - 1,62)
	CC	505 (32,2%)	380 (28,7%)	p=0,019	p=0,042	p=0,008	p=0,007	p=0,007
APOE	E2	1151 (73,5%)	1028 (77,8%)	1,26	1,27	1,22	1,23	1,34
	E3	376 (24,0%)	268 (20,3%)	(1,06 - 1,50)	(0,77 - 2,10)	(1,05 - 1,42)	(1,06 - 1,44)	(0,81 - 2,22)
	E4	39 (2,5%)	26 (2,0%)	p=0,008	p=0,344	p=0,009	p=0,008	p=0,255
ZC3	TT	171 (10,9%)	165 (12,5%)	1,16	1,23	1,16	1,16	1,29
	CT	684 (43,7%)	623 (47,1%)	(0,93 - 1,46)	(1,06 - 1,42)	(1,04 - 1,29)	(1,04 - 1,29)	(1,01 - 1,64)
	CC	711 (45,4%)	534 (40,4%)	p=0,192	p=0,007	p=0,009	p=0,009	p=0,042
ADIPOQ	CC	889 (56,8%)	811 (61,3%)	1,21	1,18	1,16	1,17	1,27
	CG	584 (37,3%)	444 (33,6%)	(1,04 - 1,40)	(0,86 - 1,63)	(1,03 - 1,32)	(1,03 - 1,32)	(0,91 - 1,76)
	GG	93 (5,9%)	67 (5,1%)	p=0,013	p=0,308	p=0,015	p=0,015	p=0,159
FTO	CC	542 (34,6%)	512 (38,7%)	1,19	1,14	1,13	1,13	1,25
	AC	763 (48,7%)	612 (46,3%)	(1,03 - 1,39)	(0,93 - 1,39)	(1,02 - 1,25)	(1,02 - 1,26)	(1,00 - 1,55)
	AA	261 (16,7%)	198 (15,0%)	p=0,022	p=0,216	p=0,025	p=0,025	p=0,051
PHACTR1	CC	294 (18,8%)	274 (20,7%)	1,13	1,17	1,11	1,11	1,23
	CT	760 (48,5%)	660 (49,9%)	(0,94 - 1,36)	(1,00 - 1,37)	(1,00 - 1,24)	(1,00 - 1,24)	(1,00 - 1,52)
	TT	512 (32,7%)	388 (29,3%)	p=0,189	p=0,053	p=0,044	p=0,043	p=0,055

Legenda: OR , IC e nível de significância para cada modelo genético.

A tabela 9 resume os resultados do estudo genético nesta população. Assinalam-se as trinta e três variantes genéticas, a posição no cromossoma, o alelo de risco e a frequência do alelo *minor*. O poder de associação (OR) com DC do genótipo e do alelo de risco são mostrados.

Tabela 9: Resumo do estudo genético das 33 variantes genéticas

SNP ID	Gene próximo	Cr	Posição	OR genotípico (IC95%)	Valor p	OR alélico (IC95%)	Valor p	FAM (%)	Mecanismo de ação
rs1333049	9p21.3	9	22125504	1,15 (1,04-1,27)*	0,008	1,16 (1,04-1,28)	0,007	45,8	Celular
rs4977574	CDKN2B	9	22098575	1,16 (1,05-1,29)*	0,004	1,17 (1,06-1,30)	0,003	42,0	Celular
rs618675	GJA4	1	34922761	1,14 (0,79-1,65)*	0,475	1,05 (0,92-1,19)	0,502	19,6	Celular
rs17228212	SMAD3	15	65245693	1,20 (0,888-1,63)*	0,234	1,03 (0,91-1,16)	0,684	25,3	Celular
rs17465637	MIA3	1	222650187	1,09 (0,97-1,22)*	0,148	1,09 (0,97-1,22)	0,147	28,6	Celular
rs12190287	TCF21	6	134256218	1,23 (1,10-1,38)*	<0,0001	1,23 (1,10-1,37)	0,0003	32,7	Celular
rs3825807	ADAMTS7	15	76876166	1,07 (0,97-1,19)*	0,185	1,07 (0,97-1,19)	0,181	41,2	Celular
rs11556924	ZC3HC1	7	130023656	1,23 (1,06-1,42)*	0,007	1,16 (1,04-1,29)	0,009	34,3	Celular
rs1332844	PHACTR1	6	12927312	1,11 (1,00-1,24)*	0,044	1,11 (1,00-1,24)	0,043	44,3	Celular
rs2114580	PCSK9	1	55167236	1,08 (0,82-1,42)*	0,587	0,97 (0,87-1,10)	0,665	26,3	Lípidos
rs3798220	LPA	6	160540105	1,48 (1,21-1,82)*	<0,0001	2,17 (1,45-3,24)	<0,0001	2,1	Lípidos
rs20455	KIF6	6	39357302	1,13 (0,90-1,42)*	0,306	1,06 (0,95-1,18)	0,302	32,8	Lípidos
rs7412/ rs429358 ¹	APOE ¹	19	44908822/ 44908684	1,26 (1,06-1,50) [#]	0,008	1,23 (1,06-1,44)	0,008	13,4	Lípidos
rs964184	ZNF259	11	116778201	1,13 (0,99-1,30)*	0,078	1,13 (0,99-1,30)	0,079	17,7	Lípidos
rs599839	PSRC1	1	109279544	1,06 (0,93-1,20)*	0,375	1,06 (0,93-1,20)	0,379	21,4	Lípidos
rs1801133	MTHFR 677	1	11796321	1,18 (1,02-1,37) [#]	0,029	1,11 (1,00-1,24)	0,055	33,5	Oxidação
rs1801131	MTHFR 1298	1	11794419	0,94 (0,82-1,09) [#]	0,443	0,96 (0,85-1,08)	0,465	28,0	Oxidação
rs705379	PON -108	7	96324583	1,14 (0,95-1,36) [#]	0,163	1,07 (0,96-1,18)	0,217	46,4	Oxidação
rs662	PON 192	7	95308134	0,84 (0,65-1,07)*	0,157	0,93 (0,83-1,04)	0,186	30,1	Oxidação
rs854560	PON 55	7	95316772	1,16 (1,04-1,29)*	0,006	1,16 (1,04-1,29)	0,006	40,4	Oxidação
rs6922269	MTHFD1L	6	150931849	1,07 (0,80-1,42)*	0,653	1,00 (0,89-1,12)	0,943	27,3	Oxidação
rs5186	AT1R	3	148742201	1,25 (0,91-1,71)*	0,177	1,06 (0,94-1,20)	0,323	24,7	SRAA
rs699	AGT	1	230710048	0,93 (0,80-1,09) [#]	0,380	0,97 (0,87-1,08)	0,552	42,9	SRAA
rs4340	ACE	17	61565892	1,17 (1,00-1,36)*	0,048	1,08 (0,97-1,21)	0,143	38,1	SRAA
rs4402960	IGF2BP2	3	185793899	1,12 (0,88-1,44)*	0,358	1,02 (0,91-1,14)	0,736	30,8	Diab/Obes
rs1326634	SLC30A8	8	117172544	1,21 (0,91-1,61) [#]	0,181	1,08 (0,96-1,22)	0,195	25,8	Diab/Obes
rs266729	ADIPOQ	3	186841685	1,21 (1,04-1,40) [#]	0,013	1,17 (1,03-1,32)	0,015	23,3	Diab/Obes
rs7903146	TCF7L2	10	112998590	0,96 (0,86-1,07)*	0,480	0,96 (0,86-1,07)	0,482	35,3	Diab/Obes
rs17782313	MC4R	18	60183864	1,31 (0,93-1,86)*	0,120	1,02 (0,90-1,15)	0,806	21,6	Diab/Obes
rs1801282	PPARG	3	12351626	1,43 (0,72-2,84) [#]	0,309	1,16 (0,97-1,40)	0,102	8,8	Diab/Obes
rs1884613	HNF4A	20	44351775	1,16 (0,99-1,36) [#]	0,072	1,11 (0,96-1,27)	0,163	16,2	Diab/Obes
rs8050136	FTO	16	53782363	1,19 (1,03-1,39) [#]	0,022	1,13 (1,02-1,26)	0,025	39,7	Diab/Obes
rs1376251	TAS2R 50	12	11030119	1,56 (0,77-3,16)*	0,217	1,08 (0,92-1,27)	0,349	11,9	Diab/Obes

Legenda: SNP – Single Nucleotide Polymorphism; Cr – Cromossoma; OR – Odds Ratio; IC – Intervalo de Confiança; FAM – Frequência do Alelo Menor; SRAA – Sistema Renina-Angiotensina-Aldosterona; Diab/Obes – Diabetes/Obesidade; *Modelo Aditivo; *Modelo Recessivo; #Modelo Dominante; ¹Resultante de um haplótipo.

O valor do OR encontrado na literatura para cada variante genética estudada variou entre 1,06 (variante APOE) e 1,73 (variante GJA4). O valor do OR alélico encontrado na nossa população variou entre 1,06 (variante PSRC1) e 2,2 (variante LPA).

Em quatro variantes genéticas TCF7L2, AGT, PON 192 e MTHFR1298 o alelo de risco na literatura, na nossa população foi protetor com OR inferiores a 1. Nas restantes vinte e nove variantes genéticas o alelo de risco descrito na literatura foi concordante com o risco para DC no nosso estudo.

Duas variantes genéticas apresentaram frequência do alelo minor <10%; a variante LPA com frequência do alelo minor de 2,1% e a variante PPARG com 8,8%. Oito variantes genéticas apresentaram frequência do alelo minor superior a 40% (as variantes 9p21, FTO, AGT, PON 108, PON 55, PHACTR1, ADAMTS7 e CDKN2B).

As variantes genéticas preditores independentes para DC por regressão logística binária multivariada foram as variantes LPA (OR=1,47), TCF21 (OR=1,23), APOE (OR=1,21), ADIPOQ (OR=1,17), CDKN2B (OR=1,17), PON55 (OR=1,16) e ZC3HC1 (OR=1,14). O alelo de risco C da variante genética LPA apenas foi encontrado em heterozigotia (Tabela 10).

Tabela 10: Estudo genético multivariado: regressão logística

Variáveis	OR (IC 95%)	Valor p
LPA	1,47 (1,20 – 1,81)	<0,0001
TCF21	1,23 (1,10 – 1,38)	<0,0001
APOE	1,21 (1,03 – 1,40)	0,017
ADIPOQ	1,17 (1,03 – 1,32)	0,013
CDKN2B	1,17 (1,05 – 1,30)	0,003
PON55	1,16 (1,04 – 1,29)	0,007
ZC3HC1	1,14 (1,02 – 1,28)	0,018
Constante	0,48	<0,0001

H.L. $\chi^2=1,0$; $p=0,998$

Legenda: A Tabela 10 mostra o resultado da Análise Multivariada por Regressão logística binária para determinar os preditores genéticos associados com DC. Método Stepwise Forward Wald (SPSS versão 19.0) em que saíram da equação os genes: Locus9p21; FTO; PHACTR1. IC – Intervalo de confiança; estatisticamente significativo para $p<0,05$. Análise do ajuste do modelo: verificou-se o pressuposto da linearidade com a função logit, através do método dos polinômios fracionários. Análise de resíduos de Pearson e Deviance para detetar outliers e aos resíduos DfBetas e à distância de Cook para detetar observações influentes, não se tendo observado nenhuma observação atípica. Tolerância e VIF dos genes CDKN2B e 9p21 de 0,18 e de 5,6 indicando multicolinearidade aceitável das variáveis. Valores de tolerância e VIF >0,2 e >5.

Estas sete variantes genéticas estão localizadas em diferentes eixos fisiopatológicos para a aterosclerose coronária (Figura 11). Apenas as variantes genéticas do eixo HTA deixam de ter significado na regressão logística multivariada.

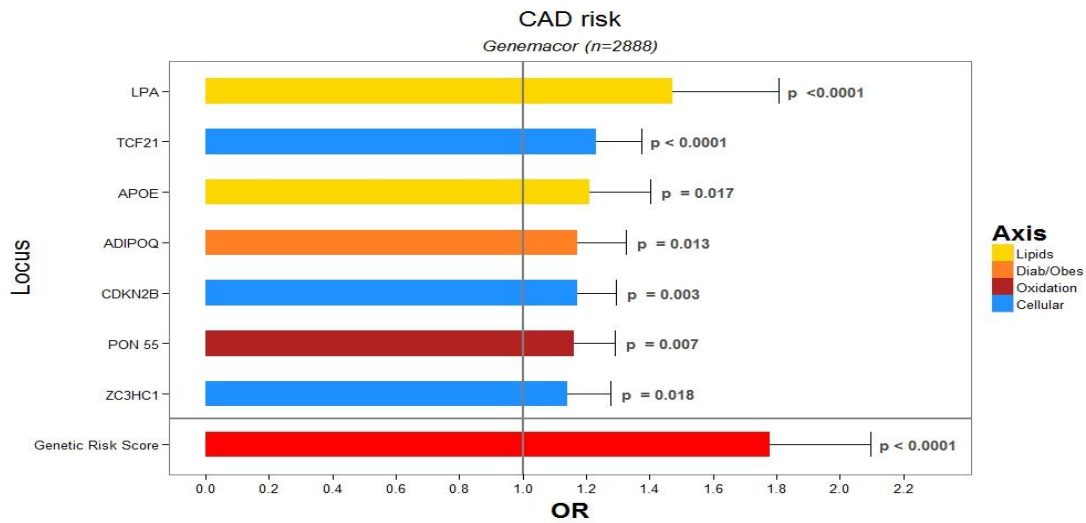


Figura 11: Presumível mecanismo de ação das 7 variantes genéticas preditores independentes por regressão logística multivariada. Legenda: Mecanismo de ação organizado por eixos fisiopatológicos. Possibilidade (OR) da ocorrência de DC para cada variante genética.

Tabela 11: Preditores não genéticos de DC: Regressão logística ajustada

Variáveis	OR	(IC 95%)	Valor p
Tabagismo	2,79	(2,45 – 3,62)	<0,0001
Diabetes	2,52	(2,02 – 3,15)	<0,0001
Clearance Cr<60ml/min	2,11	(1,40 – 3,18)	<0,0001
Sedentarismo	1,96	(1,64 – 2,35)	<0,0001
HTA	1,54	(1,27 – 1,88)	<0,0001
Obesidade Central	-----		0,007
2º tercil	1,42	(1,14 – 1,78)	0,002
3º tercil	1,28	(1,01 – 1,62)	0,044
Dislipidemia*	3,41	(2,65 – 4,38)	<0,0001
Homocisteína	-----		0,082
2º tercil	1,19	(0,95 – 1,49)	0,127
3º tercil	1,29	(1,03 – 1,62)	0,028
Lp(a)	1,01	(1,01 – 1,01)	<0,0001
Constante	0,06		<0,0001

χ^2 H.L. ⁽¹²⁾= 12,78; p=0,12

Regressão logística multivariada ajustada para as variáveis significativas determinadas após análise bivariada em que foi excluída a PCR>3 mg/l. Método Stepwise Forward Wald. (SPSS versão 19.0).

H.L. Hosmer-Lemeshow. Análise do ajuste do modelo: Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinômios fracionários. Tolerância e VIF >0,2 e >5 demonstrando ausência de multicolinearidade.

*LDL>150 ou HDL<40 ou Triglicéridos>150; Variáveis consideradas categóricas/dicotômicas exceto a Lp(a) que foi analisada como variável contínua.

Na tabela 11 da análise multivariada por regressão logística incluíram-se todos os FRT classicamente associados com DC e também outros fatores ou marcadores de aterosclerose acelerada, considerados significativos após a análise bivariada como, o sedentarismo, a obesidade de localização central, o compromisso da função renal, o terceiro tercil da Hcy, a elevação da PCR e a lipoproteína (a). Na nossa população, apenas a elevação da PCR superior a 3 mg/l, não foi preditor independente multivariado. A DM, o tabagismo e a dislipidemia foram os preditores de DC mais significativos.

A tabela 12 demonstra, de modo sumário, que das dez variantes genéticas predictoras de ocorrência de DC na análise bivariada, seis mantêm poder preditor independente, após controlo pelos FRT (dislipidemia, DM, HTA e tabagismo) na análise multivariada por regressão logística ajustada. A análise do ajuste e calibração do modelo multivariado cumpriram os pressupostos estatísticos.

Tabela 12: Preditores genéticos e convencionais, para ocorrência de DC: Regressão logística ajustada

Genes	Odds ratio	(IC 95%)	Valor p
TCF21	-----		0,014
CG	1,35	(1,02 – 1,79)	0,037
CC	1,51	(1,14 – 2,00)	0,004
LPA CT	2,32	(1,49 – 3,61)	<0,0001
CDKN2B	-----		0,005
AG	1,21	(0,97 – 1,51)	0,099
GG	1,46	(1,16 – 1,84)	0,001
PON55	-----		0,014
LM	1,15	(0,96 – 1,37)	0,140
MM	1,44	(1,13 – 1,84)	0,004
APOE	-----		0,037
E4*	1,21	(0,99 – 1,48)	0,059
E4/E4**	1,71	(0,98 – 3,00)	0,060
ZC3HC1	-----		0,008
CT	1,04	(0,79 – 1,36)	0,781
CC	1,34	(1,02 – 1,76)	0,036
Tabagismo	3,36	(2,82 – 4,02)	<0,0001
Dislipidemia	3,06	(2,46 – 3,82)	<0,0001
HTA	1,93	(1,63 – 2,30)	<0,0001
Diabetes	3,02	(2,46 – 3,71)	<0,0001
Constante	0,07		<0,0001

χ^2 H.L.⁽¹⁸⁾ = 13,63; p=0,09

Método Forward Wald (SPSS versão 19.0) em que saíram da equação os genes: Locus9p21; FTO; PHACTR1. IC– Intervalo de confiança; Estatisticamente significativo para p<0,05. Análise multivariada ajustada para genótipos de risco, FRT (como variáveis dicotômicas) e Idade e sexo (variáveis contínuas). *Inclui os genótipos E2E4 e E3E4; **Inclui os genótipos E4E4, a classe de referência para este gene inclui os genótipos E2E2, E2E3 e E3E3. Dislipidemia: LDL>150 ou HDL<40 ou Triglicéridos>150.

Análise do ajuste do modelo: Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinômios fracionários. Tolerância e VIF dos genes CDKN2B e 9p21 de 0,18 e de 5,5 indicando multicolinearidade das variáveis.

IV.3 INTERAÇÃO G-G E G-A:

As tabelas numeradas de 13 a 24 representam os três tipos de análises de interação G-G e G-A:

IV.3.1 Testes de Associação

IV.3.2 Modelos de regressão logística

IV.3.3 Modelos de *data mining* (Machine Learning) com recurso ao *Multifactor dimensionality Reduction* (MDR):

IV.3.3.1 As interações G-A são apresentadas em quatro análises, por presumíveis eixos fisiopatológicos de ação das respetivas variantes genéticas na DC.

IV.3.3.2 As interações G-G são apresentadas em duas análises.

Na tabela 13 destacamos as interações genótipo-fenótipo estatisticamente significativas entre as variantes genéticas e as variáveis ambiente.

As associações foram identificadas em maior número no eixo lipídico nomeadamente entre: a variante genética ZNF 259 e o valor de HDL e dos triglicéridos; a variante genética da LPA com o valor sérico dos triglicéridos e da Apo (B); a variante genética da APOE com os valores de HDL, a Apo (B) e com a variável dislipidemia e, ainda a variante genética PPARG com a HDL.

No eixo da DM/obesidade, identificou-se uma associação entre a variante genética TCFL2 com a variável glicose sérica. Por fim, no eixo oxidativo, as variantes genéticas MTHFR 677 com a Hcy e ácido fólico sérico apresentaram associações significativas.

Tabela 13: Variantes genéticas e variáveis ambientais: Testes de Associação

Variantes Genéticas	Ambiente	Valor p
ZNF259	HDL	0,0009
ZNF259	Triglicéridos	<0,0001
LPA	Apo B	<0,0001
LPA	Dislipidemia	<0,0001
PPARG	HDL	0,008
TCF7L2	Glicémia	0,0003
MTHFR677	Ácido fólico	<0,0001
MTHFR677	Homocisteína	<0,0001
APOE	Apo (B)	0,011
APOE	Dislipidemia	0,057
APOE	HDL	0,0002

Variáveis categóricas – Teste χ^2 . Variáveis contínuas- Teste não paramétrico de Kruskal-Wallis.

As interações mais significativas entre as variáveis modificáveis (FRT) e as variantes genéticas foram a relação da dislipidemia com a variante ZNF259 e a da diabetes com a variante genética TCF7L2 (Tabela 14).

Tabela 14: Variantes Genéticas – FRT: Testes de Associação

Interações	Valor p (Interação)
Diabetes e TCF7L2	0,0001
Dislipidemia e ZNF259	<0,0001

Estatisticamente significativo para $p < 0,05$.

Na tabela 15 verifica-se haver valor incremental em relação a um modelo de FRT, quando se adicionam as trinta e três variantes genéticas aos FRT. A adição das interações gene-gene aos modelos de FRT e variantes genéticas, não teve valor incremental na previsão de ocorrência de DC.

Tabela 15: Comparação de modelos dos FRT + Variáveis Genéticas + Interação por Regressão logística

Modelos	Valor p (Teste de likelihood ratio)
Modelo 1 versus Modelo 2	0,007
Modelo 1 versus Modelo 3	0,007
Modelo 2 versus Modelo 3	0,310

Modelo 1 incluiu os FRT, o modelo 2 incluiu os FRT + 33 variantes genéticas; modelo 3 incluiu os FRT + 33 genes + interações

A tabela 16 revela que as melhores combinações de pares de genótipos preditores da presença de DC nesta população. Foram, em três casos, combinações de variantes genéticas associadas ao eixo celular (ADAMTS7 GG, CDKN2B AG, locus9p21 GG) com variantes genéticas relacionadas com o eixo oxidativo (MHTFR677 CT e PON192 RR). O par do genótipo CC da variante genética ATIR também apresentou interação com a variante genética do KIF 6 GG, associado como atuante no eixo celular.

Tabela 16: Interação G-G (Regressão Logística)

Gene-Genótipo	Eixo	Gene-Genótipo	Eixo	Valor p
ADAMTS7:GG	Celular	MHTFR677 CT	Oxidação	0,0009
CDKN2B AG	Celular	PON192 RR	Oxidação	0,001
KIF6 GG	Celular	ATIR CC	Hipertensão	0,0003
PON192QR	Oxidação	Locus 9p21 GG	Celular	0,0009

Modelos de regressão logística, Log-likelihood ratio. Estatisticamente significativo para $p < 0,05$.

IV.3.3 MODELOS DE INTERAÇÃO COM RECURSO A ANÁLISE “DATA MINING” POR MDR EM EIXOS FISIOPATOLÓGICOS

IV.3.3.1 INTERAÇÃO G-A:

Com recurso à análise de *data mining* verifica-se que o melhor modelo de interação entre as variantes genéticas do eixo celular e os FRT foram entre a variante genética ZC3HC1 (eixo celular), tabagismo, HTA e DM com um OR de 4,30 (IC95%: 3,65-5,08) com consistente validação cruzada de 10/10 ($p < 0,0001$) (Tabela 17).

Tabela 17: Modelos de interação G-A pelo MDR: Eixo Celular e FRT

Melhores modelos	Treino			Consistência VC
	Acuidade ponderada	OR (IC 95%)	Valor p	
Tabaco	0,62	2,86 (2,41 – 3,39)	<0,0001	10/10
Tabaco; Diabetes	0,67	3,98 (3,38 – 4,69)	<0,0001	10/10
TCF21; Tabaco; Diabetes	0,67	4,06 (3,45 – 4,79)	<0,0001	6/10
ZC3HC1; Tabaco; HTA; Diabetes	0,67	4,30 (3,65 – 5,08)	<0,0001	10/10
Melhores modelos	Teste			Consistência VC
	Acuidade ponderada	OR (IC 95%)	Valor p	
Tabaco	0,62	2,86 (1,72 – 4,77)	<0,0001	10/10
Tabaco; Diabetes	0,67	3,98 (2,44 – 6,50)	<0,0001	10/10
TCF21; Tabaco; Diabetes	0,66	3,84 (2,35 – 6,27)	<0,0001	6/10
ZC3HC1; Tabaco; HTA; Diabetes	0,67	4,30 (2,62 – 7,06)	<0,0001	10/10

MDR – Multifactor dimensionality reduction; VC – Validação Cruzada; HTA – Hipertensão arterial; Melhor modelo com 4 variáveis inclui a variante ZC3HC1; Tabaco; HTA e Diabetes; Estatisticamente significativo para $p < 0,05$. Variantes genéticas: CDKN2B, Locus9p21, SMAD3, MIA3, ZC3HC1, TCF21, ADAMTS7, GJA4, PHACTR1 FRT: Hipertensão, Diabetes, Tabaco e Dislipidemia.

A análise de *Data Mining* mostra que o melhor modelo de interação entre as variantes genéticas e os FRT tradicionais foram entre o gene TAS250, TCF7L2 e a DM com um OR de 2,98 (IC95%: 2,49-3,56) com consistente validação cruzada de 10/10 ($p < 0,0001$).

A obesidade não permanece significativa por análise de *data mining* (Tabela 18).

Tabela 18: Modelos de interação G-A pelo MDR: Variantes genéticas do eixo Obesidade/DM + DM e Obesidade

Melhores modelos	Treino			Consistência VC
	Acuidade ponderada	OR (IC 95%)	Valor p	
ADIPOQ; Diabetes	0,605	3,11 (2,57 – 3,75)	<0,0001	5/10
TAS2R50; TCF7L2; Diabetes	0,613	2,98 (2,49 – 3,56)	<0,0001	10/10
ADIPOQ; TAS2R50; TCF7L2; Diabetes	0,619	2,97 (2,50 – 3,53)	<0,0001	5/10
Melhores modelos	Teste			Consistência VC
	Acuidade ponderada	OR (IC 95%)	Valor p	
ADIPOQ; Diabetes	0,593	2,66 (1,53 – 4,63)	0,0004	5/10
TAS2R50; TCF7L2; Diabetes	0,611	2,94 (1,72 – 5,01)	<0,0001	10/10
ADIPOQ; TAS2R50; TCF7L2; Diabetes	0,583	2,11 (1,28 – 3,49)	0,003	5/10

MDR – Multifactor dimensionality reduction; VC – Validação Cruzada, Melhor modelo com 3 variáveis inclui os genes TAS2R50 e TCF7L2 e a Diabetes; Estatisticamente significativo para $p < 0,05$. Interação entre as variáveis: TCF7L2, HNF4A, IGF2BP2, SLC30A8, PPARG, MC4R, ADIPOQ, FTO, TAS2R50, Obesidade e Diabetes.

Com a análise de *data mining* verifica-se que o melhor modelo de interação entre as variantes genéticas e os fatores de risco modificáveis pró-oxidativos: homocisteína >10 mg/dl e PCR >3 foi com as variantes genéticas MTHFR 677, PON108, MTHFD1L e PON 55 com um OR de 1,73 (IC95%: 1,48-2,02) com razoável validação cruzada de 9/10. (Tabela 19)

No melhor modelo preditor ficam apenas variantes genéticas, apontando na direção que, neste caso, os genótipos parecem mais consistentes que a determinação sérica da Hcy e PCR. No entanto, no caso do eixo da oxidação, nenhum dos modelos é um modelo bem ajustado, visto que o valor de teste nunca é significativo.

Tabela 19: Modelos de interação G-A pelo MDR: Variantes genéticas do eixo oxidação/Inflamação com Hcy e PCR.

Melhores modelos	Treino			Consistência VC
	Acuidade ponderada	OR (IC 95%)	Valor p	
PON55	0,525	1,22 (1,05 – 1,43)	0,011	4/10
PON55; Homocisteína	0,535	1,32 (1,13 – 1,54)	0,0004	6/10
MTHFD1L; PON1; MTHFR677	0,547	1,45 (1,25 – 1,70)	<0,0001	7/10
MTHFD1L; PON1; MTHFR677; PON55	0,568	1,73 (1,48 – 2,02)	<0,0001	9/10
Melhores modelos	Teste			Consistência VC
	Acuidade ponderada	Odds ratio (IC 95%)	Valor p	
PON55	0,482	0,86 (0,54 – 1,38)	0,542	4/10
PON55; Homocisteína	0,497	0,98 (0,62 – 1,55)	0,925	6/10
MTHFD1L; PON1; MTHFR677	0,515	1,13 (0,71 – 1,80)	0,600	7/10
MTHFD1L; PON1; MTHFR677; PON55	0,512	1,10 (0,69 – 1,75)	0,686	9/10

MDR – Multifactor dimensionality reduction; VC – Validação Cruzada. PON1=108. Melhor modelo com 3 variáveis inclui TAS2R50, TCF7L2 e a Diabetes; Estatisticamente significativo para $p < 0,05$. Interação entre as variáveis: TCF7L2, HNF4A, IGF2BP2, SLC30A8, PPARG, MC4R, ADIPOQ, FTO, TAS2R50, Obesidade e Diabetes.

Da análise de *data mining* por MDR verifica-se que o melhor modelo de interação, entre as variantes genéticas PCSK9, APOE, LPA, KIF6, PSRC1, ZNF259, e os fatores de risco modificáveis pró-lipídicos de dislipidemia, LDL > 150mg/dl, HDL <40 (Homens) e HDL <45 (Mulheres), triglicéridos > 150mg/dl foi entre as variáveis dislipidemia e o gene PSRC1 com um OR de 3,92 (IC95%: 3,16-4,87) com ótima validação cruzada de 10/10 (Tabela 20).

Contudo, também outros modelos que incluem as variantes genéticas APOE PSRC1 e dislipidemia, e o modelo PSRC1, ZNF259, PCSK9 e dislipidemia foram significativos. (Tabela 20).

Tabela 20: Modelos de interação G-A pelo MDR: variantes genéticas do eixo lipídico + dislipidemia

Melhores modelos	Treino			Consistência VC
	Acuidade ponderada	OR (IC 95%)	Valor p	
PSRC1; Dislipidemia	0,60	3,92 (3,16 – 4,87)	<0,0001	10/10
PSRC1; APOE; Dislipidemia	0,60	3,77 (3,06 – 4,65)	<0,0001	6/10
PSRC1; ZNF259; PCSK9; Dislipidemia	0,61	2,82 (2,37 – 3,36)	<0,0001	7/10
Melhores modelos	Teste			Consistência VC
	Acuidade ponderada	OR (IC 95%)	Valor p	
PSRC1; Dislipidemia	0,60	3,92 (2,05 – 7,49)	<0,0001	10/10
PSRC1; APOE; Dislipidemia	0,60	3,40 (1,83 – 6,29)	<0,0001	6/10
PSRC1; ZNF259; PCSK9; Dislipidemia	0,59	2,26 (1,35 – 3,79)	0,002	7/10

CV – Cross validation; MDR – Multifactor dimensionality reduction; Estatisticamente significativo para $p < 0.05$. Dislipidemia : LDL>150, HDL<40 para os homens e <45 para as mulheres, Triglicéridos>150. Variantes genéticas: PCSK9, APOE, LPA, KIF6, PSRC1, ZNF259.

IV.3.3.2 INTERAÇÃO GENE-GENE

No modelo com um gene, o MDR projetou o gene TCF21 como o fator de risco genético mais significativo para a DC. (Tabela 21).

O modelo com dois genes demonstrou interação sinérgica entre os genes TCF21 e APOE.

O modelo bivariado genético do TCF21 e APOE foi o melhor modelo preditor com um OR de 1,51 (IC95%: 1,29-1,76; $p < 0,0001$) e com adequada validação cruzada (10/10), sem evidência de *over fitting* do modelo.

Tabela 21: Modelos de interação G-G pelo MDR para as 10 variantes genéticas consideradas significativas na análise gene a gene

Melhores modelos	Treino		Consistência VC
	OR (IC 95%)	Valor p	
TCF21	1,27 (1,09 – 1,49)	0,002	8/10
TCF21; APOE	1,51 (1,29 – 1,76)	<0,0001	10/10
TCF21; CDKN2B; APOE	1,62 (1,39 – 1,89)	<0,0001	7/10
Melhores modelos	Teste		Consistência VC
	OR (IC 95%)	Valor p	
TCF21	1,13 (0,71 – 1,80)	0,606	8/10
TCF21; APOE	1,51 (0,94 – 2,40)	0,085	10/10
TCF21; CDKN2B; APOE	1,22 (0,77 – 1,94)	0,404	7/10

CV – Cross validation; MDR – Multifactor dimensionality reduction; significativo para $p < 0.05$.
Genes: LPA, TCF21, APOE, CDKN2B, PON55, PHACTR1, FTO, Locus9p21, ADIPOQ, ZC3HC1

Baseado nos modelos MDR, foi calculado o poder preditor ($\text{acuidade} = 1 - (S+E)/2$) do modelo e testada a sensibilidade e especificidade (Tabela 22). A acuidade do melhor modelo G-G preditor de DC foi de 0,55. Obteve-se razoável sensibilidade (60%) e razoável especificidade (50%) desse modelo.

Tabela 22- Acuidade dos modelos preditores da ocorrência de DC dados pelo MDR para 10 variantes genéticas

Melhores modelos	Treino		
	Acuidade ponderada	Sensib.	Especif.
TCF21	0,53	0,47	0,59
TCF21; APOE	0,55	0,60	0,50
TCF21; CDKN2B; APOE	0,56	0,54	0,58
Melhores modelos	Teste		
	Acuidade ponderada	Sensib.	Especif.
TCF21	0,52	0,46	0,57
TCF21; APOE	0,55	0,60	0,50
TCF21; CDKN2B; APOE	0,53	0,50	0,55

MDR – Multifactor dimensionality reduction; Estatisticamente significativo para $p < 0.05$.

O diagrama de entropia, dado pelo MDR para as dez variantes genéticas consideradas significativas na análise G-G (Figura 12), mostra que as variantes genéticas TCF21, CDKN2B, ZC3HC1, PON55, APOE e locus9p21 atuam como genes epistáticos (interagem com outras variantes genéticas) e as variantes

genéticas PHACTR1 e FTO como variantes hipostáticas.

A interação G-G entre a variante genética do APOE e a variante genética do TCF21 representa o melhor par de variantes genéticas com interação epistática elevada (a vermelho), sendo seguido do par: variante locus 9p21 e variante ZC3HC1 (interação moderada, a laranja). As restantes interações (verde/azul) não são significativas sendo por isso consideradas redundantes (figura 12 e 13).

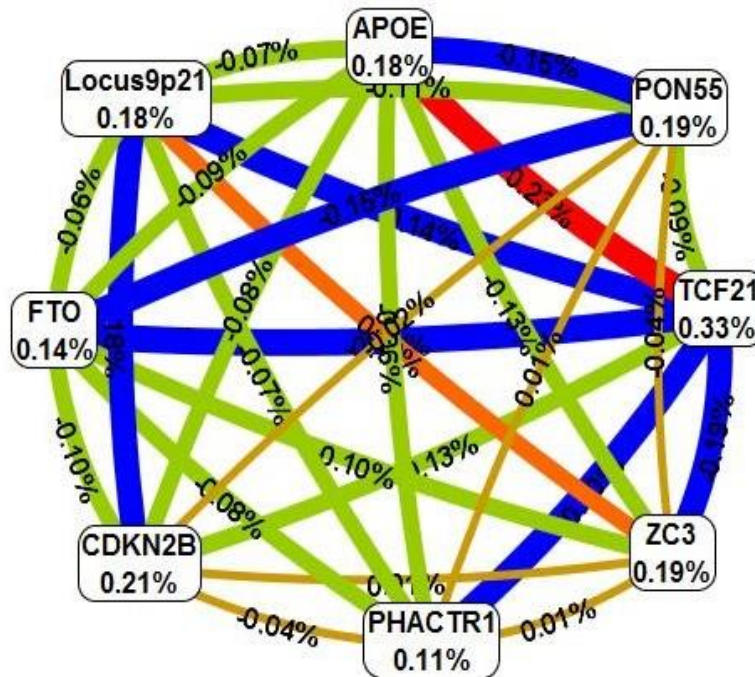


Figura 12: DIAGRAMA DE INTERAÇÕES G-G DAS 10 VARIANTES GENÉTICAS. A vermelho estão as interações sinérgicas elevadas (epistasia), a laranja as interações de grau moderado, a verde e azul- Ausência de interação ou redundância

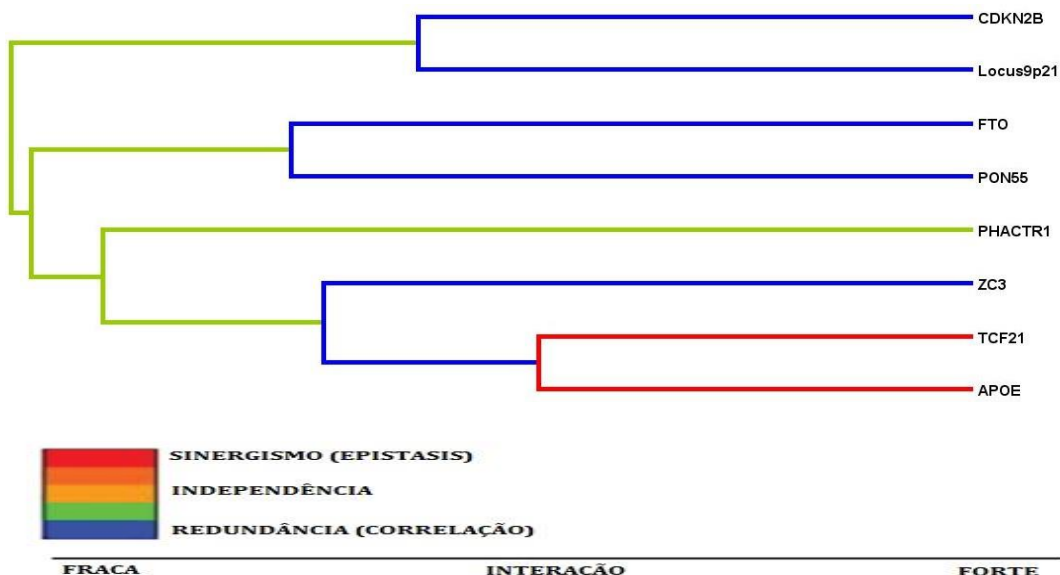


Figura 13: Dendrograma de interações G-G pelo MDR para as 10 variantes consideradas significativas na análise gene a gene.

Tabela 23: Modelos de interação G-G pelo MDR para as 33 variantes genéticas analisadas

Melhores modelos	Treino		Consistência VC
	OR (IC 95%)	Valor p	
TCF21	1,27 (1,09 – 1,49)	0,002	8/10
TCF21; APOE	1,51 (1,29 – 1,76)	<0,0001	10/10
TCF21; CDKN2B; APOE	1,63 (1,40 – 1,91)	<0,0001	3/10
TCF21; ZC3HC1; Locus9p21; MTHFR1298	2,01 (1,71 – 2,35)	<0,0001	2/10
Melhores modelos	Teste		Consistência VC
	OR (IC 95%)	Valor p	
TCF21	1,13 (0,71 – 1,80)	0,606	8/10
TCF21; APOE	1,51 (0,94 – 2,40)	0,085	10/10
TCF21; CDKN2B; APOE	1,13 (0,71 – 1,80)	0,601	3/10
TCF21; ZC3HC1; Locus9p21; MTHFR1298	0,95 (0,60 – 1,51)	0,822	2/10

CV – Cross validation; MDR – Multifactor dimensionality reduction; Estatisticamente significativo para $p < 0.05$. Genes: LPA, TCF21, APOE, CDKN2B, PON55, PHACTR1, FTO, Locus9p21, ADIPOQ, ZC3HC1

Tabela 24: Acuidade dos modelos preditores da ocorrência de DC dados pelo MDR no conjunto de treino e no conjunto de teste para as 33 variantes genéticas.

Melhores modelos	Treino		
	Acuidade ponderada	Sensib.	Especif.
TCF21	0,530	0,473	0,587
TCF21; APOE	0,551	0,600	0,502
TCF21; CDKN2B; APOE	0,561	0,587	0,534
TCF21; ZC3HC1; Locus9p21; MTHFR1298	0,586	0,591	0,581
Melhores modelos	Teste		
	Acuidade ponderada	Sensib.	Especif.
TCF21	0,515	0,463	0,567
TCF21; APOE	0,551	0,600	0,502
TCF21; CDKN2B; APOE	0,515	0,542	0,489
TCF21; ZC3HC1; Locus9p21; MTHFR1298	0,493	0,494	0,492

MDR – Multifactor dimensionality reduction; Estatisticamente significativo para $p < 0.05$.

No modelo com uma variável, o MDR projetou de novo a variante genética TCF21 como o fator de risco genético mais significativo para a DC. (tabela 23 e 24) O modelo bivariado genético demonstrou interação sinérgica entre as variantes genéticas TCF21 e APOE. O modelo bivariado genético do TCF21 e APOE foi o

melhor modelo preditor com um OR de 1,51 (IC95%: 1,29-1,76; $p < 0,0001$) com adequada validação cruzada (10/10), sem evidência de *over fitting* do modelo. O modelo com quatro variantes genéticas produziu o maior poder preditor com um OR de 2,01 (IC95%: 1,71-2,35), mas com evidência de *over fitting* do modelo.

O diagrama de entropia (Figura 14) demonstra que a variante genética TCF21 e a variante genética CDKN2B atuam como genes epistáticos e que as variantes genéticas KIF6, AGT235 e MTHFR128 atuam como variantes hipostáticas, sujeitas a epistasia. A interação epistática foi elevada entre a variante genética APOE e TCF21; a variante genética KIF6 e AGT235; a variante KIF6 e ECA e entre as variantes AGT235 e MTHFR1298 (Figura 14).

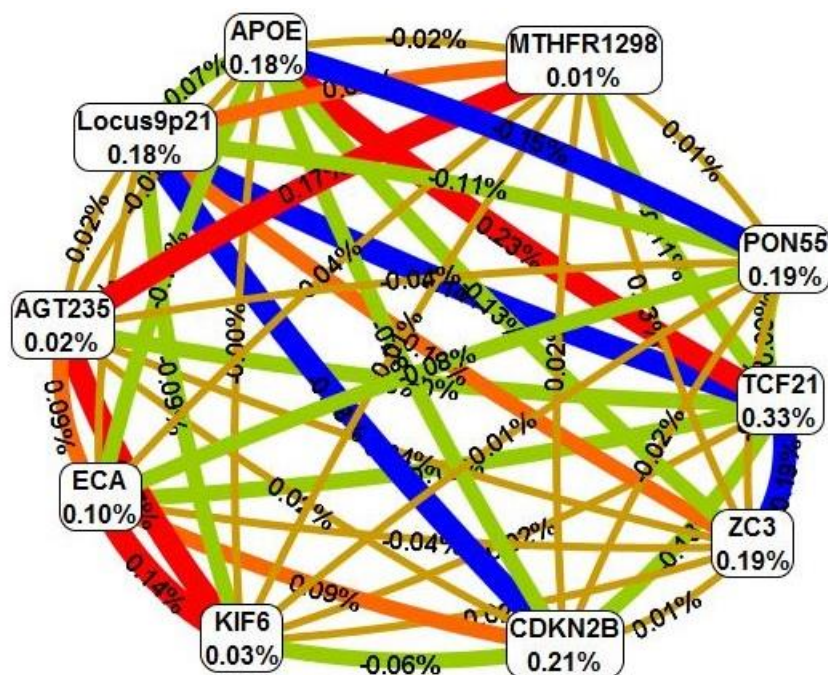


Figura 14: DIAGRAMA DE INTERAÇÕES G-G DAS 33 VARIANTES GENÉTICAS: A vermelho estão as interações (epistasia) sinérgicas elevadas, a laranja as interações (epistasia) de grau moderado, Verde-Azul- Ausência de interação ou redundância

O dendograma (Figura 15) demonstra interação epistática elevada entre a variante genética TCF21 e APOE. Existe uma interação elevada entre as quatro variantes genéticas KIF6, AGT235, ECA e MTHFR1298. A interação entre a variante genética ZC3HC1 e o par TCF21 e APOE é redundante. A variante genética CDKN2B e a variante genética locus9p21 não demonstraram ter interações significativas (Figura 15).

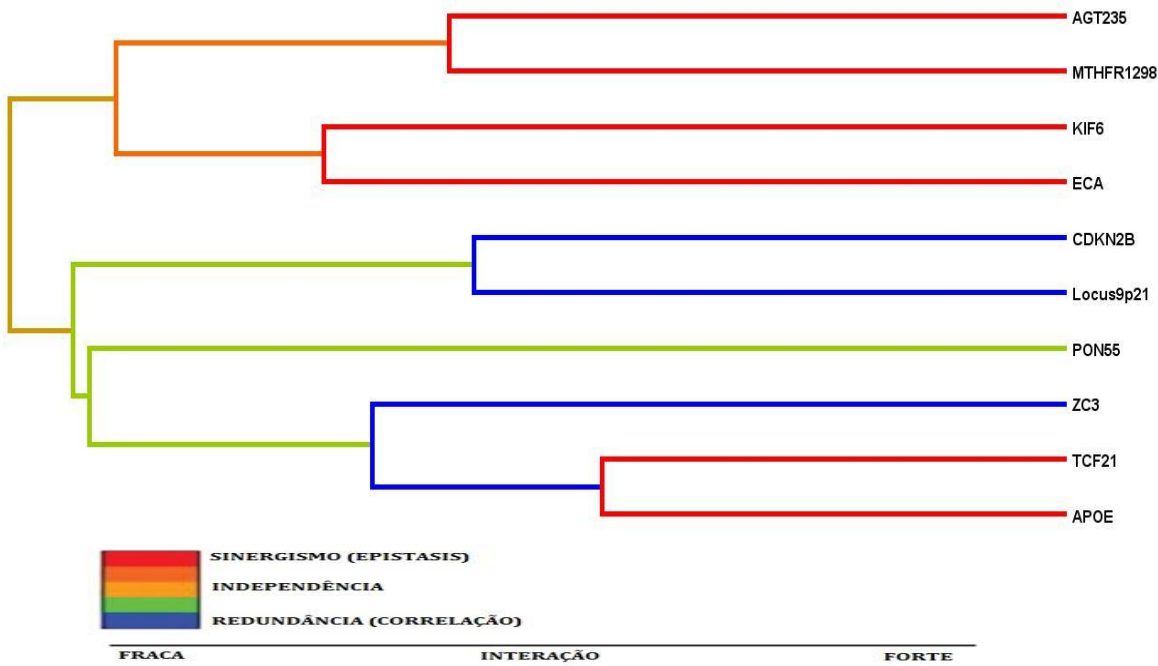


Figura 15: Dendrograma de interações G-G pelo MDR para as 33 variantes genéticas.

IV.4 SCORE GENÉTICO E DOENÇA CORONÁRIA

Os resultados deste subcapítulo serão apresentados pela seguinte ordem:

IV.4.1. Análise exploratória do SRG:

- IV.4.1.1. Nº de variantes genéticas a incluir
- IV.4.1.2. Distribuição caso-controlo dos alelos de risco, modelo aditivo
- IV.4.1.3. Distribuição caso controlo, diferentes modelos
- IV.4.1.4. Comparação de modelos aditivo/multiplicativo
- IV.4.1.5. Capacidade preditora do SRG de acordo com o número de variantes
- IV.4.1.6. Escolha do modelo do SRG
- IV.4.1.7. Validação

IV.4.2. Estudo discriminativo do SRGm:

- IV.4.2.1 Estudo comparativo em quartis
- IV.4.2.2 Distribuição do SRGm por grupo etário e género
- IV.4.2.3 SRG: Interações e correlações
- IV.4.2.4 Estudo comparativo por decis
- IV.4.2.5 Estudo do SRGm em subgrupos

IV.4.3. Valor Preditivo do SRGm

- IV.4.3.1. Valor preditor independente dos FRT
- IV.4.3.2. Valor preditor em quartis

IV.4.4. Modelo preditor combinado para a DC por data mining MDR**IV.4.5. Reclassificação média (NRI) e discriminação (IDI) com o SRG****IV.4.6. Comparação com novos marcadores de risco cardiovascular****IV.4.7 Comparação com a História Familiar de DC precoce****IV.4.8. SRG no doente com DC antes dos 45 anos:**

- IV.4.8.1 Estudo caso-controlo: características clínicas, bioquímicas e laboratoriais
- IV.4.8.2 Estudo gene a gene
- IV.4.8.3. Diferentes modelos de SRG no jovem
- IV.4.8.4. Poder preditor do SRG no jovem
- IV.4.8.5. Medidas de reclassificação média (NRI) e melhoria de discriminação (IDI) com SRG

IV.4.1 ANÁLISE EXPLORATÓRIA DO SRG

IV.4.1.1 VARIANTES A EXCLUIR

Das trinta e três variantes genéticas analisadas no estudo gene a gene, duas foram excluídas da computação do SRG: a variante genética LPA por não se encontrar em equilíbrio de *Hardy-Weinberg* (χ^2 $p < 0.0001$) e a variante genética CDKN2B- por se encontrar em *linkage* com a variante genética *locus* 9p21 ($\rho = 0.82$).

IV.4.1.2. Distribuição caso-controlo dos alelos de risco, modelo aditivo

A distribuição dos alelos de risco para a DC foi diferente na população caso e controlo (Figura 16). No grupo de indivíduos com DC, a curva que traduz o somatório dos alelos de risco para a DC encontra-se desviada para a direita da figura, implicando maior número de alelos de risco para a ocorrência de DC.

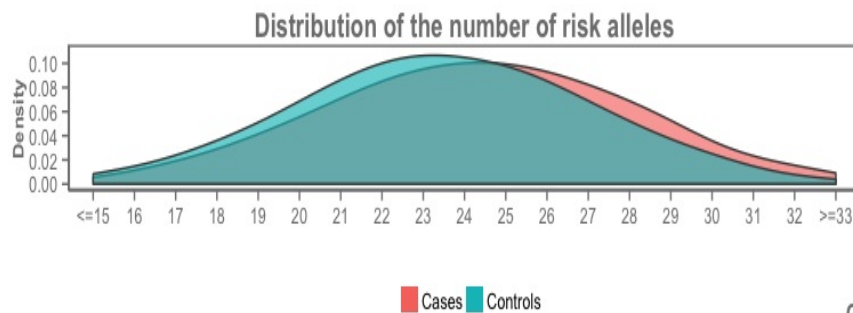


Figura 16: Distribuição caso-controlo dos alelos de risco - modelo aditivo.

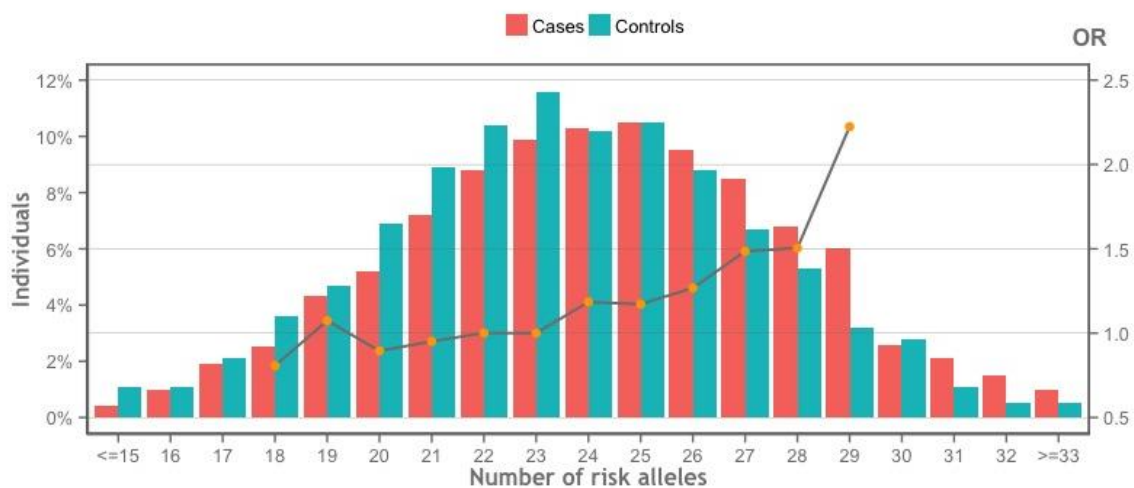


Figura 17: Modelo aditivo SRG. Distribuição dos alelos de risco na população com DC e na população controlo. Força da associação com DC de acordo com o número de alelos de risco.

IV.4.1.3. Distribuição caso controlo, diferentes modelos:

A tabela 25 mostra os valores médios do SRG de acordo com o género na população caso e controlo. Independentemente do modelo utilizado, o SRG foi sempre mais elevado na população com DC do que na população controlo. Na análise por género, demonstraram-se também valores de SRG sempre mais elevados nos doentes do género masculino ou feminino do que nos controlos do género masculino ou feminino respetivamente.

Tabela 25: Distribuição caso-controlo para os modelos de SRG

Score Genético	Casos			Controlos		
	Total (n=1566)	Masculino (n=1238)	Feminino (n=328)	Total (n=1322)	Masculino (n=1010)	Feminino (n=312)
SRG MG*	0,67 ± 0,73	0,67 ± 0,78	0,65 ± 0,58	0,48 ± 0,53	0,48 ± 0,44	0,51 ± 0,74
SRG A *	24,3 ± 3,7	24,4 ± 3,7	24,2 ± 3,6	23,5 ± 3,5	23,6 ± 3,5	23,3 ± 3,6
SRG MM *	4,01 ± 0,95	4,04 ± 0,96	3,93 ± 0,90	4 ± 0,88	4,01 ± 0,88	3,96±0,9
SRG B*	1,77 ± 0,44	1,76 ± 0,45	1,79 ± 0,41	1,61 ± 0,43	1,62 ± 0,42	1,60 ± 0,47
SRG LOGL*	1,96 ± 0,42	1,97 ± 0,42	1,94 ± 0,42	1,91 ± 0,39	1,92 ± 0,38	1,90 ± 0,39
SRG LOGP*	1,04 ± 0,31	1,04 ± 0,31	1,05 ± 0,29	0,95 ± 0,28	0,95 ± 0,28	0,93 ± 0,30

OR – Odds ratio; *valores expressos pela mediana±DP. SRG A = $\sum N^{\circ}$ ALELOS DE RISCO, SRG MG = \prod OR GENÓTIPOS, SRG MM = \sum do produto OR MELHOR MODELO GENÉTICO X N° ALELOS; SRG B = \sum produto do BETA LOGISTICA X N° ALELOS, SRG LOGL = \sum produto LOG (OR LITERATURA) X N° ALELOS; SRG LOGP = \sum produto LOG (OR MODELO) X N° ALELOS

IV.4.1.4. Comparação de modelos aditivo/multiplicativo:

As caixas de Bigodes apresentam a mediana no centro e o primeiro e terceiro quartil delimitando a caixa (Figura 18). O mínimo e máximo da distribuição do SRG estão representados por um traço horizontal. O SRG aditivo foi mais elevado nos indivíduos com DC do que nos indivíduos controlo. O SRG multiplicativo foi também mais elevado na população com DC, quando comparado com a população controlo (Figura 18).

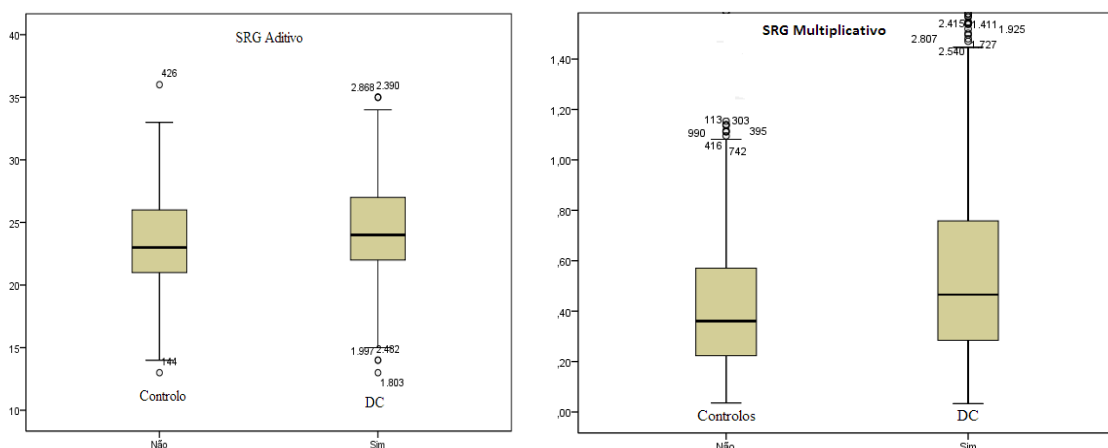


Figura 18: Distribuição do SRG aditivo e SRG multiplicativo (mediana) na população controlo á esquerda e na população com DC à direita de cada figura.

As tabelas 26 e 27 mostram a distribuição do SRG aditivo e multiplicativo em quartis. Observa-se, no geral, maior tendência para valores mais elevados do SRG nos grupos de doentes, comparado com os valores nos grupos de população controlo.

Na tabela 26 o SRG aditivo, observando os valores em cada quartil, não apresenta diferenças significativas. Só comparando os valores contínuos é que se obtém significância. No primeiro quartil do SRG multiplicativo os indivíduos com DC apresentaram valores mais elevados de SGM com resultado estatisticamente significativo.

A tabela 27 mostra a distribuição da população em quartis dos SRGs. No SRG aditivo uma parte significativa dos indivíduos com DC (28,5%) situou-se no quarto quartil. A maioria dos indivíduos controlos distribuem-se pelo primeiro e segundo quartis do SRG Aditivo (60,6%). De igual modo, 30,5% dos indivíduos com DC encontravam-se no quarto quartil do SRG Multiplicativo, enquanto 30,7% dos controlos encontrava-se no primeiro quartil do SRG demonstrando menor predisposição genética para a doença das artérias coronárias.

Tabela 26: Distribuição do SRG aditivo e multiplicativo por quartis em médias

SRG	Casos (n=1566)	Controlos (n=1322)	Valor p
SRG Aditivo	24,3 ± 3,7	23,5 ± 3,5	<0,0001
1ºQ	20,1 ± 1,8	20 ± 1,9	0,469
2ºQ	23,5 ± 0,5	23,5 ± 0,5	0,299
3ºQ	25,9 ± 0,8	25,9 ± 0,8	0,193
4ºQ	29,5 ± 1,6	29,3 ± 1,4	0,124
SRG Multiplicativo	0,67 ± 0,73	0,48 ± 0,53	<0,0001
1ºQ	0,18 ± 0,1	0,17 ± 0,1	<0,0001
2ºQ	0,33 ± 0,1	0,33 ± 0,1	0,674
3ºQ	0,52 ± 0,1	0,52 ± 0,1	0,759
4ºQ	1,36 ± 1	1,19 ± 0,9	0,022

*Valores p obtidos através do teste t de Student; Q - Quartil. Valores expressos pela média±desvio padrão; Estatisticamente significativo para p<0,05.

Tabela 27: Distribuição do SRG aditivo e multiplicativo por quartis em percentagens

SRG	Casos (n=1566) n (%)	Controlos (n=1322) n (%)	Valor p
SRG Aditivo			
1ºQ	353 (22,5)	376 (28,4)	
2ºQ	454 (29)	426 (32,2)	<0,0001
3ºQ	312 (19,9)	255 (19,3)	
4ºQ	447 (28,5)	2665 (20)	
SRG Multiplicativo			
1ºQ	300 (19,2)	406 (30,7)	
2ºQ	369 (23,6)	364 (27,5)	<0,0001
3ºQ	419 (26,8)	302 (22,8)	
4ºQ	478 (30,5)	250 (18,9)	

*Valores de p obtidos a partir do teste Qui-quadrado; Q - Quartil. Valores contínuos expressos pela média±desvio padrão; Estatisticamente significativo para p<0,05.

IV.4.1.5. Capacidade preditora do SRG de acordo com o número de variantes

Não sendo conhecido qual o número de variantes genéticas que deve incluir o SRG, calculou-se os incrementos encontrados na capacidade preditora de DC de acordo com um número crescente de variantes genéticas (Figura 19).

Gene a Gene, foram obtidas as áreas sob as curvas (AUC) sucessivas, com base no gene com maior capacidade preditora e maior poder de associação com DC. A comparação da significância do aumento da curva foi obtida pelo teste de Delong (Med Calc).

As dez variantes genéticas que mostraram maior capacidade preditiva foram as variantes genéticas LPA; TCF21; CDKN2B; PON55; Locus9p21; APOE; ZC3HC1; ADIPOQ; FTO e PHACTR1. Encontraram-se curvas AUROC em que os incrementos na capacidade preditora foram significativos. Após a incorporação da décima primeira variante genética continua a aumentar a capacidade preditora. A comparação das curvas ROC foi efetuada pelo teste de Delong (Figura 19).

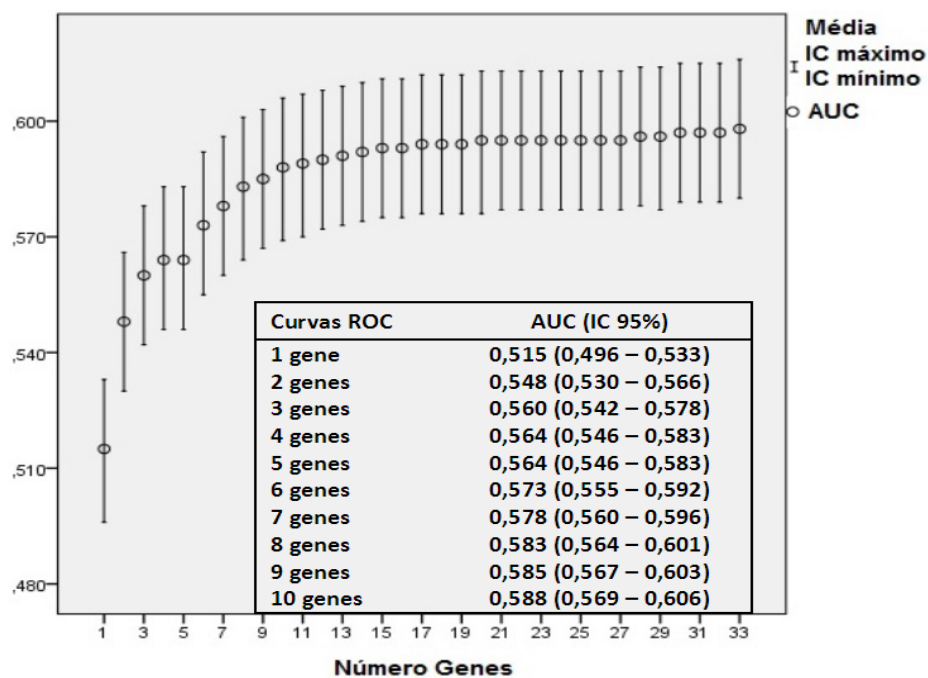


Figura 19: Capacidade preditora para a presença de DC das 31 variantes genéticas: AUC da curva ROCs

Legenda: AUC – Area under curve; IC – Intervalo de confiança

A capacidade preditora do SRG difere consoante o número de variantes genéticas (Tabela 28).

Compararam-se três Scores: um SRG com as seis variantes genéticas que permaneceram na equação final no modelo de regressão logística multivariada; um SRG com as dez variantes genéticas que mostraram

significância na análise bivariada e, um SRG, em que foram incluídas as trinta e uma variantes genéticas. A capacidade preditora, expressa pela área sob a curva das ROC, foi crescente com o aumento do número de variantes genéticas, sendo a máxima obtida com as trinta e uma variantes genéticas (AUC=0,61 (IC95%: 0,59-0,62). No global, o comportamento do SRG com as trinta e uma variantes genéticas apresentou melhor sensibilidade e melhor valor preditivo positivo.

Assim, um SRG, com trinta e uma variantes genéticas, foi utilizado na restante análise constante desta tese.

Tabela 28: Capacidade preditora e calibração do SRG de acordo com número de variantes genéticas

SRG	AUC (IC 95%)	Sensib. (%)	Especif. (%)	VPP	VPN	Valor p
6 genes	0,58 (0,56 – 0,59)	36,9	74,1	1,42	0,85	<0,0001
10 genes	0,59 (0,57 – 0,60)	54,7	58,2	1,31	0,78	<0,0001
31 genes	0,61 (0,59 – 0,62)	55,8	62,3	1,43	0,74	<0,0001

SRG – Score de risco genético (multiplicativo genotípico); AUC – Area under curve; IC – Intervalo de confiança; Valor p – significância da área abaixo da curva; Estatisticamente significativo para $p < 0,05$.

IV.4.1.6. Escolha do modelo do SRG:

A Tabela 29 demonstra a força de associação (OR), capacidade preditora, sensibilidade e especificidade para os seis diferentes modelos de SRG testados.

Todos os modelos de SRG avaliados tiveram capacidade preditora considerada significativa. No modelo de SRGMM (melhor modelo genético) a força de associação com a DC não foi significativa. Nos restantes modelos de SRG a força de associação com DC (OR) variou entre 1,06 e 3,01, sendo significativa.

O modelo multiplicativo genotípico (SRGm) apresentou melhor capacidade preditora (AUC=0,61) para a ocorrência de DC, expressa pelo valor mais elevado da AUC da curva ROC com comportamento equilibrado entre a especificidade (62,3%) e a sensibilidade (54%).

Este foi o modelo de SCORE escolhido e utilizado nas restantes análises posteriores, pelo que, sempre que ao longo desta tese não seja mencionado o modelo de score genético utilizado, pressupõe-se a utilização do modelo multiplicativo genotípico.

Tabela 29: Escolha do modelo: Capacidade preditora para a ocorrência de DC dos diferentes modelos de SRG

Modelo Score	OR (IC 95%)	Valor p ¹	AUC (IC 95%)	Sensib. (%)	Especifi. (%)	Valor p ²
SRGm	1,78 (1,52 – 2,10)	<0,0001	0,61 (0,59 – 0,62)	54,0	62,3	<0,0001
SRGa	1,06 (1,04 – 1,09)	<0,0001	0,56 (0,54 – 0,58)	58,7	50,5	<0,0001
SRG MM	1,02 (0,94 – 1,10)	0,660	0,57 (0,55 – 0,59)	41,0	70,3	<0,0001
SRG B	2,23 (1,88 – 2,65)	<0,0001	0,60 (0,58 – 0,61)	43,0	71,5	<0,0001
SRG LOGL	1,35 (1,12 – 1,62)	0,001	0,54 (0,52 – 0,55)	53,4	54,1	0,008
SRG LOGP	3,01 (2,32 – 3,89)	<0,0001	0,59 (0,57 – 0,61)	59,4	54,4	<0,0001

AUC – Area under curve; IC – Intervalo de confiança; Valor p¹ – Obtido através da regressão logística para avaliar a significância do odds ratio; Valor p² – Obtido através da curva ROC para verificar a significância da área abaixo da curva; Estatisticamente significativo para $p < 0,05$.

IV.4.1.7. Validação

As áreas sob as curvas para o SRGm na amostra global foram de 0,606 e de 0,539 na amostra de validação (Figura 20). Não foi encontrada diferença significativa no poder preditor do SRGm em ambas as amostras. (Teste de *Delong*; $p=0,676$).

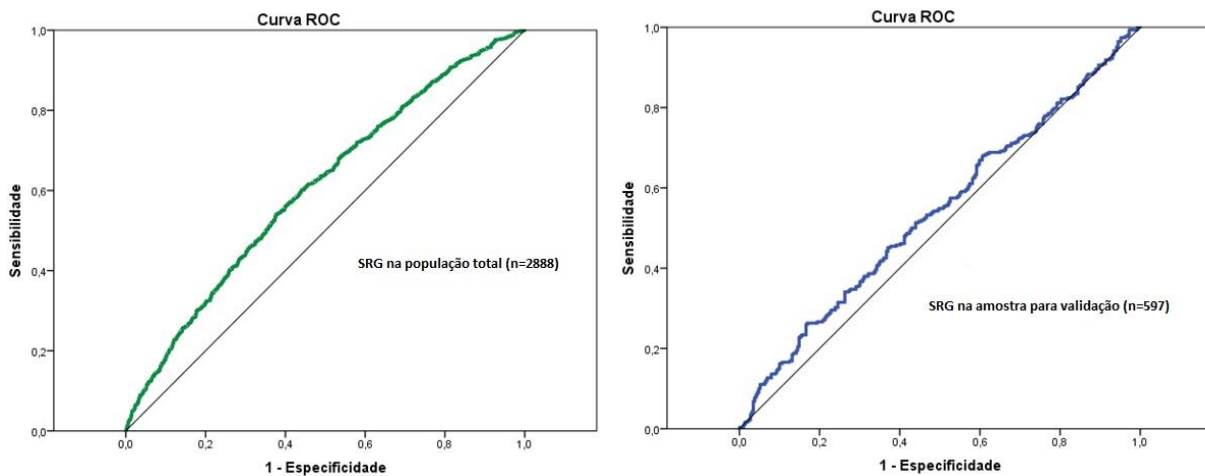


Figura 20: Gráfico das curvas ROC do SRGm para a predição de DC na população total e em uma amostra de 20% da mesma (amostra de validação).

IV.4.2 ESTUDO DISCRIMINATIVO DO SRGm

IV.4.2.1. Estudo comparativo do SRG em quartis:

A tabela 30 apresenta as características basais da população caso-controlo, comparando o primeiro com o quarto quartil do SRGm.

No subgrupo de indivíduos com risco genético mais elevado (quarto quartil do SRG), os indivíduos com DC eram mais idosos do que os controlos. Este aspeto influenciou a realização de uma análise da correlação do grupo etário com o SRG e a inclusão da idade como co-variável na regressão logística multivariada em análises posteriores. A população com DC com risco genético mais elevado (quarto quartil do SRG) era também mais sedentária (36,7% versus 55,9%), apresentava o dobro de fumadores (44,4% versus 20,8%), mais hipertensos (72,5% versus 52,7%) mais diabéticos (31,4% versus 16,7%) e mais indivíduos com dislipidemia (89,3% versus 83,7%), quando comparada com os controlos no quarto quartil do SRG ($p < 0,0001$) (Tabela 30).

No subgrupo de indivíduos com risco genético mais baixo (primeiro quartil do SRG), não se identificaram diferenças quanto à idade (Tabela 30). Neste subgrupo, 82,7% dos indivíduos com DC versus 74,9% dos controlos eram do género masculino ($p = 0,012$). Posteriormente será também, por isso, apresentada uma análise de correlação de acordo com o género. Neste grupo com menor risco genético, os indivíduos com DC apresentaram maior prevalência de DM (32,4% versus 11,2%), de tabagismo (49% versus 25,6%), de HTA (74,4% versus 52%) e de dislipidemia (88,5 versus 83,2%) ($p < 0,0001$).

Tabela 30: Características basais da população. Estudo comparativo caso-controlo no 1º e 4º quartil do SRGm

Variáveis	Score Genético Multiplicativo (quartis)				Valor p^1	Valor p^2
	1º		4º			
	Casos (n=312)	Controlos (n=410)	Casos (n=477)	Controlos (n=245)		
Género masculino, n (%)	258 (82,7)	307 (74,9)	372 (78)	190 (77,6)	0,012	0,894
Idade, anos	53,3 ± 8,4	52,6 ± 8,1	53,9 ± 7,8	52,3 ± 7,9	0,251	0,010
Exercício físico ^o , n (%)	114 (36,5)	222 (54,1)	175 (36,7)	137 (55,9)	<0,0001	<0,0001
Abuso álcool*, n (%)	64 (20,5)	62 (15,1)	84 (17,6)	33 (13,5)	0,059	0,153
Fumador, n (%)	153 (49)	105 (25,6)	212 (44,4)	51 (20,8)	<0,0001	<0,0001
HTA, n (%)	232 (74,4)	213 (52)	346 (72,5)	129 (52,7)	<0,0001	<0,0001
VOP>10m/s, n (%)	56 (17,9)	55 (13,4)	95 (19,9)	40 (16,3)	0,094	0,241
Diabetes, n (%)	101 (32,4)	46 (11,2)	150 (31,4)	41 (16,7)	<0,0001	<0,0001
Dislipidemia, n (%)	276 (88,5)	341 (83,2)	426 (89,3)	205 (83,7)	0,046	0,031
IMC>30kg/m ² , n (%)	111 (35,6)	127 (31)	154 (32,3)	79 (32,2)	0,193	0,991
Freq. cardíaca, (bpm).	68,4 ± 12,6	72,7 ± 11,2	69,1 ± 12,9	71,4 ± 11	<0,0001	0,013

^oMais do que 40min/semana; *Mais do que 300gr/semana; HTA – Hipertensão arterial; VOP – Velocidade de onda de pulso; IMC – Índice de massa corporal; Valores contínuos expressos pela média±desvio padrão; p^1 – Comparação entre casos e controlos dentro do 1º quartil do score genético; p^2 – Comparação entre casos e controlos dentro do 4º quartil. Estatisticamente significativo para $p < 0,05$.

A tabela 31 apresenta as características bioquímico-laboratoriais da população caso-controlo, comparando o primeiro com o quarto quartil do SRGm. No subgrupo de indivíduos com risco genético mais elevado (quarto quartil do SRG), os indivíduos com DC apresentaram níveis séricos mais elevados dos leucócitos. Tal como descrito nas tabelas caso-controlo gerais, também nesta sub-análise por quartis, verificaram-se nos doentes com DC de ambos os quartis, valores mais elevados de leucócitos, fibrinogénio, Lp(a), homocisteína e *clearance* creatinina ($p < 0,05$).

Tabela 31: Características Bioquímico-Laboratoriais. Estudo comparativo caso-controlo no 1º e 4º quartil do SRGm

Variáveis	Score Genético Multiplicativo (quartis)				Valor ^{p1}	Valor ^{p2}
	1º		4º			
	Casos (n=312)	Controlos (n=410)	Casos (n=477)	Controlos (n=245)		
Hemoglobina, g/dl	14,6 (13,7 – 16,6)	14,6 (14,1 – 15,4)	14,6 (13,8 – 15,4)	14,7 (14,1 – 15,5)	0,258*	0,041*
Plaquetas, 10 ³ /µl	211,5 (181 - 252)	220 (185 – 255,3)	217 (185 - 254)	222 (192,5 – 253)	0,260*	0,380*
Leucócitos, 10 ³ /µl	7 (5,9 – 8,7)	6,5 (5,6 – 7,9)	7,1 (6 – 8,2)	6,7 (5,5 – 7,9)	<0,0001*	0,001*
Fibrinogénio, mg/dl	387,5 (340 – 445,8)	365 (312 – 413,6)	394 (340,4 – 445,5)	363,1 (312 - 407)	<0,0001*	<0,0001*
Lp(a) > 25mg/dl, n (%)	125 (40,1)	116 (28,3)	234 (49,1)	86 (35,1)	0,001	<0,0001
Apo B > 90mg/dl, n (%)	157 (50,3)	208 (50,7)	295 (61,8)	139 (56,7)	0,913	0,184
Homocisteína, µmol/L	12,5 (10,5 - 15)	11,4 (9,8 – 13,6)	12,3 (9,9 – 15,2)	11,7 (9,6 – 13,7)	<0,0001*	0,018*
>10 µmol/L, n (%)	243 (77,9)	291 (71)	352 (73,8)	175 (71,4)	0,036	0,498
Creatinina, mg/dl	1 (0,9 – 1,1)	1 (0,9 – 1,1)	1 (0,9 – 1,1)	1 (0,9 – 1,1)	0,293*	0,718*
Clearance Cr.**, ml/min	94,2 (75,6 – 111,6)	92,3 (79,5 – 108,9)	92,2 (75 – 113,7)	92,6 (77,8 – 108,1)	0,872*	0,835*
Clearance Cr ^o <60, n (%)	24 (7,7)	14 (3,4)	42 (8,8)	9 (3,7)	0,011	0,011
PCR > 3mg/L, n (%)	133 (42,6)	159 (38,8)	187 (39,2)	90 (36,7)	0,297	0,518

*Valores obtidos através do teste Mann-Whitney; **Cockroft-Gault; ^o<60ml/min; Lp (a) – Lipoproteína a; Apo B – Apolipoproteína B; PCR – Proteína C-reativa; Valores contínuos expressos pela mediana (Percentil 25 – Percentil 75); ^{p1} – Comparação entre casos e controlos dentro do 1º quartil do score genético; ^{p2} – Comparação entre casos e controlos dentro do 4º quartil. Estatisticamente significativo para $p < 0,05$.

A tabela 32 e a figura 21 mostram a distribuição caso-controlo do SRGm de acordo com o género e nos diferentes grupos etários.

A mediana do SRGm foi sempre mais elevada nos casos do que nos controlos quer para o género masculino, quer para o género feminino. Detetou-se um aumento progressivo do SRG com a idade na população com DC, passando de 0,36 naqueles com idade <40 anos, para 0,47 nos com 40-59 anos e 0,49 nos indivíduos com idade superior ou igual a 60 anos.

IV.4.2.2. Distribuição do SRGm por grupo etário e género:

Independentemente do grupo etário analisado, a mediana do SRG foi mais elevado na população com DC do que na população controlo. Nestes, não houve aumento progressivo do SRG com a idade. Não se detetaram diferenças quanto à mediana do SRGm por grupo etário ($p=0,123$) no global da população. No entanto, nos diferentes grupos etários é sempre mais elevado nos casos do que na população controlo. Apenas, no grupo etário dos 30-39 anos de idade, a diferença de medianas do SRG entre casos e controlos não foi significativa (Tabela 32).

Tabela 32: Distribuição do SRGm de acordo com o género e idade

Variáveis	DC (média) (n=1566)	DC (mediana) (n=1566)	Controlos (média) (n=1322)	Controlos (mediana) (n=1322)	Valor p	Valor p*
Género						
Masculino	0,67 ± 0,8	0,45 (0,28 – 0,75)	0,48 ± 0,4	0,36 (0,23 – 0,57)	<0,0001	<0,0001
Feminino	0,65 ± 0,6	0,49 (0,30 – 0,78)	0,51 ± 0,7	0,36 (0,21 – 0,55)	0,01	<0,0001
Grupo etário						
30 a 39	0,57 ± 0,6	0,36 (0,24 – 0,70)	0,48 ± 0,6	0,32 (0,21 – 0,59)	0,358	0,177
40 a 49	0,61 ± 0,6	0,47 (0,31 – 0,69)	0,47 ± 0,4	0,34 (0,21 – 0,56)	<0,0001	<0,0001
50 a 59	0,69 ± 0,8	0,47 (0,28 – 0,79)	0,49 ± 0,5	0,38 (0,24 – 0,58)	<0,0001	<0,0001
60 a 65	0,69 ± 0,8	0,49 (0,28 – 0,78)	0,48 ± 0,6	0,34 (0,20 – 0,55)	<0,0001	<0,0001

*Estatisticamente significativo para $p<0,05$. Médias comparadas pelo teste T Student *Valores de p obtidos através do teste não paramétrico Mann-Whitney. Estatisticamente significativo para $p<0,05$.*

A média do SRG na população controlo foi semelhante entre grupos etários. Na população com DC a média do SRGm foi máxima nos grupos etários entre os 50 e os 65 anos de idade.

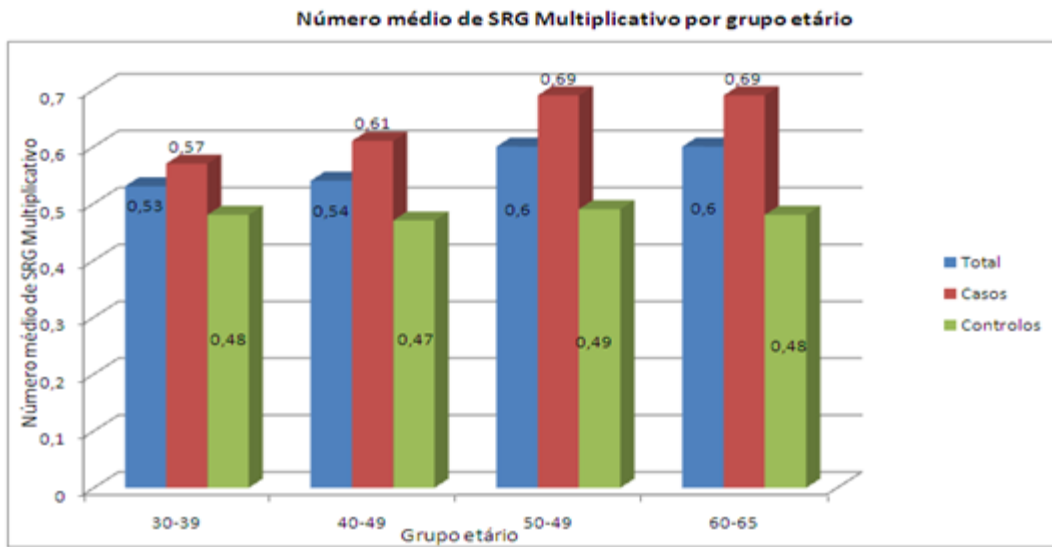


Figura 21: SRG multiplicativo na população caso, controlo, e no global (médias).

A tabela 33 demonstra ausência de diferenças significativas na mediana do SRG, entre o género masculino e feminino, $p=0,725$. Considerando como referência o género feminino, o género masculino teve maior risco de ocorrência de DC (OR=1,73; IC95%: 1,47-2,02; $p<0,0001$). Também se documentou uma interação positiva entre o grupo etário e o SRG com possibilidade (OR) de ocorrência de DC crescente, consoante um grupo etário mais idoso.

IV.4.2.3. SRG: Interações e correlações:

Tabela 33: Análises de Interação do SRGm com o grupo etário e com o género (Regressão logística)

	Género		Valor p
	Masculino	Feminino	
SRG	0,41 (0,25 – 0,67)	0,43 (0,25 – 0,68)	0,725
Género	OR (IC 95%)		Valor p
Feminino*SRGm	Referência		---
Masculino*SRGm	1,73 (1,47 – 2,02)		<0,0001
<i>Hosmer Lemeshow test: $\chi^2=17,20$; $p=0,016$. SRG – Score Risco Genético Multiplicativo. OR – Odds Ratio.</i>			
Idade (anos)	OR (IC 95%)		Valor p
30-39*SRGm	Referência		----
40-49*SRGm	1,57 (1,23 – 2,00)		<0,0001
50-59*SRGm	1,69 (1,40 – 2,05)		<0,0001
60-65*SRGm	1,93 (1,51 – 2,47)		<0,0001

Hosmer Lemeshow test: $\chi^2=17,82$; $p=0,023$. SRG – Score Risco Genético Multiplicativo. OR – Odds Ratio.

Não sendo clara a relação entre o SRGm e a idade e o género, realizamos uma análise de correlação e interação para estas variáveis (Tabela 34). O SRGm apresentou uma correlação muito fraca (desprezível) com a idade, como variável contínua embora significativa nos doentes ($p=0,027$) e no global da população ($p=0,021$). Não se demonstrou correlação significativa entre a idade e o SRGm no grupo controlo. O coeficiente de determinação para a correlação SRGm com a idade na população global foi de 0,2% e de 0,4% na população de doentes coronários.

Tabela 34: Correlação de Pearson entre o SRGm e a variável idade (contínua)

Correlação de Pearson		R (ρ)	R ²	Valor p
Total				
SRGm	Idade (contínua)	0,04	0,002	0,021
Doentes				
SRGm	Idade (contínua)	0,06	0,004	0,027
Controlos				
SRGm	Idade (contínua)	0,01	0,0001	0,723

Estatisticamente significativo para $p < 0,05$. R coeficiente de Correlação, R²= Coeficiente de determinação

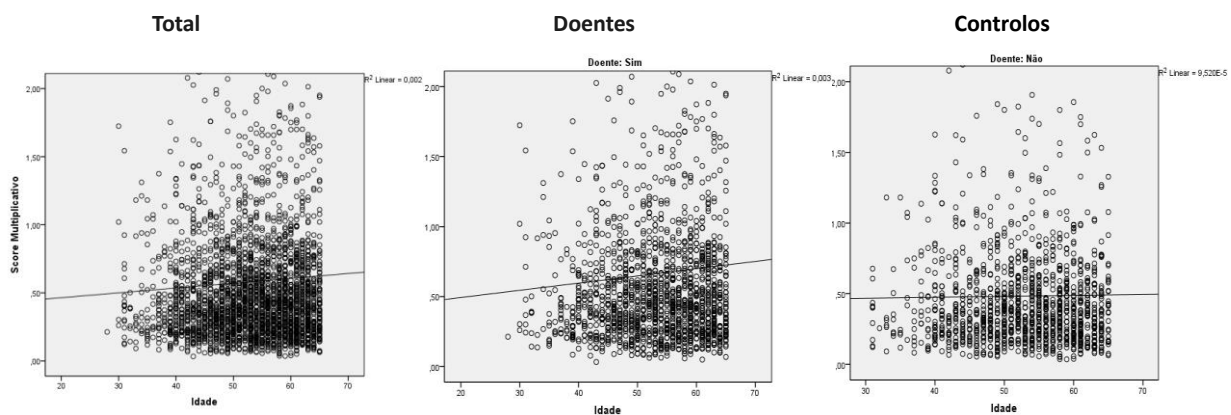


Figura 22: Correlações bivariadas significativas para o SRGm e a variável idade (contínua). Análise nos doentes com DC, nos controlos e na população total.

Para as variáveis categóricas (Tabela 35), obteve-se uma correlação significativa com o grupo etário, com a dislipidemia e com a história familiar de DC precoce. Não se verificou correlação significativa com o género. No seu conjunto, o grupo etário, o género, os FRT e a história familiar de DC precoce, explica 11,2% da variância do SRG.

Tabela 35: Regressão linear múltipla entre o SRGm com as variáveis categóricas

Variáveis	B	Beta	t	Valor p
Grupo etário	0.033	0.043	2.162	0.031
Género	0.018	0.011	0.590	0.555
Tabagismo	-0.026	-0.019	-0.987	0.324
HTA	-0.023	-0.017	-0.875	0.385
Dislipidemia	0.089	0.053	2.788	0.005
Diabetes	-0.005	-0.003	-0.171	0.865
História familiar DC	0.117	0.070	3.733	<0.0001
Constante	0.371	----	5.307	<0.0001

Método Enter da regressão linear (SPSS versão 19.0). B – Coeficiente de regressão não standardizado; Beta- Coeficiente standardizado. Estatisticamente significativo para $p < 0,05$.

Para além de haver correlação entre o SRGm e o grupo etário, verificou-se que o grupo dos doentes tinha maior número de FRT, do que os controlos em qualquer grupo etário (Figura 23). A diferença foi máxima no grupo etário entre os 30 e 39 anos, no qual os controlos apresentavam 1,14 FRT em média e os doentes apresentavam 2,3 FRT em média.

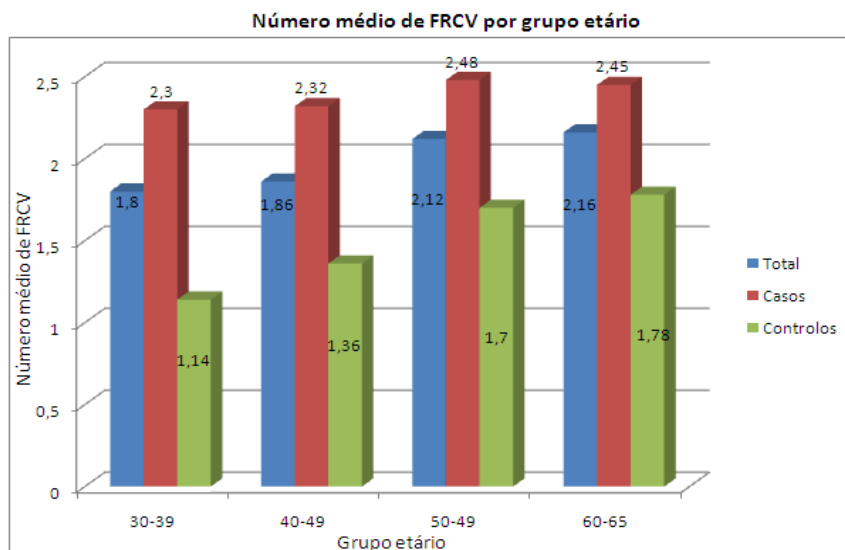


Figura 23: Distribuição dos FRT em cada grupo grupo etário. Número médio de FRT na população caso, controlo e no global da população.

IV.4.2.4. Estudo comparativo do SRGm, análise por decis:

Na tabela 36 está representada a distribuição em decis do SRGm. Considerou-se como classe de referência o primeiro decil na parte superior da tabela e, na parte inferior da tabela, considerou-se a mediana do SRG na população como classe de referência. Neste caso, a mediana da população situou-se no quinto decil.

Na parte superior da tabela verifica-se que a força de associação do SRGm com a DC expressa em OR apresenta valores crescentes de 1,42 (IC95%: 1,02-1,98) no primeiro decil até 4,04 (IC95%: 2,85-5,72) no décimo decil. Com efeito, um indivíduo no décimo decil apresenta quatro vezes maior probabilidade de ter DC, quando comparado com um indivíduo no primeiro decil.

Quando comparado com um indivíduo no quinto decil (mediana), um indivíduo situado no décimo decil apresenta um risco (OR) 2,47 vezes superior (Tabela 36).

Tabela 36: Estudo discriminativo do SRGm por decis

SRGm	Odds ratio (IC 95%)	Valor p
Referência (1º decil)		
2º	1,42 (1,02 – 1,98)	0,038
3º	1,59 (1,14 – 2,51)	0,006
4º	1,80 (1,29 – 2,51)	0,001
5º	1,63 (1,17 – 2,28)	0,004
6º	2,32 (1,66 – 3,23)	<0,0001
7º	2,34 (1,67 – 3,26)	<0,0001
8º	2,40 (1,72 – 3,35)	<0,0001
9º	3,20 (2,28 – 4,49)	<0,0001
10º	4,04 (2,85 – 5,72)	<0,0001
Referência (5º decil)		
1º	0,61 (0,44 – 0,85)	0,004
2º	0,87 (0,63 – 1,21)	0,405
3º	0,97 (0,70 – 1,35)	0,870
4º	1,10 (0,80 – 1,53)	0,559
6º	1,42 (1,02 – 1,97)	0,037
7º	1,43 (1,03 – 1,98)	0,033
8º	1,47 (1,06 – 2,04)	0,022
9º	1,96 (1,40 – 2,73)	<0,0001
10º	2,47 (1,76 – 3,48)	<0,0001

SRGm- Score de risco genético Multiplicativo (genotípico); IC – Intervalo de confiança; Estatisticamente significativo para $p < 0,05$.

A figura 24 mostra que a população controlo representada a azul está mais frequentemente representada nos primeiros decis do SRG. A partir do sexto decil do SRG, estão mais frequentemente representados indivíduos com DC. A mediana do SGM na população foi de 0,36 (0,04-9,29) (quinto decil).

Quando dividido em decis, os participantes do primeiro decil de SRGm estavam geneticamente protegidos para DC (OR=0,62; IC 95%: 0,44-0,85; $p=0,004$) e os restantes decis, acima da mediana, foram considerados de risco para DC. Assim, a análise dos OR e respetiva significância revelou que a partir do sexto decil existe uma associação com DC de modo significativo. Os primeiros decis do SRG foram protetores de DC com OR inferiores a um. Os doentes no décimo decil do SGM apresentaram um OR de 2,47 (IC95%: 1,75-3,48; $p < 0,0001$) para DC (Figura 24).

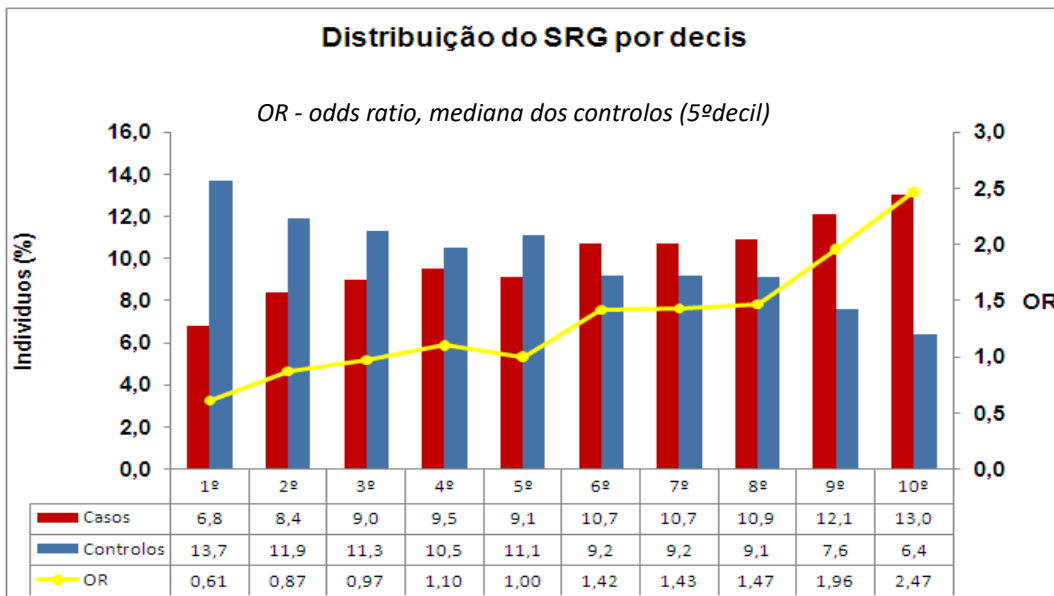


Figura 24: Distribuição do SRGm na população com DC e na população controlo: Análise por Decis

IV.4.2.5. Estudo do SRGm em subgrupos:

O poder preditivo do SRGm em subgrupos de indivíduos com DM, dislipidemia, fumadores ou com HTA foi semelhante (Tabela 37). Os valores da área debaixo da curva (AUC) variam entre 0,57 e 0,61. No grupo com história familiar de DC precoce o SRG teve uma sensibilidade elevada (87,9%) e reduzida especificidade (24,6%).

Tabela 37: Poder preditor do SRGm em subgrupos com HTA, DM, Dislipidemia, Tabagismo, HF de DC precoce.

FRCV	Casos	Controlos	AUC (IC 95%)	Sensib. (%)	Especif. (%)	VPP	VPN
Diabetes	533	175	0,57 (0,53 – 0,61)	58,3	56	1,31	0,75
Dislipidemia	1398	1103	0,61 (0,59 – 0,63)	54,1	62,6	1,44	0,73
Hipertensão	1114	700	0,60 (0,58 – 0,63)	54,4	61,9	1,43	0,74
Tabaco	730	309	0,61 (0,58 – 0,64)	50,7	65,7	1,48	0,75
História familiar DC	373	167	0,57 (0,53 – 0,61)	87,9	24,6	1,17	0,49

FRCV - Fatores de risco CV; AUC – Area under Curve; IC – Intervalo de confiança; VPP - Valor preditivo positivo; VPN - Valor preditivo negativo.

IV.4.3 VALOR PREDITIVO DO SCORE DE RISCO GENÉTICO

Neste subcapítulo, pretende-se determinar o valor preditivo do SRGm para a presença de DC, submetendo a diferentes análises de regressão logística multivariada.

IV.4.3.1. Valor preditor independente dos FRT:

Por análise multivariada por regressão logística (Tabela 38) demonstra-se que todos os FRT são preditores independentes para a DC com poder de associação (OR) entre os 1,90 e os 3,23 ($p < 0,0001$).

Entre os FRT, o preditor de DC com maior poder de associação foi o tabagismo (OR=3,23; IC 95%: 2,71-3,85) seguido pela dislipidemia (OR=3,10; IC95%: 2,49-3,85) e pela DM (OR=2,92; IC95%: 2,38-3,58).

Tabela 38: Capacidade preditora dos FRT para a presença de DC por análise de regressão logística multivariada ajustada para gênero e idade

FRT	OR	IC 95%	Valor p
HTA	1,90	1,60 – 2,26	<0,0001
Diabetes	2,92	2,38 – 3,58	<0,0001
Dislipidemia	3,10	2,49 – 3,85	<0,0001
Tabagismo	3,23	2,71 – 3,85	<0,0001
Constante	0,17	-----	<0,0001

H.L. $X^2 = 4,04$; $p = 0,672$

Regressão logística multivariada. Método Stepwise Forward Wald (SPSS versão 19.0); HTA – Hipertensão; OR – Odds ratio; IC – Intervalo de confiança; Valores significativos para $p < 0,05$.

Análise do ajuste do modelo: Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinômios fracionários. Análise de resíduos de Pearson e Deviance para detetar outliers e aos resíduos DfBetas e à distância de Cook para detetar observações influentes, não se tendo observado nenhuma observação atípica. Tolerância e VIF $> 0,2$ e < 5 demonstrando ausência de multicolinearidade. Variáveis consideradas categóricas/dicotômicas e definidas previamente.

O SRGm entrando como co-variável na análise de regressão logística multivariada (Tabela 39) permanece um fator independente de risco para DC com o OR de 1,80 (IC95%: 1,52-2,21; $p < 0,0001$). Os FRT mantêm também a sua significância ($p < 0,0001$).

Tabela 39: Associação do Score de Risco Genético na DC por Análise de Regressão logística multivariada

Variáveis	OR	(IC 95%)	Valor p
Tabagismo	3,37	2,82 – 4,03	<0,0001
Hipertensão	1,93	1,63 – 2,30	<0,0001
Diabetes	2,97	2,42 – 3,65	<0,0001
Dislipidemia	3,01	2,42 – 3,76	<0,0001
SRGm	1,80	1,52 – 2,12	<0,0001
Constante	0,12	-----	<0,0001

H.L. $X^2= 15,63$; $p=0,05$

Regressão logística multivariada. Método Stepwise Forward Wald (SPSS versão 19.0); OR – Odds ratio; IC – Intervalo de confiança; SRGm – SRG multiplicativo-genótipo. Valores significativos para $p<0,05$.

Análise do ajuste do modelo: Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinômios fracionários. Análise de resíduos de Pearson e Deviance para detetar outliers e aos resíduos DfBetas e à distância de Cook para detetar observações influentes, não se tendo observado nenhuma observação atípica. Tolerância e VIF > 0.2 e <5 demonstrando ausência de multicolinearidade. Variáveis consideradas categóricas/dicotômicas e definidas previamente.

As AUCs foram de 0,738 para os FRT e de 0,758 para o modelo do SRGm adicionado aos FRT (Figura 25). O incremento da AUC com a adição do SRGm foi significativo (*Delong* $p<0,0001$), refletindo uma melhoria da sensibilidade e especificidade.

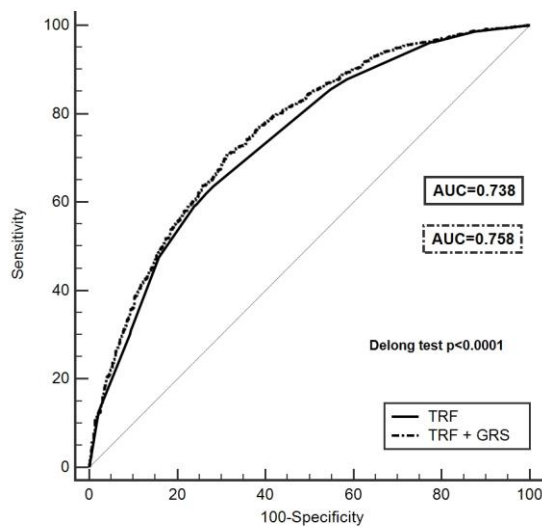


Figura 25: Gráfico das curvas ROC dos FRT e SRGm para a predição de DC.

IV.4.3.2. Valor preditor dos quartis do SRG:

Tabela 40: Poder da Associação dos Quartis do SRGm com a DC por Regressão logística multivariada

SRG (quartis)	OR (IC 95%)	Valor p
1º	Referência	---
2º	1,29 (1,05 – 1,59)	0,015
3º	1,82 (1,48 – 2,24)	<0,0001
4º	2,56 (2,07 – 3,17)	<0,0001
Constante	0,76	<0,0001

Regressão logística multivariada não ajustada. Método Stepwise Forward Wald (SPSS versão 19.0); IC – Intervalo de confiança; Score - Modelo Multiplicativo Genótipo. Estatisticamente significativo para $p < 0.05$. Análise de resíduos de Pearson e Deviance para detetar outliers e aos resíduos DfBetas e à distância de Cook para detetar observações influentes, não se tendo observado nenhuma observação atípica Análise do ajuste do modelo: Tolerância e VIF > 0.2 e < 5 demonstrando ausência de multicolinearidade. Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinómios fracionários. Hosmer Lemeshow test: $\chi^2=0,0$; $p= 1.0$. Variável score considerada categórica em quartis.

Na tabela 40 constata-se um aumento do risco para DC com o aumento do risco genético analisado em quartis. O indivíduo no 4º quartil apresenta 2,56 vezes maior risco de DC que o indivíduo no 1º quartil. Apesar do estudo estar ajustado em virtude do seu desenho para o género e idade, procedemos a novo ajuste para essas variáveis no modelo multivariado por regressão logística, representado na coluna da esquerda (Tabela 41). Na coluna da direita fez-se um ajuste adicional para outras co-variáveis.

Tabela 41: Preditores independentes para DC: SRG e FRT: Análise multivariada ajustada

Variáveis	Modelo Ajustado		Modelo Multi-Ajustado	
	OR* (IC 95%)	Valor p	OR* (IC 95%)	Valor p
SRGM (Quartis)	-----	----	-----	----
2 nd	1,3 (1,04 – 1,64)	0,024	1,34 (1,05 – 1,71)	0,018
3 rd	1,83 (1,45 – 2,3)	<0,0001	1,67 (1,46 – 2,39)	<0,0001
4 th	2,61 (1,06 – 3,3)	<0,0001	2,5 (1,96 – 3,21)	<0,0001
Hipertensão	1,91 (1,61 – 2,27)	<0,0001	2,01 (1,7 – 2,48)	<0,0001
Diabetes	2,93 (2,38 – 3,6)	<0,0001	3,26 (2,62 – 4,07)	<0,0001
Dislipidemia	2,99 (2,4 – 3,73)	0,032	3,16 (2,49 - 4)	<0,0001
Tabagismo	3,37 (2,82 – 4,03)	<0,0001	3,56 (2,95 – 4,3)	<0,0001
Constante	0,11	<0,0001	2,33	<0,0001
H.L.	$\chi^2=4,4$	0,82	$\chi^2=5,69$	0,681

Estatisticamente significativo para $p < 0.05$. Regressão logística multivariada pelo método Stepwise Forward Wald (SPSS versão 19.0); SRGM – Score de risco genético multiplicativo; OR – Odds ratio; IC – Intervalo de confiança; OR – Odds ratio ajustados para a idade e para o sexo; OR+ – Odds ratio ajustado para idade, género, frequência cardíaca, VOP, sedentarismo, abuso de álcool, IMC e história familiar DC.*

H.L. – Hosmer Lemeshow test. Análise do ajuste do modelo. Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinómios fracionários. Análise de resíduos de Pearson e Deviance para detetar outliers e aos resíduos DfBetas e à distância de Cook para detetar observações influentes, não se tendo observado nenhuma observação atípica. Análise do ajuste do modelo: Tolerância e VIF $> 0,2$ e < 5 demonstrando ausência de multicolinearidade. Variáveis na tabela consideradas categóricas.

Estas co-variáveis foram selecionadas por serem potenciais fatores para a aterosclerose coronária e por terem sido encontradas diferenças significativas na comparação entre casos e controles.

Assim, se demonstra que o SRG é um preditor independente para DC, mesmo considerando todos os potenciais fatores de risco (Tabela 41). Considerando o primeiro quartil do SRG como referência, todos os quartis do SRG associaram-se com DC com força de associação (OR) crescente em qualquer dos modelos. Um indivíduo no quarto quartil do SRGm apresenta um risco de 2,61 vezes mais e de 2,5 vezes mais, após controlo por todos os fatores de risco num modelo ajustado e multi-ajustado respetivamente.

Como referido anteriormente, quando a complexidade do modelo multivariado é elevada, por incluir um elevado número de variáveis categóricas e contínuas, foi necessário realizar a análise de resíduos para detetar observações influentes, não se tendo observado nenhuma observação atípica.

Não se observou também colinearidade significativa entre as variáveis, uma vez que a tolerância e o fator de inflação da variância (VIF) das variáveis são >0,1 e <5. O valor de p encontrado no teste de Hosmer-Lemeshow foi de 0,68, mostrando assim haver uma boa calibração do modelo usado. Para todos os fatores de risco verificamos um ligeiro aumento no OR, quando ajustado, mantendo assim a sua significância.

IV.4.4. Melhor modelo preditor para DC com recurso a MDR:

O melhor modelo preditor para a DC dado pelo MDR, incluiu o tabagismo, a diabetes, a dislipidemia, o sedentarismo, o marcador de risco Lp(a) e o SRG (Tabela 42).

Tabela 42: Poder preditor para DC de modelos interativos por *Data mining* (MDR) combinando os FRT e o SRGm

Melhor modelo	Treino		Teste		Consistência VC
	OR (IC 95%)	Valor p	OR (IC 95%)	Valor p	
Tabagismo	2,86 (2,41 – 3,39)	<0,0001	2,86 (1,72 – 4,77)	<0,0001	10/10
Diabetes; Tabagismo	3,98 (3,38 – 4,69)	<0,0001	3,98 (2,44 – 6,50)	<0,0001	10/10
Dislipidemia; Diabetes; Tabagismo	4,52 (3,82 – 5,35)	<0,0001	4,52 (2,73 – 7,48)	<0,0001	10/10
Diabetes; Tabag; Sedentarismo; Lp(a)	4,79 (4,06 – 5,67)	<0,0001	4,19 (2,55 – 6,86)	<0,0001	7/10
Dislipidemia; Diabetes; Tabagismo; Sedentarismo; Lp(a)	5,29 (4,46 – 6,27)	<0,0001	4,90 (2,95 – 8,13)	<0,0001	10/10
Dislipidemia; Diabetes; Tabagismo; Sedentarismo; Lp(a); SRG	5,70 (4,81 – 6,76)	<0,0001	4,61 (2,80 – 7,59)	<0,0001	10/10

MDR – Multifactor dimensionality reduction; Estatisticamente significativo para p<0.05. VC – Validação cruzada. Dislipidemia : LDL>150, HDL<40 para os homens e <45 para as mulheres, Triglicéridos>150. Variantes genéticas: PCSK9, APOE, LPA, KIF6, PSRC1, ZNF259. Obesidade IMC>30.

Este modelo combinado demonstrou uma forte associação com DC, expressa por um OR de 5,70 (IC95%: 4,81 – 6,76; $p < 0,0001$). A acuidade ponderada foi de 0,71 com 69% de sensibilidade e 72% de especificidade (Tabela 43). Este modelo apresentou comportamento similar nas amostras de teste geradas pelo MDR, revelando uma consistência de validação cruzada de 10/10 com significância estatística ($p < 0,0001$).

Tabela 43: Acuidade dos modelos preditores da ocorrência de DC dados pelo MDR para os FRT + Marcadores de Risco + SRG

Melhor modelo	Treino			Teste		
	Acuidade ponderada	Sens.	Especif.	Acuidade ponderada	Sens.	Especif.
Tabaco	0,62	0,47	0,77	0,62	0,47	0,77
Diabetes; Tabaco	0,67	0,67	0,66	0,67	0,67	0,66
Dislipidemia; Diabetes; Tabaco	0,68	0,62	0,74	0,68	0,62	0,74
Diabetes; Tabaco; Sedentarismo; Lp(a)	0,69	0,66	0,71	0,67	0,65	0,69
Dislipidemia; Diabetes; Tabaco; Sedentarismo; Lp(a)	0,7	0,65	0,74	0,69	0,63	0,74
Dislipidemia; Diabetes; Tabaco; Sedentarismo; Lp(a); SRG	0,71	0,69	0,72	0,68	0,67	0,7

MDR – Multifactor dimensionality reduction; Estatisticamente significativo para $p < 0,05$.

O dendograma de interações demonstra que o SRG foi independente dos restantes constituintes do modelo preditor (Figura 26). A dislipidemia, o sedentarismo e a DM apresentam correlação entre elas.



Figura 26: Dendograma de interações obtido pelo MDR para os FRT+ Novos Marcadores + SRG.

IV.4.5. Reclassificação média (NRI) e discriminação (IDI) com o SRG

A tabela 44 mostra a utilidade do SRGm na reclassificação de indivíduos, dispostos em categorias de risco pré-definidas de acordo com os FRT. A maioria dos doentes está situada no risco intermédio-baixo e intermédio-alto (25-50% e 50-75%) de acordo com a estratificação tradicional. A maioria da população controlo ($n=504$) encontra-se na categoria de risco intermédio baixo (25-50%).

O maior impacto da incorporação do SRGm na população com DC traduz-se na subida de categoria do risco 50-75% de 305 em 1566 doentes, para a categoria de muito alto risco (>75%), traduzindo uma reclassificação clínica (NRI) de 19,5%. Na população controlo, o maior impacto da incorporação do SRG traduziu-se na descida de categoria de 186 doentes de 1322 (14,1%) da categoria de risco intermédio-baixo para baixo risco de DC. Na média das reclassificações, 20,9% dos doentes foram reclassificados e 3,3% dos indivíduos da população controlo. Admitindo-se, por isso, que no global da nossa amostra a reclassificação média do NRI situou-se nos 24,2% com a inclusão do SRGm.

Tabela 44: Melhoria da Reclassificação média do risco (NRI) para DC com a inclusão do SRG

RISCO DC (SEM SRGm)	RECLASSIFICAÇÃO DO RISCO PARA DC (COM SRGm)				% AUMENTO	% REDUÇÃO
	<25%	25%-50%	50%-75%	75%-100%		
Doentes (n=1566)						
<25%	6	11	0	0	0,7%	0%
25%-50%	44	335	123	0	7,9%	2,8%
50%-75%	0	59	471	305	19,5%	3,8%
75%-100%	0	0	9	203	0%	0,6%
<i>NRI Doentes</i>	20,9%					
Controlos (n=1322)						
<25%	65	36	0	0	2,7%	0%
25%-50%	186	504	88	0	6,7%	14,1%
50%-75%	0	60	268	79	6%	4,5%
75%-100%	0	0	1	35	0%	0,1%
<i>NRI controlos</i>	3,3%					
<i>NRI total</i>	24,2%					

NRI, Net Reclassification Improvement (categorical NRI); FRT, Fatores de risco tradicionais (Dislipidemia, Tabagismo, Diabetes e Hipertensão); SRG (quartis) Score de risco genético.

A tabela 45 mostra os resultados do NRI computado de forma contínua e aplicado ao estudo caso controlo. Obteve-se uma reclassificação média (NRI) de 14,6% nos indivíduos com DC com a introdução do SRG e uma reclassificação sobreponível de 16,4% dos indivíduos da população controlo, ambas com significado estatístico (p<0,0001).

Tabela 45: Tabela de Reclassificação média do Risco (NRI) para DC com a inclusão do SRGm

	AUMENTO RISCO, n (%)	REDUÇÃO RISCO, n (%)	P (CFNRI)	CFNRI (%)	CFNRI (IC 95%)
DC (n=1566)	897 (57,3)	669 (42,7)	<0,0001	14,6	(9,7 - 19,5)
Controlos (n=1322)	553 (41,8)	769 (58,2)	<0,0001	16,4	(11,2 - 21,8)
Total (n=2888)	-----	-----	<0,0001	31	(23,8 - 38,3)

IC – Intervalo de confiança; cfNRI – category-free net reclassification index.

A tabela 46 mostra os resultados obtidos na avaliação da discriminação da reclassificação de casos e controlos, considerando o SRG para além dos FRT. O IDI mede a melhoria da diferença de probabilidades predita pelos FRT entre casos e controlos, quando incorporado o SRG. Obteve-se um IDI de 2,50% (IC95%: 1,90-3,10%) com significado estatístico ($p < 0,0001$).

Tabela 46: Índice de melhoria de discriminação (IDI) para Doença coronária com a inclusão do SRGm

Modelos	IDI (%)	(IC 95%)	Valor p
FRT versus FRT+SRGm	2,50	(1,90 - 3,10)	<0,0001

IC – Intervalo de confiança; IDI – Integrated Discrimination Improvement; FRT – Fatores de risco tradicionais (Dislipidemia, Tabagismo, Diabetes e Hipertensão); SRG quartis – Score de risco genético.

IV.4.6. Comparação do valor preditivo do SRGm com outros marcadores CV:

Foram incluídos na análise de regressão logística multivariada (Tabela 47), para além dos FRT (categóricos) e do SRGm, um ajuste adicional pelos marcadores de risco CV que tenham sido também associados com DC no estudo de regressão logística bivariada, tais como a Hcy, a PCR, a VOP e a Lp (a). Apesar de, na população com DC se ter demonstrado diferenças quanto à PCR, à Lp (a), à Hcy e à VOP, por análise bivariada, apenas o SRGm com um OR de 1,72 (IC95%: 1,46-2,04) e a Lp(a) com um OR de 1,90 (IC 95%: 1,60-2,26) estiveram associados de forma independente com a DC para além dos FRT ($p < 0,0001$).

Tabela 47: Comparação do valor preditivo do SRGm com marcadores de risco CV: Regressão Logística multivariada ajustada

Variáveis	OR	(IC 95%)	Valor p
Tabaco	3,39	(2,83 – 4,05)	<0,0001
Hipertensão	1,89	(1,58 – 2,25)	<0,0001
Dislipidemia	2,94	(2,35 – 3,67)	<0,0001
Diabetes	2,95	(2,40 – 3,63)	<0,0001
Lpa > 25mg/dl	1,90	(1,60 – 2,26)	<0,0001
SRGm	1,72	(1,46 – 2,04)	<0,0001
Constante	0,099		<0,0001

H.L. $\chi^2 = 10,7$; $p = 0,217$.

Regressão logística multivariada ajustada para os marcadores de risco: PCR, Lp(a), Homocisteína, VOP e SRGm. Método Stepwise Forward Wald (SPSS versão 19.0); OR – Odds ratio; IC – Intervalo de confiança; Estatisticamente significativo para $p < 0,05$.

Ajuste do modelo: Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinómios fracionários. Análise de resíduos de Pearson e Deviance para detetar outliers e aos resíduos DfBetas e à distância de Cook para detetar observações influentes, não se tendo observado nenhuma observação atípica. Tolerância e VIF > 0,2 e < 5 demonstrando ausência de multicolinearidade.

A adição do SRGm ao modelo basal dos FRT (Tabela 48) incrementa de modo significativo o poder preditor dos FRT ($p < 0,0001$). A adição da Lp(a) ao modelo dos FRT com SRGm também aumentou de modo significativo o poder de predição para a DC ($p = 0,003$) com acréscimo de especificidade (Tabela 48).

Tabela 48: Comparação do Valor Preditivo do SRGm com novos marcadores de risco CV

Modelos	H.L.		AUC (IC 95%)	Sensib. (%)	Especif. (%)	VPP	VPN	Teste Delong (p)
	χ^2	Valor p						
FRT	4,04	0,672	0,74 (0,72 – 0,75)	61,9	73,5	2,34	0,52	
FRT + SRG	15,63	0,05	0,76 (0,74 – 0,77)	70,6	68,8	2,26	0,43	<0,0001*
FRT + SRG + Lp(a)	10,74	0,217	0,77 (0,75 – 0,78)	64,9	75	2,60	0,47	0,003**

FRT – Factores de risco tradicionais; H.L. – Hosmer Lemeshow; AUC – Area under curve; IC – Intervalo de confiança; VPP – Valor preditivo positivo; VPN – Valor preditivo negativo; *Comparação das curvas FRT versus FRT+SRG; **Comparação das curvas FRT+SRG versus FRT+SRG+Lp(a); Estatisticamente significativo para $p < 0,05$.

As AUCs foram de 0,738 para os FRT, de 0,758 adicionando o SRGm aos FRT e de 0,767 com a inclusão da Lp(a) ao modelo anterior (Figura 27). O incremento da área sob a curva com a adição quer do SRGm quer da Lp(a) foi significativa (Delong $p < 0,0001$).

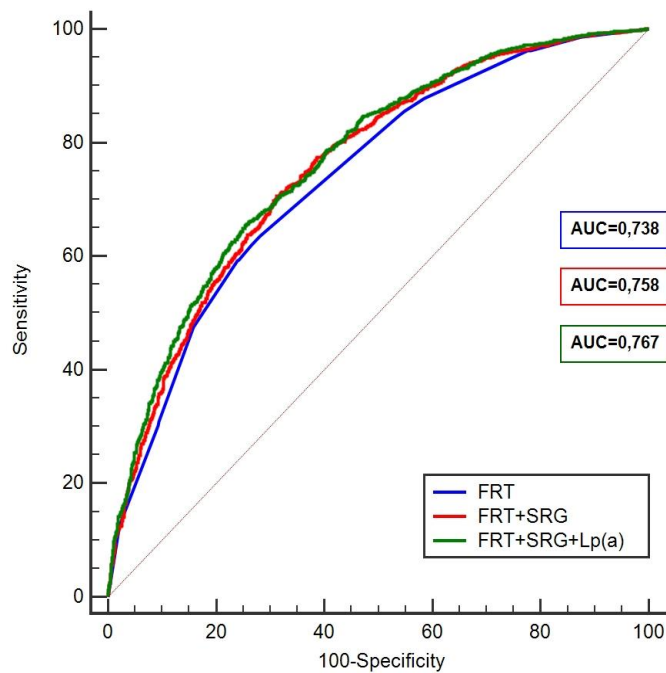


Figura 27: Gráfico das curvas AUC das ROC com os FRT, SRGm e Lp(a) para a predição de DC.

IV.4.7. SRG E HISTÓRIA FAMILIAR DE DOENÇA CORONÁRIA PRECOCE

A tabela 49 mostra o poder preditor basal, expresso pela área sob a curva ROC que é máxima para os FRT (AUC=0,74), intermédia para o SRGm (AUC=0,61) e reduzida para a HF de DC precoce (AUC=0,56). A associação do SRGm aos FRT permitiu incrementar o poder preditor de DC com melhor sensibilidade (AUC=0,76). A associação dos FRT com o SRG e com a HFDC precoce, não melhorou o poder preditivo para DC em relação ao modelo dos FRT com o SRGm (AUC=0,76).

As AUCs foram de 0,74 para os FRT, 0,56 para a HFDC precoce e de 0,61 para o SRGm. A AUC dos FRT com o SRG (0,76) foi igual à AUC dos FRT+SRG + HF (0,76) (Figura 28).

Tabela 49: Comparação do valor preditivo para DC do SRGm com os FRT e com a História Familiar de DC Precoce: Curvas AUC

Modelos	AUC (IC 95%)	Sensib. (%)	Especifi. (%)	VPP	VPN	Teste Delong (p)
FRCV	0,74 (0,72 – 0,75)	61,9	73,5	2,34	0,52	-----
HF DC	0,56 (0,54 – 0,57)	23,8	87,4	1,89	0,87	-----
SRGm	0,61 (0,59 – 0,62)	54	62,3	1,43	0,74	-----
FRCV+HF DC	0,75 (0,73 – 0,76)	65,1	71,9	2,32	0,48	0,001
FRCV+SRGm	0,76 (0,74 – 0,77)	70,6	68,8	2,26	0,43	<0,0001
FRCV+HF DC+SRGm	0,76 (0,75 – 0,78)	60,3	78,7	2,84	0,50	<0,0001

AUC – Area under the curve; IC – Intervalo de confiança; VPP – Valor preditivo positivo; VPN – Valor preditivo negativo; FRCV – Fatores de risco CV; HF – História familiar de doença coronária; SRGm – Score de risco genético multiplicativo; Estatisticamente significativo para $p < 0,05$.

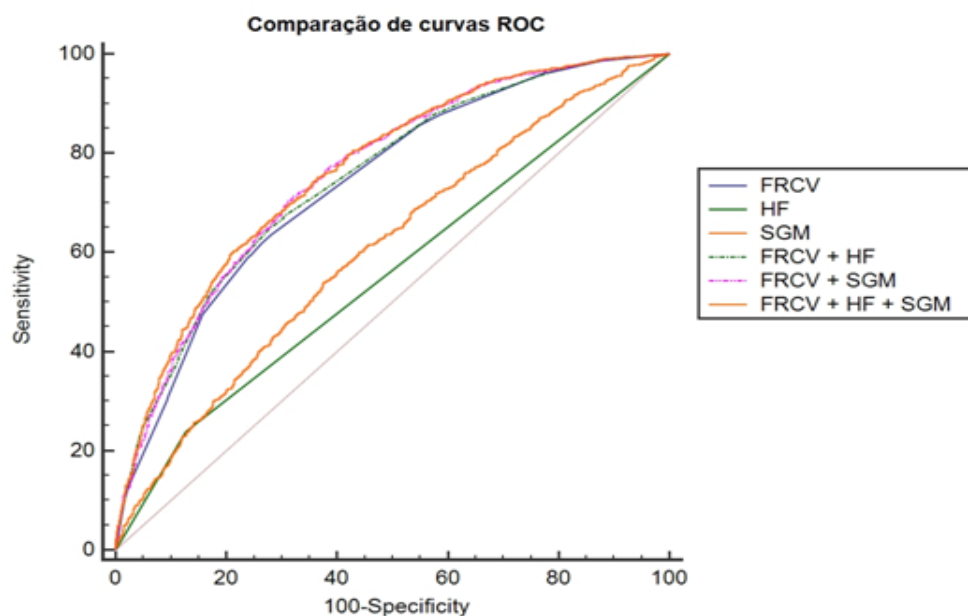


Figura 28: Gráfico das curvas ROC dos FRT, HFDC precoce e SRGm para a predição de DC

O SRGm aumenta de forma significativa o grau de predição para DC, conferido pelos FRT habituais (Tabela 50). Com efeito, a utilização do SRGm permite uma mais correta reclassificação média (NRI 0,31 $p < 0,001$) e discriminação (IDI 0,02, $p < 0,0001$) do que a obtida pela HFDC.

A utilização da informação clínica dada pelos doentes sobre HF, não melhora a capacidade de reclassificação média (NRI), nem da respetiva discriminação (IDI).

Tabela 50: Reclassificação (NRI) e Discriminação (IDI) obtida pelo SRGm e História Familiar de Doença Coronária Precoce

	NRI (IC 95%)	Valor p	IDI (IC 95%)	Valor p
FRT + SRGm	0,308 (0,240 – 0,375)	<0,0001	0,020 (0,015 – 0,025)	<0,0001
FRT + HFDCP	0,224 (0,168 – 0,279)	<0,0001	0,014 (0,009 – 0,018)	<0,0001
FRT+ HFDCP + SRGm	0,290 (0,223 – 0,358)	<0,0001	0,018 (0,013 – 0,023)	<0,0001

SRG-Score de Risco Genético Multiplicativo, FRT-Fatores de risco tradicionais HFDCP- História familiar de Doença Coronária Precoce. NRI- Net Reclassification index, IDI- Discrimination index.

IV.4.8. Estudo do SRG no doente com DC antes dos 45 anos:

IV.4.8.1. Estudo caso-controlo: características clínicas, bioquímicas e laboratoriais

Os jovens com DC, documentada antes dos 45 anos, eram mais frequentemente do género masculino (86,3%) do que os do grupo controlo (76,4%) (Tabela 51).

A população com DC, documentada antes dos 45 anos de idade, realizava menos exercício físico (41,7% versus 57,6%) e apresentava maior prevalência de tabagismo (69,2% versus 23,4%) e dislipidemia (87,1% versus 70,3%) relativamente aos controlos.

Não se encontraram diferenças com significado estatístico, quanto aos restantes fatores de risco CV, quanto à VOP e nem quanto ao IMC. O grupo dos doentes exibiu melhor controlo da frequência cardíaca, possivelmente relacionado com o uso de beta-bloqueantes neste grupo.

Tabela 51: Estudo do SRG no jovem (<45anos) com DC, características demográficas e FR modificáveis para DC

Variáveis	DC <45 anos (n=240)	Controlos (n=1322)	Valor p
Género masculino, n (%)	207 (86,3)	1010 (76,4)	0,001
Idade, anos	39,7 ± 3,8	52,7 ± 7,8	<0,0001
Exercício físico ^o , n (%)	100 (41,7)	761 (57,6)	<0,0001
Abuso álcool*, n (%)	48 (20)	181 (13,7)	0,011
Fumador [†] , n (%)	166 (69,2)	309 (23,4)	<0,0001
HTA, n (%)	132 (55)	700 (53)	0,558
VOP>10m/s, n (%)	10 (4,2)	180 (13,6)	<0,0001
Diabetes, n (%)	38 (15,8)	175 (13,2)	0,281
Dislipidemia ^o , n (%)	209 (87,1)	930 (70,3)	<0,0001
IMC>30kg/m ² , n (%)	71 (29,6)	386 (29,2)	0,904
Frequência cardíaca, bpm	69,5 ± 11,9	72,3 ± 11,5	<0,0001

oMais do que 40min/semana; *Mais do que 300gr/semana; †Fumadores ou <5 anos de cessação; oHDL<40 ou LDL>140 ou Triglicéridos>150; Valores contínuos expressos pela média±desvio padrão; estatisticamente significativo para p<0.05.

A tabela 52 mostra diferenças encontradas nos dados laboratoriais dos doentes jovens com DC. Esta amostra de indivíduos jovens com DC apresentou valores mais elevados de leucócitos, plaquetas (variáveis contínuas) e Lp(a) superior a 25 mg/dl (variável categórica).

Não se verificaram diferenças com significância quanto: à elevação da Apo B (>90 mg/dl), elevação da homocisteína (variável contínua ou categórica), insuficiência renal estimada pela clearance creatinina, nem da PCR (variável contínua ou categórica).

Tabela 52: Características bioquímicas-laboratoriais do jovem com DC antes dos 45 anos

Variáveis	Casos (n=240)	Controlos (n=1322)	Valor p
Hemoglobina, g/dl	14,9 (14,3 – 15,8)	14,7 (14 – 15,4)	0,001*
Plaquetas, 10 ³ /μl	220,5 (196 – 273,3)	220 (190 - 252)	0,033*
Leucócitos, 10 ³ /μl	7,4 (6,1 – 8,9)	6,6 (5,6 – 7,8)	<0,0001*
Fibrinogénio, mg/dl	364,5 (316 – 420,8)	361 (315 - 409)	0,423*
Lp(a)>25mg/dl, n (%)	100 (41,7)	374 (28,3)	<0,0001
Apo B>90mg/dl, n (%)	139 (57,9)	695 (52,6)	0,127
Homocisteína, μmol/L	11,4 (9,6 - 14)	11,4 (9,7 – 13,6)	0,932*
>10 μmol/L, n (%)	166 (69,2)	932 (70,5)	0,678
Creatinina, mg/dl	0,96 (0,85 – 1,08)	0,99 (0,89 – 1,1)	0,062*
Clearance Cr,**, ml/min	112,4 (99,3 – 127,8)	92,7 (78,9 – 108,7)	<0,0001*
Clearance Cr,<60ml/min, n (%)	5 (2,1)	49 (3,7)	0,205
PCR, mg/L	2,7 (1,49 – 6,3)	2,7 (1,2 – 5,3)	0,104*
>3mg/L, n (%)	93 (38,8)	496 (37,5)	0,717

Lp(a) – Lipoproteína (a); Apo B – Apolipoproteína B; **Cockcroft-Gault; Cr. – Creatinina; PCR – Proteína C reativa; Valores contínuos expressos pela mediana (Percentil 25 – Percentil 75); *Valores obtidos através do teste não paramétrico Mann-Whitney. Estatisticamente significativo para p<0.05.

IV.4.8.2. Estudo gene a gene

Por análise gene a gene, tendo em conta os possíveis modelos genéticos, encontraram-se quatro variantes genéticas associadas de forma estatisticamente significativa com a DC (Tabela 53); nomeadamente a variante TCF21 com um OR de 1,24 (IC95%: 1,01-1,54); a variante ZC3HC1 com um OR de 1,72 (IC95%: 1,03-2,85); a variante ZNF259 com um OR de 1,35 (IC95%: 1,02-1,80) e a variante da ECA com um OR de 1,36 (IC95%: 1,03-1,80). Estas variantes genéticas atuam em três diferentes eixos fisiopatológicos, o do SRAA (HTA), o do eixo lipídico e o do ciclo celular.

Tabela 53: Estudo Gene a Gene (bivariado) na subpopulação de doentes com DC antes dos 45 anos

Genes		Casos (<45 anos) (n=240)	Controlos (n=1322)	Modelos				
				Dominante	Recessivo	Aditivo	Multiplicativo	Co-dominante
TCF21	GG	21 (8,8%)	160 (12,1%)	1,44	1,28	1,24	1,24	1,58
	CG	104 (43,3%)	609 (46,1%)	(0,89 - 2,31)	(0,97 - 1,69)	(1,01 - 1,54)	(1 - 1,53)	(0,96 - 2,61)
	CC	115 (47,9%)	553 (41,8%)	p=0,135	p=0,080	p=0,044	p=0,045	p=0,070
ZC3HC1	TT	20 (8,3%)	165 (12,5%)	1,57	1,27	1,26	1,25	1,72
	CT	109 (45,4%)	623 (47,1%)	(0,97 - 2,55)	(0,96 - 1,67)	(1,02 - 1,56)	(1,02 - 1,54)	(1,03 - 2,85)
	CC	111 (46,3%)	534 (40,4%)	p=0,067	p=0,090	p=0,033	p=0,035	p=0,037
ZNF259	CC	150 (62,5%)	916 (69,3%)	1,35	0,92	1,25	1,25	1,02
	CG	84 (35,0%)	370 (28,0%)	(1,02 - 1,80)	(0,38 - 2,20)	(0,98 - 1,61)	(0,97 - 1,59)	(0,42 - 2,46)
	GG	6 (2,5%)	36 (2,7%)	p=0,038	p=0,844	p=0,077	p=0,080	p=0,969
ECA	II	27 (11,3%)	188 (14,2%)	1,31	1,36	1,26	1,25	1,52
	ID	109 (45,4%)	658 (49,8%)	(0,85 - 2,01)	(1,03 - 1,80)	(1,02 - 1,55)	(1,02 - 1,53)	(0,96 - 2,40)
	DD	104 (43,3%)	476 (36,0%)	p=0,219	p=0,031	p=0,030	p=0,033	p=0,071

IV.4.8.3. Diferentes modelos de SRG no jovem

Todos os modelos de SRG calculados apresentaram valores mais elevados nos doentes com DC, antes dos 45 anos do que na população controlo (Tabela 54).

Tabela 54: Diferentes modelos de score genético no jovem com DC antes dos 45 anos

Scores	Total (n=1562)	DC<45anos (n=240)	Controlos (n=1322)	Valor p
SRGm	0,37 (0,22 – 0,59)	0,42 (0,26 – 0,69)	0,36 (0,22 – 0,57)	<0,0001
SRGa	24 (21 - 26)	24 (21 – 26)	23 (21 – 26)	0,073
SRG MM	26,1 (23,33 – 28,85)	26,8 (23,53 – 29,39)	26 (23,3 – 28,76)	0,028
SRG B	1,64 (1,34 – 1,92)	1,71 (1,42 – 1,97)	1,62 (1,32 – 1,9)	0,001

SRG MG – Score de risco genético multiplicativo (Genotípico); SRG A – Score de risco genético aditivo; SRG MM – Score de risco genético (Melhor modelo); SRG B – Score de risco genético Beta; Estatisticamente significativo para p<0,05.

IV.4.8.4. Poder preditor do SRG no jovem

Na tabela 55 pode comparar-se o poder preditor dos diferentes modelos de SRG na população jovem com DC. O SRGm obteve o melhor poder preditor com AUC de 0,58 (IC95%: 0,55-0,60) e OR para DC de 1,30 (IC95%: 1,05-1,61) com equilíbrio entre sensibilidade (65%) e razoável especificidade (46%). Os restantes modelos produziram menores incrementos da AUC, sendo no entanto também significativos.

Tabela 55: Poder preditor dos MODELOS de SCORE GENÉTICO NO JOVEM COM DC antes dos 45 anos

Modelos Scores	OR (IC 95%)	Valor p ¹	AUC (IC 95%)	Sensib. (%)	Especif. (%)	VPP	VPN	Valor p ²
SRG MG	1,30 (1,05 – 1,61)	0,018	0,58 (0,55 – 0,60)	65,4	46,2	1,22	0,75	0,0001
SRG A	1,04 (1,00 – 1,08)	0,049	0,54 (0,41 – 0,56)	68,7	38,8	1,12	0,81	0,077
SRG MM	1,04 (1,01 – 1,08)	0,017	0,55 (0,52 – 0,57)	44,6	63,7	1,23	0,87	0,029
SRG B	1,79 (1,30 – 2,45)	<0,0001	0,57 (0,54 – 0,59)	37,9	72,9	1,40	0,85	0,0006

SRG MG – Score de risco genético multiplicativo (Genotípico); SRG A – Score de risco genético aditivo; SRG MM – Score de risco genético (Melhor modelo); SRG B – Score de risco genético Beta; OR – Odds ratio; AUC – Area under curve; IC – Intervalo de confiança; Valor p¹ – Obtido através da regressão logística; Valor p² – Obtido através da curva ROC para avaliar a significância da AUC; Estatisticamente significativo para p<0,05.

A AUC dos FRT foi de 0,76 e a do FRT com SRG foi de 0,79. O incremento da AUC com a adição do SRG foi significativo (p=0,004)

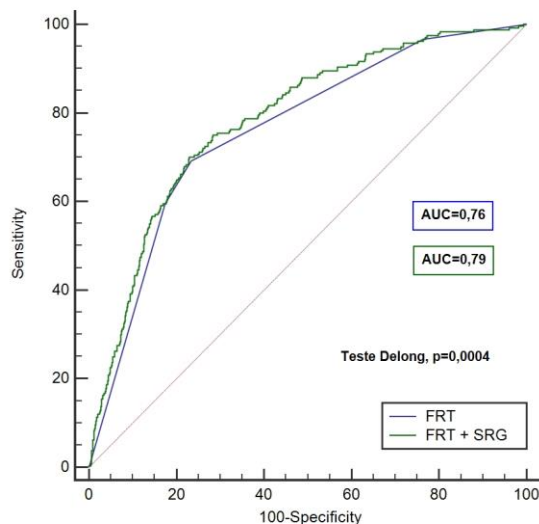


Figura 29: Gráfico das curvas ROC dos FRT e SRG para a predição de DC no jovem (<45 anos)

No subgrupo de doentes jovens com DC a incorporação do SRG também provou incrementos positivos na curva AUC de 0,76 para 0,79 com acréscimo da sensibilidade (70%) e especificidade (77%), em relação a um modelo basal de FRT ($p=0,0004$) (Tabela 56).

Tabela 56: Poder preditor do SRG no jovem com DC antes dos 45 anos

Modelos	AUC (IC 95%)	Sensib. (%)	Especif. (%)	VPP	VPN	Teste Delong (p)
FRT	0,76 (0,74 – 0,78)	69,2	76,6	2,96	0,40	0,0004
FRT+SRG	0,79 (0,76 – 0,81)	70	77,1	3,05	0,39	

FRT – Fatores de risco tradicionais; SRG – Score de risco genético; AUC – Area under the curve; IC – Intervalo de confiança; VPP – Valor preditivo positivo; VPN – Valor preditivo negativo; Estatisticamente significativo para $p<0,05$.

IV.4.8.5. Medidas de reclassificação média (NRI) e melhoria de discriminação (IDI) com SRG:

A tabela 57 demonstra um NRI contínuo global de 26%, sendo que ocorre *net* reclassificação com interesse clínico em 47,6% dos controlos. Pelo contrário, na população jovem com DC, obteve-se um NRI de -21,6%. Este NRI negativo indica maior capacidade em reclassificar doentes num risco mais baixo (60,8%) do que num aumento de risco (39,2%).

Tabela 57: Tabela de Reclassificação do Risco (NRI) para Doença coronária com a inclusão do SCORE Risco Genético (Risco Contínuo)

	Risco		Valor p (cfNRI)	cfNRI (%)	(IC 95%)
	Aumento, n (%)	Diminuição, n (%)			
DC<45 anos (n=240)	94 (39,2)	146 (60,8)	<0,0001	-21,6	(-34-(-9))
Controlos (n=1322)	346 (26,2)	976 (73,8)	<0,0001	47,6	(42,8-52,2)
Total (n=1562)	-----	-----	<0,0001	26	(12,6-39,1)

IC – Intervalo de confiança; cfNRI – category-free net reclassification index. Estatisticamente significativo para $p<0,05$.

O SRG produziu melhoria das medidas de discriminação do modelo basal com um IDI de 0,6% (0,1-1) de modo estatisticamente significativo.

Tabela 58: Índice de melhoria de discriminação (IDI) para DC com a inclusão do SRG (Risco Contínuo)

Modelos	IDI (%)	(IC 95%)	Valor p
FRT versus FRT+SRG	0,6	(0,1-1)	0,017

IDI – Integrated Discrimination Improvement; FRT – Fatores de risco tradicionais (Dislipidemia, Tabagismo, Diabetes e Hipertensão); SRG – Score de risco genético; Estatisticamente significativo para $p < 0,05$.

IV.5 RECLASSIFICAÇÃO DO RISCO CARDIOVASCULAR COM O SRGm:

As tabelas 59 a 67 mostram os resultados obtidos na estimativa do risco predita pelo SRG por comparação com:

- Risco a dez anos de doença CV fatal estimado pelo Score Europeu (SE).
- Risco a dez anos de DC estimado pelo calculador DC de Framingham. (Framingham DC)
- Risco a dez anos de doença CV estimado pelo calculador CV de Framingham (Framingham CV)

Em cada análise apresentam-se os seguintes resultados:

- Análise de regressão logística multivariada por subgrupos de risco com e sem incorporação do SRGm,
- Curvas ROC e AUC para cada modelo, com e sem a inclusão do SRGm em cada subgrupo,
- Índices de reclassificação (NRI) e índices de discriminação (IDI) por subgrupos de risco.

IV.5.1 RISCO CARDIOVASCULAR PELO SCORE EUROPEU

Foi computada a função de risco de mortalidade CV a dez anos, dada pelo SE, utilizando como referência as tabelas para países de baixo risco CV, onde atualmente Portugal é contemplado. Foram divididos em três grupos de acordo com o risco calculado: baixo risco CV <2% (n=1710), risco intermédio CV 2-5% (n=859) e risco elevado >5% (n=319) (Tabela 59).

Em todos os subgrupos, os FRT foram preditores independentes para morte CV no modelo basal ajustado para a idade e o géneros, sem o SRGm (resultados dos modelos basais não descritos na tabela). O modelo combinado de FRT e SRG foi mais preditivo que o modelo basal (AUC=0,76; IC 95% 0,74 – 0,77; teste de Delong $p < 0,0001$) (Tabela 59). Em todos os subgrupos o SRGm foi preditor independente para DC ajustando para os FRT em análise multivariada ($p < 0,0001$). Aplicando a correcção de Bonferroni, o SRGm permanece significativo para $p < 0,01$ (0,05/5) (Tabela 59).

Nos 1710 indivíduos com risco CV a dez anos, estimado pelo SE, inferior a 2%, todos os FRT foram preditores independentes do evento com associação mais evidente pelo tabagismo com um OR de 4,22 (IC95%: 3,28-5,43), pela dislipidemia com um OR de 3,47 (IC95%: 2,63-4,57) e pela DM com um OR de 3,24 (IC95%: 2,40-4,37) (Tabela 59).

No grupo de indivíduos considerado de baixo risco pelo HEARTSCORE, quando se adiciona o SGRm, todas as variáveis mantêm-se como preditores independentes para o evento fatal, sendo que há um discreto acréscimo do valor do OR para os FRT. O SRG, neste subgrupo de baixo risco, foi preditor independente do evento fatal com um OR de 1,82 (IC95%: 1,46-2,27; $p < 0,0001$) (Tabela 59).

Nos 859 indivíduos, com risco CV fatal a dez anos, estimado pelo SE, entre 2 a 5%, todos os FRT foram preditores independentes do evento fatal, com associação mais evidente pelo tabagismo com um OR de 3,67 (IC95%: 2,63-5,13) e pela DM com um OR de 3,40 (IC95%: 2,41-4,80). Quando se adiciona o SGRm, todas as variáveis mantêm-se como preditores independentes para a DC. Com efeito, há um discreto acréscimo do valor do OR para os FRT. O SRGm neste subgrupo de risco foi preditor independente do evento fatal com um OR de 1,57 (IC 95%: 1,22– 2,01; $p < 0,0001$) (Tabela 59).

Num modelo multivariado, contendo os FRT e o SRG, apenas no subgrupo de risco superior a 5% ($n=319$), a HTA deixa de ter significado multivariado.

Todos os outros FRT mantêm-se preditores independentes e com ligeiro acréscimo da força de associação. Neste subgrupo, o SRG apresenta maior força de associação com o evento fatal, expresso por um OR de 3,39 (IC 95%: 1,78 – 6,45; $p < 0,0001$) (Tabela 59).

Tabela 59: Análise de regressão logística multivariada, por subgrupos de risco do score Europeu com e sem incorporação do SRGm.

Variáveis	SE <2% (n=1710)		SE 2-5% (n=859)		SE > 5% (n=319)		Total	
	OR (95% CI)	Valor p	OR (95% CI)	Valor p	OR (95% CI)	Valor p	OR (95% CI)	Valor p
HTA	2,54 (2,03 – 3,17)	<0,0001	1,84 (1,27 – 2,67)	0,001	-----	-----	1,93 (1,63 – 2,3)	<0,0001
DM	3,24 (2,40 – 4,37)	<0,0001	3,40 (2,41 – 4,80)	<0,0001	2,14 (1,27 – 3,62)	0,004	2,97 (2,42 – 3,65)	<0,0001
Dislipid.	3,47 (2,63 – 4,57)	<0,0001	2,67 (1,72 – 4,15)	<0,0001	2,24 (1,02 – 4,95)	0,046	3,01 (2,42 – 3,76)	0,030
Tabag.	4,22 (3,28 – 5,43)	<0,0001	3,67 (2,63 – 5,13)	<0,0001	2,50 (1,51 – 4,14)	<0,0001	3,37 (2,82 – 4,03)	<0,0001
SRGm	1,82 (1,46 – 2,27)	<0,0001	1,57 (1,22 – 2,01)	<0,0001	3,39 (1,78 – 6,45)	<0,0001	1,8 (1,52 – 2,12)	<0,0001
Constante	0,10	<0,0001	0,12	<0,0001	0,174	<0,0001	0,12	<0,0001
H.L. (FRT)	$\chi^2=3,04$	0,882	$\chi^2=1,67$	0,947	$\chi^2=0,98$	0,914	$\chi^2=4,04$	0,672
H.L. (FRT+ SRGm)	$\chi^2=8,76$	0,363	$\chi^2=5,14$	0,742	$\chi^2=7,99$	0,435	$\chi^2=15,63$	0,05

*Método stepwise Forward Wald (SPSS versão 19.0); SE – Score Europeu risco a 10 anos; OR – Odds ratio; SRGm – Score de risco genético multiplicativo; Estatisticamente significativo para $p<0.05$ IC – Intervalo de confiança; Análise multivariada ajustada para os FRT (como variáveis dicotômicas). dislipidemia= LDL>150, HDL<40 para os homens e <45 para as mulheres, triglicéridos>150.

Análise do ajuste do modelo: Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinómios fracionários. Análise de resíduos de Pearson e Deviance para detetar outliers e aos resíduos DfBetas e à distância de Cook para detetar observações influentes, não se tendo observado nenhuma observação atípica. Tolerância e VIF >0.2 e <5 demonstrando ausência de multicolinearidade.

Em todos os subgrupos de risco, o modelo combinado (FRT+SRGm) produziu incrementos significativos da AUC basal (Tabela 60). No global, a curva AUC incrementou de 0,74 (0,72-0,75) para 0,76 (0,74-0,77) com sensibilidade de 71% e 69% de especificidade.

O SRGm, em todos os subgrupos, produz o aumento da especificidade que é máxima no grupo de risco intermédio (74,5%) à custa de discreta redução da sensibilidade. O valor preditivo positivo aumenta em todos os subgrupos e também é maior no subgrupo de risco intermédio (2,47%) (Tabela 60).

No subgrupo de menor risco CV, estimado pelo SCORE, a AUC produzida pelo conjunto dos FRT situou-se nos 0,77 com incremento discreto, mas significativo ($p<0,0001$) para 0,79, quando incorporado no modelo o SRGm (Tabela 60).

No subgrupo de risco CV intermédio, estimado pelo SCORE, a AUC produzida pelo conjunto dos FRT situou-se nos 0,71 com incremento discreto, mas significativo ($p<0,0001$) para 0,73, quando incorporado no modelo preditor o SRGm (Tabela 60).

No subgrupo de risco CV elevado, estimado pelo SCORE, a AUC produzida pelo conjunto dos FRT situou-se nos 0,67 com incremento significativo ($p < 0,0001$) para 0,73, quando incorporado no modelo preditor o SRGm (Tabela 60).

Tabela 60: Comparação de curvas AUROC antes e após inclusão do SRG em subgrupos de risco (SE)

Score Europeu	Modelos	AUC (IC 95%)	Sensib. (%)	Especif. (%)	VPP	VPN	Teste Delong
<2%	FRT	0,77 (0,74 – 0,79)	79,1	60,3	1,99	0,35	<0,0001
	FRT + SRG	0,79 (0,76 – 0,80)	75,1	68,6	2,40	0,36	
2%-5%	FRT	0,71 (0,68 – 0,74)	68	65,9	1,99	0,49	0,012
	FRT + SRG	0,73 (0,70 – 0,76)	63	74,5	2,47	0,50	
>5%	FRT	0,67 (0,61 – 0,72)	79,5	45,2	1,45	0,45	0,004
	FRT + SRG	0,73 (0,68 – 0,78)	76,4	60,5	1,93	0,39	
Total	FRT	0,74 (0,72 – 0,75)	61,9	73,5	2,34	0,52	<0,0001
	FRT + SRG	0,76 (0,74 – 0,77)	70,6	68,8	2,26	0,43	

FRT – Fatores de risco tradicionais; SRGm – Score de risco genético; AUC – Area under curve; IC – Intervalo de confiança; VPP – Valor preditivo positivo; VPN – Valor preditivo negativo; Estatisticamente significativo para $p < 0,05$.

A adição do SRGm ao modelo basal dos FRT melhorou a capacidade discriminativa e reclassificativa dos modelos em todos os subgrupos (Tabela 61). O SRGm produziu um NRI não categórico de 32,2% (IC95%: 23,4-41,0%; $p < 0,0001$) no grupo de baixo risco; 26,9% (IC95%: 14,8-39,0; $p < 0,0001$) no grupo de risco intermédio e 45,4% (IC95%: 24,7-66,1; $p < 0,0001$) no grupo de elevado risco.

A inclusão do SRGm no modelo dos FRT produziu um IDI de 2% (IC95%: 1,4-2,6%; $p < 0,0001$) no grupo de menor risco; 1,7 % (IC95%: 0,8-2,5%; $p < 0,0001$) no grupo de risco intermédio e 5,7% (IC95%: 3,4- 8,1%; $p < 0,0001$) no grupo de risco mais elevado (Tabela 61).

O valor do R quadrado do Nagelkerke aumentou em cada subgrupo, com a inclusão do SRG, isto é, a nova variável melhorou a proporção da capacidade preditiva no modelo. Assim, 30%, 22% e 20% da mortalidade CV são explicadas pelas variáveis independentes (FRT e SRG) no grupo de risco baixo, intermédio e elevado, respetivamente.

Tabela 61: Reclassificação (NRI) e Discriminação (IDI) obtida pelo SRGm em relação ao SE

	SCORE EUROPEU*			
	Total (n=2888)	<2% (n=1710)	2-5% (n=859)	>5% (n=319)
Casos (n)	1566	884	487	195
Controlos (n)	1322	826	372	124
NRI contínuo (%)	30,8 (24-37,5)	32,2 (23,4-41,0)	26,9 (14,8-39,0)	45,4 (24,7-66,1)
	<0,0001	<0,0001	<0,0001	<0,0001
IDI (%)	2,0 (1,5-2,5)	2,0 (1,4-2,6)	1,7 (0,8-2,5)	5,7 (3,4-8,1)
	<0,0001	<0,0001	<0,0001	<0,0001
	Nagelkerke R Square			
FRT	0,231	0,281	0,195	0,123
FRT + SRG	0,256	0,303	0,216	0,199

SE* % risco a 10 anos. Intervalo de confiança a 95% apresentado entre parêntesis; IDI – Integrated discrimination improvement; NRI – Net reclassification improvement; FRT – Fatores de risco tradicionais (Dislipidemia, Tabagismo, Diabetes e Hipertensão); SRG – Score de risco genético.

IV.5.2 RISCO CORONÁRIO PELO SCORE DE FRAMINGHAM

Foi computada a função de risco de DC proposta por D’Agostino *et al.* (92) que estima o risco a dez anos de DC.

Em todos os subgrupos, o SRGm foi um preditor independente dos eventos coronários com significado estatístico ($p < 0,0001$) (Tabela 62). Aplicando a correção de Bonferroni, o SRG permanece significativo para $p = 0,01$ (0,05/5).

O modelo multivariado, contendo os FRT e o SRG no subgrupo de risco intermédio (Framingham DC 10-20%), revela que a dislipidemia deixa de ter significado multivariado. No subgrupo de elevado risco (Framingham DC >20%) a HTA e a dislipidemia deixam de ter significado multivariado, sendo que todos os outros FRT mantêm-se preditores independentes e com ligeiro acréscimo do seu OR basal (modelo basal não expresso).

No subgrupo de risco elevado, o SRGm apresenta maior força de associação com um OR de 2,5 (IC 95%: 1,5 – 4,1).

Tabela 62: Análise de Regressão Logística Multivariada, por subgrupos de risco de Framingham DC, com e sem incorporação do SRGm.

Variáveis	Fr. DC <10% (n=1312)		Fr. DC 10%-20% (n=1049)		Fr. DC > 20% (n=527)		Total	
	OR (IC 95%)	Valor p	OR (IC 95%)	Valor p	OR (IC 95%)	Valor p	OR (IC 95%)	Valor p
HTA	3,25 (2,52 – 4,18)	<0,0001	1,88 (1,39 – 2,56)	<0,0001	-----	-----	1,93 (1,63 – 2,3)	<0,0001
Diabetes	2,93 (1,90 – 4,52)	<0,0001	4,34 (3,14 – 6)	<0,0001	3,56 (2,34 – 5,40)	<0,0001	2,97 (2,42 – 3,65)	<0,0001
Dislipidemia	1,71 (1,28 – 2,29)	<0,0001	-----	-----	-----	-----	3,01 (2,42 – 3,76)	0,030
Tabagismo	4,92 (3,66 – 6,62)	<0,0001	3,57 (2,65 – 4,80)	<0,0001	3,11 (2,04 – 4,73)	<0,0001	3,37 (2,82 – 4,03)	<0,0001
SRGm	1,76 (1,41 – 2,20)	<0,0001	1,86 (1,39 – 2,49)	<0,0001	2,49 (1,54 – 4,05)	<0,0001	1,8 (1,52 – 2,12)	<0,0001
Constante	0,16	<0,0001	0,24	<0,0001	0,342	<0,0001	0,12	<0,0001
H.L.(FRT)	$X^2=7,82$	0,251	$X^2=1,08$	0,956	$X^2=0,11$	0,946	$X^2=4,04$	0,672
H.L. (FRT + GRS)	$X^2=6,92$	0,545	$X^2=8,23$	0,411	$X^2=3,60$	0,891	$X^2=15,63$	0,05

*Método stepwise Forward Wald (SPSS versão 19.0); OR – Odds ratio; DC – % de risco a 10 anos de Doença Coronária; SRGm – Score Risco Genético, $p<0,05$. IC – Intervalo de confiança; Análise multivariada ajustada para os FRT (como variáveis dicotômicas). Dislipidemia=LDL>150, HDL<40 para os homens e <45 para as mulheres, Triglicéridos>150,

HL- Hosmer Lemeshow χ^2 . Análise do ajuste do modelo: Verificou-se o pressuposto da linearidade com a função logit, através do método dos polinômios fracionários. Análise de resíduos de Pearson e Deviance para detetar outliers e aos resíduos DfBetas e à distância de Cook para detetar observações influentes, não se tendo observado nenhuma observação atípica Análise do ajuste do modelo: Tolerância e VIF >0.2 e <5 demonstrando ausência de multicolinearidade.

A comparação do poder preditor do modelo combinado de FRT com o SRGm e com o modelo basal de FRT é demonstrada na tabela 63. Nesta demonstram-se os valores das AUC das curvas ROC em cada subgrupo, de acordo com a estratificação a dez anos do risco para DC em três subgrupos (baixo risco <10%; risco intermédio 10-20%, risco elevado>20%).

Em cada subgrupo, constatou-se que a inclusão do SRGm no modelo basal dos FRT leva a um incremento preditor do novo modelo combinado, demonstrando que o SRGm acresce em poder preditivo independente do risco basal estimado. Concretamente, no subgrupo de baixo risco, a AUC elevou-se de 0,72 para 0,75 ($p<0,0001$). Este aumento de capacidade preditiva, embora pequeno em termos absolutos (2,5% IC 1,4-3,6%), é significativo quando olhamos o valor p do teste de Delong para comparação das curvas ROC (Tabela 63).

Do mesmo modo, no subgrupo de risco intermédio a AUC elevou-se de 0,70 para 0,73 com 2,6% de incremento na capacidade preditiva (IC 95% 1,2-3,9%) para um valor de $p=0,0002$. Por último, no subgrupo de risco elevado o SRGm adicionado ao modelo basal também produziu incrementos significativos (AUC basal 0,68 para AUC final 0,72, com 4,4% de incremento da capacidade preditiva (IC 95% 1,5-7,3%) com

teste de Delong significativo $p=0,003$. Em todos os subgrupos, o IC não incluiu o zero sugerindo que o aumento documentado nas AUC é estatisticamente significativo (Tabela 63).

Tabela 63: Comparação das curvas AUROC antes e após inclusão do SRGm em subgrupos de risco (Framingham DC)

Fr DC (%)	Modelos	AUC (IC 95%)	Sensib. (%)	Especif. (%)	VPP	VPN	Teste Delong
<10%	FRT	0,72 (0,70 – 0,75)	71,5	60,8	1,82	0,47	<0,0001
	FRT + SRGm	0,75 (0,73 – 0,77)	71,9	64,3	2,01	0,44	
10%-20%	FRT	0,70 (0,67 – 0,73)	60,2	71,5	2,11	0,56	0,0002
	FRT + SRGm	0,73 (0,70 – 0,75)	67,1	68,7	2,14	0,48	
>20%	FRT	0,68 (0,63 – 0,72)	62,7	61,3	1,62	0,61	0,003
	FRT + SRGm	0,72 (0,68 – 0,76)	70,1	64,7	1,99	0,46	
Total	FRT	0,74 (0,72 – 0,75)	61,9	73,5	2,34	0,52	<0,0001
	FRT + SRGm	0,76 (0,74 – 0,77)	70,6	68,8	2,26	0,43	

FR DC – % risco de DC a 10 anos de Framingham. FRT- Fatores de risco tradicionais; SRGm – Score de risco genético ; AUC – Area under curve; IC – Intervalo de confiança; VPP – Valor preditivo positivo; VPN – Valor preditivo negativo; Utilizado o teste Delong para a comparação das AUC quando adicionado o SRG aos FRT; Estatisticamente significativo para $p<0,05$.

A tabela 64 revela os valores de reclassificação e discriminação obtidos pelo SRGm nos subgrupos de risco de DC a dez anos de Framingham. O SRGm produziu um NRI não categórico (contínuo) de 32,3% (IC95%: 22,4-42,3%; $p<0,0001$) no grupo de baixo risco; 30,4% (IC95%: 19-41,8%; $p<0,0001$) no grupo de risco intermédio e 29,8% (IC 95%: 13,1-46,6%; $p<0,0001$) no grupo de risco elevado.

A inclusão do SRG no modelo dos FRT produziu um IDI no grupo de menor risco de 2,4% (IC95%: 1,6-3,2%; $p<0,0001$); no grupo de risco intermédio de 2,1 % (IC95%: 1,3-2,9%; $p<0,0001$) e no grupo de risco mais elevado 3,3% (IC 95%: 1,9- 4,7%; $p<0,0001$) (Tabela 64).

O valor do R quadrado do Nagelkerke aumentou em cada subgrupo com a inclusão do SRG, isto é, a nova variável melhorou a proporção da capacidade preditiva no modelo. Assim, 24,2%, 19,1 % e 18,4% do risco para DC a dez anos, são explicadas pelas variáveis independentes (FRT e SRG), respetivamente, nos grupos de baixo, intermédio e risco elevado (Tabela 64).

Tabela 64: Reclassificação (NRI) e Discriminação (IDI) obtida pelo Score Genético em subgrupos de risco Framingham DC

	FR DC			
	Total (n=2888)	<10% (n=1312)	10-20% (n=1049)	>20% (n=527)
DC (n)	1566	629	583	354
Controlos (n)	1322	683	466	173
NRI contínuo (%)	31,7 (25,0-38,4)	32,3 (22,4-42,3)	30,4 (19,0-41,8)	29,8 (13,1-46,6)
p	<0,0001	<0,0001	<0,0001	0,0005
IDI (%)	2,3 (1,8-2,9)	2,4 (1,6-3,2)	2,1 (1,3-2,9)	3,3 (1,9-4,7)
p	<0,0001	<0,0001	<0,0001	<0,0001
Nagelkerke R Square				
FRT	0,190	0,213	0,166	0,138
FRT + SRGm	0,219	0,242	0,191	0,184

Intervalos de confiança a 95% apresentados em parêntesis. Framingham DC – Score de risco Framingham para DC (fatal e não fatal) a 10 anos; AUC – Area under curve; IDI – Integrated discrimination improvement; NRI – Net reclassification improvement; FRT – Fatores de risco tradicionais (Dislipidemia, Tabagismo, Diabetes e Hipertensão); SRGm – Score de risco genético; estatisticamente significativo para $p < 0,05$.

IV.5.3 RISCO CARDIOVASCULAR PELO SCORE DE FRAMINGHAM

Foi computada a função de risco de doença CV, proposta por D'Agostino *et al.* (92) que estima o risco de eventos CV a dez anos.

Tabela 65: Análise de regressão logística multivariada por subgrupos de risco do score Framingham CV com e sem inclusão do SRGm

Variáveis	FR CV <10%		FR CV 10%-20%		FR CV > 20%	
	OR (IC 95%)	Valor p	OR (IC 95%)	Valor p	OR (IC 95%)	Valor p
Tabagismo	5,57 (3,81– 8,15)	<0,0001	3,47 (2,45 – 4,90)	<0,0001	3,34 (2,53 – 4,42)	<0,0001
Hipertensão	2,63 (1,89 – 3,65)	<0,0001	2,40 (1,73 – 3,32)	<0,0001	2,00 (1,39 – 2,89)	<0,0001
Dislipidemia	1,85 (1,30 – 2,64)	0,001	-----	-----	-----	-----
Diabetes	2,68 (1,26 – 5,72)	0,011	3,69 (2,36 – 5,75)	<0,0001	3,74 (2,85 – 4,92)	<0,0001
SRGm	1,92 (1,44 – 2,55)	<0,0001	1,93 (1,40 – 2,66)	<0,0001	1,82 (1,40 – 2,37)	<0,0001
Constante	0,138	<0,0001	0,292	<0,0001	0,225	<0,0001
H.L. (FRT)	$\chi^2=2,28$	0,809	$\chi^2=1,66$	0,798	$\chi^2=1,39$	0,925
H.L. (FRT + SRG)	$\chi^2=6,96$	0,541	$\chi^2=5,75$	0,675	$\chi^2=8,10$	0,424

*Método Stepwise Forward Wald (SPSS versão 19.0); FR CV – Risco de eventos CV a 10 anos pelo Score de risco Framingham; SRGm – Score de risco genético; H.L. – Hosmer Lemeshow; FRT – Fatores de risco tradicionais; OR – Odds ratio; IC – Intervalo de confiança; estatisticamente significativo para $p < 0,05$.

Quando analisado em subgrupos (Tabela 65), o modelo multivariado contendo os FRT e o SRG, no subgrupo de risco intermédio e elevado mostra que a dislipidemia deixa de ter significado multivariado, sendo que todos os outros FRT se mantêm preditores independentes e com um ligeiro acréscimo do seu OR basal (modelo basal não expresso).

O SRGm manteve um poder preditor sobreponível nos três subgrupos, sendo discretamente mais elevado no subgrupo de risco intermédio com um OR de 1,93 (IC95%: 1,40 – 2,66) (Tabela 65).

Para a ocorrência de eventos CV a dez anos, os FRT produziram AUC das curvas ROC (Tabela 66) entre 0,67 e 0,70, que demonstram, como esperado, uma boa capacidade preditiva basal. O modelo combinado de FRT e SRGm foi mais preditivo que o modelo basal em todos os subgrupos ($p < 0,0001$).

Em cada subgrupo, constatou-se que a inclusão do SRG no modelo basal dos FRT leva a um incremento preditor do novo modelo combinado, demonstrando que o SRGm acresce em poder preditivo independente do risco basal estimado. Concretamente, a AUC no subgrupo de baixo risco elevou-se de 0,70 para 0,74 ($p < 0,0001$), no de risco intermédio de 0,67 para 0,71 ($p < 0,0001$) e no de risco elevado de 0,70 para 0,72 ($p < 0,0001$) todos com significância pelo teste de Delong.

É no subgrupo de risco intermédio que a adição do SRG produz aumentos mais significativos da sensibilidade e especificidade (64% e 69% respetivamente) (Tabela 66).

Tabela 66: Comparação de curvas AUROC antes e após inclusão do SRGm em subgrupos de risco Framingham CV.

FR CV	Modelos	AUC (IC 95%)	Sensib. (%)	Especif. (%)	VPP	VPN	Teste Delong (p)
<10%	FRT	0,70 (0,67 – 0,73)	58,9	72,1	2,11	0,57	<0,0001
	FRT + SRG	0,74 (0,71 – 0,77)	73	62,8	1,96	0,43	
10%-20%	FRT	0,67 (0,63 – 0,70)	57,4	67,4	1,76	0,63	<0,0001
	FRT + SRG	0,71 (0,68 – 0,74)	63,8	68,7	2,04	0,53	
>20%	FRT	0,70 (0,67 – 0,73)	77,4	54,3	1,69	0,42	<0,0001
	FRT + SRG	0,72 (0,70 – 0,75)	65,8	68,5	2,09	0,50	

FR CV – Risco de eventos CV a 10 anos pelo Score de Risco Framingham CV; FRT – Fatores de risco tradicionais; SRGm – Score de risco genético; AUC – Area under curve; IC – Intervalo de confiança; VPP – Valor preditivo positivo; VPN – Valor preditivo negativo; Utilizado o teste Delong para a comparação das AUC quando adicionado o SRG aos FRT; Estatisticamente significativo para $p < 0,05$.

A inclusão do SRGm no modelo dos FRT (Tabela 67) produziu um IDI de 2,9% (IC95%: 1,7-4,1%; $p < 0,0001$) no grupo de menor risco; 2,7% (IC95%: 1,7-3,7%; $p < 0,0001$) no grupo de risco intermédio e 2% (IC95%: 1,2-2,7%; $p < 0,0001$) no grupo de risco mais elevado.

O valor do R quadrado do Nagelkerke aumentou em cada subgrupo com a inclusão do SRG, isto é, a nova variável melhorou a proporção da capacidade preditiva no modelo. Assim, 21% da DCV são explicadas pelas variáveis independentes (FRT e SRG) no grupo de baixo risco, 16% no intermédio e 19% no grupo de maior risco CV (Tabela 67).

Tabela 67: Reclassificação (NRI) e Discriminação (IDI) obtida pelo SRGm em subgrupos de risco de Framingham CV

	FRAMINGHAM CV			
	Total (n=2888)	<10 (n=852)	10-20 (n=849)	>20 (n=1187)
Casos (n)	1566	333	481	752
Controlos (n)	1322	519	368	435
NRI contínuo (%)	31,7% (25-38,4%)	33% (20-45,9%)	39,6% (27,2-52%)	30,3% (19,5-41%)
Valor p	<0,0001	<0,0001	<0,0001	<0,0001
IDI (%)	2,3% (1,8-2,9%)	2,9% (1,7-4,1%)	2,7% (1,7-3,7%)	2% (1,2-2,7%)
Valor p	<0,0001	<0,0001	<0,0001	<0,0001
	Nagelkerke R Square			
TRF	0,19	0,177	0,126	0,161
TRF+GS	0,219	0,213	0,158	0,187

Intervalos de confiança a 95% apresentados em parêntesis. SRF CV – Score de risco Framingham CV; AUC – Area under curve; IDI – Integrated discrimination improvement; NRI – Net reclassification improvement; FRT – Fatores de risco tradicionais (Dislipidemia, Tabagismo, Diabetes e Hipertensão); SRG – Score de risco genético; Estatisticamente significativo para $p < 0,05$.

IV.6 ESTUDO PROSPETIVO DOS EVENTOS CARDIOVASCULARES ADVERSOS

Este subcapítulo divide-se em duas sub-análises:

-Avaliação da Mortalidade CV - Tabelas 68 a 73

-Avaliação do Evento Global (Mortalidade CV + EAM+ Revascularização) - Tabelas 74 a 80

Nos 1549 doentes com DC documentada, após angiografia e com um seguimento médio de 55,8 meses (4,5 – 184,2 meses), ocorreram 156 mortes, das quais 107 de causa CV (69,1 por 1000 pessoas/ano) (Tabela 68). Um evento adverso ocorreu em 466 indivíduos (300,1 por 1000 pessoas/ano), incluindo 155 Síndromes coronárias agudas (100,1 por 1000 pessoas/ano), 295 revascularizações (190,1 por 1000 pessoas/ano) e 45 Acidentes Vasculares Cerebrais (29,1 por 1000 pessoas/ano). O evento global ocorreu em 233,1 por 1000 pessoas/ano e correspondeu a 380 eventos de Morte CV ou EAM ou Revascularização (Tabela 68).

O SRG utilizado no estudo prospetivo foi um modelo de score genético aditivo (SRGa). Não se encontraram diferenças quanto à mediana do SRGa na maioria dos eventos CV adjudicados. Na mortalidade CV, no EAM e no evento global observou-se uma tendência para o SRGa ser mais elevado no grupo com eventos CV ($p=0,075$, $p=0,095$ e $p=0,078$ respetivamente). Os 53 doentes submetidos a revascularização cirúrgica apresentaram SRGa mais elevado ($25,3 \pm 3,7$ versus $24,3 \pm 3,7$; $p=0,044$) (Tabela 68).

Tabela 68: Distribuição do SRGa de acordo com a ocorrência dos diversos eventos CV

Eventos CV	Casos (n=1549)	SRGa		VALOR P
		Evento	Sem evento	
Global de eventos, n (%)	466 (30,1)	24,6 ± 3,5	24,2 ± 3,7	0,115
Morte CV, n (%)	107 (6,9)	24,9 ± 3,3	24,3 ± 3,7	0,075
EAM, n (%)	155 (10)	24,8 ± 3,1	24,3 ± 3,7	0,095
Angina instável, n (%)	182 (11,7)	24,7 ± 3,5	24,3 ± 3,7	0,144
TVR, n (%)	112 (7,2)	24,5 ± 3,7	24,3 ± 3,7	0,751
TLR, n (%)	101 (6,5)	24,2 ± 3,5	24,4 ± 3,7	0,647
PTCA, n (%)	219 (14,1)	24,5 ± 3,4	24,3 ± 3,7	0,494
CABG, n %	53 (3,4)	25,3 ± 3,7	24,3 ± 3,7	0,044
Revascularização*, n (%)	295 (19)	24,6 ± 3,5	24,3 ± 3,7	0,220
AVC, n (%)	45 (2,9)	24,6 ± 3,6	24,3 ± 3,7	0,602
IC, n (%)	101 (6,5)	24,7 ± 3,4	24,3 ± 3,7	0,346
Evento global+, n (%)	380 (24,5)	24,6 ± 3,5	24,3 ± 3,7	0,078

CV – Cardiovascular; EAM – Enfarte anterior do miocárdio; TVR – Target vessel revascularization; TLR – Target lesion revascularization; PTCA- Percutaneous Angioplasty. CABG – Coronary artery bypassgraft, Revasc* = (TVR+TLR+PTCA+CABG), AVC – Acidente vascular cerebral. IC – Insuficiência Cardíaca. Evento global+ = EAM+Revasc+MorteCV. Valores contínuos expressos pela média±desvio padrão; Estatisticamente significativo para $p<0,05$.

IV.6.1 MORTALIDADE CARDIOVASCULAR

IV.6.1.1 Características clínicas, do exame objetivo e bioquímico laboratoriais:

A população com DC em que ocorreu o evento de mortalidade CV tinha idade média mais avançada (55,5 ± 7,6 versus 53,1 anos; p=0,002). Não se observaram diferenças significativas quanto ao gênero entre os grupos (Tabela 69). O perfil lipídico dos doentes com evento era também mais adverso com valores séricos mais elevados de colesterol total, Apo B e triglicéridos e menores de HDL.

Tabela 69: Características basais e do exame objetivo da população de acordo com o evento (mortalidade CV)

Variáveis	Morte CV		Valor p
	Sim (n=107)	Não (n=1442)	
Gênero masculino, n (%)	85 (79,4)	1140 (79,1)	0,925
Idade, anos	55,5 ± 7,6	53,1 ± 8	0,002
Dislipidemia [†] , n (%)	102 (95,3)	1298 (90)	0,072
Colesterol total (mg/dl)	193 (155 – 229)	179 (153 – 211)	0,041*
HDL (mg/dl)	38 (33,4 – 44,6)	41 (35 – 49)	0,002*
LDL (mg/dl)	111,8 (87 – 138,4)	104,3 (82,6 – 128,4)	0,135*
Lipoproteína (a) (mg/dl)	25,9 (11,9 – 62)	20,1 (9 – 62,7)	0,291*
Apolipoproteína B (mg/dl)	99,3 (81,9 – 124)	93,5 (75,4 – 112,5)	0,003*
Triglicéridos (mg/dl)	155 (121 – 230)	139 (101 – 209)	0,033*
Abuso de álcool**, n (%)	29 (27,1)	230 (16)	0,003
Exercício físico ^o , n (%)	29 (27,1)	539 (37,4)	0,033
Fumador, n (%)	47 (43,9)	673 (46,7)	0,583
HTA, n (%)	77 (72)	1026 (71,2)	0,858
PAS	135,5 ± 20,5	138,1 ± 20,8	0,200
PAD	80,6 ± 11,2	82,7 ± 11,8	0,070
VOP> 10m/s, n (%)	28 (26,2)	280 (19,4)	0,091
VOP (m/s)	8,8 ± 2,1	8,6 ± 1,9	0,372
Diabetes, n (%)	54 (50,5)	473 (32,8)	<0,0001
Glicose (mg/dl)	120 (102 – 176)	105 (96 – 127)	<0,0001*
Índices Antropométricos			
IMC (kg/m ²)	29 ± 4,4	28,5 ± 4,2	0,306
IMC> 30kg/m ² , n (%)	42 (39,3)	477 (33,1)	0,192
FC, bat./min.	75,5 ± 13,6	68,3 ± 12,3	<0,0001
Cintura/Altura	0,62 ± 0,07	0,61 ± 0,06	0,004
DC 3 vasos, n (%)	51 (47,7)	379 (26,3)	<0,0001
Angina estável, n (%)	25 (23,4)	306 (21,2)	0,602
Terapêutica			
Estatinas, n (%)	21 (77,8)	466 (72,5)	0,544
Dupla AA, n (%)	26 (24,3)	622 (43,3)	<0,0001
IECA/ARAs	57 (53,3)	647 (45,1)	0,100
SRG			
SRGa	25 (23 - 27)	24 (22 - 27)	0,069*
SRGa ≥mediana	74 (69,2)	837 (58)	0,024*
SRGm	0,51 (0,31 – 0,80)	0,46 (0,28 – 0,75)	0,444*
SRGm ≥mediana	59 (55,1)	716 (49,7)	0,273*

CV – Cardiovascular; ^oMais de 40min/semana; **Mais de 300gr/semana; HTA – Hipertensão; VOP – Velocidade de onda de pulso; IMC – Índice de massa corporal; FC – Frequência cardíaca; SRG – Score de risco genético; PAS – Pressão arterial sistólica; PAD – Pressão arterial diastólica; CAT – Cateterismo; AA – Anti agregação; [†]LDL>140 ou HDL <40 ou Triglicéridos>150; Valores contínuos expressos pela média±desvio padrão ou mediana (Percentil 25 – Percentil 75); *Valores obtidos através do teste não paramétrico Mann-Whitney; Estatisticamente significativo para p <0,05.

Verificou-se também que o grupo com mortalidade CV era mais diabético (50,5 versus 30,2%), praticava menos exercício (27,1% versus 37,4), apresentava maior consumo de álcool e tinha um índice cintura-altura mais elevado ($p < 0,05$).

Quanto à terapêutica medicamentosa, não se encontraram significâncias quanto à utilização de inibidores da enzima de conversão angiotensina/antagonistas dos recetores da angiotensina (Tabela 69).

O grupo em que ocorreu mortalidade CV estava menos duplamente anti-agregado (24,3% versus 43,3%, $p < 0,0001$) e apresentava menor controlo da frequência cardíaca. Este subgrupo apresentava ainda maior gravidade angiográfica com 47,7% versus 26,3% dos doentes com DC de três vasos.

Num maior número de indivíduos (69,2% versus 58%) com SRG acima da mediana ocorreu mortalidade CV ($p = 0,02$) (Tabela 69).

Quanto às características bioquímicas da população (Tabela 70), salienta-se que os doentes no grupo com mortalidade CV apresentaram valores séricos mais elevados de plaquetas, leucócitos e Hcy ($p < 0,05$). Os doentes em que ocorreu a mortalidade CV apresentaram mais frequentemente compromisso da função renal, níveis séricos de PCR superior a 3mg/l e de Apo B superior a 90mg/dl.

Tabela 70: Características bioquímicas da população de acordo com o evento de mortalidade CV

Variáveis	Mortalidade CV		Valor p
	Sim (n=107)	Não (n=1442)	
Hemoglobina, g/dl	14,6 (14,6 – 14,9)	14,6 (13,8 – 15,4)	0,743*
Plaquetas, $10^3/\mu\text{l}$	226 (193 - 264)	214 (182 - 253)	0,041*
Leucócitos, $10^3/\mu\text{l}$	7,5 (6,7 – 8,7)	7 (5,9 – 8,2)	0,004*
Fibrinogénio, (mg/dl)	388 (324 - 479)	387 (337,5 - 442)	0,572*
Lp (a)>25 (mg/dl), n (%)	54 (50,5)	656 (45,5)	0,319
Apo B>90 (mg/dl), n (%)	75 (70,1)	801 (55,5)	0,003
Homocisteína, $\mu\text{mol/L}$	13 (10,6 – 17,3)	12,2 (9,9 – 14,7)	0,003*
>10, n (%)	88 (82,2)	1070 (74,2)	0,065
Creatinina, mg/dl	1 (0,8 – 1,2)	1 (0,9 – 1,1)	0,448
Clearance Cr.**, ml/min	86,6 (74,6 – 112,1)	94,2 (77,3 – 112,8)	0,077*
Clearance Cr. <60 ml/min, n (%)	14 (13,1)	108 (7,5)	0,038
PCR, mg/L	7,1 (2,4 - 39)	2,7 (1,5 – 5,4)	<0,0001*
>3, n (%)	72 (67,3)	565 (39,2)	<0,0001

CV – Cardiovascular; Lp (a) – Lipoproteína a; Apo B – Apolipoproteína B; Cr. – Creatinina; **Cockcroft-Gault; PCR – Proteína C reativa; Valores contínuos expressos pela mediana (Percentil 25 – Percentil 75); *Valores obtidos através do teste não paramétrico Mann-Whitney; Estatisticamente significativo para $p < 0,05$.

IV.6.1.2. Estudo gene a gene e mortalidade CV:

Das trinta e três variantes analisadas neste grupo, apenas se encontrou diferenças significativas na variante genética do ADAMTS7. Os doentes com mortalidade CV apresentaram mais frequentemente o genótipo de risco AA do ADAMTS7, relativamente aos indivíduos em que não ocorreu o evento (Tabela 71).

Assim, a curva de *Kaplan-Meier* mostra a separação das curvas de sobrevida com redução da mortalidade CV nos doentes com genótipo protetor GG (Figura 30). Com efeito, 47,4% dos doentes eram heterozigotos (AG), 36% eram homozigóticos para o alelo selvagem A e apenas 16,5% eram homozigóticos para o alelo G protetor. Durante o seguimento clínico ocorreram 47 eventos (8,4%) entre os 557 pacientes, que possuíam o genótipo de risco (AA) e 11 eventos (4,3%) entre os 257 doentes com o genótipo protector (GG). No fim do seguimento, a probabilidade de sobrevivência dos pacientes com o genótipo AA, foi de 70% e 85,2% nos doentes com o genótipo GG.

Pela regressão de Cox ajustada para os FRT confirmou-se que o genótipo AA foi um fator de risco independente para a mortalidade CV, quando comparado com o genótipo GG protetor (HR=1,89; IC95%: 1-3,67; p=0,05).

Tabela 71: Análise gene a gene: Variantes genéticas e mortalidade CV:

ADAMTS7	Mortalidade CV		Valor p
	Sim (n=58), %	Não (n=756), %	
GG	11 (19)	246 (32,5)	0,032
AA	47 (81)	510 (67,5)	

CV – Cardiovascular; Estatisticamente significativo para $p < 0,05$.

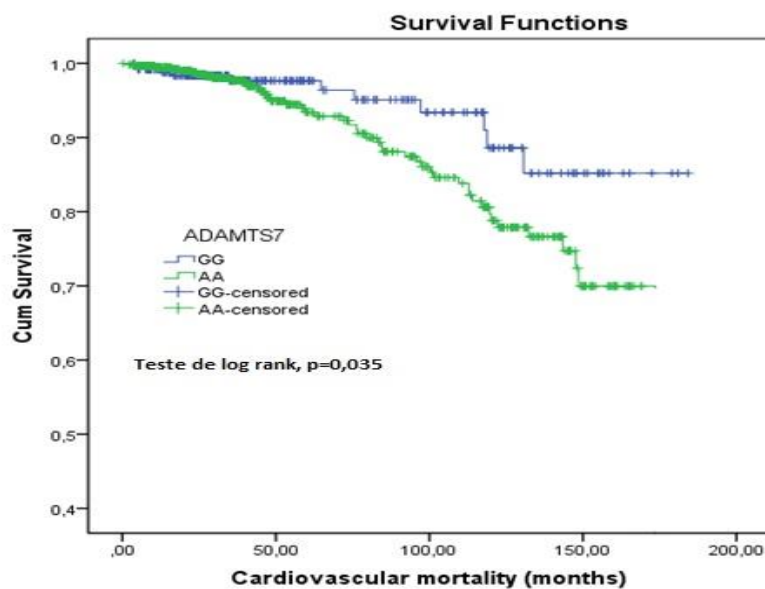


Figura 30: Gráfico de Kaplan-Meier para a mortalidade CV de acordo com a variante genética ADAMTS7 (AA versus GG). O teste de *Log-rank* foi de 4,4 (p=0,035)

IV.6.1.3. Preditores independentes de mortalidade CV:

A tabela 72 revela os resultados da análise de regressão logística multivariada de acordo com dois modelos: o modelo ajustado pelos FRT, SRG e três co-variáveis de interesse para a ocorrência de mortalidade CV (Dupla AA, DC de três vasos, e angina versus EAM) e o modelo multi-ajustado para a idade e por todas as outras variáveis significativas na análise bivariada.

No modelo ajustado verificou-se que o SRGa > 24 (OR=1,57), a DC de três vasos (OR=2,97), a dislipidemia (OR=3,47), e a diabetes (OR=1,80) foram preditores independentes para a mortalidade CV ($p < 0,05$) (Tabela 72).

No modelo multi-ajustado encontraram-se, para além da idade (OR=1,06), os seguintes preditores de mortalidade CV estatisticamente significativos ($p < 0,05$): o SRGa > 24 (OR=1,52); a DC de três vasos (OR=1,51); o consumo de álcool (OR=1,77) e a frequência cardíaca (OR=1,03).

Tabela 72: Preditores independentes de mortalidade CV: regressão logística multivariada

Variáveis	Modelo Ajustado		Modelo Multi-Ajustado		
	OR (95% IC)	Valor p	Variáveis	OR (95% IC)	Valor p
SRGa > 24	1,57 (1,05 – 2,35)	0,028	SRGa > 24	1,52 (1-2,31)	0,05
DC 3 vasos	2,97 (2,01 – 4,38)	<0,0001	DC 3 vasos	1,51 (1,02-2,24)	0,04
Dislipidemia	3,47 (1,27 – 9,49)	0,015	Álcool > 300gr./sem.	1,77 (1,13-2,78)	0,012
Diabetes	1,80 (1,21 – 2,66)	0,003	Freq. cardíaca	1,03 (1,02-1,05)	<0,0001
Hipertensão	1,03 (0,66 – 1,6)	0,905	Idade	1,06 (1,03-1,09)	<0,0001
Tabagismo	1,34 (0,9 – 2,01)	0,152	-----	-----	-----
Angina estável	1,16 (0,73 – 1,84)	0,53	-----	-----	-----
Dupla AA	1,17 (0,73 – 1,87)	0,514	-----	-----	-----

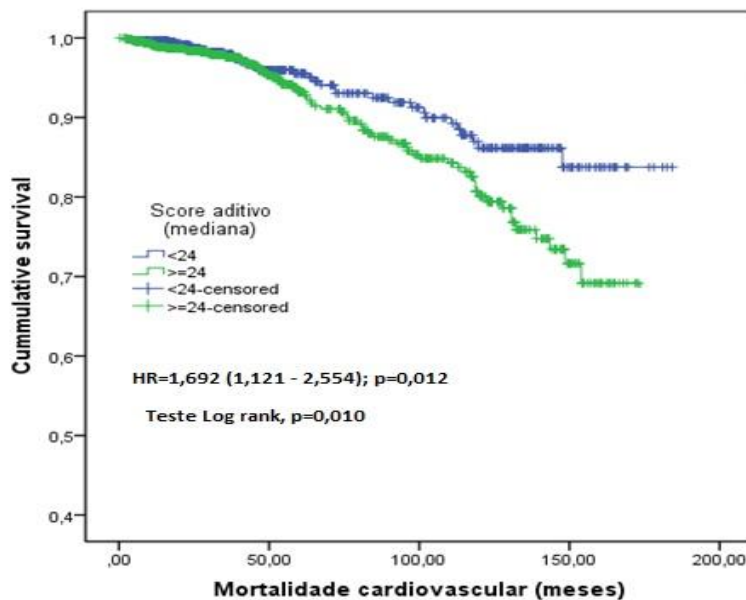
**Método Enter ajustado para as variáveis Angina estável e Dupla AA- Dupla Anti-agregação. O modelo multi-ajustado foi ajustado para as variáveis: Tabaco, Hipertensão, Dislipidemia, Diabetes, Sedentarismo, Dupla AA, Angina estável, Homocisteína, Plaquetas, Leucócitos, ApoB>90, Clearance Cr. <6 e PCR>3. IC – Intervalo de confiança; estatisticamente significativo para $p < 0,05$.*

Tabela 73: FRT e SRGa como preditores de mortalidade CV: regressão de Cox

Variáveis	HR	(IC 95%)	Valor p
SRGa > 24	1,69	(1,10–2,50)	0,012
DM	1,83	(1,30–2,70)	0,002
Idade	1,05	(1,03-1,08)	<0,0001
Abuso de álcool	1,88	(1,22-2,90)	0,004
DC 3 vasos	1,52	(1,03-2,24)	0,035
Freq. cardíaca	1,04	(1,25-1,05)	<0,0001
PCR > 3mg/L	1,52	(1,00-2,31)	0,05

*Método Forward Wald (SPSS versão 19) em que saíram da equação as variáveis: Tabaco, HTA, Dislipidemia, Exercício físico, Dupla AA, Angina estável, Homocisteína, Plaquetas, Leucócitos, Apo B > 90, Clearance Cr. < 60; CV – Cardiovascular; IC – Intervalo de confiança; PCR-Proteína C reativa, DM-Diabetes mellitus, Estatisticamente significativo para $p < 0,05$.

Através da regressão de Cox (Tabela 73) verificou-se que, para além da idade (HR 1,05), um valor de SRGa > 24 associou-se com maior incidência da mortalidade CV com um HR de 1,69 (IC95%: 1,10-2,50; $p=0,012$). De igual modo, a DM, o elevado consumo de álcool, a DC de três vasos, a frequência cardíaca e a PCR > 3mg/dl foram preditores independentes para a mortalidade CV no modelo de Cox, com significância estatística ($p < 0,05$) (Tabela 73).

**Figura 31:** Curva de sobrevivência de Kaplan-Meier para a Mortalidade CV de acordo com o SRGa acima e abaixo da mediana ($=24$).

As curvas de sobrevivência de KM (Figura 31) mostram uma separação significativa das curvas de sobrevivência de acordo com o SRGa constatando-se que a população com risco genético acima da mediana teve maior mortalidade CV. Durante o seguimento médio de 55,8 meses, ocorreram 74 eventos (8,1%) entre os 911 pacientes com SRGa acima da mediana e 33 eventos (5,2%) entre os 638 doentes que se apresentavam abaixo da mediana. No fim do seguimento a probabilidade de sobrevivência dos doentes com SRGa > 24 foi de 69,2% e de 83,7% nos doentes com SRGa < 24.

A probabilidade de mortalidade CV nos doentes com SRG acima da mediana foi de 30,8% com um HR de 1,69 (IC95%: 1,12-2,55; $p=0,012$) e de 16,3% nos doentes com SRGa abaixo da mediana. O teste de *Log-rank* foi de 6,6 com significância estatística ($p=0,010$) (Figura 31)

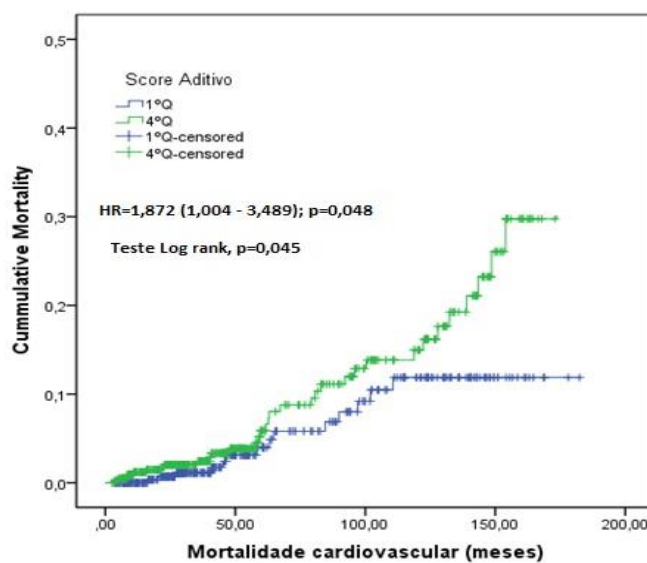


Figura 32: Incidência cumulativa da mortalidade CV de acordo com quartis do SRGa (1º versus 4º quartil)

Analisando por quartis do SRGa, o gráfico das curvas de KM (Figura 32) mostra que, durante o seguimento médio de 55,8 meses, também ocorre separação das curvas de sobrevivência. Os doentes no quarto quartil do SRG (8,7%; $n=43$) apresentaram maior mortalidade CV do que os situados no primeiro quartil (5,6%; $n=26$), com uma incidência do evento expressa por um HR de 1,87 (IC95% 1-3,49; $p=0,048$), Log-rank; $p=0,045$.

Na figura 33, as curvas de sobrevivência média preditas $S(t)$, para cada um dos grupos de risco, são comparadas com as respectivas curvas de *Kaplan-Meier*, demonstrando uma boa calibração pela boa coincidência entre cada par de curva predita. Exceção feita para o grupo de risco intermédio no fim do seguimento, em que se observa uma separação de curvas.

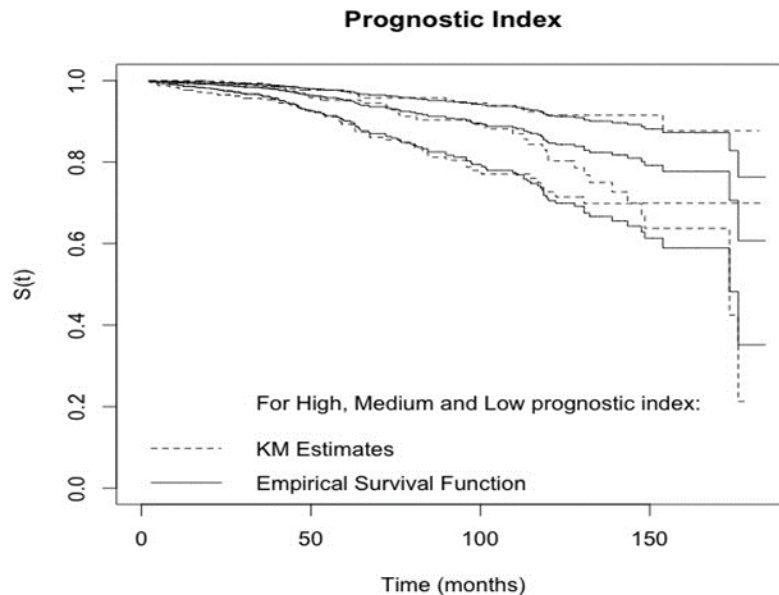


Figura 33: Sobrevida estimada pelo modelo de Cox e Curva de *Kaplan-Meier* (KM) para índice prognóstico elevado, médio e baixo.

IV.6.2 EVENTO GLOBAL

IV.6.2.1 Características clínicas, do exame objetivo e bioquímico laboratoriais:

Não existiram diferenças significativas, quanto à idade e ao género entre os grupos (Tabela 74). No perfil lipídico documentaram-se diferenças, quanto ao valor sérico das HDL mais reduzido nos doentes com evento adverso e os valores séricos da Lp(a) e da Apo (B) mais elevados nos doentes com eventos adversos. Os doentes com eventos eram também mais sedentários, mais diabéticos, consumiam mais álcool e tinham um índice cintura-altura mais elevado e PAD mais baixa ($p < 0,05$). Quanto à terapêutica medicamentosa, o grupo dos doentes com eventos adversos apresentaram menor grau de dupla anti-agregação mantida (29% versus 46%, $p < 0,0001$) e menor controlo da frequência cardíaca (70 bpm versus 68 bpm) à data da última consulta de seguimento. Este grupo com eventos apresentou também maior gravidade angiográfica da sua DC (DC 3 vasos 37,1 versus 24,9%; $p = 0,007$) (Tabela 74)

Tabela 74: Características gerais da população de acordo com a ocorrência do evento global (Morte CV + EAM + Revascularização)

Variáveis	Evento Global		Valor p
	Com evento (n=380)	Sem evento (n=1169)	
Género masculino, n (%)	302 (79,5)	923 (79)	0,829
Idade, anos	53,5 ± 7,7	53,2 ± 8,1	0,548
Dislipidemia [†] , n (%)	349 (91,8)	1051 (89,9)	0,266
Colesterol total, mg/dl	180,5 (153 – 215,5)	179 (154 - 211)	0,722*
HDL, mg/dl	39 (34 – 47,7)	42 (35,8 - 49)	0,004*
LDL, mg/dl	103,5 (82,7 – 125,8)	105,1 (82,8 – 129,6)	0,389*
Lipoproteína (a), mg/dl	26,3 (10,5 – 65,6)	18,8 (8,8 – 58,9)	0,043*
Apolipoproteína B, mg/dl	98,2 (79,5 – 117,2)	92,7 (73,2 – 111,4)	<0,0001*
Triglicéridos, mg/dl	145 (105 – 209,5)	140 (101 - 212)	0,306*
Abuso de álcool**, n (%)	89 (23,4)	170 (14,5)	<0,0001
Exercício físico ^o , n (%)	110 (28,8)	458 (39,2)	<0,0001
Fumador*, n (%)	160 (42,1)	560 (47,9)	0,049
HTA, n (%)	277 (72,9)	826 (70,7)	0,403
PAS, mmHg	137 ± 19,7	138,2 ± 21,1	0,334
PAD, mmHg	81,2 ± 11,3	83 ± 11,9	0,009
VOP, m/s	8,6 ± 2	8,6 ± 1,9	0,775
VOP> 10 m/s, n (%)	84 (22,1)	224 (19,2)	0,212
Diabetes, n (%)	153 (40,3)	374 (32)	0,003
Glicose, mg/dl	110 (98 - 138)	105 (96 - 126)	<0,0001*
IMC, kg/m ²	28,7 ± 4	28,5 ± 4,3	0,586
IMC> 30 kg/m ² , n (%)	137 (36,1)	382 (32,7)	0,226
FC, bpm	70,2 ± 12,9	68,3 ± 12,3	0,010
Cintura/Altura	0,61 ± 0,1	0,60 ± 0,1	0,042
DC 3 vasos, n (%)	148 (38,9)	282 (24,1)	<0,0001
Angina estável, n (%)	84 (22,1)	247 (21,1)	0,687
Estatinas, n (%)	91 (66,4)	396 (74,3)	0,065
Dupla AA, n (%)	111 (29,2)	537 (45,9)	<0,0001
SRGa	25 (22 - 27)	24 (22 - 27)	0,071*
SRGa ≥mediana	237 (62,4)	674 (57,7)	0,105
SRGm	0,48 (0,30 – 0,78)	0,46 (0,28 – 0,74)	0,337*
SRG ≥mediana	195 (51,3)	580 (49,6)	0,565

Valores contínuos expressos pela média±desvio padrão ou pela mediana (Percentil 25 – Percentil 75); [†]LDL>140 ou HDL <40 ou Triglicéridos>150; ^oMais de 40min/semana; **Mais de 300gr/semana; *Fumadores ou <5 anos de cessação; HTA – Hipertensão; VOP – Velocidade de onda de pulso; IMC – Índice de massa corporal; FC – Frequência cardíaca; SRG – Score de risco genético. HDL – High density lipoprotein; LDL – Low density lipoprotein; PAS – Pressão arterial sistólica; PAD – Pressão arterial diastólica; CAT – Cateterismo; AA – Anti agregação. *Valores obtidos através do teste não paramétrico Mann-Whitney. Estatisticamente significativo para p<0,05.

Os doentes com qualquer evento adverso apresentaram mais frequentemente valores de PCR superior a 3mg/l (p<0,0001). (Tabela 75). Não se encontraram diferenças significativas em outros marcadores de risco CV, como a elevação da homocisteína, o valor sérico dos leucócitos e o valor sérico do fibrinogénio.

Tabela 75: Características bioquímico-laboratoriais de acordo com o evento global: análise bivariada

Variáveis	Evento Global		Valor p
	Com evento (n=380)	Sem evento (n=1169)	
Hemoglobina, g/dl	14,6 (13,8 – 15,2)	14,6 (13,8 – 15,4)	0,483*
Plaquetas, 10 ³ /µl	216 (182 – 261,5)	214 (184 - 251)	0,406*
Leucócitos, 10 ³ /µl	7,1 (6,1 – 8,4)	7,1 (5,9 – 8,2)	0,197*
Fibrinogênio, mg/dl	393 (334,9 - 456)	385 (337 - 440)	0,113*
Homocisteína, µmol/L	12,8 (10,3 – 16,1)	12,1 (9,8 – 14,5)	0,002*
>10, n (%)	296 (77,9)	862 (73,7)	0,105
Creatinina, mg/dl	1 (0,9 – 1,1)	1 (0,9 – 1,1)	0,859*
Clearance Cr.**, ml/min	92,3 (76,1 – 113,2)	94,1 (77,1 – 112,3)	0,743*
Clearance Cr. <60 ml/min, n (%)	38 (10)	84 (7,2)	0,077
PCR, mg/L	3,4 (1,8 – 13,5)	2,7 (1,5 - 5)	<0,0001*
>3, n (%)	202 (53,2)	435 (37,2)	<0,0001

*Evento global (EAM+Revascularização+ Morte CV – Cardiovascular); EAM – Enfarte anterior do miocárdio; Revascularização= (TVR – Target vessel revascularization+ TLR – Target lesion revascularization+ OLR – Other revascularization+ CABG); Lp (a) – Lipoproteína a; Apo B – Apolipoproteína B; Cr. – Creatinina; **Cockcroft-Gault; PCR – Proteína C reativa; Valores contínuos expressos pela mediana (Percentil 25 – Percentil 75); *Valores obtidos através do teste não paramétrico Mann-Whitney; Estatisticamente significativo para p<0,05.*

Dos FRT, apenas a DM mostrou estar associada com a ocorrência de qualquer evento no seguimento, com um OR de 1,42 (IC95%: 1,11-1,81; p=0,005) (Tabela 76). O SRGa e os outros FRT não demonstraram ter significância estatística multivariada.

Tabela 76: FRT e SRG como Preditores independentes de qualquer evento adverso: análise de regressão logística multivariada

Variáveis	Odds ratio (IC 95%)	Valor p
Diabetes	1,42 (1,11 – 1,81)	0,005
Constante	0,27	<0,0001

Método Forward Wal (SPSS versão 19.0), ajustada para os FRT (variáveis dicotômicas) e SRG. IC – Intervalo de confiança; Estatisticamente significativo para p<0,05.

No modelo logístico multi-ajustado pelas variáveis significantes prévias (Tabela 77), encontraram-se os seguintes preditores do evento global: o sedentarismo (OR=1,57); o consumo de álcool > 300gr/semana (OR=1,83); a frequência cardíaca (OR=1,01); a DC de três vasos (OR=1,57) e a PCR>3 (OR=1,76). Foram protetores da ocorrência do evento adverso: a PAD (OR=0,98) e a Dupla Anti-agregação (OR=0,49) (Tabela 77).

Tabela 77: Preditores independentes de qualquer evento adverso: análise de regressão logística multivariada ajustada

Variáveis	Odds ratio (IC 95%)	Valor p
Sedentarismo	1,57 (1,2 – 2,06)	0,001
Abuso álcool*	1,83 (1,35 – 2,48)	<0,0001
Frequência cardíaca	1,01 (1 – 1,02)	0,048
PAD	0,98 (0,97 – 1)	0,001
DC 3 vasos	1,57 (1,21 – 2,04)	0,001
Dupla AA	0,49 (0,38 – 0,64)	<0,0001
PCR>3	1,76 (1,37 – 2,26)	<0,0001
Constante	0,4	0,081

H.L. $\chi^2=6,08$; $p=0,639$

*>300gr./sem. Regressão logística com o Método Stepwise Forward Wald (SPSS versão 19.0) ajustada para: Diabetes; Lp (a); Apo B; Homocisteína; SRGa. Evento global= Morte CV +EAM+Revascularização. PAD – Pressão arterial diastólica; DC – Doença coronária; AA – Anti-agregação; PCR – Proteína C-reativa; IC – Intervalo de confiança; Estatisticamente significativo para $p<0,05$.

Através da regressão de Cox ajustada (Tabela 78), verificou-se que a variante genética TCF21 associou-se com uma maior incidência de eventos no seguimento com um HR de 1,55 ($p=0,06$) por um alelo C de risco e um HR de 1,73 por dois alelos C de risco ($p=0,02$). A variante genética HNF4, em homozigotia, conferiu proteção para a ocorrência do evento global (HR=0,26; $p=0,059$). Os 4 preditores ambientais, identificados previamente, mantêm poder preditor significativo: a gravidade angiográfica (HR=1,34), a elevação da PCR (HR=1,35), o sedentarismo (HR=1,30) e o consumo de álcool (HR=1,54). Não interferiram de forma significativa na ocorrência de eventos: a idade, a diabetes, a dupla antigregação, a angina estável versus enfarte, a frequência cardíaca, nem outros marcadores como a Hcy, a Lp (a) e a Apo (B).

Tabela 78: Preditores genéticos e não genéticos de qualquer evento adverso: regressão de Cox ajustada

Variáveis	HR (95% IC)	Valor p
DC 3 vasos	1,34 (1,08 – 1,66)	0,008
PCR> 3mg/L	1,35 (1,1 – 1,67)	0,005
Sedentarismo	1,30 (1,04 – 1,64)	0,024
Álcool> 300gr./sem.	1,54 (1,21 – 1,96)	<0,0001
TCF21	-----	0,052
CG	1,55 (0,98 – 2,45)	0,060
CC	1,73 (1,10 – 2,72)	0,018
HNF4A	-----	0,046
CG	0,81 (0,64 – 1,03)	0,089
GG	0,26 (0,07 – 1,06)	0,059

*Método Forward Wald (SPSS versão 19.0) em que saíram da equação as variáveis: Idade, Dupla AA, Angina estável, PAD, Diabetes, Freq. Cardíaca, Lp (a)>25, Apo B>90mg, Homocisteína, ADIPOQ e APOE; IC – Intervalo de confiança; Estatisticamente significativo para $p<0,05$.

A variante genética TCF21 (CC) associou-se de forma independente com a ocorrência de eventos adversos (C-Index 0,52, $p=0,018$) (Figura 34).

De igual modo, a DC de três vasos (C-Index=0,56, $p=0,008$), a PCR> 3mg/L (C-Index=0,55, $p=0,005$), o

sedentarismo (C-Index=0,53, p=0,024), o alcoolismo>300 gr/semana (C-Index=0,532, p<0,0001) foram os preditores independentes não genéticos para a ocorrência do evento adverso (Figura 34).

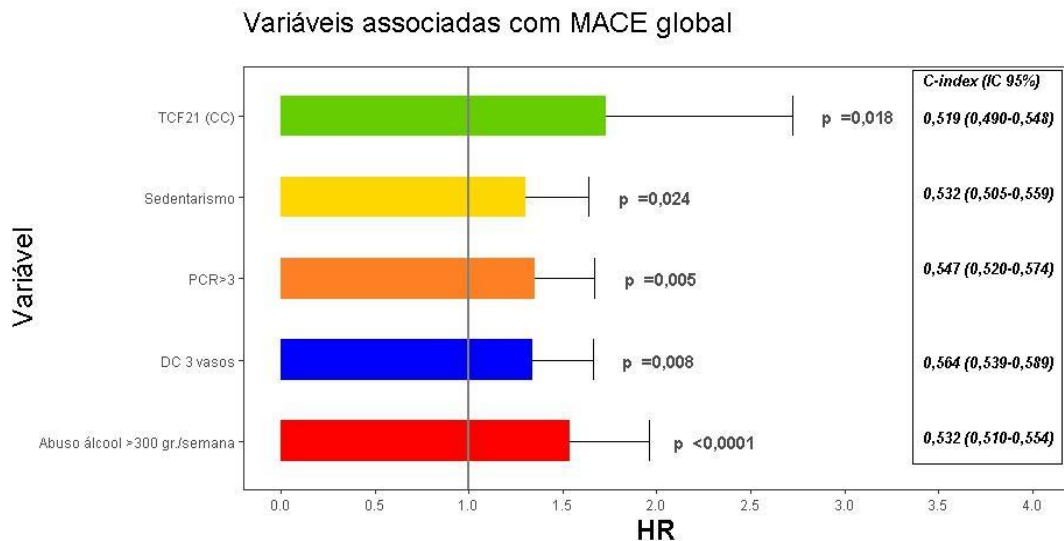


Figura 34: Preditores multivariados independentes do evento global: análise de regressão de Cox ajustada e C-Index para a ocorrência de qualquer evento adverso no seguimento.

A tabela 79 mostra a associação do SRGa com a ocorrência dos eventos adversos analisados previamente, mortalidade CV e evento global (mortalidade CV + EAM + revascularização) em dois modelos de regressão de Cox, um ajustado para a idade e para o gênero e um multi-ajustado para a idade e gênero, para os FRT, e para co-variáveis de interesse para a ocorrência de eventos em doentes com DC, como a dupla AA e a DC3 vasos.

Comparando com o grupo de menor risco genético (SRGa ≤ 21), os doentes com risco genético moderado a elevado (SRGa > 22) tinham 1,96 vezes maior probabilidade de sofrerem morte CV (p=0,019) e 1,34 vezes maior probabilidade de ocorrência do evento global (p=0,025), ajustado para a idade e o gênero (Tabela 79).

Após ajuste adicional do modelo preditor controlando com os FRT, observa-se um aumento do risco de ocorrência de mortalidade CV com um HR de 2,03 (p=0,014) no grupo com SRGa moderado a elevado. De igual modo, após o ajuste multifatorial, manteve-se a força da associação do SRGa com o evento global com um HR de 1,35 (p=0,023) (Tabela 79).

Tabela 79: Associação entre o SRGa baixo (n=347, SRG≤21) e moderado ou elevado (n= 1222, SRGa ≥22-27) e a Incidência de eventos adversos no seguimento

SRG	Modelo Ajustado		Modelo Multi-Ajustado	
	HR (95% IC)	Valor p	HR (95% IC)	Valor p
Mortalidade CV				
Baixo	1	----	1	----
Moderado/Elevado	1,96 (1,12 – 3,44)	0,019	2,03 (1,16 – 3,56)	0,014
Evento GLOBAL: Morte CV+ EAM+ Revascularização				
Baixo	1	----	1	----
Moderado/Elevado	1,34 (1,04 – 1,74)	0,025	1,35 (1,04 – 1,75)	0,023

*Método Enter ajustado para as variáveis Idade e Género. O modelo multi-ajustado foi ajustado para as variáveis Idade, género, Tabagismo, HTA, Dislipidemia, Diabetes, DC>3 vasos, Dupla AA. IC – Intervalo de confiança; Estatisticamente significativo para $p<0,05$. HR reportado por comparação com o baixo risco (classe de referência).

O C-Index do modelo preditor da mortalidade CV, incluindo a dislipidemia, o tabagismo, a DM, a HTA e o SRGa superior à mediana foi de 0,62, demonstrando capacidade discriminativa discreta entre doentes que tiveram ou não o evento fatal em estudo (Tabela 80).

O C-index no modelo preditor do evento global adverso (mortalidade CV, EAM, necessidade de revascularização) foi de 0,54 demonstrando reduzida capacidade discriminativa para o evento global.

A capacidade reclassificativa do modelo final produziu um NRI mais elevado para a mortalidade CV (12%) e menor para o evento global (7,3%) sem atingirem significado estatístico (Tabela 80).

Tabela 80: Discriminação e Reclassificação do modelo preditor (FRT+SRG) na ocorrência do evento global: NRI e IDI

Comparação de modelos	C-INDEX (95% IC)	NRI (95% IC)	IDI (95% IC)	Likelihood ratio test (p)
Mortalidade Cardiovascular				
FRT + SRG elevado	0,62 (0,56-0,69)	0,01 (-0,37-0,41)	0,0002 (-0,001-0,001)	0,008
Valor p	-----	0,757	0,591	
Evento Global: Morte CV+EAM+Revascularização				
FRT + SRG elevado	0,54 (0,51-0,58)	0,07 (0,08-0,18)	0,001 (-0,001-0,003)	0,041
Valor p	-----	0,239	0,405	

FRT – Fatores de risco tradicionais (Dislipidemia, Diabetes, Tabagismo e Hipertensão); SRG – Score de risco genético elevado (> mediana).

V. DISCUSSÃO

7- DISCUSSÃO:

O capítulo da discussão será organizado nos seguintes subcapítulos:

V.1. TIPO DE ESTUDO E METODOLOGIA

V.2. ESTUDO GENÉTICO E DOENÇA CORONÁRIA

V.3. INTERACÇÃO GENE-GENE E GENE-AMBIENTE

V.4. SCORE DE RISCO GENÉTICO E DOENÇA CORONÁRIA

V.5. RECLASSIFICAÇÃO DO RISCO CARDIOVASCULAR COM O SCORE DE RISCO GENÉTICO

V.6. ANÁLISE PROSPETIVA DOS EVENTOS CARDIOVASCULARES

V.1 TIPO DE ESTUDO E METODOLOGIA

V.1.1 ESTUDO CASO-CONTROLO

No estudo das associações de fatores com a doença são, habitualmente, utilizados estudos caso-controlo que expressam aquelas interações em *Odds ratios*, uma vez que, os estudos prospetivos requerem longos períodos de seguimento e são muito dispendiosos. Por isso optámos por um estudo tipo caso-controlo em que um grupo de indivíduos com doença (casos) é comparado com um grupo de indivíduos considerados livres da doença (controlos).

Nestes estudos, os fatores de risco tradicionais são inquiridos oralmente ou documentados num exame objetivo ou laboratorial. Este facto condiciona alguma limitação, uma vez que, alguns destes fatores ficam sujeitos à memória dos doentes e controlos, do qual é exemplo a história de DC precoce na família e a própria história pessoal. De igual modo, há outros fatores de risco que podem alterar-se ao longo do tempo, com ou sem intervenção terapêutica.

Uma limitação do nosso estudo, advém do fato de não incluir os doentes falecidos no período pré-hospitalar, falecidos durante o internamento ou que não tenham feito angiografia coronária. Apenas se incluíram doentes com alta hospitalar, com a DC documentada angiograficamente, sem doença valvular grave. Eventualmente, podemos ter excluído os doentes com pior perfil genético. Admitimos que pode existir uma auto-selecção negativa dos casos, por óbito dos doentes com pior perfil genético que faleceram durante o evento indexado. Contudo, os resultados mantêm a sua validade. Quando muito, a importância do perfil genético na DC pode ter ficado mais diluída na população em estudo.

Uma limitação já referida, advém do fato de que, em ambos os grupos, ocorreram intervenções várias sobre os FRT que são em muitos casos os fenótipos dos polimorfismos em estudo. Assim, a inferência de conclusões deve ter em linha de conta estas potenciais intervenções.

No global, contando com as limitações, podemos concluir que um estudo caso-controlo é um método adequado e válido no estudo da associação do perfil genético e da DC.

Na avaliação dos doentes com DC, um modelo de coorte prospetiva, afigura-se exequível, adequado e sem limitações significativas para estudo das diferenças na sobrevida e na ocorrência de eventos cardiovasculares adversos, após o diagnóstico da DC por causa da elevada incidência de eventos neste subgrupo da população.

V.1.1.1 SELEÇÃO DOS CASOS

Todos os doentes com alta clínica, do único hospital da Região Autónoma da Madeira, com a DC significativa diagnosticada por cateterismo foram incluídos no estudo. A seleção foi feita na alta hospitalar, tendo-se excluído os doentes que apresentavam patologia valvular grave. Deste modo, não há enviesamento na seleção dos doentes, uma vez que, representam toda a incidência de casos existentes na população durante o período de estudo.

Foi considerada DC significativa a existência de pelo menos uma artéria epicárdica com 70% de estenose ou de 50% no tronco comum. Embora este seja o critério mais consensual para Doença Coronária hemodinamicamente significativa (112), reconhece-se que doentes com estenoses intermédias (30-70%), que foram excluídos desta análise, são também portadores de aterosclerose coronária não limitante do fluxo.

Foi discutida a possibilidade de utilizar as amostras de sangue colhidas no Laboratório de Hemodinâmica durante a fase aguda do evento coronário para estudo genético, após obtenção do consentimento informado. Contudo, com base na premissa que o estudo genético não era decisivo para a sobrevivência imediata, a Comissão de Ética não considerou adequado o pedido de consentimento informado durante a fase aguda do EAM ou mesmo numa fase posterior de instabilidade clínica. Assim, o consentimento informado para o estudo genético, foi dado na altura da alta clínica ou na primeira consulta na Unidade de Investigação, não dispondo, por isso, de dados dos doentes que faleceram antes da alta clínica. É possível que os doentes mais graves, e com mortalidade intra-hospitalar com um perfil genético mais pro-aterosclerótico, tenham sido excluídos do estudo, sendo esta uma das suas limitações.

V.1.1.2 SELEÇÃO DOS CONTROLOS

Uma limitação do estudo caso-controlo é a definição de controlo como ausência de DC. Esta, foi definida como ausência de sintomas ou de alterações no exame objetivo ou exames complementares de diagnóstico compatíveis com cardiopatia isquémica. Estando sujeita à possibilidade de erro é a única abordagem viável por não ser eticamente responsável a realização de angiografia coronária para documentação de possível DC no grupo controlo.

Assim, os controlos foram selecionados a partir dos cadernos eleitorais representativos de toda a população residente na RAM para o grupo etário em estudo. Foram recrutados aleatoriamente, e emparelhados de acordo com o grupo etário e o género, pela mesma ordem de inclusão dos doentes.

Mais uma vez, este método de seleção dos controlos permitiu uma amostra não enviesada representativa da população da RAM. Pensamos ter um grupo de estudo de doentes e de controlos muito robusto para o trabalho que nos propusemos realizar. Por outro lado, após estudo de haplótipos ligados ao Y e do ADN mitocondrial, está documentado por Gonçalves *et al*, que os portugueses continentais, os da Região Autónoma da Madeira e os dos Açores são uma população geneticamente homogénea e representativa dos países do sul da Europa. (110)

V.1.1.3 CRITÉRIOS DE INCLUSÃO E DE EXCLUSÃO

Numa população homogénea e representativa do sul da Europa, para que fizesse sentido estudar genes preditores de DC, os critérios de inclusão tentaram minimizar erros de seleção, de modo a garantir o fenótipo de doença. O critério de inclusão é robusto: lesão angiográfica significativa, determinada por angiografia coronária. Excluimos os mais novos, com menos de 30 anos, com baixa incidência de DC, e os mais velhos, com idade acima dos 65 anos, com muitas co-morbilidades associadas. Pelo mesmo motivo excluimos a DC associada com doença valvular grave.

V.1.1.4 CRITÉRIOS E FATORES DE RISCO UTILIZADOS NA CARACTERIZAÇÃO DA POPULAÇÃO

A doença aterosclerótica coronária permanece uma das principais causas de morbilidade e mortalidade no Mundo Ocidental. Os mecanismos multifatoriais subjacentes à sua fisiopatologia são de natureza ambiental, de natureza genética e da interação entre os dois.

Os FRT estão comprovadamente associados à DC. Não podiam deixar de ter uma atenção muito especial no estudo desta população. É necessária uma boa caracterização fenotípica para documentar os preditores de DC e para efetuar a sua comparação robusta com o valor do SRG, como preditor independente e adicional aos FRT.

A nível do fenótipo, avaliamos os FRT, os fatores modificáveis e os marcadores de doença CV já referidos. Estes foram avaliados de forma dicotómica e também como variáveis discretas em categorias de risco. Para esse efeito foram introduzidos *cut-points* ou limites emergentes dos consensos alargados da Sociedade Europeia de Cardiologia (ESC).

Na definição de tabagismo incluímos os ex-fumadores com menos de 5 anos de cessação, pelo facto da diminuição do risco para os eventos CV ser comparável com um não fumador apenas após os 5 anos. (121) Fomos exaustivos na avaliação antropométrica, utilizando vários índices de avaliação de obesidade central, cujo valor prognóstico é superior ao IMC isolado.

A avaliação sistemática com eletrocardiograma permitiu excluir a cicatriz de necrose miocárdica e submeter a avaliação por prova de esforço e ecocardiograma de todos os controlos com evidência de isquémia sub-endocárdica.

À variável dislipidemia, escrutinada por categorias de LDL, HDL e triglicéridos, segundo as orientações da ESC, adicionamos a Apo (B) e a lipoproteína (a), cruciais na fisiopatologia da placa aterosclerótica para construir um painel lipídico mais completo.

Na avaliação laboratorial também foram incluídas variáveis bioquímicas para uma avaliação dos potenciais marcadores de aterosclerose, como os leucócitos, o fibrinogénio, a hemoglobina e a PCR.

Avaliámos a homocisteína sérica em toda a população, uma vez que, incluímos no estudo genético variantes genéticas associadas ao eixo oxidativo, sendo assim possível inferir a interação genótipo-fenótipo.

Consideramos ter conseguido uma adequada caracterização fenotípica para documentar os preditores de DC.

V.1.1.5 CARACTERIZAÇÃO DA AMOSTRA PELOS FATORES DE RISCO TRADICIONAIS

Na caracterização basal da população, como seria de esperar, encontramos uma maior prevalência de FR nos doentes versus controlos. Por ordem decrescente, na dislipidemia 90,3 versus 70,3%; na HTA 71 versus 53%; no tabagismo 47 versus 23%; no sedentarismo 73 versus 45%; na obesidade 34 versus 29% e na DM 34 versus 13% (Tabelas 5-7).

Destes resultados, podemos inferir que nestes grupos etários o diagnóstico e controle da dislipidemia na prevenção da DC é ainda reduzido.

O abuso de álcool, muitas vezes omisso nos estudos de risco CV, continua a ser muito elevado na população portuguesa. No presente estudo 16,9% dos doentes e 14% dos controlos excediam as 300gr/semana. Se níveis baixos <30gr/dia podem estar associados com proteção CV, níveis elevados de consumo são considerados fator de risco para a ocorrência de DC. (212)

O excesso ponderal e obesidade central elevada foram documentados em mais de 90% dos indivíduos desta amostra, independentemente dos índices utilizados. Assim, em mais de 95% dos indivíduos, os índices estavam aumentados com significância: o perímetro abdominal (97 versus 95% nos homens e 99 versus 98% nas mulheres); a relação cintura-anca (96 versus 92% nos homens) e o índice cintura-altura (>0,5 – 96 versus 92%). Nesta população portuguesa confirma-se a elevada prevalência do excesso ponderal e da

obesidade central, reflexo da epidemia que atinge as sociedades desenvolvidas e em desenvolvimento. Os doentes, quando comparados aos controlos, apresentam resultados laboratoriais com níveis séricos mais elevados de leucócitos, fibrinogénio e PCR e mais baixos de hemoglobina e de clearance da creatinina, inferior a 60ml/min (Tabela 6).

No global, os doentes apresentavam características compatíveis com um fenótipo pró-aterogénico, com valores mais elevados da pressão arterial sistólica, da VOP e dos índices antropométricos. Curiosamente, a população doente apresentou melhor controlo da frequência cardíaca, por possível efeito de medicação cronotrópica negativa, habitual no doente com DC. (Tabela 5)

No global, a população em estudo apresenta características semelhantes aos de outros registos europeus como o *Malmö Diet and Cancer study* (213), *REGICOR study* (214) e o *Wellcome Trust Case Control Consortium* (WTCCC). (182)

Contudo, embora a média de idade da nossa amostra seja mais baixa que a dos registos, a prevalência de dislipidemia, DM, excesso ponderal e obesidade foi mais elevada. Isto, associado ao consumo exagerado de álcool, pode levar a que o risco CV na nossa população seja superior ao dos restantes registos. Estes dados põem em evidência o trabalho que temos que desenvolver para que haja um melhor controlo dos FRT na nossa população.

V.1.1.6 SELEÇÃO DE SNPs E GENOTIPAGEM

A seleção das variantes obedeceu a uma combinação de variantes genéticas causais implicadas em mecanismos fisiopatológicos conhecidos e de variantes genéticas emanadas dos GWAS, sem mecanismo fisiopatológico atribuído. Esta decisão esteve de acordo com a literatura da altura, sendo mesmo, nessa data, a forma mais comum de inclusão das variantes genéticas nos estudos de risco genético.

Assim, seleccionámos 33 variantes genéticas que pensávamos que poderiam vir a ter algum impacto na DC. No início, algumas destas variantes genéticas não tinham mecanismo fisiopatológico conhecidos. Passados cinco anos, o “estado da arte” atual atribui os seguintes mecanismos ainda em estudo: (Figura 8)

- seis polimorfismos genéticos no eixo lipídico (APOE, LPA, ZNF259, KIF6, SORT1 e PSCK9).
- nove polimorfismos genéticos nos eixos da diabetes, obesidade e insulinoresistência (ADIPOQ, FTO, HNF4A, IGF2BP2, MC4R, TCF7L2, SLC30A8, PPAR Gama, TAS250).

- três polimorfismos genéticos nos eixos da Renina-Angiotensina-Aldosterona, HTA (ECA I/D, ATIR1 e AGT 235)
- seis polimorfismos genéticos no eixo pró-oxidativo (MTHFD1L, MTHFR 1296, MTHF677, PON108, PON 192 e PON 55)
- nove polimorfismos genéticos, cujo mecanismo fisiopatológico não se encontra totalmente esclarecido. Aceita-se que estão envolvidos no ciclo celular, na transcrição de sinais, na diferenciação de células musculares ou que atuam como *binders* de co-transporte (SMAD3, GJA4, ADAMTS7, MIA3, 9p21, CDKN2B, PHACTR1, o ZC3HC1 e TCF21).

Embora posteriormente haja uma tendência, numa maioria dos estudos genéticos, em incluir exclusivamente um número cada vez maior de variantes genéticas emanadas dos GWAS, a nossa decisão de incluir variantes genéticas relacionadas com fenótipos intermediários (Obesidade, DM, HTA) reforça a importância que damos a esses fenótipos intermédios na etiopatogénese da DC na prática clínica. Na verdade, recomeça-se a encontrar esta atitude nas atuais linhas de investigação genética.

Se aceitarmos que o valor do risco associado à variante genética de um fenótipo intermédio é inferior ao valor associado ao de uma variante do fenótipo principal, neste caso a DC, compreenderemos que, para este tipo de análise, serão necessárias amostras populacionais cada vez maiores. É nossa convicção que muitas das variantes genéticas dos GWAS, cujo mecanismo não é conhecido, estarão relacionadas com os fenótipos intermediários para DC. Contudo, reconhecemos que, por vezes, se pode encontrar uma variante e que ao estudar o seu mecanismo de atuação na DC, poderá encontrar-se alvos terapêuticos com impacto prognóstico. Este foi o caso de sucesso dos inibidores do PCSK9, sobejamente conhecidos.

Tendo em conta este facto e considerando as frequências alélicas e os resultados reportados na literatura em europeus caucasianos para as 33 variantes genéticas analisadas no nosso estudo, 23 dos SNPs tem um poder expectável de 80%, e 10 dos SNPs de 90% considerando em erro alfa de 0,05.

Todas as análises genéticas deste estudo foram efetuadas num único Laboratório com certificação em Genética Humana. Isto, representa uma mais-valia do trabalho. A técnica da genotipagem com recurso a *primers* representa o *gold standard* na determinação de genótipos; seja isoladamente, gene por gene, ou com *chips* disponíveis. O controlo de qualidade no laboratório com 99% de *call rates* dá-nos garantias para os resultados encontrados.

Apenas determinamos as sequências de ADN das variantes genéticas de interesse no estudo. Pode ser redutor, limitar o risco genético para DC às variantes genéticas comuns do ADN. Embora o ADN não codificante possa desempenhar o seu papel, ele não foi o âmbito deste estudo. Há um entusiasmo científico, com recurso a técnicas de imputação, ampliando o conhecimento genético alargado do ADN (codificante, não codificante e *copy-variant numbers*) que poderá também aumentar o conhecimento sobre a aterosclerose coronária. Este caminho tem as suas limitações: a tecnologia de genotipagem pode permitir processar um grande número de pares-bases, mas teoricamente quanto maior o seu número, menor será o risco conferido por cada fragmento do ADN isolado, tornando difícil a sua valorização.

V.2 ESTUDO GENÉTICO e DOENÇA CORONÁRIA

V.2.1 ESTUDO GENÉTICO GENE A GENE E MULTIVARIADO

Para fazer o estudo genético numa população temos que caracterizar a amostra, relativamente aos genes em estudo e estudar o seu poder de associação com a DC. Num estudo genético tipo caso-controlo compara-se a frequência de alelos e genótipos, em diferentes *loci*, entre indivíduos com e sem fenótipo de doença. Estes *loci* são geralmente polimorfismos de um único nucleótido ou SNP.

Estudamos, na nossa amostra, as diferenças gene a gene, por análise bi e multivariada, com ajuste para os FRT (Tabelas 8-12).

Genotipámos os polimorfismos genéticos e identificámos o alelo de risco, a frequência do alelo *minor*, e o poder de associação com DC, para o genótipo de risco e para o alelo de risco. As frequências alélicas variaram entre os valores de 2,1% (gene LPA) e 8,8% (gene PPARG) para as menos frequentes e 45,8% (gene 9p21) e 46,4% (gene PON 108) para as mais frequentes. Em 18 dos 33 SNPs variantes, o alelo de risco, descrito na literatura, foi o alelo mais frequente. Nos restantes 15 SNPs, o alelo de risco correspondeu ao alelo *minor* da literatura (Tabela 9).

Em 10 SNPs, o efeito do alelo de risco na literatura foi semelhante ao descrito em estudos europeus de referência. Em 19 dos SNPs, o efeito demonstrado esteve de acordo com o descrito na literatura sem, no entanto, se atingir diferenças estatisticamente significativas. Em 4 SNPs, o efeito do alelo de risco na literatura foi contrário ao efeito protetor, encontrado na nossa população (Tabela 9).

Na análise bivariada gene a gene foram identificados 10 SNPs com significância, destes, sete mantiveram a significância na análise multivariada gene a gene e seis na análise multivariada com os fatores de risco (Tabelas 8, 10 e 12).

Avaliamos o poder da associação dos 10 genes com significância com a DC de acordo com os diferentes modos de hereditariedade (Tabelas 8).

Destes, a variante genética da LPA demonstrou ser a única a não estar em equilíbrio de *Hardi-Weinberg*. Devido à raridade do alelo de risco (1:1000), que apenas se encontrou em heterozigotia na nossa população, o resultado dos presentes dados é apresentado utilizando o alelo de risco em heterozigotia. O poder de associação com a DC avaliado pelo OR foi de 2,2 (IC95%: 1,47-3,2), pressupondo a codominância. Na nossa população, 5,5% dos casos e 2,6% controles apresentavam o alelo de risco em heterozigotia CT. A variante da Apo E conferiu um aumento do risco para DC de 1,21 (IC95%: 1,06-1,44) por cada alelo de risco, pondo em evidência a importância do eixo lipídico na aterosclerose coronária. Este achado foi já alvo de meta-análise. (126)

As variantes 9p21 e CDKN2B demonstraram estar em desequilíbrio de *linkage*. O alelo de risco da variante do CDKN2B conferiu um OR de 1,17 (IC95%: 1,06-1,3) e o alelo de risco da variante do 9p21 evidenciou um OR de 1,16 (IC95%: 1,04-1,28). Este efeito é similar ao efeito observado em outros estudos europeus, como o descrito no estudo *CARDIoGRAM* em 2008 (79). O WTCCC e três outras coortes europeias juntaram-se ao projeto *CARDIoGRAM* obtendo, numa amostra com mais de 22 000 casos, um aumento de cerca de 29% no risco de EAM por cada cópia da variante rs1333049 (9p21) ($P = 2 \times 10^{-20}$). (215)

A associação, entre a doença aterosclerótica e o *locus* 9p21, foi já reproduzida por vários grupos a nível mundial, independentemente da raça, da geografia e dos fatores de risco tradicionais. (216,217) Outros autores demonstraram que o *locus* 9p21, não só está associado à gravidade da DC, como também à sua extensão e progressão. (217,218) Contudo, em diversos estudos os resultados nem sempre foram concordantes quanto à sua associação com EAM no seguimento. (219,220) Com efeito, Horne *et al.* (221) demonstraram a existência de uma associação entre a presença do genótipo 9p21 e a ocorrência de DC, mas não com a ocorrência de EAM. (221) A variante do *locus* 9p21 parece estar também associada à mortalidade de qualquer causa, após cirurgia de revascularização miocárdica em doentes de raça branca, melhorando significativamente o valor preditivo do *EUROSCORE* logístico. (222)

O nosso grupo replicou a análise desta variante (9p21) na população portuguesa e encontrou uma prevalência de 35,7% do genótipo CC em pacientes com DC, com um OR de 1,34 ($p = 0,010$). (223) Quando ajustado para os FRT, o genótipo CC apresentou um OR de 1,7 ($p = 0,018$) e o genótipo CG obteve um OR de

1,5 ($p = 0,048$). Concluímos que, embora o mecanismo subjacente ao risco ainda seja desconhecido, a robustez deste alelo de risco na estratificação do risco para a DC tem sido consistente, mesmo em populações diferentes. A presença do genótipo CC ou CG pode, assim, ser útil para prever o risco de desenvolver a DC na população portuguesa. (223) O estudo da variante 9p21 foi alvo da primeira publicação que consta em anexo.

O polimorfismo PON 55M confere à isoforma da paraoxonase uma menor capacidade hidrolítica na esterificação dos ácidos gordos. A isoforma variante tem sido implicada na gênese da placa aterosclerótica, por tornar menos eficiente a ação antioxidante das HDL que é o local principal de ação das paraoxonases. Os resultados da literatura têm sido díspares, quanto ao seu efeito no risco para a DC, o que conduziu à não inclusão sistemática deste gene em muitos estudos genéticos de base populacional. Na nossa população, este gene esteve associado com a DC, quer por análise bivariada, quer por análise multivariada (OR=1,35; IC 95%: 1,08-1,69) enfatizando que os estados de oxidação-redução são uma possível via etiopatogénica independente para a aterosclerose. A meta-análise mais recente, publicada em 2016, por Hernandez Dias, suporta os nossos resultados (224). Ela inclui 64 estudos caso-controlo, envolvendo 19 715 casos e 33 397 controlos, onde se confirma a associação positiva do polimorfismo PON 55M com a DC (OR=1,44; IC95%= 1,33-1,56). (224) Recentemente, este gene tem sido alvo de investigação, com estudos caso-controlo de base populacional noutros fenótipos em que a capacidade antioxidante possa ter implicações. Assim, este polimorfismo tem demonstrado associação com o risco oncológico, principalmente com a neoplasia da mama e com a obesidade. Não se demonstrou associação com as doenças degenerativas do sistema nervoso central, como a doença de *Alzheimer* ou esclerose lateral amiotrófica.

Os polimorfismos ADIPOQ e os FTO têm sido consistentemente associados com a obesidade, insulinoresistência e aterosclerose acelerada. (225,226) Estudos realizados sobre a associação do polimorfismo FTO e o risco CV apresentaram resultados variáveis. No nosso estudo, o polimorfismo FTO demonstrou ser significativo na análise univariada com OR alélico de 1,13 (IC95%: 1,02-1,26). Confirmando a sua importância, a independência deste polimorfismo no risco CV foi objeto de uma meta-análise em que foi associado com aumento do risco para a DC com OR = 1,18, IC 95%: 1,07-1,30, independente do IMC e dos fatores de risco cardiovascular. (227) O polimorfismo da variante genética da adiponectina (ADIPOQ) tem sido relacionado com a desregulação do metabolismo glicídico e de forma menos direta com o metabolismo lipídico e a lesão vascular mediada pela adiposidade.

Na nossa população, demonstrou-se que aumenta o risco de DC com o risco por alelo de 1,16 (IC95%: 1,03-1,32) quer na análise bivariada gene a gene quer na análise multivariável (OR=1,2; IC95%: 1,03-1,42).

Neste estudo, os polimorfismos nas variantes genéticas TCF21, ZC3HC1 e PHACTR1, cujos mecanismos ainda não se encontram totalmente esclarecidos, apresentaram um risco elevado para DC OR entre 1,11 e 1,23 por alelo de risco. A variante genética TCF21 é um fator de transcrição que se julga ser crítico no desenvolvimento das células progenitoras epicárdicas que se diferenciam em células musculares lisas coronárias e fibroblastos cardíacos. Assim, esta variante genética, que na nossa população se associou a maior risco de eventos adversos, é apontada como possível mediadora da migração de células progenitoras para as lesões vasculares iniciais, contribuindo para a cápsula fibrosa das placas ateroscleróticas. Este poderá ser o seu mecanismo de ação na aterosclerose coronária.

Na análise multivariável por regressão logística das variantes genéticas associadas com DC foram encontradas 7 variantes genéticas significativas: TCF21 (OR=1,23), ZC3HC1 (OR=1,14), CDKN2B (OR=1,17), ADIPOQ (OR=1,17), LPA (OR=1,47), APOE (OR=1,21), PON55 (OR=1,16). Nesta análise multivariável, embora o eixo celular apresente o maior peso relativo, todos os eixos fisiopatológicos com exceção do eixo da hipertensão arterial mantiveram-se (Tabela 10).

Após identificarmos os preditores não genéticos da nossa população para a DC: o tabagismo (OR=3,1; IC 95%: 2,6-3,7); a diabetes (OR=2,7; IC95%: 2,2-3,3); a insuficiência renal (OR=2,2; IC95%: 1,5-3,2); o sedentarismo (OR=2,0; IC 95%: 1,7-2,4); a HTA (OR=1,6; IC95%: 1,4-2,0); a Lp (a) (OR=1,01; IC95%:1,01-1,01) e a dislipidemia (OR=1,3; IC95%: 1,0-1,6); o terceiro tercil da obesidade central (OR=1,4; IC95%: 1,1-1,8) e a Homocisteína (OR=1,3; IC95%: 1,0-1,6), fizemos a análise de regressão logística multivariada dos preditores genéticos e não genéticos, podendo concluir-se que na nossa população as seis variantes genéticas TCF21, LPA, CDKN2B, PON55, APOE 4, ZC3HC1 foram preditores do risco independentes para a DC após ajuste para os FRT (Tabelas 11 e 12).

Com este estudo genético da população demonstramos as associações estatisticamente significantes entre as variantes genéticas e o fenótipo de doença. No global, isto merece algumas considerações.

Primeiro, há que referir que as análises genéticas dependem da homogeneidade e do tamanho da amostra populacional. Uma vantagem é que o estudo ocorreu entre indivíduos provenientes duma região geográfica bem definida com *pool* genético estável, representativo da população da europa do sul e assumiu o desenho dum estudo caso-controlado, a forma mais comum de estudo genético. Outra vantagem é ser um estudo de associação com base populacional, em oposição aos estudos genéticos de base familiar onde são incluídos indivíduos com laços familiares.

No fim, 23 SNPs não atingiram significância estatística. Contudo, pelo tamanho da amostra e pelos OR encontrados, o estudo não apresenta poder suficiente para descartar a possível associação destes SNPs com a DC. Como referido anteriormente, seria necessária uma amostra populacional maior e mais robusta para ter poder suficiente para avaliar associações, na nossa população, de alelos com menor prevalência individual e com menores ORs para a DC, do que os encontrados nos estudos prévios. Pelo mesmo motivo, também não podemos excluir de forma inequívoca a sua não associação com a DC.

A segunda consideração relaciona-se com o facto de o tipo de associação ser diferente, consoante o tipo de polimorfismo testado, se localizado em gene candidato ou se emanado dos GWAS.

Inicialmente, as primeiras associações com a DC foram efetuadas em estudos orientados para polimorfismos alvo, já identificados e considerados como marcadores implicados na etiopatogénese da DC. Esta abordagem tradicional do gene candidato identificou cerca de 30 *loci* com efeitos significativos na aterosclerose. (228) Contudo, geraram resultados contraditórios ou com associações fracas e os estudos de replicação necessários para sua validação, não foram consistentes.

Na população portuguesa, foi exemplo do estudo em genes alvo, a publicação do nosso grupo de investigação em 2004 sobre a genotipagem dos polimorfismos I/D do gene ECA com resultados significativos semelhantes aos descritos na literatura. (229)

Nos estudos GWAS por outro lado, ocorre uma associação que é estatística, por inferência numérica entre grandes populações, estando limitada a sua associação conclusiva apenas à altura em que se passa a conhecer a fisiopatologia associada. Em 2007, Samani *et al.* identificaram pela primeira vez *loci* fortemente associados à DC no estudo WTCCC com 1 926 casos e 2 938 controlos, replicadas consistentemente no *German MI Family Study*. (182) Após o desenvolvimento de *arrays* de alta capacidade em 2008, (230) os GWAS simultaneamente, já examinaram milhões de polimorfismos em estudos de casos-controlos em subpopulações de diferentes etnias. Muitas variantes genéticas foram avaliadas de forma consistente nos anos seguintes, sendo a variante 9p21 a mais estudada em muitas populações de diferentes etnias. Na população portuguesa, como referido anteriormente, a variante do 9p21 foi também comprovada estar associada com a DC de acordo com o nosso estudo com publicação em anexo. (223)

Nos estudos tipo GWAS, os marcadores genéticos encontrados são submetidos a um adequado controlo estatístico, a fim de prevenir os erros de amostragem múltipla. Proceder-se desta forma para impedir a probabilidade de encontrar um falso positivo, resultado dum número elevado de repetições. Estes polimorfismos genéticos devem demonstrar independência de associação com genes próximos

(desequilíbrio de *linkage*), especialmente se estes estiverem associados com o fenótipo. Finalmente, os achados não devem refletir apenas o património genético duma população restrita, na medida em que pode existir um erro na seleção de casos e controlos, denominado de estratificação populacional ou *admixture*.

O número de variantes genéticas, encontradas com este tipo de associação, encontra-se ainda em crescendo, sendo que a meta-análise mais recente dos estudos de mapeamento genómico amplo (GWAS) para a DC identificou já 58 *loci* significantes e 104 *loci* sugestivos de associação, com o risco aumentado para a DC. (82, 87,231)

Se por um lado o nível de significância mínimo de 5×10^{-8} trouxe consistência e futuro na genética CV, por outro lado trouxe a necessidade de replicação em diferentes populações e a busca da ligação do polimorfismo à via fisiopatológica implicada na aterosclerose, para validação dos resultados. (232)

Após todas as considerações, apraz-nos concluir que no presente estudo, a maioria dos polimorfismos estudados demonstraram efeitos similares aos observados na literatura, e que, na nossa amostra 10 SNPs evidenciaram associação com significância estatística com a DC.

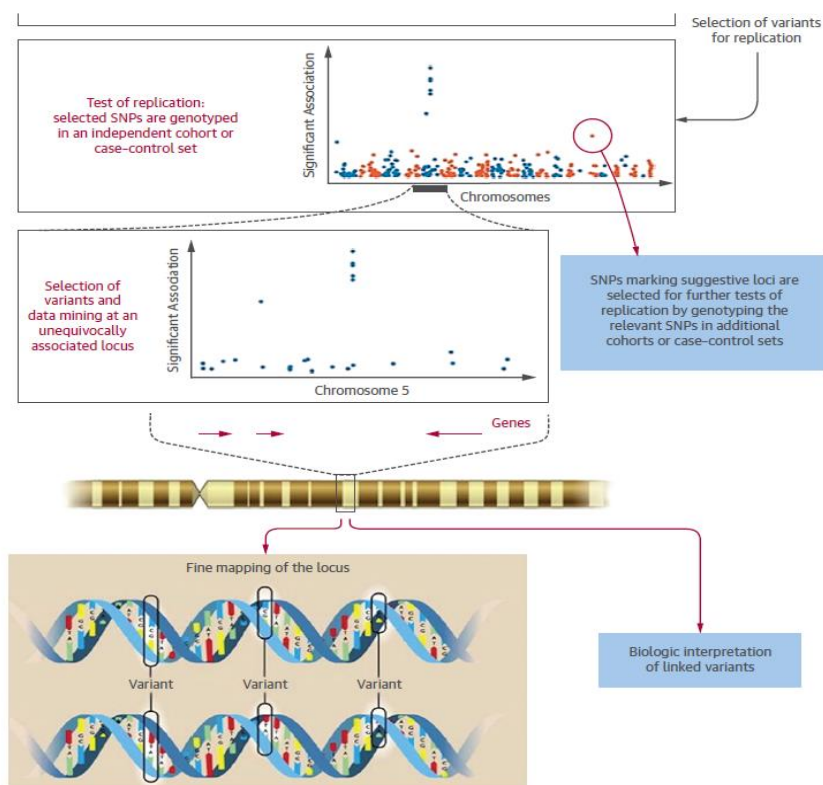


Figura 35: As variantes genéticas identificadas em estudos GWAS são posteriormente confirmadas em estudos de replicação. Estudos adicionais incluindo o mapeamento fino do locus pretendem conduzir à sua implicação numa via biológica. *Adaptado de Assimes & Roberts, 2016. (233)*

Sabendo que um resultado significativo, num estudo de associação genética, por si só, não determina uma implicação direta do SNP no risco de DC, ficamos a aguardar que, muitos destes alelos, em *locus* marcadores, venham um dia a ser diretamente implicados numa via fisiopatológica da DC.

V.4 INTERAÇÃO GENE-GENE E GENE-AMBIENTE

A interação G-G e G-A parece um campo promissor na genética cardiovascular atual. Após um período de excecional entusiasmo com a identificação de novas variantes genéticas nos estudos GWAS, concluiu-se que uma parte significativa do componente genético da DC, está ainda por identificar. Esse componente "escondido" é atribuído às interações complexas gene-gene e gene ambiente.

A análise destas interações apresenta dificuldades de ordem metodológica e de limitação computacional, dado o elevado número de combinações a avaliar.

Muitos métodos exploratórios foram desenvolvidos para a análise multivariável de dados de elevada ordem, que vão desde os simples testes de associação, passam pelos métodos de regressão logística e vão até aos vários métodos de *machine learning* ou de reconhecimento de padrões.

O método de *machine learning* mais popular para estudar as interações é o *Multifactor Dimension Reduction* (MDR). (234-236)

Utilizando os testes de associação, identificámos 11 interações significativas entre as variantes genéticas e as variáveis bioquímicas correspondentes; demonstrando que, na nossa amostra, podemos melhorar a predição do modelo dos FRT, se tivermos em linha de conta as possíveis interações G-G e G-A. (Tabela 13)

No caso da variante genética LPA e APOE observou-se uma interação significativa com a dislipidemia. A dislipidemia, categorizada de acordo com as indicações da Sociedade Europeia de Cardiologia, contempla o LDL, o HDL, e os triglicéridos. Da interação, encontrada neste estudo, depreende-se a importância da caracterização do perfil lipídico nos doentes. Pensamos que em centros sem possibilidade de estudo genético da LPA, para além dos LDL, os doseamentos de Apo B e lipoproteína(a) podem vir a ajudar na quantificação do risco lipídico.

Consideramos interessante a interação encontrada entre o gene PPARG e o HDL pelo seu potencial terapêutico já em estudo por várias empresas farmacêuticas. Os PPARG são fatores de transcrição, reguladores do metabolismo lipídico intra e extracelular.

O PPARG é expresso ao nível do tecido adiposo e regula a expressão da lipoproteína lipase, estimula o catabolismo lipídico e melhora a sensibilidade à insulina. É o local de ação de fármacos como as tiazolidinonas. Publicações semelhantes, com estudos das interações G-G e G-A são escassas na literatura. Contudo, numa população chinesa foram documentadas Interações G-A com os triglicéridos e com as LDL. (237,238)

Confirma-se o que se esperava da interação G-A entre a variante genética TCF7L2 e a glicémia com significância pela sua fisiopatologia.

Salientamos, na nossa análise, as interações entre as variantes genéticas da meta-tetra-hidrofolato redutase com os valores séricos da homocisteína (mais significativa) e do ácido fólico (menos significativa), embora o valor da homocisteína como fator de risco direto para a DC seja ainda considerado controverso.

Recorrendo à regressão logística através do teste de razão de verosimilhança analisamos as interações mais significativas G-G e G-A do nosso estudo. Compara-se o modelo de regressão logística que inclui apenas os genótipos e outro que inclui também as interações entre as variantes. Verifica-se que o padrão de interação G-G, mais frequente ocorre entre os genótipos de variantes genéticas ligadas ao ciclo celular (ADAMTS7, CDKN2B, KIF6 e *locus* 9p21) e os genótipos de variantes genéticas ligadas à oxidação-redução (MTHFR677 e PON192) (Tabela 16).

Sabemos que as variantes genéticas CDKN2B e *locus* 9p21 são importantes reguladores da lipotoxicidade e da expansão do tecido celular subcutâneo. (239) Considerando a relação das variantes ligadas ao ciclo celular com as variantes da oxidação-redução, é possível que exista uma fisiopatologia comum, a investigar, relacionada com os níveis de LDL oxidados e com a redução de HDL associadas às variantes da MTHFR e da PON. (240,241)

A variante genética ADAMTS7 foi associada à DC por condicionar uma resposta alterada na lesão vascular, com menor capacidade de migração celular das células musculares lisas do endotélio coronário. Encontrámos uma interação G-G significativa *de novo* entre o ADAMTS7 e a variante MTHFR677 na nossa população. No futuro, para explorar esta interação, o estudo dum modelo de célula muscular oxidada parece interessante.

A variante genética KIF6 é responsável pelo transporte celular de vários componentes moleculares, organelos e mRNAs. É conhecida a sua interação com o parâmetro lipídico LDL e com a resposta da terapêutica à estatina. (242) A interação da variante KIF6 com o gene do recetor do angiotensina,

encontrada no nosso estudo, era desconhecida. Espera-se que estudos futuros possam confirmar esta interação, e se for o caso, que seja encontrado o mecanismo fisiopatológico responsável (Tabela 16).

Por testes de associação identificamos a interação das variantes genéticas com os FRT. Das 33 variantes testadas encontraram-se dois modelos de interação significativa: a variante ZNF259 associou-se com a dislipidemia e a variante TCF7L2 com a diabetes (Tabela 14).

A variante genética ZNF259, pertencente à família dos anéis de Zinco, está genericamente implicada em vários processos de proliferação celular e transdução de sinal, e está também associada a alterações do perfil lipídico em populações europeias e orientais. (243,244) No presente estudo valida-se a sua significância num grupo representativo da população portuguesa e do sul da Europa. O seu mecanismo de ação fisiopatológico ainda não está estabelecido.

A interação positiva da variante genética TCF7L2 e a diabetes, já conhecida das grandes meta-análises com predomínio de população europeia (245), também se valida nesta amostra. Não só está envolvida no controlo da secreção da insulina através dum recetor da célula β pancreática, pela via *wingless*. A variante implica um risco aumentado de 50% para a DM. (159) Como também está associada a um risco independente para DC com um OR 1,9 a 2,3. (170)

Para avaliar melhor os modelos de variantes genéticas em cada eixo fisiopatológico com os FRT utilizou-se o modelo de *data mining* MDR.

A nível celular, a interação da variante genética ZC3HC1 com o tabagismo, a HTA e a DM permitiram obter um modelo interativo com OR de 4,3 (IC95%: 3,7-5,1) com consistente validação cruzada de 10/10 ($p < 0.0001$). A variante genética ZC3HC1, da família dos anéis de zinco, é responsável pela dinâmica regulatória do ciclo celular e curiosamente na literatura veio inicialmente descrita como sendo independente dos fatores de risco como o colesterol, a HTA, a obesidade, o tabagismo e a DM. (193) Contudo, recentemente, numa população finlandesa foi constatada uma interação da variante ZC3HC1 com a hipertensão. (246) Esta interação com os restantes FRT não se encontra descrita, mas devido à sua consistência pode ser um motivo promissor de investigações futuras (Tabela 17).

Ao nível do eixo obesidade e DM, nesta amostra, a interação da DM e das variantes de risco genéticas TCF7L2 e TAS2R50 constituíram o melhor modelo para a predição de DC com um OR de 2,98 (IC95%: 2,5-3,6) com consistente validação cruzada 10/10. A variante genética TAS2R50 funciona como um homólogo do

recetor do paladar e foi associada em várias populações ao risco de DC e EAM por influenciar as escolhas dietéticas dos portadores e por estes apresentarem menor saciedade após ingestão calórica. (247) O TCF7L2, como foi referido, está envolvido no controlo da secreção da insulina através dum recetor da célula β pancreática, pela via *wingless*. A variante implica um risco aumentado de 50% para DM (159). Não existem ainda outras publicações com recurso a este tipo de análise estatística para encontrar modelos preditores para DC no eixo da diabetes e obesidade (Tabela 18).

Nos modelos de interação G-A associadas ao eixo da oxidação-inflamação, o melhor modelo de interação G-A é composto pelas 4 variantes genéticas associadas com a oxidação: MTHFR677, PON108, PON55, MTHFD1L (OR=1,45; IC95%: 1,25-1,7; $p < 0,0001$), com validação cruzada de 9/10. As restantes variáveis, homocisteína e PCR > 3, não demonstraram melhorar o valor preditivo. Estes dados confirmam que a variabilidade da concentração sérica de marcadores do risco de CV como a PCR e a homocisteína são menos atrativas na estratificação do risco de CV. Contudo, nenhum dos modelos é um modelo bem ajustado, visto que, o valor do p teste, 0,686, não atinge significância estatística (Tabela 19).

O melhor modelo de interação G-A associado ao eixo lipídico foi o da interação G-A composto pela variante genética PSRC1 e a dislipidemia (LDL, HDL, Triglicéridos categorizados) com risco conjunto para a DC com o OR de 3,92 (IC95%: 3,2-4,9; $p < 0,0001$) e com adequada consistência do modelo de teste 10/10.

A variante genética PSRC1, responsável por cerca de 1% da variação dos níveis circulantes de LDL, tem sido associada com o risco para a DC por condicionar a elevação sérica das LDL (Tabela 20).

No global, neste estudo encontraram-se importantes interações G-G e G-A evidenciadas por MDR.

A demonstração da utilidade do MDR nesta análise não pretende ser exaustiva, pois como se pode depreender da literatura, não é ainda consensual qual o método de *data mining* a aplicar às interações biológicas de elevada ordem. Apesar de ser o método mais comum, estes resultados devem apontar metas para novos estudos e para novos modelos, não devendo ser redutores em relação a outras potenciais interações.

Para ter sucesso, as novas descobertas na genética cardiovascular devem basear-se e não ignorar a complexidade da relação entre o genótipo e o fenótipo. Este permanece um desafio significativo, dada a natureza complexa da arquitetura genética.

V.4 SCORE DE RISCO GENÉTICO E DOENÇA CORONÁRIA

Mais do que excluir potenciais associações, a constelação das variantes genotipadas reunidas sob a forma de um *score* é com efeito a uma forma eficaz de quantificar o risco genético para a DC.

Após o aparecimento entusiástico da associação do gene 9p21 com a DC, em 2007, na maioria das populações mundiais genotipadas, rapidamente se entendeu que o peso relativo de cada variante genética para a DC seria pequeno.

Assim, o risco genético individual em cada doente seria proporcional ao todo (conjunto dos riscos versus proteções de todas as variantes genéticas) e não às partes (cada variante genética). Simultaneamente, na literatura, surgiam mais variantes genéticas, cada vez com peso relativo menor e apareciam as primeiras formas de congregar a informação genética, denominadas de SCORE de RISCO GENETICO (SRG).

V.4.1 ANÁLISE EXPLORATÓRIA DOS MODELOS DE SCORE DE RISCO GENÉTICO

A computação do SRG apresenta essencialmente três considerações metodológicas que são ainda motivo de debate científico não conclusivo: a seleção, o número e o modelo genético das variantes a incluir no SRG.

Neste estudo, optámos por incluir variantes genéticas replicadas para a DC e para fenótipos intermediários, significativos para a doença, tais como a EAM, a diabetes, a obesidade e a dislipidemia. Assim, foram incluídas variantes de genes candidatos e variantes genéticas dos GWAS.

Incluir genes candidatos da era pré-GWAS, admitindo níveis de significância inferior ao $p=5 \times 10^{-8}$ ou restringir a sua inclusão a variantes emanadas e replicadas em GWAS é uma dúvida que se coloca. Uma grande desvantagem, de só incluir SNPs altamente significativos e replicados, é existirem outros que apesar duma verdadeira associação com a DC, não atingem o limiar da significância GWAS e, por isso, não são objeto de estudo. (248) Assim, como parte da nossa seleção, mantivemos genes candidatos em eixos fisiopatológicos plausíveis que, nos estudos genéticos do nosso trabalho, tinham demonstrado valor estatístico na nossa amostra.

De referir que, de início, dado o alto valor de p imposto nos GWAS, após a correção de Bonferroni ($p=5 \times 10^{-8}$), a maioria das variantes associadas com os FRT não atingiam significância estatística para DC. Análises posteriores, menos restritivas, passaram a associar as variantes com o risco de DC de fenótipos intermediários (a dislipidemia, a obesidade, a diabetes, a oxidação – FRT), legitimando assim a nossa assunção inicial de as incluir no estudo. (249,250)

De facto, parecem existir determinantes genéticas partilhadas entre os FRT e a DC, reforçando a ideia comum de que as doenças complexas são o produto de múltiplas variantes genéticas – que podem ser comuns a vários fenótipos – pelas interações entre si e com o meio. Do mesmo modo, se depreende existirem correlações genéticas que acompanham as correlações fenotípicas, por nós já extensamente conhecidas, e valorizadas clinicamente. (251)

Finalmente, com o aparecimento dos estudos de randomização mendeliana, é confirmado o papel comum dos determinantes genéticos da diabetes no risco para a DC. (252) Para outros marcadores de risco como a HDL e a PCR essa relação ainda não está estabelecida. (253,254)

Queremos ainda chamar a atenção para o facto de haver uma grande sobreposição poligénica entre a DC e os fatores de risco, uma vez que, os métodos estatísticos utilizados não permitem distinguir entre pleiotropia ou mediação. Neste tipo de investigação não nos foi possível tirar conclusões sobre a relação causal ou não dos genes com a DC.

Foram estudadas 33 variantes, destas duas foram excluídas do SRG:

a variante do gene da LPA (por não se encontrar em equilíbrio de Hardy-Weinberg) e a variante CDKN2 (por se ter encontrado em desequilíbrio de *linkage* com a variante genética 9p21 e ter menos significado estatístico que esta na análise bivariada (gene a gene).

Assim procedemos à computação do SRG com 31 variantes genéticas, e realizamos uma avaliação exaustiva de todos os modelos habitualmente considerados para construir os SRG.

Existe uma certa homogeneidade nos resultados, sendo o modelo multiplicativo o escolhido (produto do OR dos genótipos) por ter demonstrado uma curva AUC mais discriminativa de 0,61 (IC95%: 0,59-0,62). Este modelo demonstrou uma associação com a DC com um OR de 1,78 (IC95%: 1,5-2,1), 54% de sensibilidade e 62,3% de especificidade (Tabela 29).

Na análise exploratória efetuada para determinar o número de variantes a incluir, as curvas AUROC demonstram incrementos positivos apenas para os 10 primeiros genes. (Figura 19) Contudo, a inclusão das restantes variáveis permitiu elevar a curva AUC de 0,59 (IC95%: 0,57-0,60) referente às 10 variantes genéticas, para 0,61 (IC95%: 0,59-0,62). Esse incremento foi estatisticamente significativo (analisado pelo teste de DeLong ($p < 0,0001$)) e com um acréscimo de especificidade (de 58% para 62%) razão pela qual mantivemos todas as variantes genéticas em análises posteriores (Tabelas 28).

Voltando a recalcular a discriminação do SRG numa amostra aleatória de 20% da população, obtiveram-se AUCs sobreponíveis. A comparação dessas curvas (pelo teste de DeLong) não foi por isso significativa, indicando que o SRG é reprodutível (Figura 20).

V.4.2 ESTUDO DISCRIMINATIVO DO SCORE DE RISCO GENÉTICO

No conjunto, o SRG é mais elevado nos doentes do que na população controlo, sendo 24,3 versus 23,5 no modo aditivo e 0,67 versus 0,48 no modelo multiplicativo. Na distribuição por quartis, observam-se diferenças significativas nos quartis dos extremos. Uma grande percentagem de doentes está no quarto quartil do SRG (30,5%) e o grupo controlo apresenta a maior percentagem de população no primeiro quartil do SRG (30,7%). Como seria de esperar, os doentes estavam geneticamente mais predispostos para a DC e o SRG possui a sensibilidade e especificidade para os classificar corretamente (Tabelas 26 e 27).

Como esperávamos, observamos níveis crescentes de risco para a DC quer com a distribuição crescente do número de alelos de risco quer com a distribuição crescente do SRG em decis (Figuras 16, 17 e 24).

Se considerarmos o quinto decil (mediana) como classe de referência, os decis mais baixos são protetores do risco de ocorrência de DC e os doentes no décimo decil apresentam um risco com OR=2,5 (IC 95%: 1,8-3,5). Se considerarmos o primeiro decil como classe de referência, documentam-se níveis crescentes de risco para a DC, desde o segundo decil com OR = 1,4 (IC95%: 1,0-2,0) até ao décimo decil com OR= 4,04 (IC95%: 2,9-5,7), mantendo-se sempre a sua significância. Constata-se que acima do sexto decil ocorre risco para a DC, independente da classe de referência utilizada (Figura 24 e Tabela 36).

Embora a distribuição do *score* se sobreponha entre casos e controles, o SRG está significativamente associado ao risco de DC e, como tal, poderia ser usado como uma ferramenta independente para identificar indivíduos com maior risco, com vista à modificação do estilo de vida ou intervenção terapêutica.

Nas tabelas 30 e 31 apresentam-se os resultados do estudo comparativo caso-controlo dentro do primeiro e quarto quartil do SRG, analisados os extremos dos quartis do SRG, quanto às características: basais, antropométricas do exame objetivo e bioquímicas.

Tendo em conta que os doentes no quarto quartil do SRG apresentavam idade mais avançada $53,9 \pm 7,8$ anos comparados com os controlos no mesmo quartil ($52,3 \pm 7,9$ anos), tivemos o cuidado de incluir a variável idade em análises posteriores de regressão logística, apesar de, por desenho do estudo, haver já emparelhamento por idade e género entre os grupos.

No primeiro quartil, não se identificaram diferenças quanto à idade, mas como o gênero masculino era preponderante, subtemos o SRG a uma análise de acordo com o gênero.

Como esperávamos, independentemente do quartil do SRG, os doentes comparados aos controles apresentavam sempre maior prevalência de FRT, nomeadamente DM, tabagismo, dislipidemia, HTA e sedentarismo e, no perfil bioquímico-laboratorial tinham valores séricos mais elevados de leucócitos, fibrinogénio, Lp (a) e homocisteína e valores reduzidos da clearance da creatinina. Estas diferenças têm significância estatística.

Como descrito na literatura, o risco de DC foi superior no gênero masculino em relação ao gênero feminino (OR=1,7; IC95%: 1,5-2,0), independente do SRG. Contudo, não se observam correlações significativas do SRG em relação ao gênero (Tabela 33).

Devemos ter em atenção que os genes do SRG pertencem aos cromossomas autossómicos e por isso a sua distribuição deve ser homogénea em ambos os géneros. Quando incluímos o gênero na análise, na verdade, estamos a incluir um conjunto de genes desconhecidos, protetores ou promotores da DC que se pensam presentes nos cromossomas sexuais. Isto pode vir a explicar parcialmente a maior incidência da DC no sexo masculino.

Contudo, como seria de esperar, nos diferentes grupos etários, o SRG é sempre mais elevado nos casos do que na população controlo. Mas, se incluirmos o gênero, o SRG dos casos foi mais elevado no sexo masculino. Nos controlos, o sexo feminino apresentou valores superiores aos do sexo masculino (Tabela 33). Mais uma vez, isto leva-nos a especular sobre o possível valor genético deletério do cromossoma Y e do provável valor protetor de um segundo X.

Associado ao risco crescente para DC, verificou-se uma interação significativa entre a idade e o SRG. Assim, após estratificação por décadas de idade, o risco para DC dos indivíduos com mais de 60 anos foi quase duplo (OR máximo=1,93; IC95%: 1,5-2,5) do risco para DC dos indivíduos com 30 a 39 anos de idade (Tabela 34). Isto compreende-se, se tivermos em linha de conta que a DC é uma doença progressiva, que começa na infância e se manifesta nas décadas mais avançadas da vida. A presença de um SRG aumentado numa baixa idade, não deve apresentar um risco significativamente elevado para essa idade. Pelo contrário, ao fim de algumas décadas, aumenta a probabilidade da DC se vir a manifestar em indivíduos com SRG mais elevado, daí que para o mesmo SRG, o risco associado a DC seja mais elevado na população mais idosa. Isto ao contrário de nos deixar despreocupados em relação aos jovens com SRG elevados, talvez nos deva levar a uma intervenção preventiva na saúde desses indivíduos, na tentativa de modificar o aparecimento tardio da

DC. Preocupamo-nos com o facto de o SRG poder traduzir só um componente genético de um fenótipo intermédio. Para esclarecer esta relação dos FRT com o SRG, calcularam-se então as *area under the curve* e respetivos IC do SRG em cada subgrupo de doentes com fenótipo da diabetes, da dislipidemia, da HTA e da história de DC familiar precoce.

A discriminação do SRG para a DC foi máxima nos doentes com tabagismo e dislipidemia (AUC=0,61) e menor nos doentes com DM ou hipertensão (AUC de 0,57 e 0,60 respetivamente). Na análise efetuada, não se obtiveram diferenças significativas dos valores AUC para especificidade e sensibilidade para os subgrupos estudados, pelo que o SRG mantém a sua robustez independente dos subgrupos de FRT.

Curiosamente, o subgrupo de doentes com história familiar de DC precoce apresenta uma sensibilidade elevada (87,9%) e uma especificidade baixa (24,6%) e por isso tem um reduzido valor preditivo negativo (Tabela 37).

Isto significa que nos indivíduos com HF de DC, o SRG elevado identifica corretamente a doença, dada a elevada sensibilidade da análise. Contudo, dos indivíduos sem SRG elevado, uma elevada proporção também pode ter DC, daí a baixa especificidade. Contudo, raciocinando sobre o que achamos, podemos afirmar que os indivíduos com HF de DC têm uma elevada representação no nosso SRG, o que nos leva a especular que os genes que constituem o nosso SRG devem participar de forma significativa na DC, poligénica e multifatorial, de incidência familiar. Mais, como o SRG tem um OR de 1,8 na amostra, significativamente associado a DC, parece que o naipe genético responsável pela HF de DC e pela DC habitual na população pode ser semelhante, com uma maior concentração de alelos de risco nos indivíduos com HF de DC.

Ao comparar as medianas do SRG nos diferentes grupos etários, não se detetaram diferenças quanto à mediana do SRG ($p=0,123$) no global da população.

Nos casos, não se detetaram diferenças na distribuição do número médio de FRT. Nos controlos, os FRT aumentaram gradualmente com a idade de 1,14 para 1,78, mas sem significado estatístico (Figura 23).

Este facto justificou a sub-análise dos doentes muitos jovens com DC, cuja menor prevalência teórica dos FRT, torna interessante estudar o valor preditivo do SRG em estudos prospetivos.

V.4.3 VALOR PREDITIVO DO SCORE DE RISCO GENÉTICO

A importância do SRG para a prática clínica dependerá de o SRG for preditor independente para a DC, comparando-o com os FRT. Nesta população, todos os FRT foram preditores independentes para a DC por análise multivariada. O tabagismo apresentou um OR de 3,3 (IC95%: 2,8-3,9); a DM um OR de 3,1 (IC95%: 2,6-3,8); a HTA um OR de 2,1 (IC95%: 1,8-2,4) e a dislipidemia um OR de 1,3 (IC 95%: 1,0-1,6) (Tabela 38).

Dada a natureza poligénica e interdependência dos vários FRT, foi importante avaliar a calibração e ajustamento do modelo. Neste estudo, obteve-se uma adequada calibração e ajustamento do modelo logístico, sem evidência de colinearidade (sobreposição) entre os FRT e o SRG (Tabela 39).

Verificamos que o SRG é preditor independente dos FRT para a DC, com um risco para DC de 1,8 (IC95%: 1,5-2,1) $p < 0,0001$. Com a adição do SRG, os FRT mantêm as suas associações com a DC com um poder sobreponível. Mantem-se o tabagismo e a DM como preditores principais para a DC (Tabela 39).

Nesta amostra, confirma-se que os FRT apresentam uma boa capacidade discriminativa para a predição da DC, tendo-se encontrado uma curva AUC de 0,74 (Figura 25).

Como veremos, quanto mais aferida esta curva basal for, maior dificuldade em obter resultados favoráveis com a introdução de um novo marcador de risco. Na amostra, após a introdução do SRG, ocorreu uma elevação da curva discriminativa para 0,76, refletindo melhoria da sensibilidade e especificidade. A comparação das curvas, realizada pelo teste de Delong, demonstrou que o aumento foi estatisticamente significativo.

Por regressão logística multivariada todos os quartis do SRG foram preditores significativos para DC, sendo que no quarto quartil o OR foi de 2,6 (IC95%: 2,1-3,2) (Tabela 40).

Para confirmar a robustez do SRG fizeram-se duas análises multivariadas por regressão logística do SRG em quartis com dois graus de ajuste: um ajuste para os FRT, idade e género e um modelo multi-ajustado com todas as outras co-variáveis, potencialmente associadas a DC (sedentarismo, VOP, álcool, IMC, frequência cardíaca). Pelo teste de *Hosmer Lemeshow*, os dois modelos encontram-se calibrados. Assim, demonstra-se que, mesmo considerando todos os potenciais fatores de risco, em todos os quartis do SRG, este mantém-se como preditor independente para DC. Após o ajuste, ocorreu um ligeiro acréscimo do valor da associação com a DC, para todos os preditores de DC: tabagismo com um OR de 3,6 (IC95%: 3-4,3); a DM com um OR

de 3,3 (IC95%: 2,6-4,1) e o quarto quartil do SRG com um de OR= 2,5 (IC95%: 2-3,2) foram os preditores independentes que demonstraram maior força de associação com a DC nesta população (Tabela 41).

A utilidade do SRG na prática clínica pode medir-se pela capacidade de discriminação e na reclassificação de indivíduos dispostos em categorias de risco pré-definidas de acordo com os FRT. Calcula-se o movimento para uma categoria superior no caso da população com DC e o movimento para uma categoria inferior na população controlo. Quando adicionado um novo marcador de risco – neste caso o SRG – NRI, define-se como a percentagem de indivíduos que muda de categoria em cada subgrupo de risco.

Analisando os resultados do NRI, computado de forma categórica e aplicado ao estudo, verificamos que a incorporação do SRG com os FRT melhorou a classificação de risco dos modelos basais. O NRI demonstrou melhor incremento na reclassificação dos doentes da categoria 50-75% para a categoria de risco máximo (>75%). Do mesmo modo 14,1% dos controlos saudáveis foram reclassificados numa categoria inferior de risco (de 25-50% para <25%). No conjunto, houve um NRI de 20,9% (28,1% - 7,2%) para doentes e NRI de 3,3% (18,7%-15,4%) para controlos. Assim, se a nossa amostra tivesse incorporado o SRG, uma proporção significativa dos seus indivíduos seria classificada numa categoria de risco diferente (Tabelas 44 e 45).

Do mesmo modo, verifica-se que considerando a reclassificação como contínua e não em categorias de risco, na amostra houve nos doentes um NRI de 31% (NRI 14,6% (57,3% - 42,7%) e nos controlos um NRI de 16,4% (58,2%- 41,8%), com uma vantagem para os doentes de ter um score diferente do seu risco de CV. Quando se incorpora o SRG com os FRT, dá-se uma pequena melhoria na discriminação da predição, com um IDI de 2,5% (1,9-3,1%), com significado estatístico (Tabela 46).

Estes resultados são semelhantes a outros estudos em populações Caucasianas, cujo SRG com 13, 29 ou 109 SNPs (97,255,98) aumentaram marginalmente o poder preditivo dos FRT. Este aumento do poder preditivo foi conferido tanto pelo aumento da AUC, como por alterações do índice C ou por métodos estatísticos discriminativos mais recentes, como as medidas de reclassificação e discriminação.

Na coorte de Ripatti *et al.*, o SRG ponderado baseado em 13 SNPs também teve um valor preditor independente, mesmo depois de ajustado para o sexo, idade e fatores de risco tradicionais. Similarmente, o risco relativo do SRG foi menor que o conferido pela dislipidemia e comparável aos efeitos da hipertensão.

(97)

Ao comparar o peso relativo do SRG na análise logística multivariada, obtivemos um OR ligeiramente inferior ao tabagismo, à hipertensão e à dislipidemia. De qualquer modo, obtivemos um OR de 2,6 para o quarto quartil do SRG, comparativamente superior ao OR de 1,66 descrito por Ripatti *et al.* no quinto quintil. (97)

Thanassoulis *et al.* (98) demonstraram que, adicionando 89 SNPs associados a fatores de risco modificáveis a um SRG baseado em 13 SNPs GWAS, o poder do SRG não aumentou, indicando um risco de 1,01 (IC95%: 0,99-1,03; $p=0,48$). Isto revelou que, na análise de SRG, os polimorfismos associados aos FRT podem ter um efeito marginal, sendo mascarados pelo risco de polimorfismos com mais peso.

Uma vez que é difícil aceitar que variantes genéticas importantes para FRTs não atingissem uma associação estatisticamente significativa com o risco de DC, em 2015, Jansen *et al.* observaram uma associação significativa entre vários SNPs, associados à DM e ao risco de DC. (249) Recentemente, Webb *et al.* identificaram 6 novos *loci* relacionados com os FRT que se associaram com a DC. (256) Este resultado confirmou que a pleiotropia está presente entre os perfis lipídicos, os fenótipos de pressão arterial, o índice de massa corporal, a diabetes e os hábitos tabágicos. (256)

Atualmente, nos estudos GWAS, há 58 SPNs com significado estatístico e, é provável que este número continue a aumentar. (87)

O SRG neste trabalho, é uma associação de SNPs, alguns relacionados com fatores de risco, outros não, que no seu conjunto reforçam as interações genótipo-fenótipo. Demostramos que um SRG, calculado com os 31 SNPs estudados, associa-se com significância à DC. Demostramos que 25% dos indivíduos, que apresentam os alelos de maior risco, situados no quarto quartil, têm 2,5 vezes mais risco de DC, quando comparados aos 25% do primeiro quartil. Este SRG, quando integrado com os FRT, proporciona uma ligeira melhoria da capacidade preditiva em relação ao modelo dos FRT. Contudo, mantém-se como preditor independente e com um OR de 1,8. Estes resultados destacam o potencial valor de incluir a informação genética nos modelos habituais, ponderando vir a melhorar a estratificação de risco individual.

Apesar das conhecidas limitações das métricas do NRI categórico e contínuo e do IDI, ainda em debate aberto na comunidade científica, fica bem expressa a capacidade do SRG de reclassificar doentes e controlos nos sentidos com maior impacto na prática clínica. O aumento de categoria no grupo dos doentes pode ampliar e reforçar as medidas de prevenção secundária, quer farmacológicas, quer ambientais e a descida de categoria nos controlos permite que, no contexto de prevenção primária, nos indivíduos de baixo risco, não sejam feitas intervenções para além da manutenção de um estilo de vida saudável.

V.4.4 COMPARAÇÃO DO VALOR PREDITIVO DO SCORE DE RISCO GENÉTICO COM NOVOS MARCADORES DE RISCO CARDIOVASCULAR

Existem novos marcadores de risco de CV que podem apresentar um papel significativo na predição de DC. Nesta população com DC, demonstrou-se na análise bivariada diferenças, quanto aos valores séricos dos leucócitos, da homocisteína, da PCR, da VOP e da Lp (a). Na Tabela 48 testamos a independência do SRG em relação aos novos marcadores, comparando o seu valor preditivo adicional. Os FRT foram incluídos na análise de regressão logística multivariada e fez-se um ajuste adicional pelos marcadores de risco CV associados com DC no estudo bivariado. Assim, apesar de nos doentes com DC haver diferenças quanto à leucocitose, Hcy, PCR, VOP e Lp (a) por análise bivariada, apenas o SRG com um OR de 1,72 (IC95%: 1,46-2,04) e a Lp(a) com um OR de 1,90 (IC95%: 1,6-2,26) estiveram associados de forma independente com a DC, quando comparados aos FRT.

Os níveis séricos da Lp(a) são determinados geneticamente, são relativamente estáveis na circulação, e não são influenciados pela dieta, exercício físico ou por fatores ambientais. (257,258) A Lp(a) demonstrou já, em vários estudos, estar associada com maior risco de CV e poder melhorar a estimativa do mesmo. (259). Num estudo publicado em 2014, Willeit *et al.* demonstraram que a Lp(a) melhorou a reclassificação de casos (17,1%) e controlos (22,5%) num grupo de risco intermédio, estimado pelo score de risco de Framingham. (260)

A principal limitação à sua utilização é a técnica bioquímica, utilizada na sua determinação sérica. A falta de homogeneidade, entre os diferentes testes diagnósticos e a ausência de testes independentes do peso molecular, tem tornado a sua utilidade limitada. (123)

A inclusão da Lp(a) nos modelos de estratificação de risco já foi proposta pela ESC para indivíduos considerados de alto risco CV ou com significativa história familiar de aterosclerose prematura. Até à data, ainda não existem recomendações para a sua utilização generalizada. (119-120) No presente estudo, a adição da Lp(a) ao modelo dos FRT e do SRG proporcionou ainda a melhoria da capacidade discriminativa do modelo na predição da DC.

A inclusão dos FRT (AUC=0,74) com o SRG (+0,02) e da Lp(a) (+0,01), permitiu um incremento da AUC da ROC para 0,77, revelando boa discriminação do modelo.(Figura 27)

V.4.5 SCORE DE RISCO GENÉTICO E HISTÓRIA FAMILIAR DE DOENÇA CORONÁRIA PRECOCE

A história familiar de DC precoce (HF de DC precoce) representa uma estimativa clínica do componente genético individual. É conhecido que o risco para DC aumenta, quer pelo número de progenitores com passado de DC, quer pela precocidade da mesma nos familiares. Sendo o SRG também um fiel representante do componente hereditário, é concebível que exista um benefício em associarmos o conhecimento de ambos para melhorar a nossa capacidade preditiva de DC.

Calculámos os AUCs – valores preditivos da capacidade discriminativa – do modelo usual dos FRT, da HF de DC precoce e do SRG. A curva mais baixa correspondeu à curva da HF de DC (AUC=0,56) com valores muito reduzidos de sensibilidade (23,8%) e com boa especificidade (87,4%), indicando que a discriminação do risco para DC pelo conhecimento da HF do indivíduo é reduzida. O SRG demonstrou, na nossa população, uma razoável capacidade discriminativa para o risco de DC (AUC 0,61) com 54% de sensibilidade e 62% de especificidade. A associação do SRG ao modelo basal do FRT permitiu melhorar a nossa capacidade discriminativa para o risco de DC, com um incremento positivo da AUC para 0,76, com melhor sensibilidade (70,6%) e especificidade (68,8%). Pelo contrário, a capacidade preditiva para o risco de DC não melhorou com a adição da HF ao modelo combinado do SRG e FRT (AUC=0,76, $p>0.05$) (Tabela 49 e Figura 28).

Também, as medidas de reclassificação e discriminação numéricas (NRI e IDI) são mais favoráveis à utilização do SRG, com NRI de 30% e com IDI de 2%. A incorporação do valor preditivo da HF ao modelo do SRG e FRT reduz a reclassificação (NRI 22,4%) e a discriminação (IDI 1,1%) (Tabela 50).

Embora, potencialmente, possa haver um erro grosseiro na informação oral transmitida pelo doente do passado de DC nos familiares, os nossos resultados estão em linha com outras publicações, surgidas em 2016 com resultados similares, em que a adição de SRG ao modelo dos FRT conduziu à melhoria da reclassificação e discriminação em relação à ocorrência de eventos CV, sem melhoria, após ajuste adicional da HF. (261-263)

V.4.5.1 SCORE DE RISCO GENÉTICO E DOENÇA CORONÁRIA NO JOVEM

Nos doentes com diagnóstico de DC em idade jovem (<45 anos), ainda sem muito tempo para os FRT atuarem, seria de esperar um modelo de aterosclerose acelerada, atribuído à predisposição genética, parecendo por isso um grupo atrativo para avaliação da *performance* do SRG.

Analisou-se o subgrupo de doentes, com idade inferior a 45 anos, com o diagnóstico de DC por Angiografia. Nesta sub-análise pretendemos verificar o poder preditor do SRG na amostra de doentes com DC, antes dos 45 anos, comparando-a com o nosso grupo controle (Tabelas 51 a 58).

Os nossos resultados demonstram que na população muito jovem (dos 30 aos 45 anos) o SRG tem um baixo desempenho. Quando se comparam os doentes com os controlos, na estimativa do risco para DC, o SRG tem mesmo uma performance negativa, quanto à capacidade de reclassificação de doentes com um NRI - 21,6%. Já na capacidade de reclassificação dos controlos, este apresenta um NRI de 47,6%.

É possível especular que, a DC do jovem, é uma identidade diferente, não detetada pelos genes em análise. Tada *et al.* (261) observaram um melhor poder do SRG na predição de eventos CV numa população jovem, dos 45 aos 73 anos de idade, com uma média de 57 anos e com um seguimento de 14 anos. Os indivíduos situados no quinto quintil do SRG apresentaram um risco 2,4 vezes maior do que os indivíduos no primeiro quintil do SRG. Essa diferença, entre predição estimada de risco para DC do primeiro para o quinto quintil do SRG, ultrapassa a diferença esperada de risco pela menor prevalência nesse grupo jovem dos FRT. Nesta população de 23,595 indivíduos sem DC e com baixa prevalência dos FRT ocorreram 2213 eventos. Na população do estudo o SRG analisado em quintis foi independente dos FRT, e com valor preditivo. Também Hughes *et al.* (264) e o *CARDIoGRAM* (76) reportaram dados semelhantes.

No nosso estudo devemos questionar as características encontradas no primeiro quartil. Doentes jovens, com uma incidência de FRT, semelhantes às amostras de doentes do quartil superior. O SRG não tem significado estatístico relevante quer nos doentes quer nos controlos deste quartil. Este é um estudo caso-controlo observacional. Comparando com a incidência da DC na população em qualquer grupo etário da amostra, os doentes estão sempre sobre representados, ainda mais nos grupos etários mais jovens. Com o peso confirmado dos FRT (AUC de 0,74) na DC, não nos surpreende que um doente em qualquer grupo etário tenha uma prevalência elevada de FRT. Pelo contrário, a DC significativa sem fatores de risco é que seria surpreendente, principalmente numa população jovem. Fica por explicar porque é que o SRG não apresenta discriminação significativa neste grupo etário, por oposição aos grupos etários mais velhos. Podemos especular que a DC no jovem poderá ter uma identidade genética diferente dos genes integrados no SRG do estudo.

É provável que, dependendo da nossa carga hereditária, do ambiente em que vivemos e das interações dos dois, a DC tenha início pouco depois do nascimento. É um processo lento, no qual as coronárias, para além de um possível processo próprio de doença, parecem ser, principalmente, um órgão alvo de várias doenças.

Dependendo do genótipo com que nascemos a DC manifesta-se habitualmente após a sexta década de vida. Compreende-se que, não se tratando de genética mendeliana, o valor preditivo de um SRG pode não ter valor preditivo num grupo etário em que o seu fenótipo não se manifesta. Pelo contrário, esse valor preditivo para o grupo etário, só se pode estabelecer a partir da idade em que o seu fenótipo se começa a manifestar. Este facto acarreta uma importante consideração. A não relação do SRG com os doentes jovens pode dever-se ao facto de as variantes genéticas incluídas estarem relacionadas com a DC habitual e mais tardia, e não com a sua forma jovem, mais rara e agressiva.

Pode ser antecipado que, em estudos futuros, investigando painéis alargados de SNPs em indivíduos de meia-idade e de risco intermédio, se possa adequar o conhecimento genético a decisões com impacto clínico. Pelo contrário, pode não ser compensador utilizar SRG em idosos cuja a idade e o efeito ponderado dos FRT ao longo da vida será o maior responsável pelo risco para DC. Faz sentido que, fatores genéticos que existem desde a nascença, sejam por isso mais importantes nos indivíduos mais jovens.

V.5 RECLASSIFICAÇÃO DO RISCO CARDIOVASCULAR COM O SRG

Na prática clínica, a estratificação do risco cardiovascular é de primeira importância. Desde há várias décadas que investigadores têm tentado criar um modelo preditor do risco de desenvolvimento de DC. Assim, foram criados vários *Scores* de Risco (SR).

Os SR mais utilizados na prática clínica para a estratificação de risco cardiovascular são: o calculador *Heart-Score* do score europeu para cálculo de risco da mortalidade CV; o calculador de risco DC *Framingham* para o risco para DC; e o calculador de risco CV *Framingham* para o risco CV. Destes, o calculador *Heart-Score* do Score Europeu (SE) adequa-se melhor à nossa população. Foi elaborado numa população europeia e é adaptado ao risco cardiovascular do país, onde se utiliza.

O score de risco *Framingham* teve por base uma população mais reduzida, de origem norte americana e cujos hábitos dietéticos e de influência ambiental se reportam às décadas de 60/70, o que pode constituir uma limitação à sua aplicação. Embora existam várias funções estimadas a partir do estudo inicial de Framingham, apenas se reportam neste estudo as duas funções de risco para DC e para Risco CV.

Estes scores integram a informação de variáveis clínico – laboratoriais e FRT para a DC. Contudo, a estratificação de risco CV mantém-se pouco precisa em vários contextos, como, por exemplo: nos grupos de indivíduos com poucos FRT e no subgrupo intermédio dos grupos de risco do *Heart-Score* ou de *Framingham*, nos quais a estratificação de risco está sujeita a erros de amostragem.

Numa tentativa de dar uma boa resposta à necessidade clínica, o perfil genético apresenta-se extremamente atrativo como medida de estratificação e preditiva do risco CV e do risco da DC. Embora o seu verdadeiro poder preditivo ainda seja pouco conhecido, o fato de ser individual e imutável ao longo do tempo torna-o estimulante.

Assim, avaliamos a contribuição do SRG na reclassificação do risco nos três subgrupos dos SR mais significativos na prática clínica, já referidos: o calculador *Heart-Score* para a mortalidade CV estimada pelo Score Europeu; o *Framingham DC* para a DC; e o *Framingham CV* para risco CV.

Demonstramos que, no global, todos os FRT foram preditores independentes no modelo basal ajustado para a idade e género. O tabagismo e a DMs foram os que condicionaram maior risco. Em qualquer das sub-análises, o valor basal AUC da curva ROC dos FRT foi alto, acima dos 0,7. Este elevado valor basal deve-se à esperada boa discriminação dos FRT. Tendo em conta este valor basal, a capacidade do SRG incrementar a curva torna-se desafiante. No conjunto da amostra, em todos os subgrupos de risco calculado, o SRG foi preditor independente, ajustado para os FRT em análise multivariada (Tabelas 59, 62 e 65).

Em todos os subgrupos de risco, o modelo combinado que incorpora o SRG produziu incrementos significativos do poder preditivo, mostrando que o componente genético pode aumentar a estimativa do risco calculado. O incremento global do valor da AUC em cada subgrupo, embora pareça modesto é significativo pelo teste de Delong. Nos subgrupos de risco intermédio a melhoria do poder preditivo variou entre 2 e 4%, o mesmo valor encontrado nos restantes subgrupos (Tabelas 60, 63 e 66).

No global, a incorporação da informação genética sob a forma do SRG no modelo preditor habitual mostrou a capacidade do SRG em reclassificar e discriminar indivíduos em relação ao evento (Mortalidade CV, DC e DCV) a 10 anos, estimados pelo Score Europeu, *Framingham DC* e *Framingham CV*. Nesta amostra, a reclassificação apresentou um NRI de 30%, com 2% de melhoria do declive da curva preditora do evento de morte cardiovascular a 10 anos. Em todos os subgrupos, o SRG permite melhorar a capacidade preditora do modelo basal de modo significativo, apresentando um NRI e IDI máximos nos grupos de elevado e de baixo

risco (Tabela 61, 64 e 67).

Assim, confirma-se que o SRG melhora a reclassificação de risco basal em todos os subgrupos, mas é precisamente nos subgrupos de risco intermédio, que não atinge um valor suficiente alto para aumentar o seu poder discriminativo. Por outro lado, os resultados do NRI e IDI no subgrupo de risco elevado, estimado pelo Score Europeu, parecem-nos demasiado otimistas e possivelmente inflacionados, pelo número reduzido (n=316) de doentes nesse grupo.

Na literatura, a capacidade de reclassificação dos SRG em relação aos modelos habituais dos FRT tem sido avaliada pela sua capacidade de reclassificação e tem sido expressa pelo NRI.

No presente estudo, observaram-se tendências favoráveis para ambas as medidas de melhoria da reclassificação, o NRI e o IDI após a inclusão do SRG no modelo basal. Os resultados, de NRI e IDI nesta amostra, referentes à reclassificação e incremento discriminativo com o SRG no risco de Mortalidade CV, de DC, e de doença CV estão em linha com as publicações existentes. Estas abrangeram maioritariamente populações europeias e norte-americanas e exploram o impacto da adição do SRG – compostos por vários números de variantes genéticas – aos FRT na estratificação do risco, na melhoria da capacidade preditiva e do índice de discriminação. (133,261,265)

No global, encontraram-se boas capacidades de reclassificação do SRG na amostra o que proporcionou uma melhoria média de reclassificação do risco de 30%. Porém, a melhoria da capacidade discriminativa expressa pelo IDI inferior a 3%, foi modesta.

É importante salientar que o cálculo do NRI foi feito de acordo com as categorias de risco preditas pelo *Score Europeu* e de *Framingham*, de modo a terem impacto clínico. Contudo, há limitações à sua interpretação. Por um lado, estes calculadores não devem ser aplicados em contexto de prevenção secundária (doentes), onde a prevalência de DC é muito maior e por isso com frequências e riscos subestimados; por outro lado o cálculo do NRI pode ser efetuado de forma contínua ou em categorias de risco, o que impõe outras questões importantes. No primeiro caso (NRI contínuo) o resultado expressa qualquer mudança de risco, mesmo se dentro do mesmo subgrupo de risco estimado, podendo por isso ser criticado por não se traduzir, no imediato, em mudança de atitudes terapêuticas. Estas iriam ser executadas, apenas se houvesse mudança de categoria de risco.

No segundo caso (NRI categórico) os resultados são altamente dependentes do número de categorias e dos limites impostos a essas categorias. No global, tendo em conta estas considerações, adotamos a estratégia de NRI contínuo, aplicado às classes de risco (NRI categórico) como referido por Pencina *et al.* (204).

Os resultados modestos, na reclassificação do risco estimado, indicam dois fatos incontornáveis: primeiro, a curva basal dos FRT, neste estudo entre 68-72%, representa uma boa capacidade discriminativa para a DC, sendo difícil melhorá-la; segundo, os genes isolados ou agregados em score não explicam, ainda, a maior parte do componente hereditário da DC.

Assim, com variantes genéticas apresentando frequência alélicas baixas entre 5% e 10% e risco relativo para DC e EAM entre 1,1 e 1,5% não se conseguirão produzir grandes incrementos na capacidade preditiva, se estas forem adicionadas aos modelos de FRT. (133,266)

Não estando descoberto o mistério da hereditariedade escondida, deposita-se atualmente muita esperança na análise gene-gene e gene-ambiente. Porém, é possível que estratégias adicionais, recorrendo quer a biomarcadores, modalidades de imagem ou a perfis genéticos, venham a ser utilizadas.

Apesar da reclassificação modesta, é notório que o doente de baixo risco CV com apenas 1 ou dois FRT, é frequentemente excluído de estratégias mais intensivas. Nestes subgrupos, a reclassificação do risco pelo SRG pode permitir o benefício de uma atitude mais pró-ativa, iniciar, por exemplo: terapêutica farmacológica com estatina, inclusão em programas de exercício, ou medidas ativas de evicção tabágica.

De qualquer modo, estes resultados confirmam a importante contribuição da genética para a DC, mesmo ajustando para a estratificação de FRT.

Acreditamos que o SRG, no futuro, pode melhorar estratégias preventivas desenhadas individualmente, provavelmente dirigidas a indivíduos com elevado risco genético, mesmo que este seja em pequeno número. (267)

O SRG já está consagrado como um preditor de risco independente, não só como valor preditivo para a ocorrência de DC como para a ocorrência de vários eventos CV. Surpreendente é também o seu valor como estratificador para prescrição de estatinas no contexto da prevenção primária. (100,233,268-271, 273,277)

Neste contexto, o NNT (*number needed to treat*), necessário para reduzir um primeiro evento em indivíduos com SRG elevado, é de apenas 13 comparado com os 37 indivíduos que tenham SRG baixo. (272-276) Já este ano, também se documentou que o benefício da terapêutica com estatina, de acordo com o SRG, leva a uma redução de risco absoluto diferencial, e é independente do valor de LDL sérico. Também se

documentou que para além da redução do risco relativo de 44% para DC nos doentes com elevado risco genético versus restantes, há uma redução do risco absoluto de DC de 3,6% (IC95%: 2,0-5,1) para 1,3% (IC 95%: 0,6-1,9) (Figura 36). (277)

Pensamos que estes dados vão influenciar a nossa perceção de risco CV, num futuro próximo. É possível que ao invés das atuais indicações da ESC, a utilização do SRG possa vir a ser uma realidade para identificar indivíduos que, independentemente do valor sérico do LDL, possam beneficiar da introdução de estatinas para evitar a eclosão de DC. (277)

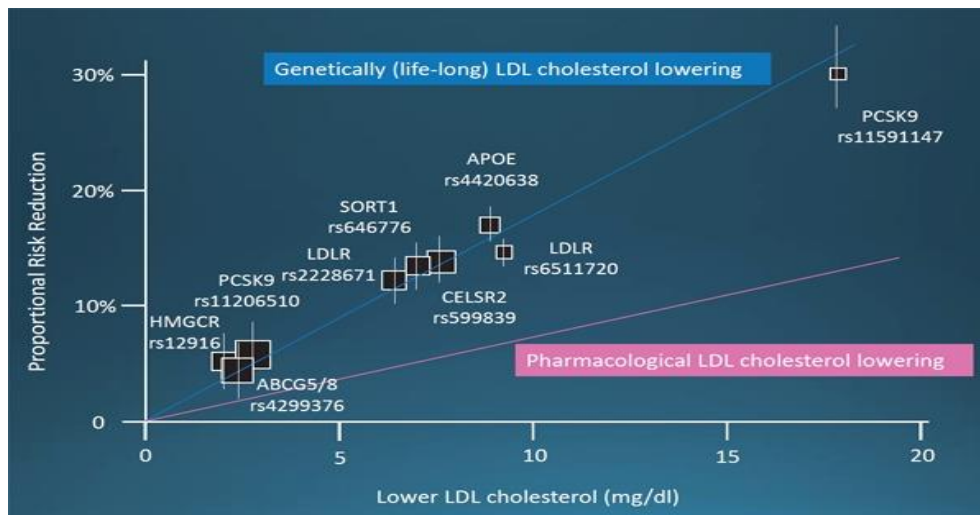


Figura 36: Relação entre a redução do Risco de DC (em %) e a redução do LDL. Adaptado com permissão de: Tada et al. 2017 (278). Linha azul indica a relação entre a redução de risco para DC e a redução do LDL guiada pelo risco genético. A cor de Rosa está representada a redução do risco de DC se a introdução de fármacos hipolemiantes for guiada pelos valores de LDL séricos e de acordo com as *Guidelines*.

V.6 VARIANTES GENÉTICAS E EVENTOS ADVERSOS, UM ESTUDO PROSPETIVO

Fez-se um estudo prospetivo da coorte de 1549 doentes com DC, com uma taxa de retenção de 98,9%, com um seguimento médio de 55,2 meses. O objetivo foi avaliar o impacto do componente genético na ocorrência de eventos adversos. Caracterizaram-se os grupos com e sem eventos CV, quanto aos dados clínicos do exame objetivo e bioquímico-laboratoriais. Determinaram-se as variantes genéticas associadas aos eventos adversos, mortalidade CV, EAM ou necessidade de revascularização. Não era um objetivo inicial do estudo, a análise isolada da mortalidade CV. Contudo, como a amostra apresentou robustez necessária para permitir inferir conclusões sobre a mortalidade CV, passamos a considerar dois eventos adversos, a mortalidade CV e um global com a mortalidade CV, EAM ou necessidade de revascularização.

Numa análise gene a gene fez-se uma análise de sobrevivência para a mortalidade CV de acordo com o genótipo para a variante genética ADAMTS7. Nos doentes com risco genético elevado (quarto quartil versus primeiro quartil e acima versus abaixo da mediana do SRG aditivo), foram feitas duas análises comparativas da sobrevivência livre de eventos CV. Foram identificados em modelos ajustados os preditores independentes da mortalidade CV, de um evento global de EAM, necessidade de revascularização e ou mortalidade CV. O seu poder preditor foi discriminado pelo C-estatística. O potencial do SRG aditivo para a reclassificação dos doentes na predição de eventos foi expresso pelo NRI e pelo IDI.

Na amostra de doentes coronários a prevalência da Mortalidade CV, EAM ou necessidade de revascularização foi de 6,9%, 20,7% e 23,3%. Tiveram pelo menos um evento, 30,1% dos doentes. Estes, são dados habituais nas *coortes* de doentes com DC documentada, nas quais a ocorrência de eventos CV é mais elevada do que na população em geral (Tabela 68).

O tamanho da amostra com 1 549 indivíduos, dividida em dois grupos: com eventos (n=466) e sem eventos (n=1083), é reduzida para encontrar diferenças imediatas no SGR. Contudo, existe uma tendência para o SRG ser mais elevado no subgrupo dos doentes que, no seguimento apresentaram morte CV e EAM, (p=0,07 e p=0,09) (Tabela 68).

Os doentes submetidos a revascularização cirúrgica apresentam um SRG aditivo mais elevado (SRG 25,3±3,7 versus 24,3±3,7, p=0,044), porém a interpretação conclusiva destes dados requeria uma amostra de maior dimensão. De igual modo, optámos por não fazer a análise isolada da ocorrência de AVC porque o número de eventos foi reduzido.

Ao comparar as características demográficas e clínicas dos doentes, de acordo com a ocorrência de eventos adversos, verificamos que aqueles que sofreram eventos têm um perfil mais desfavorável, quer genético, quer dos FRT. Este grupo, também tem uma idade média mais avançada (55,5 versus 53,1 anos, p=0,002) (Tabela 69).

No grupo com mortalidade CV documenta-se um maior número de indivíduos com SRG, acima da mediana (69,2% versus 58%, p=0,02). Do ponto de vista fenotípico, este grupo tem um perfil lipídico com valores médios mais elevados de colesterol total, Apo B e triglicéridos e mais baixos de HDL. Também tem maior prevalência de diabetes, sedentarismo, consumo de álcool, e alguns índices antropométricos de obesidade central em que o índice cintura-altura foi o mais elevado. Com significado estatístico, foram encontrados níveis mais elevados de homocisteína, PCR, leucócitos e plaquetas (Tabela 70).

Ainda no grupo de maior mortalidade CV foram encontradas diferenças significativas em relação à terapêutica farmacológica, a longo prazo. Tinham um menor grau de dupla antiagregação mantida (24,3% versus 43,3%, $p < 0,0001$) e menor controle da frequência cardíaca à data da última consulta de seguimento. Este subgrupo apresentava igualmente maior gravidade angiográfica com envolvimento de 3 vasos (47,7% versus 26,3%) (Tabela 69).

Realizamos dois tipos de análises para avaliação dos preditores independentes da mortalidade CV, nas quais foi possível incluir todas as possíveis variáveis implicadas na ocorrência do evento em análise. Uma baseada em modelos de regressão logística, com vários graus de ajuste para as co-variáveis potencialmente implicadas na ocorrência da morte CV. Outra com base em modelos de regressão de COX, que permite a ponderação do tempo de sobrevida, em relação ao poder de cada variável potencialmente implicada (Tabelas 72 e 73). (279)

No modelo de regressão logística, ajustamos os preditores de mortalidade CV base, para as variáveis dos FRT, do SRG e para três variáveis altamente determinantes dos eventos, nas amostras com DC: a gravidade angiográfica da DC, o tipo de evento inicial (Angina estável versus EAM), e a terapêutica com dupla antiagregação.

Com base neste modelo, o SRG foi preditor independente da mortalidade CV, com um HR de 1,57 (IC95%: 1,1-2,4) no modelo ajustado e um HR de 1,52 (IC95%: 1-2,3) no modelo multi-ajustado. No modelo multi-ajustado foram também significativas a DC de três vasos (HR=2,97; IC95%: 2-4,4), a dislipidemia (HR=3,47; IC 95%: 1,3-9,5) e a DM (HR=1,8; IC95%: 1,2-2,7). Nas outras variáveis, foram identificadas como preditores multivariados independentes: a frequência cardíaca (HR=1,03; IC95%: 1,03-1,05) e o consumo elevado de álcool (OR=1,77; IC95%: 1,1-2,8) (Tabela 72).

Com base no modelo de Regressão de Cox, o SRG mantém-se como preditor independente da mortalidade CV com um HR de 1,69 (IC95%: 1,1-2,5), e dos FRT, a DM aumentou o risco de morte CV com um HR de 1,83 vezes (IC95%: 1,3-2,7%) (Tabela 73).

Os FRT estão bem estabelecidos como preditores para a ocorrência de eventos cardiovasculares. (280-283) Na amostra, contudo, o tabagismo, a HTA e a dislipidemia não foram preditores independentes para a mortalidade CV. Esta ausência, pode ser parcialmente explicada pelo “controle” desses fenótipos após o diagnóstico da DC, com a cessação tabágica, um melhor controle da HTA e o uso de estatinas. A redução dos valores médios de LDL de 141mg/dl na primeira consulta para 105 mg/dl no seguimento, deve ser a resposta terapêutica à escalação ou introdução de estatinas.

Note-se que, ainda assim, os valores médios de LDL estão bem acima do máximo de 70mg/dl propostos para a prevenção secundária pela ESC. (119-120)

A PCR, também não foi preditora independente da mortalidade nesta amostra. Na literatura é aceite que a PCR é um potente preditor de eventos CV. Contudo, como a terapêutica com estatina também pode beneficiar os doentes com valores elevados de PCR (283), a terapêutica com hipolipidemiantes pode ter anulado parte do efeito deletério no risco de morte cardiovascular. Assim, embora ainda não existam terapêuticas farmacológicas dirigidas ao eixo inflamatório, são depositadas grandes esperanças na interação entre inflamação, perfil lipídico e aterosclerose na redução o risco CV residual. (284-285)

Nesta amostra a DM mantêm-se como preditor independente com qualquer modelo. Este facto sublinha a relevância da inclusão de variantes genéticas, associadas à diabetes no SRG.

Na análise gene a gene verificou-se um aumento da mortalidade CV associada à variante genética do ADAMTS7. Isto confirma-se na análise de regressão de COX com a respetiva curva de *Kaplan Meyer* em que, a variante do ADAMTS7 em homozigotia (AA) confere um risco de mortalidade CV mais elevado (HR=1,89 IC95%: 1-3,7), quando comparada com a do alelo G protetor (Tabela 71 e Figura 30). Estes resultados foram objeto de uma publicação feita pelo nosso grupo em 2016. (286)

Estudos genómicos, recentes, estabeleceram uma relação entre polimorfismos do gene ADAMTS7 e a aterosclerose, a doença arterial periférica, o risco cardiovascular e uma maior gravidade angiográfica. (196,197,286) Uma redução da função do gene ADAMTS 7 pelo alelo G pode traduzir-se numa redução dos eventos CV e melhor sobrevida. As proteínas codificadas por este gene inserem-se na família das “metzincinas”, uma família de metaloproteinases de zinco. Estas metaloproteinases são capazes de processar componentes da matriz extracelular. As alterações da matriz induzidas por estas “metzincinas” induzem mudanças no comportamento das células endoteliais e musculares lisas da parede vascular, promovendo a formação de neoíntima. Assim, como importantes reguladoras do *remodelling* tecidual, a nível vascular, promovem a aterosclerose (Figura 37).

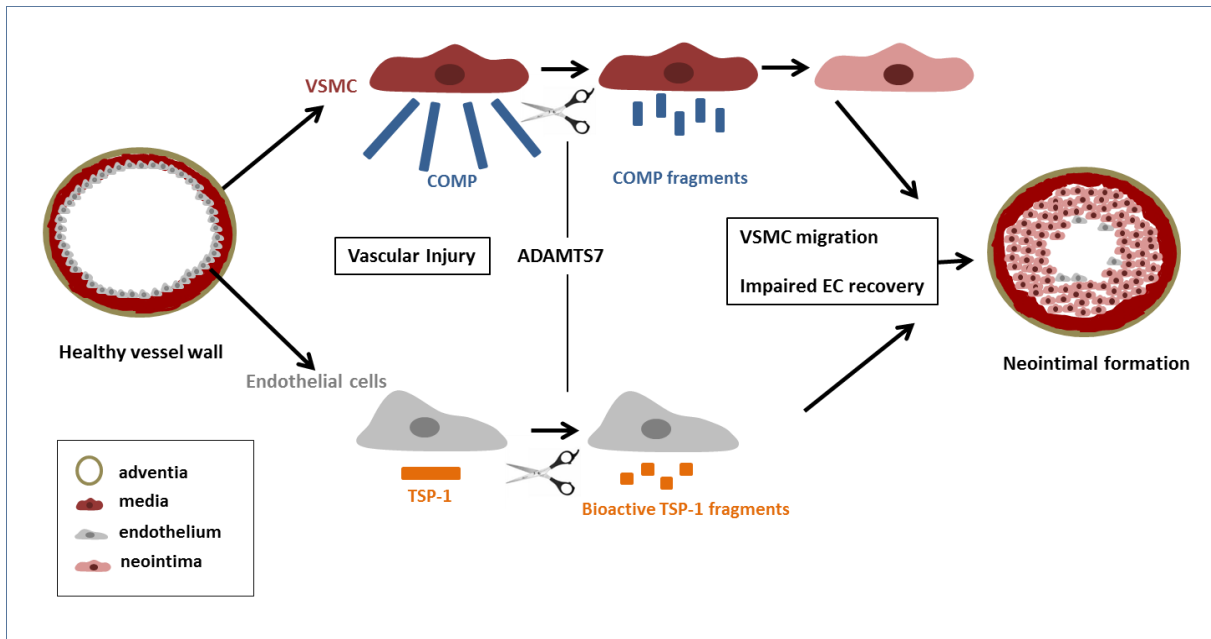


Figura 37: Papel da variante genética ADAMTS7 como mediador na etiopatogênese da aterosclerose e progressão da placa. Adaptado de “Arroyo & Andrés, 2015 (288)

Perante a lesão vascular, as células musculares lisas respondem com maior expressão de ADAMTS7 e maior processamento da proteína COMP (*α 7B1 integrin ligand cartilage oligomeric matrix protein*). Esta promove a migração e recrutamento das células musculares lisas – com a clivagem da trombospondina-1 (TSP-1) em fragmentos bioativos – que em última instância, reduzem a migração e proliferação das células endoteliais, comprometendo a reendotelização da parede vascular e passando a haver um predomínio de formação de neointima. Pensa-se que estes mecanismos poderão estar implicados na gênese da aterosclerose e progressão da placa. (195)

Em 2015 foi publicado num modelo animal com ratos um possível mecanismo etiopatogénico, para o gene ADAMT7, pela relação à trombospondina um, em que o gene foi inibidor da reendotelização. (289) No mesmo ano, também num modelo animal com ratos, o gene ADAMT7 foi localizado em células das lesões das artérias coronárias positivas para marcadores de tecido muscular liso. (290) No modelo animal, os ratos mutados para o gene ADAMT7 revelaram reduções significativas na formação de lesões na aorta, quando comparados com o grupo de ratos não mutados. (290)

Após a nossa publicação de 2016, já foi demonstrado numa população europeia, que a expressão de mRNA ADAMTS7 e de metaloproteínases degradadoras da matriz (MMP) eram mais elevadas em doentes com doença arterial periférica com a variante genética TT (SNP rs3825807), em comparação com controlos

saudáveis. (291) Já nos fumadores parece haver um aumento da expressão vascular do ADAMT7, podendo ser este o possível mecanismo de perda do efeito protetor do alelo T do ADAMTS7 neste subgrupo de doentes. (292)

Esta variante genética foi posteriormente avaliada como preditora de eventos cardiovasculares, em doentes com DC angiográfica numa publicação de 2017. (293) Apesar de nesse estudo com pequena dimensão de amostra, não se ter conseguido demonstrar associação com eventos adversos, consideramos que, a serem validados os nossos resultados, a variante ADAMT7 é muito atrativa para uma linha de investigação celular com implicação fisiopatológica na aterosclerose coronária.

Na nossa amostra, um SRG elevado está associado a maior mortalidade CV. Nas curvas de sobrevivência de *Kaplan Meyer* dos doentes, de acordo com a mediana do SRG e os quartis do SRG, fica bem expressa a separação das curvas que corresponde a maior mortalidade no grupo com um SRG superior à mediana. Um SRG elevado (>mediana) associou-se a um risco 1,7 vezes maior de mortalidade CV (HR=1,69; IC95%: 1,1-2,6, p=0,012) (*log-rank* 5,6 p=0.01). De igual modo, quando analisado em quartis, um indivíduo no quarto quartil do SRG tem um risco 1,9 vezes superior de sofrer morte CV no seguimento, quando comparado com um indivíduo no primeiro quartil. (HR=1,87; IC95%: 1-3,5, p=0,048) (*log-rank* p=0,045) (Figuras 31 e 32)

Assim, conclui-se que o SRG, imutável património do indivíduo, é preditor multivariado independente da mortalidade CV em qualquer cenário estatístico. Pelo contrário, à medida que melhora o controlo dos fenótipos associados com DC, diminui a significância dos FRT para a ocorrência de eventos no seguimento clínico.

Nas tabelas 74 a 78 estão descritos os resultados referentes à análise do evento adverso global. A avaliação dos eventos adversos abre outra questão pertinente e ainda em discussão científica, que reflete a atual ausência de conhecimento sobre o mecanismo de ação de muitas das variantes genéticas. Assim, quando analisamos eventos relacionados com EAM e revascularização, não sabemos se cada variante só predispõe para DC ou se também predispõe para eventos relacionados com a placa aterosclerótica, fundamentais na eclosão da aterosclerose subclínica. Por este motivo, realizámos uma análise univariada gene a gene para encontrar variantes genéticas associadas a eventos na placa aterosclerótica, que podem indicar novas linhas de investigação de mecanismos fisiopatológicos.

As variantes genéticas TCF21, KIF6, HNF4A e APOE associaram-se de modo significativo com EAM ou revascularização. Os genes 9p21 e TCF21, já amplamente replicados como de risco para DC, na nossa amostra, também foram discriminadores de nova angioplastia no seguimento (não expressos na tabela).

A variante genética TCF21 (OR=1,3; IC95%: 1,1-1,6) esteve associada ao risco para ocorrência do evento de EAM ou revascularização e a variante HNF4 foi protetora da ocorrência (OR=0,7; IC95%: 0,6-0,9).

As outras duas variantes genéticas (KIF6 e APOE), associadas a risco para EAM ou revascularização, foram apenas significativas no modelo de penetrância recessivo. Contudo, a maioria das publicações “aceita” a generalização do modelo de penetrância aditivo. Assim, estas duas variantes têm sido publicadas em populações caucasianas com respostas diferentes à terapêutica farmacológica com estatinas nos estudos de associação a eventos cardiovasculares. Isto salienta a importância de avaliar o modelo de hereditariedade, quando se pesquisa significância estatística, entre variantes genéticas.

A variante genética KIF6, presente em 30% das populações caucasianas, foi associada com eventos coronários em vários estudos observacionais e em estudos randomizados controlados por placebo. (132,294-296). Contudo, o seu papel não foi replicado noutros estudos sendo ainda controversa a sua associação definitiva com eventos coronários. (297) As diferenças encontradas parecem devidas a diferenças populacionais. Na população caucasiana é mais aceite o envolvimento desta variante na ocorrência de eventos. (135)

A variante APOE tem sido associada ao maior risco de eventos coronários por induzir alterações conhecidas no painel lipídico. Estas alterações são, maioritariamente, a elevação da LDL e dos triglicéridos e a redução das HDL. Num estudo prospetivo, que incluiu 10 035 indivíduos do sexo masculino e 12 134 do sexo feminino, os portadores do alelo E2 tiveram um risco de ocorrência de eventos 12% menor que os homozigóticos para o alelo E3, e os portadores do alelo E4 tiveram um risco 9% maior que os homozigóticos para o alelo E3. (298) No entanto, após ajuste para a relação LDL/HDL, essa associação deixou de ser significativa. (298)

No desenho do estudo, o risco da APOE foi ajustado para vários fatores de risco CV, incluindo o sedentarismo, o consumo de álcool e o IMC, fazendo com que estas variáveis (que podem confundir) expliquem parte dessa atenuação. De qualquer modo o risco manteve-se no subgrupo de doentes portadores do alelo E4. Esta constatação deixa em aberto a discussão sobre o seu real valor.

Avaliamos as características basais da amostra para análise do evento adverso global – mortalidade CV, EM ou necessidade de revascularização, mas não existiram diferenças significativas quanto à idade e ao género. No conjunto, voltámos a obter um perfil mais desfavorável nos casos que tiveram pelo menos um evento. Com efeito, os doentes com eventos eram mais sedentários, mais diabéticos, consumiam mais álcool,

apresentavam índice cintura-altura mais elevado e a pressão arterial diastólica mais baixa. Tal como previamente citado, em relação à terapêutica farmacológica a longo prazo, o grupo dos doentes com eventos era caracterizado por um menor grau de dupla antiagregação mantida (29% versus 46%, $p < 0.0001$) e por um menor controlo da frequência cardíaca (70 versus 68 bpm) à data da última consulta de seguimento. Este grupo, apresentou ainda maior gravidade angiográfica (DC3 vasos 37,1% versus 24,9%, $p = 0,007$). Também o perfil lipídico encontrado era mais adverso com menores valores de HDL e valores mais elevados de Apo (B) e Lp (a). De igual modo, os doentes com eventos no seguimento apresentavam níveis mais elevados de PCR e de homocisteína sérica (Tabelas 74 e 75).

Por análise de regressão logística multivariada ajustada para a idade e género, FRT, e para o SRG, apenas a DM confere risco multivariado com OR 1,4 (IC95%: 1,1-1,8). Contudo, a DM perde significado quando é realizado o ajuste para as variáveis significativas, previamente encontradas. Neste contexto, as variantes genéticas previamente descritas, e o SRG, não atingiram significância estatística (Tabela 76).

Assim, no modelo multi-ajustado de regressão logística, foram preditores para o evento global o sedentarismo, a PCR, a DC de três vasos, a frequência cardíaca, o álcool, a PAD (protetora) e a dupla anti-agregação (protetora) (Tabela 77).

No modelo multivariado, com os preditores ajustados para o evento global, confirma que as variantes genéticas TCF21 e HNF4 são preditores multivariados independentes para a ocorrência de EAM ou revascularização. A variante genética TCF21 conferiu um aumento do risco de 1,55 vezes por um alelo de risco e de 1,73 vezes por dois alelos. O gene HNF4, para dois alelos protetores, conferiu uma redução do risco com um OR de 0,26 (IC95%: 0,07-1,06) para a ocorrência do evento (Tabela 78).

De igual modo, a PCR elevada, o consumo de álcool e o sedentarismo foram preditores multivariados para a ocorrência do evento.

Assim, depreende-se que o risco residual, conferido pela elevação da Apo B, representa a ausência de controlo eficaz do fenótipo da dislipidemia. O risco conferido pelo sedentarismo e pelo consumo de álcool salienta a importância dos programas de exercício físico e de evicção alcoólica no contexto da prevenção secundária.

Nesta análise, depreende-se novamente que ficam quatro preditores finais do evento em análise: dois fatores imutáveis, o património genético e a gravidade angiográfica (a variante TCF21 e a DC de 3 três vasos) e dois modificáveis, ainda com pouca eficácia da intervenção em prevenção secundária, o sedentarismo e o consumo de álcool.

Em linha com o previamente escrito, para além do património genético e da gravidade angiográfica (DC3 vasos HR=1,3; IC95%: 1,1-1,7), foram preditores na regressão de Cox ajustada, variantes ainda pouco trabalhadas no contexto da prevenção secundária. Assim, a PCR com HR de 1,35 (IC95%: 1,1-1,7), o sedentarismo com HR de 1,3 (IC95%: 1,0-1,6) e o alcoolismo com HR de 1,5 (IC95%: 1,2-2), na nossa amostra, também se associaram à ocorrência do evento em análise. Como vimos, a PCR tem um papel de marcador do eixo inflamatório na aterosclerose. Esta tem sido apontada como potencial alvo terapêutico para redução do risco CV residual.

Na avaliação do risco para eventos adversos, conferido por comparação entre um SRG ≤ 21 (n=347) e um SRG > 22 (n= 1202), verificamos que o grupo com SRG ≥ 22 tem mais 1,96 probabilidade de mortalidade CV e 1,34 probabilidade de ocorrência do evento global. Ajustando para a idade, género, FRT, DC 3 vasos e Dupla AA temos um HR 2,03 para a mortalidade CV (IC95%: 1,16-3,56) e um HR de 1,35 (IC95%: 1,04-1,75) para o evento global. Estes resultados põem em evidência a importância do componente genético na DC. (tabela 79)

Um SRG, independente da idade e do género, poder ser uma arma importante na estratificação dos doentes para a ocorrência de eventos (Tabela 79).

Analisamos a capacidade de reclassificação e discriminativa do SRG com os modelos habituais dos FRT com recurso ao *C-Index*, NRI e IDI para as duas variáveis estudadas.

A *area under the curve (C-index)* para as variáveis estudadas que incluiu, a dislipidemia, o tabagismo, a diabetes e a hipertensão, foi de 0,62, demonstrando uma discreta capacidade discriminativa entre os doentes que tiveram, ou não, os eventos em estudo. Para ter significância estatística o C-index deve possuir um valor acima de 0,5. Assim, a incorporação do SRG num modelo basal multi-ajustado, não produziu incrementos significativos da capacidade preditiva do modelo (Tabela 80).

Da análise global dos NRI e IDI, depreende-se que a inclusão do SRG no modelo, apesar de incrementar as curvas basais referentes aos FRT, aumento da discriminação, não produz NRIs muito robustos na reclassificação.

Este aumento, de discriminação com reduzida reclassificação, foi também comunicado por outros grupos de investigação. (97,100,101) Em alguns desses estudos, com associação positiva do SRG com os eventos cardiovasculares adversos no seguimento, o NRI não foi sequer positivo, o que não constitui motivo suficiente para a sua não utilização. (97,267)

Vários estudos abordam a genética e a ocorrência de eventos cardiovasculares adversos tanto em populações europeias como americanas. (100,133, 299-301) Estes diferem em dois aspectos: primeiro quanto ao tipo de informação genética incluída, desde gene a gene (220,302,303) a scores genéticos ponderados ou não ponderados (305) desde 13 a várias centenas de SNPs incorporados e; segundo quanto aos eventos analisados, sendo a forma mais frequente referente à incidência de qualquer evento cardiovascular. (302-304) Recentemente, surgiram análises para a mortalidade cardiovascular (301) e, em menor número de estudos, foram analisados os eventos isolados de EAM (301), ou de revascularização. (293)

É irrefutável que o GRS está associado com o pior prognóstico e com eventos CV adversos em múltiplas coortes, em diferentes contextos, incluindo na prevenção primária e secundária. Estudos prospectivos, a longo prazo em prevenção primária, validaram o SRG como marcador para a ocorrência de eventos adversos em todas as categorias de risco. (253,300)

Da análise individual, G-G, para a incidência de eventos, sabe-se que tendo em conta as diferenças do património genético em diferentes populações, as meta-análises já publicadas para o 9p21 são pouco esclarecedoras, pois a associação com a incidência de eventos desta variante produziu um modesto HR de 1,01 (IC95%: 0,97-1,06) por alelo de risco. Esta constatação foi atribuída à heterogeneidade genética das populações incluídas nesta meta-análise. (220)

Neste tipo de análise G-G, consideramos importante salientar que o gene TCF21 foi consistente com a ocorrência dos eventos analisados, sendo promissora a sua utilização como alvo terapêutico. Este gene foi recentemente associado à ocorrência de eventos numa população europeia similar à nossa. (293) Esta variante genética, que foi consistente ao longo da nossa análise, parece implicada na regulação da diferenciação dos precursores de células musculares lisas, que migram para as lesões vasculares, podendo contribuir para a constituição fibrosa das placas ateroscleróticas. Este novo alvo, na parede vascular, representa um modelo de mecanismo fisiopatológico atrativo para reduzir o risco cardiovascular.

FORÇAS E LIMITAÇÕES

Uma mais-valia deste trabalho foi a sua realização numa ilha com uma população social e culturalmente homogénea, com estilos de vida e hábitos alimentares semelhantes e com interações com o meio ambiente similares. Dado que os movimentos populacionais são limitados, este facto permitiu-nos fazer o seu estudo com uma taxa de retenção de 98,9%. O facto de viverem numa ilha, o passado genético uniforme, confere uma representatividade e estabilidade genética à amostra. Outra vantagem, é a nossa população já se encontrar geneticamente bem caracterizada e ser representativa das populações do sul da Europa: Portugal, Espanha, Itália e Grécia. Este contexto, de estabilidade genética, é uma mais-valia nos estudos genéticos das Doenças Complexas como a DC, porque permite a sua melhor caracterização e, ao tornar as interações genes ambiente mais estáveis e fortes, torna-as melhores objetos de estudo.

O facto de haver um só hospital e termos acesso aos cadernos eleitorais facilitou o recrutamento de doentes sem viés, e uma seleção randomizada de controlos emparelhados para a idade e para o género, que fez a nossa amostra bastante robusta e consistente da sua representatividade.

Relativamente ao número e tipo de variantes genéticas usadas, a nível internacional há estudos com mais variantes, aparentemente, ainda sem grandes ganhos preditivos. A nível nacional, este é o estudo com maior número de variantes genéticas utilizadas e, que seja do nosso conhecimento, o único na área da DC. Na altura em que começámos, com o nível de conhecimento que havia e com as restrições orçamentais existentes, acabamos por escolher 33 SNPs suspeitos, dos quais acabamos por incluir 31 no nosso SRG. Olhando retrospectivamente, entre uma infinidade de boas opções que podíamos ter feito, não fizemos uma má opção.

Assim, no decorrer do presente estudo fomos os primeiros a confirmar a maior ou menor a prevalência de alguns genes relevantes na amostra e a identificar os alelos de risco e os alelos protetores característicos da nossa população.

Outro ponto forte do presente trabalho corresponde ao estudo prospetivo da associação do SRG e dos eventos cardiovasculares no subgrupo dos doentes. Na verdade, a grande maioria dos eventos cardiovasculares ocorre em doentes já com o diagnóstico de Doença Coronária e para este subgrupo de doentes não existe ainda uma estratificação do risco residual adequada. A maioria dos estudos referentes a SRGs têm-se centrado na associação de risco, em populações caso-controlo, ou seja, no estudo do primeiro evento cardiovascular. A confirmar os nossos achados está o facto da informação genética poder ser fundamental para o prognóstico individual do doente que, apesar da terapêutica médica otimizada, pode ter pior prognóstico se tiver um SRG individual mais elevado.

Nesta era de elevada análise computacional de grande volume de dados, genéticos e outros, consideramos uma mais-valia a aplicação no nosso estudo de métodos de *machine learning* para escrutinar novas modalidades de risco cardiovascular. O uso atual dos modelos de regressão logística multivariada para testar incrementos de novos biomarcadores um por um, necessita de novos impulsos que virão da introdução da inteligência artificial na biologia celular.

Na nossa população, a utilização da métrica do NRI categórico e contínuo melhorou a capacidade preditiva do modelo com o SRG, em cada categoria de risco. Contudo, dadas as atuais críticas em debate, referentes às suas limitações, aguardam-se novas medidas métricas, com melhor capacidade para traduzir o poder do SRG na predição do risco de DC e de eventos cardiovasculares.

Qualquer estudo deste tipo gostaria de ter sempre um maior grupo de variantes genéticas em estudo. Contudo, mais do que aumentar o número de variantes genéticas num estudo, pensamos que estas devem ser testadas na população e só após a determinação das frequências *in loco*, devem ser incluídas num SRG. De qualquer modo, pensamos que o aumento do poder discriminativo se vai dever ao estudo das interações G-G e G-A, e não ao aumento do número de variantes de baixa frequência e de baixo valor preditor individual.

O nosso estudo, embora incluindo 1566 doentes e 1322 controlos, pelo número da população ficou limitado na possibilidade de estudar subgrupos de interesse.

CONCLUSÕES

CONCLUSÕES e PERSPETIVAS FUTURAS

Nesta amostra da população portuguesa existem diferenças genéticas significativas, entre a população com doença coronária e a população controlo. As variantes genéticas que influenciam o aparecimento da doença coronária nesta amostra foram 10, a saber: LPA (OR=2,2), APOE (OR=1,26), ADIPOQ (OR=1,21), FTO (OR=1,19), PON 55 (OR=1,35), CDKN2B (OR=1,36), *locus* 9p21 (OR=1,32), TCF721 (OR=1,54), ZC3 (OR=1,29) e PHACTR1 (OE=1,1) por análise bivariada. Por análise multivariada foram as seguintes seis: TCF21, LPA, CDKN2B, PON 55, APOE e ZC3, que se mantiveram como preditores independentes da doença coronária após o controlo com os FRT.

Com o presente estudo validamos, pela primeira vez, numa população portuguesa a associação da variante do *locus* 9p21 por análise uni e multivariada com a DC.

Sendo a doença coronária uma doença multifatorial, avaliamos as interações gene-gene e gene-ambiente para o aparecimento da DC. As quatro interações G-G com significado estatístico encontradas foram a ADAMTS7 GG–MTHFR677 CT; a CDKN2BAG–PON192RR; a KIF6 GG–ATIRCC e a PON192QR–*locus* 9p21 GG. Das interações G-A encontradas duas tiveram significado: a DM e o TCF7L2 e a dislipidemia e o ZNF259.

A interação epistástica com *data mining* foi elevada na interação G-G das variantes APOE–TCF21, do KIF6–ECA e AGT235 –MTHFR1298. O modelo APOE–TCF21 foi o melhor preditor de DC com OR 1,5. No grupo da interação G-A, o melhor modelo no eixo celular foi o do gene ZC3HC1 com Tabagismo, HTA e DM com uma acuidade de 0,67 e um OR de 4,3. No eixo da DM a variante TAS2R50, a variante TCF7L2 e a DM tiveram uma acuidade de 0,613 e um OR 2,98. No eixo oxidativo, o modelo composto pelas variantes genéticas MTHFD1L; PON1; MTHFR677; PON55 produziu uma acuidade ponderada de 0,568 e um OR de 1,73. No eixo lipídico o PSRC1 e dislipidemia mostraram maior interação com uma acuidade ponderada de 0,6 e um OR de 3,92.

Na constituição do SRG, estudado o equilíbrio Hardy-Weinberg e o desequilíbrio de *linkage* na população, tivemos que excluir dois SNPs ficando com 31 variantes. Optámos por um score de risco genético baseado no método multiplicativo, produto dos ORs dos génotipos individuais, que representa a melhor forma de quantificar o risco do componente genético individual. O SRG mostrou poder de associação com a doença coronária. O risco de DC aumenta de acordo com a elevação do número de variantes genéticas e foi significativamente mais elevado nos doentes que nos controlos. O SRG associa-se ao aumento de 2,56 vezes do risco de DC para os indivíduos no quarto quartil, com SRG mais elevado, quando comparados aos do quartil mais baixo. Este SRG foi independente dos FRT, de outros marcadores de risco cardiovascular e da HF de DC precoce.

Embora no grupo etário abaixo dos 45 anos o SRG tenha também estado associado com significância à DC, o SRG apresenta um poder preditor moderado com sensibilidade e especificidade mais baixas e com um NRI negativo. Isto leva-nos a especular que os genes do SRG não serão bons marcadores da DC precoce.

Eventualmente, este tipo de doença mais agressiva e fortemente associada aos FRT poderá ser geneticamente modelada por genes, ainda não identificados. Assim, os genes do SRG parecem mais associados à DC habitual na nossa população.

Contudo, na população jovem, a análise G-G identificou as variantes TCF21, ZC3HC1, ZNF259 e ECA com significado estatístico para DC. A variante ZNF259 do eixo lipídico e a variante ECA do eixo da HTA, só tiveram significado neste grupo. Estas, por estarem associadas às variantes do eixo celular, poderão ter algum significado que a nossa amostra, por ser pequena, não pode identificar.

Os FRT com uma AUC de 0,74 apresentam um valor preditor para a DC elevado. O valor preditivo adicional do SRG com os FRT era desconhecido.

No nosso estudo encontramos uma capacidade preditiva adicional do SRG, quando associada aos métodos de estratificação cardiovascular usuais. Este SRG adicionado aos FRT produziu uma AUC de 0,76, e aumentou a capacidade preditiva em todos os subgrupos do score europeu, do Framingham DC e do Framingham CV. Destes destacam-se o aumento da AUC nos subgrupos de risco elevado do score europeu e do Framingham DC de 0,67 para 0,73 e de 0,68 para 0,72 respetivamente. Nos subgrupos de risco intermédio e baixo do Framingham CV, são de 0,67 para 0,71 e de 0,70 para 0,74 respetivamente.

Na prática clínica, qualquer novo modelo deve trazer uma melhoria na acurácia e capacidade de classificação dos doentes. Na nossa população a inclusão do SRG apresentou um NRI de 30,8% e um IDI de 2% para o score europeu e um NRI de 31,7% e um IDI de 2,3% para os Framingham DC e CV. A inclusão do componente genético no modelo preditivo, avaliado pelos scores convencionais, poderá melhorar a estratificação de risco individual, proporcionando uma adequada reclassificação de doentes numa categoria de risco superior e dos controlos numa categoria de risco inferior. Esta reclassificação é de interesse para a prática clínica.

Estes resultados podem ainda ser mais expressivos em indivíduos sem fatores de risco para doença cardiovascular onde o score pode ainda ter mais impacto clínico. É possível que, nessa população, o score de risco genético resulte numa melhor predição, comparada com os nossos achados.

Nos doentes, o SRG foi preditor independente da mortalidade CV para qualquer grau de ajuste com um Hazard ratio de 1,52 (IC 95% 1-2,3) no modelo multi-ajustado para potenciais co-variáveis.

Após o diagnóstico de DC, nos doentes, duas variantes genéticas influenciaram as complicações: a TCF21 e a HNF4A que estiveram associadas à ocorrência de EAM ou revascularização ou mortalidade CV num modelo ajustado para co-variáveis de interesse. Estas variantes do eixo celular podem estar intimamente relacionadas com a fisiopatologia, ainda não descoberta da placa instável e fornecerem bases para o estudo da parede vascular. De realçar, de igual modo, o impacto da variante genética do ADAMTS7 na mortalidade CV destes doentes, evidenciado pelo nosso estudo.

Os resultados, obtidos neste estudo com significância estatística, provam o potencial de incluir informação genética nos modelos habituais de estratificação cardiovascular.

No global, estes dados devem influenciar a nossa percepção de risco cardiovascular num futuro próximo. Possivelmente, a utilização do SRG pode vir a ser uma realidade em dois contextos: na prevenção primária ao identificar indivíduos com elevado risco genético e, que independentemente do valor sérico do LDL, podem beneficiar da introdução de estatinas, de forma a reduzir a eclosão de DC; e na prevenção secundária para melhorar estratégias dirigidas a doentes com elevado risco genético, independentemente da categoria de risco basal dada pelos FRT, mesmo que seja num menor número de doentes do subgrupo.

Pensamos que a nossa aposta inovadora no estudo da DC na interação G-G e G-A com modelos do MDR foi correta. Nesta amostra da população portuguesa encontraram-se importantes interações G-G e G-A, cujos resultados podem apontar novas direções para mecanismos fisiopatológicos na aterosclerose coronária. Para ter sucesso, as novas descobertas na genética cardiovascular devem basear-se, e não ignorar a complexidade da relação entre o genótipo e o fenótipo, resultante da interação G-G e G-A.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Townsend N, Wilson L, Bhatnagar P, Wickramasinghe K, Rayner M, Nichols M. Cardiovascular disease in Europe: epidemiological update 2016. *Eur Heart J*. 2016;37(42):3232-45.
2. 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.
3. Forouzanfar MH, Liu P, Roth GA, Ng M, Biryukov S, Marczak L, et al. Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm hg, 1990-2015. *JAMA*. 2017;317(2):165-82.
4. Prospective Studies Collaboration, Lewington S, Whitlock G, Clarke R, Sherliker P, Emberson J, et al. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Prospective Studies Collaboration*. *Lancet*. 2007 Dec 1; 370(9602):1829-39.
5. Staessen JA, Gasowski J, Wang JG, Thijs L, Den Hond E, Boissel JP, et al. Risks of untreated and treated isolated systolic hypertension in the elderly: meta-analysis of outcome trials. *Lancet*. 2000;355(9207):865-72.
6. Turnbull F, Neal B, Ninomiya T, Algert C, Arima H, Barzi F, et al. Effects of different regimens to lower blood pressure on major cardiovascular events in older and younger adults: meta-analysis of randomised trials. *BMJ*. 2008;336(7653):1121-3.
7. Macedo ME, Lima MJ, Silva AO, Alcantara P, Ramalhinho V, Carmona J. Prevalence, awareness, treatment and control of hypertension in Portugal: the PAP study. *J Hypertens*. 2005;23(9):1661-6.
8. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. *Lancet*. 2016;387(10027):1513-30.
9. Perk J, De Backer G, Gohlke H, Graham I, Reiner Z, Verschuren M, et al. European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). *Eur Heart J*. 2012;33(13):1635-701.
10. Haffner SM, Lehto S, Rönnemaa T, Pyörälä K, Laakso M. Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. *N Engl J Med*. 1998 Jul 23;339 (4):229-34.
11. Huxley R, Barzi F, Woodward M. Excess risk of fatal coronary heart disease associated with diabetes in men and women: meta-analysis of 37 prospective cohort studies. *BMJ*. 2006;332(7533):73-8.
12. Lee C, Joseph L, Colosimo A, Dasgupta K. Mortality in diabetes compared with previous cardiovascular disease: a gender-specific meta-analysis. *Diabetes Metab*. 2012;38(5):420-7.
13. Hu G, Jousilahti P, Qiao Q, Katoh S, Tuomilehto J. Sex differences in cardiovascular and total

- mortality among diabetic and non-diabetic individuals with or without history of myocardial infarction. *Diabetologia*. 2005;48(5):856-61.
14. Farzadfar F, Finucane MM, Danaei G, Pelizzari PM, Cowan MJ, Paciorek CJ, et al. National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3.0 million participants. *Lancet*. 2011;377(9765):578-86.
 15. Nordestgaard BG, Langsted A. Lipoprotein (a) as a cause of cardiovascular disease: insights from epidemiology, genetics, and biology. *J Lipid Res*. 2016;57(11):1953-75.
 16. Erqou S, Kaptoge S, Perry PL, Di Angelantonio E, Thompson A, White IR, et al. Lipoprotein(a) concentration and the risk of coronary heart disease, stroke, and nonvascular mortality. *JAMA*. 2009;302(4):412-23.
 17. Prospective Studies Collaboration. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet*. 2007;370(9602):1829-39.
 18. Baigent C, Blackwell L, Emberson J, Holland LE, Reith C, Bhalra N, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet*. 2010;376(9753):1670-81.
 19. Cannon CP, Blazing MA, Giugliano RP, McCagg A, White JA, Theroux P, et al. Ezetimibe Added to Statin Therapy after Acute Coronary Syndromes. *N Engl J Med*. 2015;372(25):2387-97.
 20. NCD Risk Factor Collaboration (NCD-RisC). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet*. 2016;387(10026):1377-96.
 21. Ashwell M, Gunn P, Gibson S. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. *Obes Rev*. 2012;13(3):275-86.
 22. Seidell JC. Waist circumference and waist/hip ratio in relation to all-cause mortality, cancer and sleep apnea. *Eur J Clin Nutr*. 2010;64(1):35-41.
 23. Lee CM, Huxley RR, Wildman RP, Woodward M. Indices of abdominal obesity are better discriminators of cardiovascular risk factors than BMI: a meta-analysis. *J Clin Epidemiol*. 2008;61(7):646-53.
 24. Wormser D, Kaptoge S, Di Angelantonio E, Wood AM, Pennells L, Thompson A, et al. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet*. 2011;377(9771):1085-95.
 25. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3(11):e442.
 26. GBD 2015 Tobacco Collaborators. Smoking prevalence and attributable disease burden in 195

- countries and territories, 1990-2015: a systematic analysis from the Global Burden of Disease Study 2015. *Lancet*. 2017;389(10082):1885-906.
27. World Health Organization. WHO report on the global tobacco epidemic, 2015.
28. Csordas A, Bernhard D. The biology behind the atherothrombotic effects of cigarette smoke. *Nat Rev Cardiol*. 2013;10(4):219-30.
29. Messner B, Bernhard D. Smoking and cardiovascular disease: mechanisms of endothelial dysfunction and early atherogenesis. *Arterioscler Thromb Vasc Biol*. 2014;34(3):509-15.
30. Ockene IS, Miller NH. Cigarette smoking, cardiovascular disease, and stroke: a statement for healthcare professionals from the American Heart Association. American Heart Association Task Force on Risk Reduction. *Circulation*. 1997;96(9):3243-7.
31. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ*. 2004;328(7455):1519.
32. Tverdal A, Thelle D, Stensvold I, Leren P, Bjartveit K. Mortality in relation to smoking history: 13 years' follow-up of 68,000 Norwegian men and women 35-49 years. *J Clin Epidemiol*. 1993;46(5):475-87.
33. Arem H, Moore SC, Patel A, Hartge P, Berrington de Gonzalez A, Visvanathan K, et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. *JAMA internal medicine*. 2015;175(6):959-67.
34. Ding D, Lawson KD, Kolbe-Alexander TL, Finkelstein EA, Katzmarzyk PT, van Mechelen W, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. *Lancet*. 2016;388(10051):1311-24.
35. Owen N, Leslie E, Salmon J, Fotheringham MJ. Environmental determinants of physical activity and sedentary behavior. *Exerc Sport Sci Rev*. 2000;28(4):153-8.
36. Biswas A, Oh PI, Faulkner GE, Bajaj RR, Silver MA, Mitchell MS, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Ann Intern Med*. 2015;162(2):123-32.
37. Wilmot EG, Edwardson CL, Achana FA, Davies MJ, Gorely T, Gray LJ, et al. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. *Diabetologia*. 2012;55(11):2895-905.
38. Ford ES, Caspersen CJ. Sedentary behaviour and cardiovascular disease: a review of prospective studies. *Int J Epidemiol*. 2012;41(5):1338-53.
39. Wijndaele K, Brage S, Besson H, Khaw KT, Sharp SJ, Luben R, et al. Television viewing and incident cardiovascular disease: prospective associations and mediation analysis in the EPIC Norfolk Study. *PLoS One*. 2011;6(5):e20058.
40. Kaptoge S, Di Angelantonio E, Lowe G, Pepys MB, Thompson SG, Collins R, et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-

analysis. *Lancet*. 2010;375(9709):132-40.

41. Watanabe H, Ohtsuka S, Kakihana M, Sugishita Y. Coronary circulation in dogs with an experimental decrease in aortic compliance. *J Am Coll Cardiol*. 1993;21(6):1497-506.
42. Van Bortel LM, Laurent S, Boutouyrie P, Chowienczyk P, Cruickshank JK, De Backer T, et al. Expert consensus document on the measurement of aortic stiffness in daily practice using carotid-femoral pulse wave velocity. *J Hypertens*. 2012;30(3):445-8.
43. Townsend RR, Wilkinson IB, Schiffrin EL, Avolio AP, Chirinos JA, Cockcroft JR, et al. Recommendations for Improving and Standardizing Vascular Research on Arterial Stiffness: A Scientific Statement From the American Heart Association. *Hypertension*. 2015;66(3):698-722.
44. Vlachopoulos C, Aznaouridis K, Stefanadis C. Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-analysis. *J Am Coll Cardiol*. 2010;55(13):1318-27.
45. World Health Organization. WHO Global Status Report on Alcohol and Health. Geneva, 2014.
46. Kiechl S, Willeit J, Rungger G, Egger G, Oberhollenzer F, Bonora E. Alcohol consumption and atherosclerosis: what is the relation? Prospective results from the Bruneck Study. *Stroke*. 1998;29(5):900-7.
47. Roerecke M, Rehm J. Alcohol consumption, drinking patterns, and ischemic heart disease: a narrative review of meta-analyses and a systematic review and meta-analysis of the impact of heavy drinking occasions on risk for moderate drinkers. *BMC Med*. 2014;12:182.
48. Au Yeung SL, Jiang C, Cheng KK, Cowling BJ, Liu B, Zhang W, et al. Moderate alcohol use and cardiovascular disease from Mendelian randomization. *PLoS One*. 2013;8(7):e68054.
49. Leong DP, Smyth A, Teo KK, McKee M, Rangarajan S, Pais P, et al. Patterns of alcohol consumption and myocardial infarction risk: observations from 52 countries in the INTERHEART case-control study. *Circulation*. 2014;130(5):390-8.
50. Hankey GJ, Eikelboom JW. Homocysteine and vascular disease. *Lancet*. 1999;354(9176):407-13.
51. Neves LB, Macedo DM, Lopes AC. Homocisteína. *J Bras Patol Med Lab*. 2004; 40(5):311-20.
52. Tsai MY, Hanson NQ, Schwichtenberg K, Garg U. Amplification refractory mutation system to identify mutations in cystathionine beta-synthase deficiency. *Clin Chem*. 1995;41(12 Pt 1):1775-7.
53. Jacques PF, Bostom AG, Williams RR, Ellison RC, Eckfeldt JH, Rosenberg IH, et al. Relation between folate status, a common mutation in methylenetetrahydrofolate reductase, and plasma homocysteine concentrations. *Circulation*. 1996;93(1):7-9.
54. Gambhir DS, Gambhir JK. Clinical spectrum and diagnosis of homocysteinemia. *Indian Heart J*. 2000;52(7 Suppl):S27-30.
55. Ganguly P, Alam SF. Role of homocysteine in the development of cardiovascular disease. *Nutr J*. 2015;14:6.
56. Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a

meta-analysis. *BMJ*. 2002;325(7374):1202.

57. Homocysteine Studies Collaboration. Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. *JAMA*. 2002;288(16):2015-22.

58. Hou X, Chen X, Shi J. Genetic polymorphism of MTHFR C677T and premature coronary artery disease susceptibility: A meta-analysis. *Gene*. 2015;565(1):39-44.

59. White PD. Genes, the heart and destiny. *N Engl J Med*. 1957;256(21):965-9.

60. Lloyd-Jones DM, Nam BH, D'Agostino RB Sr, Levy D, Murabito JM, Wang TJ, et al. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA*. 2004;291(18):2204-11.

61. Murabito JM, Pencina MJ, Nam BH, D'Agostino RB Sr, Wang TJ, Lloyd-Jones D, et al. Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. *JAMA*. 2005;294(24):3117-23.

62. Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med*. 1994;330(15):1041-6.

63. Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *The Lancet*. 2004;364(9438):937-52.

64. Bachmann JM, Willis BL, Ayers CR, Khera A, Berry JD. Association between family history and coronary heart disease death across long-term follow-up in men: the Cooper Center Longitudinal Study. *Circulation*. 2012;125(25):3092-8.

65. Cuchel M, Bruckert E, Ginsberg HN, Raal FJ, Santos RD, Hegele RA, et al. Homozygous familial hypercholesterolaemia: new insights and guidance for clinicians to improve detection and clinical management. A position paper from the Consensus Panel on Familial Hypercholesterolaemia of the European Atherosclerosis Society. *Eur Heart J*. 2014;35(32):2146-57.

66. Innerarity TL, Weisgraber KH, Arnold KS, Mahley RW, Krauss RM, Vega GL, et al. Familial defective apolipoprotein B-100: low density lipoproteins with abnormal receptor binding. *Proc Natl Acad Sci USA*. 1987;84(19):6919-23.

67. Lehnart SE, Ackerman MJ, Benson DW, Jr, Brugada R, Clancy CE, Donahue JK, et al. Inherited arrhythmias: a National Heart, Lung, and Blood Institute and Office of Rare Diseases workshop consensus report about the diagnosis, phenotyping, molecular mechanisms, and therapeutic approaches for primary cardiomyopathies of gene mutations affecting ion channel function. *Circulation*. 2007;116(20):2325-45.

68. Morgan TH. SEX LIMITED INHERITANCE IN DROSOPHILA. *Science*. 1910;32(812):120-2.

69. Cardon LR, Bell JI. Association study designs for complex diseases. *Nat Rev Genet*. 2001;2(2):91-9.

70. Terwilliger JD, Goring HH. Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. 2000. *Hum Biol*. 2009;81(5-6):663-728.

71. Avery OT, MacLeod CM, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribosenucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* 1944; 79:137-158.
72. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature.* 1953;171(4356):737-8.
73. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* 1986;51 Pt 1:263-73.
74. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409(6822):860-921.
75. Couzin J. Human genome. HapMap launched with pledges of \$100 million. *Science.* 2002;298(5595):941-2.
76. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* 2011;43(4):333-8.
77. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747–753.
78. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet.* 2011;43(4):339-44.
79. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, et al. Repeated Replication and a Prospective Meta-Analysis of the Association Between Chromosome9p21.3 and Coronary Artery Disease. *Circulation.* 2008;117:1675-1684.
80. Anderson JL, Horne BD. The 9p21 Locus and Coronary Heart Disease: Initiator, Promoter, or Precipitator? *J Am Coll Cardiol.* 2010;56:487-489.
81. Borg A, Sandberg T, Nilsson K, Johannsson O, Klinker M, Måsbäck A, et al. High Frequency of Multiple Melanomas and Breast and Pancreas Carcinomas in CDKN2A Mutation-Positive Melanoma Families. *JNCI J Natl Cancer Inst.* 2000;92(15):1260-1266.
82. CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet.* 2013;45(1):25-33.
83. Erdmann J, Grosshennig A, Braund PS, König IR, Hengstenberg C, Hall AS, et al. New susceptibility locus for coronary artery disease in chromosome 3q22.3. *Nat Genet.* 2009;41:280-282.
84. Pearson TA, Manolio TA. How to Interpret a Genome-wide Association Study. *JAMA.* 2008;299(11):1335-1344.
85. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev.* 2005;6:95–108.
86. Hunter DJ. Gene-environment interactions in human diseases. *Nat Rev Genet.* 2005;6:287-298.

87. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015;47(10):1121-30.
88. Terwilliger JD, Goring HH. Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. *Hum Biol.* 2000;72(1):63-132.
89. Cardon LR, Bell JI. Association study designs for complex diseases. *Nat Rev Genet.* 2001;2:91–9.
90. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J. Factors of Risk in the development of Coronary Heart Disease-six year follow-up experience. The Framingham study. *Annals Intern Med.* 1961;55:33-50
91. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation.* 1998;97:1837-47.
92. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P, CHD Risk Prediction Group. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA.* 2001;286(2):180-7.
93. Brindle P, Emberson J, Lampe F, Walker M, Whincup P, Fahey T, et al. Predictive accuracy of the Framingham coronary heart risk score in British men: prospective cohort study. *BMJ.* 2003;327:1267-70.
94. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J.* 2003;24(11):987-1003.
95. O'donnell CJ, Cupples LA, D'Agostino RB, Fox CS, Hoffmann U, Hwang SJ, et al. Genome-wide association study for subclinical atherosclerosis in major arterial territories in the NHLBI's Framingham Heart Study. *BMC Med Genet.* 2007;8 (Suppl 1):S4.
96. Van der Net JB, Janssens AC, Sijbrands EJ, Steyerberg EW. Value of genetic profiling for the prediction of coronary heart disease. *Am Heart J.* 2009;158(1):105-10.
97. Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet.* 2010;376(9750):1393-400.
98. Thanassoulis G, Peloso GM, Pencina MJ, Hoffman U, Fox CS, Cupples LA. A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham Heart Study. *Circ Cardiovasc Genet.* 2012;5(1):113-21.
99. Ganna A, Magnusson PK, Pedersen NL, de Faire U, Reilly M, Arnlov J, et al. Multilocus Genetic Risk Scores for Coronary Heart Disease Prediction. *Arterioscler Thromb Vasc Biol.* 2013;33(9):2267-72.
100. de Vries PS, Kavousi M, Ligthart S, Uitterlinden AG, Hofman A, Franco OH Incremental predictive value of 152 single nucleotide polymorphisms in the 10-year risk prediction of incident coronary heart disease: the Rotterdam Study. *Int J Epidemiol.* 2015;44(2):682-88.
101. Morris RW, Cooper JA, Shah T, Wong A, Drenos F, Engmann J, et al (UCLEB Consortium). Marginal

- role for 53 common genetic variants in cardiovascular disease prediction. *Heart*. 2016;102:1640-47.
102. European Cardiovascular Disease Statistics 2017
103. Campbell MJ, Julious SA, Altman DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *BMJ*. 1995;311(7013):1145-8.
104. Portal de estatísticas oficiais, Direção Regional de Estatística da Madeira.
105. Gould SJ. *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press. 2002; xxii:1433.
106. Tian C, Gregersen PK, Seldin MF. Accounting for ancestry: population substructure and genome-wide association studies. *Hum Mol Genet*. 2008;17(R2):R143-50.
107. Abdi H, Williams LJ. Principal component analysis. *WIREs Comp Stat*. 2010;2:433-459.
108. Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo Á, Lareu MV. An overview of STRUCTURE: applications, parameter settings, and supporting software. *Frontiers in Genetics*. 2003;4:98.
109. Tian C, Plenge RM, Ransom M, Lee A, Villoslada P, Selmi C, et al. Analysis and Application of European Genetic Substructure Using 300K SNP Information. *PLoS Genet*. 2008;4(1):e4.
110. Gonçalves R, Freitas A, Branco M, Rosa A, Fernandes AT, Zhivotovsky LA, et al. Y-chromosome Lineages from Portugal, Madeira and Açores Record Elements of Sephardim and Berber Ancestry. *Ann Hum Genet*. 2005;69(4):443-54.
111. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-9.
112. Tonino PA, De Bruyne B, Pijls NH, Siebert U, Ikeno F, van' t Veer M, et al. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. *N Engl J Med*. 2009;360(3):213-24.
113. National Institute of Health, National Heart, Lung, and Blood Institute North American Association for the Study of Obesity. *The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults*. NHLBI Obesity Education Initiative, NIH Publication Number: 00-4084, 2000.
114. Alberti KG, Zimmet P, Shaw J. Metabolic syndrome--a new world-wide definition. A Consensus Statement from the International Diabetes Federation. *Diabet Med*. 2006;23(5):469-80.
115. McGuire S. U.S. Department of Agriculture, U.S. Department of Health and Human Services, *Dietary Guidelines for Americans, 2010*. 7th Edition, Washington, DC: US Government Printing Office, January 2011. *Adv Nutr*. 2011;2(3):293-4.
116. 2003 European society of Hypertension-European society of cardiology guidelines for the management of arterial hypertension. *J Hypertens*. 2003;21:1011-1053.
117. Asmar R, Benetos A, Topouchian J, Laurent P, Pannier B, Brisac A-M, Target R, Levy BI: Assessment of arterial distensibility by automatic pulse wave velocity measurement: validation and clinical application studies. *Hypertension* 1995; 26:485-490.
118. American Diabetes Association. *Executive Summary: Standards of medical care in diabetes – 2010*.

Diabetes Care. 2010; 3(Suppl 1):S4-S10.

119. Catapano AL, Graham I, De Backer G, Wiklund O, Chapman MJ, Drexel H, et al. ESC/EAS Guidelines for the Management of Dyslipidaemias: The Task Force for the Management of Dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS) Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J*. 2016;37(39):2999-3058.
120. Piepoli MF, Hoes Aw, Agewall S. et al. 2016 European Guidelines on Cardiovascular disease prevention in clinical practice. *European Heart Journal*. 2016; 37(29)2315-81.
121. Mons U, Müezzinler A, Gellert C, Schöttker B, Abnet CC, Bobak M, et al. Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. *BMJ*. 2015; 350:h1551.
122. Scanu AM, Kurschinski D, Conn M, Garcia M. Important limiting features of the precipitation method in estimating LDL cholesterol: interference by IDL and Lp(a). *Atherosclerosis*.1988;8:57.
123. Crawford DC, Peng Z, Cheng JF, Boffelli D, Ahearn M, Nguyen D, et al. LPA and PLG sequence variation and kringle IV-2 copy number in two populations. *Hum Hered*. 2008;66(4):199-209.
124. Marcovina SM, Albers JJ, Gabel B, Koschinsky ML, Gaur VP. Effect of the number of apolipoprotein(a) kringle 4 domains on immunochemical measurements of lipoprotein(a). *Clin Chem*. 1995;41(2):246-55.
125. Greenow K, Pearce NJ, Ramji DP. The key role of apolipoprotein E in atherosclerosis. *J Mol Med (Berl)*. 2005;83(5):329-42.
126. Xu M, Zhao J, Zhang Y, Ma X, Dai Q, Zhi H, et al. Apolipoprotein E Gene Variants and Risk of Coronary Heart Disease: A Meta-Analysis. *BioMed Research International*. 2016;2016:3912175.
127. Kraft HG, Kochl S, Menzel HJ, Sandholzer C, Utermann G. The apolipoprotein (a) gene: a transcribed hypervariable locus controlling plasma lipoprotein (a) concentration. *Hum Genet*. 1992;90(3):220-30.
128. Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, Heath SC, et al. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med*. 2009;361(26):2518-28.
129. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. *JAMA*. 2009;301(22):2331-9.
130. Tregouet DA, König IR, Erdmann J, Munteanu A, Braund PS, Hall AS, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat Genet*. 2009;41(3):283-5.
131. Koschinsky ML. Lipoprotein(a) and atherosclerosis: new perspectives on the mechanism of action of an enigmatic lipoprotein. *Curr Atheroscler Rep*. 2005;7(5):389-95.
132. Bare LA, Morrison AC, Rowland CM, Shiffman D, Luke MM, Iakoubova OA, et al. Five common gene variants identify elevated genetic risk for coronary heart disease. *Genet Med*. 2007;9(10):682-9.
133. Morrison AC, Bare LA, Chambless LE, Ellis SG, Malloy M, Kane JP, et al. Prediction of coronary heart

- disease risk using a genetic risk score: the Atherosclerosis Risk in Communities Study. *Am J Epidemiol.* 2007;166(1):28-35.
134. Li Y, Iakoubova OA, Shiffman D, Devlin JJ, Forrester JS, Superko HR. KIF6 polymorphism as a predictor of risk of coronary events and of clinical event reduction by statin therapy. *Am J Cardiol.* 2010;106(7):994-8.
135. Peng P, Lian J, Huang RS, Xu L, Huang Y, Ba Y, et al. Meta-analyses of KIF6 Trp719Arg in coronary heart disease and statin therapeutic effect. *PLoS One.* 2012;7(12):e50126.
136. Galcheva-Gargova Z, Gangwani L, Konstantinov KN, Mikrut M, Theroux SJ, Enoch T, et al. The cytoplasmic zinc finger protein ZPR1 accumulates in the nucleolus of proliferating cells. *Mol Biol Cell.* 1998;9(10):2963-71.
137. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 2010;466(7307):707-13.
138. Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet.* 2009;41(3):334-41.
139. Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature.* 2010;466(7307):714-9.
140. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet.* 2008;40(2):161-9.
141. Nielsen MS, Jacobsen C, Olivecrona G, Gliemann J, Petersen CM. Sortilin/neurotensin receptor-3 binds and mediates degradation of lipoprotein lipase. *J Biol Chem.* 1999;274(13):8832-6.
142. Kjolby M, Andersen OM, Breiderhoff T, Fjorback AW, Pedersen KM, Madsen P, et al. Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. *Cell Metab.* 2010;12(3):213-23.
143. Seidah NG, Awan Z, Chretien M, Mbikay M. PCSK9: a key modulator of cardiovascular health. *Circ Res.* 2014;114(6):1022-36.
144. Seidah NG, Benjannet S, Wickham L, Marcinkiewicz J, Jasmin SB, Stifani S, et al. The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): Liver regeneration and neuronal differentiation. *Proc Natl Acad Sci U S A.* 2003;100(3):928-33.
145. Cohen JC, Boerwinkle E, Mosley TH, Jr., Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med.* 2006;354(12):1264-72.
146. Susan-Resiga D, Girard E, Kiss RS, Essalmani R, Hamelin J, Asselin MC, et al. The Proprotein Convertase Subtilisin/Kexin Type 9-resistant R410S Low Density Lipoprotein Receptor Mutation: A NOVEL MECHANISM CAUSING FAMILIAL HYPERCHOLESTEROLEMIA. *J Biol Chem.* 2017;292(5):1573-90.
147. Cambien F, Costerousse O, Tiret L, Poirier O, Lecerf L, Gonzales MF, et al. Plasma level and gene polymorphism of angiotensin-converting enzyme in relation to myocardial infarction. *Circulation.*

1994;90(2):669-76.

148. Gardemann A, Fink M, Stricker J, Nguyen QD, Humme J, Katz N, et al. ACE I/D gene polymorphism: presence of the ACE D allele increases the risk of coronary artery disease in younger individuals. *Atherosclerosis*. 1998;139(1):153-9.

149. Seckin D, Ilhan N, Ilhan N, Ozbay Y. The relationship between ACE insertion/deletion polymorphism and coronary artery disease with or without myocardial infarction. *Clin Biochem*. 2006;39(1):50-4.

150. Sayed-Tabatabaei FA, Oostra BA, Isaacs A, van Duijn CM, Witteman JC. ACE polymorphisms. *Circ Res*. 2006;98(9):1123-33.

151. Olivieri O, Stranieri C, Girelli D, Pizzolo F, Grazioli S, Russo C, et al. Homozygosity for angiotensinogen 235T variant increases the risk of myocardial infarction in patients with multi-vessel coronary artery disease. *J Hypertens*. 2001;19(5):879-84.

152. Zafarmand MH, van der Schouw YT, Grobbee DE, de Leeuw PW, Bots ML. The M235T polymorphism in the AGT gene and CHD risk: evidence of a Hardy-Weinberg equilibrium violation and publication bias in a meta-analysis. *PLoS One*. 2008;3(6):e2533.

153. Wang YJ, Pan Y. Angiotensinogen gene M235T polymorphism and risk of coronary artery disease: a meta-analysis. *Mol Med Rep*. 2012;6(4):884-8.

154. Sui X, Gao C. The angiotensinogen gene M235T polymorphism and acute myocardial infarction risk: a meta-analysis of 22 studies. *Mol Biol Rep*. 2013;40(7):4439-45.

155. Joshi PH, Xu H, LeStrange R, Flockhart N, Kirkland B, Vazquez G, et al. The M235T single nucleotide polymorphism in the angiotensinogen gene is associated with coronary artery calcium in patients with a family history of coronary artery disease. *Atherosclerosis*. 2006;192(2):433-9.

156. Feng X, Zheng BS, Shi JJ, Qian J, He W, Zhou HF. A systematic review and meta-analysis of the association between angiotensin II type 1 receptor A1166C gene polymorphism and myocardial infarction susceptibility. *J Renin Angiotensin Aldosterone Syst*. 2014;15(3):307-15.

157. Zhang K, Zhou B, Zhang L. Association study of angiotensin II type 1 receptor: A1166C (rs5186) polymorphism with coronary heart disease using systematic meta-analysis. *J Renin Angiotensin Aldosterone Syst*. 2013;14(2):181-8.

158. Jones BK, Levorse J, Tilghman SM. Deletion of a nuclease-sensitive region between the IGF2 and H19 genes leads to IGF2 misregulation and increased adiposity. *Hum Mol Genet*. 2001;10(8):807-14.

159. Saxena R, Elbers Clara C, Guo Y, Peter I, Gaunt Tom R, Mega Jessica L, et al. Large-Scale Gene-Centric Meta-Analysis across 39 Studies Identifies Type 2 Diabetes Loci. *Am J Hum Genet*. 2012;90(3):410-25.

160. Rodriguez S, Gaunt TR, O'Dell SD, Chen XH, Gu D, Hawe E, et al. Haplotypic analyses of the IGF2-INS-TH gene cluster in relation to cardiovascular risk traits. *Hum Mol Genet*. 2004;13(7):715-25.

161. Chimienti F, Favier A, Seve M. ZnT-8, a pancreatic beta-cell-specific zinc transporter. *Biometals*.

2005;18(4):313-7.

162. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. 2007;445(7130):881-5.
163. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science*. 2007;316(5829):1341-5.
164. Cauchi S, Meyre D, Durand E, Proenca C, Marre M, Hadjadj S, et al. Post genome-wide association studies of novel genes associated with type 2 diabetes show gene-gene interaction and high predictive value. *PLoS One*. 2008;3(5):e2031.
165. Hou H, Ge S, Zhao L, Wang C, Wang W, Zhao X, et al. An Updated Systematic Review and Meta-analysis of Association Between Adiponectin Gene Polymorphisms and Coronary Artery Disease. *OMICS*. 2017;21(6):340-51.
166. Janani C, Ranjitha Kumari BD. PPAR gamma gene-a review. *Diabetes Metab Syndr*. 2015;9(1):46-50.
167. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Multiple type 2 diabetes susceptibility genes following genome-wide association scan in UK samples. *Science*. 2007;316(5829):1336-41.
168. Qian Y, Li P, Zhang J, Shi Y, Chen K, Yang J, et al. Association between peroxisome proliferator-activated receptor-alpha, delta, and gamma polymorphisms and risk of coronary heart disease: A case-control study and meta-analysis. *Medicine*. 2016;95(32):e4299.
169. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet*. 2006;38(3):320-3.
170. Sousa AG, Marquezine GF, Lemos PA, Martinez E, Lopes N, Hueb WA, et al. TCF7L2 polymorphism rs7903146 is associated with coronary artery disease severity and mortality. *PLoS One*. 2009;4(11):e7697.
171. Vaisse C, Clement K, Durand E, Hercberg S, Guy-Grand B, Froguel P. Melanocortin-4 receptor mutations are a frequent and heterogeneous cause of morbid obesity. *J Clin Invest*. 2000;106(2):253-62.
172. Farooqi IS, Keogh JM, Yeo GS, Lank EJ, Cheetham T, O'Rahilly S. Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N Engl J Med*. 2003;348(12):1085-95.
173. Wankhade UD, Good DJ. Melanocortin 4 receptor is a transcriptional target of nescient helix-loop-helix-2. *Mol Cell Endocrinol*. 2011;341(1-2):39-47.
174. Silander K, Mohlke KL, Scott LJ, Peck EC, Hollstein P, Skol AD, et al. Genetic Variation Near the Hepatocyte Nuclear Factor-4 α Gene Predicts Susceptibility to Type 2 Diabetes. *Diabetes*. 2004;53(4):1141-9.
175. Palu RAS, Thummel CS. Sir2 Acts through Hepatocyte Nuclear Factor 4 to maintain insulin Signaling and Metabolic Homeostasis in *Drosophila*. *PLoS Genetics*. 2016;12(4):e1005978.
176. Chamouton J, Latruffe N. PPAR α /HNF4 α interplay on diversified responsive elements. Relevance in the regulation of liver peroxisomal fatty acid catabolism. *Curr Drug Metab*. 2012;13(10):1436-53.

177. Ronkainen J, Mondini E, Cinti F, Cinti S, Sebert S, Savolainen MJ, et al. Fto-Deficiency Affects the Gene and MicroRNA Expression Involved in Brown Adipogenesis and Browning of White Adipose Tissue in Mice. *Int J Mol Sci.* 2016;17(11).
178. Yan Y, Hu Y, North KE, Franceschini N, Lin D. Evaluation of population impact of candidate polymorphisms for coronary heart disease in the Framingham Heart Study Offspring Cohort. *BMC Proc.* 2009;3 Suppl 7:S118.
179. Akao H, Polisecki E, Kajinami K, Trompet S, Robertson M, Ford I, et al. KIF6, LPA, TAS2R50, and VAMP8 genetic variation, low density lipoprotein cholesterol lowering response to pravastatin, and heart disease risk reduction in the elderly. *Atherosclerosis.* 2012;220(2):456-62.
180. Shiffman D, Ellis SG, Rowland CM, Malloy MJ, Luke MM, Iakoubova OA, et al. Identification of four gene variants associated with myocardial infarction. *Am J Hum Genet.* 2005;77(4):596-605.
181. Hou X, Chen X, Shi J. Genetic polymorphism of MTHFR C677T and premature coronary artery disease susceptibility: A meta-analysis. *Gene.* 2015;565(1):39-44.
182. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide Association Analysis of Coronary Artery Disease. *N Engl J Med.* 2007;357(5):443-53.
183. Huen K, Yousefi P, Street K, Eskenazi B, Holland N. PON1 as a model for integration of genetic, epigenetic, and expression data on candidate susceptibility genes. *Environ epigen.* 2015;1(1).
184. Schrader C, Rimbach G. Determinants of paraoxonase 1 status: genes, drugs and nutrition. *Curr Med Chem.* 2011;18(36):5624-43.
185. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science.* 2007;316(5830):1488-91.
186. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science.* 2007;316(5830):1491-3.
187. Palomaki GE, Melillo S, Bradley LA. Association between 9p21 genomic markers and heart disease: a meta-analysis. *JAMA.* 2010;303(7):648-56.
188. Broadbent HM, Peden JF, Lorkowski S, Goel A, Ongen H, Green F, et al. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. *Hum Mol Genet.* 2008;17(6):806-14.
189. Ardissino D, Berzuini C, Merlini PA, Mannuccio Mannucci P, Surti A, Burti N, et al. Influence of 9p21.3 genetic variants on clinical and angiographic outcomes in early-onset myocardial infarction. *J Am Coll Cardiol.* 2011;58(4):426-34.
190. Congrains A, Kamide K, Oguro R, Yasuda O, Miyata K, Yamamoto E, et al. Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of ANRIL and CDKN2A/B. *Atherosclerosis.* 2012;220(2):449-55.

191. Bujak M, Ren G, Kweon HJ, Dobaczewski M, Reddy A, Taffet G, et al. Essential role of Smad3 in infarct healing and in the pathogenesis of cardiac remodeling. *Circulation*. 2007;116(19):2127-38.
192. Wilson DG, Phamluong K, Li L, Sun M, Cao TC, Liu PS, et al. Global defects in collagen secretion in a Mia3/TANGO1 knockout mouse. *J Cell Biol*. 2011;193(5):935-51.
193. Jones PD, Kaiser MA, Ghaderi Najafabadi M, McVey DG, Beveridge AJ, Schofield CL, et al. The Coronary Artery Disease-associated Coding Variant in Zinc Finger C3HC-type Containing 1 (ZC3HC1) Affects Cell Cycle Regulation. *J Biol Chem*. 2016;291(31):16318-27.
194. Nurnberg ST, Cheng K, Raiesdana A, Kundu R, Miller CL, Kim JB, et al. Coronary Artery Disease Associated Transcription Factor TCF21 Regulates Smooth Muscle Precursor Cells That Contribute to the Fibrous Cap. *PLoS Genet*. 2015;11(5).
195. Pu X, Xiao Q, Kiechl S, Chan K, Ng F L, Gor S, et al. ADAMTS7 Cleavage and Vascular Smooth Muscle Cell Migration Is Affected by a Coronary-Artery-Disease-Associated Variant. *Am J Hum Genet*. 2013;92(3):366-74.
196. Bengtsson E, Hultman K, Dunér P, Ascitutto G, Almgren P, Orho-Melander M, et al. ADAMTS-7 is associated with a high-risk plaque phenotype in human atherosclerosis. *Sci Rep*. 2017;7.
197. Reilly MP, Li M, He J, Ferguson JF, Stylianou IM, Mehta NN, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet*. 2011;377(9763):383-92.
198. Kelloniemi A, Szabo Z, Serpi R, Näpänkangas J, Ohukainen P, Tenhunen O, et al. The Early-Onset Myocardial Infarction Associated PHACTR1 Gene Regulates Skeletal and Cardiac Alpha-Actin Gene Expression. *PLoS One*. 2015;10(6).
199. You W, Smith SA, Coughlin SS. Evaluation of genetic risk scores for prediction of dichotomous outcomes. *Int J Mol Epidemiol Genet*. 2015;6(1):1-8.
200. Goldstein BA, Knowles JW, Salfati E, Ioannidis JP, Assimes TL. Simple, standardized incorporation of genetic risk into- non genetic risk prediction tools for complex traits: coronary artery disease as an exemple. *Front Genet*. 2014;5:254.
201. Hosmer, D., Lemeshow, S. (2013). *Applied Logistic Regression (3rd Edition)*. John Wiley & Sons, New York.
202. Hosmer, D., Lemeshow, S. (2008). *Applied Survival Analysis (3rd Edition)*. John Wiley & Sons, New York.
203. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988; 44:837-845.
204. Pencina MJ, Steyerberg EW, D'Agostino RB Sr. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30(1):11–21.
205. Kerr KF, McClelland RL, Brown ER, Lumley T. Evaluating the incremental value of new biomarkers

- with integrated discrimination improvement. *Am J Epidemiol.* 2011;174 (3):364–74.
206. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med.* 2008;27(2):157–172.
207. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics in Medicine.* 2004;23(13):2109–2123.
208. Harrell FE, Lee KL, Mark DB. Tutorial in Biostatistics: Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996;15:361–87.
209. Simon PH, Sylvestre MP, Tremblay J, Hamet P. Key Considerations and Methods in the Study of Gene-Environment Interactions. *Am J Hypertens.* 2016;29(8):891-9.
210. Gui J, Andrew AS, Andrews P, Nelson HM, Kelsey KT, Karagas MR, et al. A simple and computationally efficient sampling approach to covariate adjustment for multifactor dimensionality reduction analysis of epistasis. *Hum Hered.* 2010;70(3):219-25.
211. Yang CH, Lin YD, Yang CS, Chuang LY. An efficiency analysis of high-order combinations of gene-gene interactions using multifactor-dimensionality reduction. *BMC Genomics.* 2015;16:489.
212. Roerecke M, Rehm J. Alcohol consumption, drinking patterns, and ischemic heart disease: a narrative review of meta-analyses and a systematic review and meta-analysis of the impact of heavy drinking occasions on risk for moderate drinkers. *BMC Med.* 2014;12:182.
213. Berglund G, Elmstahl S, Janzon L, Larsson SA. The Malmo Diet and Cancer Study. Design and feasibility. *J Intern Med.* 1993;233(1):45-51.
214. Sala J, Marrugat J, Masia R, Porta M. Improvement in survival after myocardial infarction between 1978-85 and 1986-88 in the REGICOR study. (Registre Gironi del COR) registry. *Eur Heart J.* 1995;16(6):779-84.
215. Preuss M, Konig IR, Thompson JR, Erdmann J, Absher D, Assimes TL, et al. Design of the Coronary ARtery Disease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: A Genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ Cardiovasc Genet.* 2010;3(5):475-83.
216. Szpakowicz A, Kiliszek M, Pepinski W, Waszkiewicz E, Franaszczyk M, Skawronska M, et al. Polymorphism of 9p21.3 locus is associated with 5-year survival in high-risk patients with myocardial infarction. *PLoS One.* 2014;9(8):e104635.
217. Patel RS, Su S, Neeland IJ, Ahuja A, Veledar E, Zhao J, et al. The chromosome 9p21 risk locus is associated with angiographic severity and progression of coronary artery disease. *Eur Heart J.* 2010;31(24):3017-23.
218. Munir MS, Wang Z, Alahdab F, Steffen MW, Erwin PJ, Kullo IJ, et al. The association of 9p21-3 locus

- with coronary atherosclerosis: a systematic review and meta-analysis. *BMC Med Genet.* 2014;15:66.
219. Buyschaert I, Carruthers KF, Dunbar DR, Peuteman G, Rietzschel E, Belmans A, et al. A variant at chromosome 9p21 is associated with recurrent myocardial infarction and cardiac death after acute coronary syndrome: the GRACE Genetics Study. *Eur Heart J.* 2010;31(9):1132-41.
220. Patel RS, Asselbergs FW, Quyyumi AA, Palmer TM, Finan CI, Tragante V, et al. Genetic variants at chromosome 9p21 and risk of first versus subsequent coronary heart disease events: a systematic review and meta-analysis. *J Am Coll Cardiol.* 2014;63(21):2234-45.
221. Horne BD, Carlquist JF, Muhlestein JB, Bair TL, Anderson JL. Association of variation in the chromosome 9p21 locus with myocardial infarction versus chronic coronary artery disease. *Circ Cardiovasc Genet.* 2008;1(2):85-92.
222. Muehlschlegel JD, Liu KY, Perry TE, Fox AA, Collard CD, Shernan SK, et al. Chromosome 9p21 variant predicts mortality after coronary artery bypass graft surgery. *Circulation.* 2010;122(11 Suppl):S60-5.
223. Mendonca I, dos Reis RP, Pereira A, Cafe H, Serrao M, Sousa AC, et al. Independent association of the variant rs1333049 at the 9p21 locus and coronary heart disease. *Rev Port Cardiol.* 2011;30(6):575-91.
224. Hernandez-Diaz Y, Tovilla-Zarate CA, Juarez-Rojop IE, Gonzalez-Castro TB, Rodriguez-Perez C, Lopez-Narvaez ML, et al. Effects of paraoxonase 1 gene polymorphisms on heart diseases: Systematic review and meta-analysis of 64 case-control studies. *Medicine (Baltimore).* 2016;95(44):e5298.
225. Tong G, Wang N, Leng J, Tong X, Shen Y, Yang J, et al. Common variants in adiponectin gene are associated with coronary artery disease and angiographical severity of coronary atherosclerosis in type 2 diabetes. *Cardiovasc Diabetol.* 2013;12:67.
226. Mofarrah M, Ziaee S, Pilehvar-Soltanahmadi Y, Zarghami F, Boroumand M, Zarghami N. Association of KALRN, ADIPOQ, and FTO gene polymorphism in type 2 diabetic patients with coronary artery disease: possible predisposing markers. *Coron Artery Dis.* 2016;27(6):490-6.
227. Liu C, Mou S, Pan C. The FTO gene rs9939609 polymorphism predicts risk of cardiovascular disease: a systematic review and meta-analysis. *PLoS One.* 2013;8(8):e71901.
228. Humphries SE, Morgan L. Genetic risk factors for stroke and carotid atherosclerosis: insights into pathophysiology from candidate gene approaches. *Lancet Neurol.* 2004;3(4):227-35.
229. Mendonca I, Freitas IA, Sousa CA, Gomes S, Faria P, Drumond A, et al. Angiotensin converting enzyme gene polymorphisms and coronary risk in a Portuguese population. *Rev Port Cardiol.* 2004;23(12):1593-601.
230. Keating BJ, Tischfield S, Murray SS, Bhangale T, Price TS, Glessner JT, et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS One.* 2008;3(10):e3583.
231. Roberts R. Genetics of coronary artery disease. *Circ Res.* 2014;114(12):1890-903.
232. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide

- association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet.* 2008;9(5):356-69.
233. Assimes TL, Roberts R. Genetics: Implications for Prevention and Management of Coronary Artery Disease. *J Am Coll Cardiol.* 2016;68(25):2797-818.
234. Cordell HJ. Detecting gene-gene interactions that underlie human diseases. *Nature reviews Genetics.* 2009;10(6):392-404.
235. Moore JH, Williams SM. Epistasis and its implications for personal genetics. *Am J Hum Genet.* 2009;85(3):309-20.
236. Ritchie MD, Moutsinger AA. Multifactor dimensionality reduction for detecting gene-gene and gene-environment interactions in pharmacogenomics studies. *Pharmacogenomics.* 2005;6(8):823-34.
237. Gu SJ, Liu MM, Guo ZR, Wu M, Chen Q, Zhou ZY, et al. Gene-gene interactions among PPARalpha/delta/gamma polymorphisms for hypertriglyceridemia in Chinese Han population. *Gene.* 2013;515(2):272-6.
238. Fan W, Shen C, Wu M, Zhou ZY, Guo ZR. Association and interaction of PPARalpha, delta, and gamma gene polymorphisms with low-density lipoprotein-cholesterol in a Chinese Han population. *Genet Test Mol Biomarkers.* 2015;19(7):379-86.
239. Svensson PA, Wahlstrand B, Olsson M, Froguel P, Falchi M, Bergman RN, et al. CDKN2B expression and subcutaneous adipose tissue expandability: possible influence of the 9p21 atherosclerosis locus. *Biochem Biophys Res Commun.* 2014;446(4):1126-31.
240. Pitsavos C, Panagiotakos D, Trichopoulou A, Chrysohoou C, Dedoussis G, Chloptsios Y, et al. Interaction between Mediterranean diet and methylenetetrahydrofolate reductase C677T mutation on oxidized low density lipoprotein concentrations: the ATTICA study. *Nutr Metab Cardiovasc Dis.* 2006;16(2):91-9.
241. Real JT, Martinez-Hervas S, Garcia-Garcia AB, Chaves FJ, Civera M, Ascaso JF, et al. Association of C677T polymorphism in MTHFR gene, high homocysteine and low HDL cholesterol plasma values in heterozygous familial hypercholesterolemia. *Journal of atherosclerosis and thrombosis.* 2009;16(6):815-20.
242. Ference BA, Yoo W, Flack JM, Clarke M. A common KIF6 polymorphism increases vulnerability to low-density lipoprotein cholesterol: two meta-analyses and a meta-regression analysis. *PLoS One.* 2011;6(12):e28834.
243. Aung LH, Yin RX, Wu DF, Wang W, Liu CW, Pan SL. Association of the variants in the BUD13-ZNF259 genes and the risk of hyperlipidaemia. *J Cell Mol Med.* 2014;18(7):1417-28.
244. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 2010;466(7307):707-13.
245. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat*

Genet. 2008;40(5):638-45.

246. Kunnas T, Nikkari ST. Association of Zinc Finger, C3HC-Type Containing 1 (ZC3HC1) rs11556924 Genetic Variant With Hypertension in a Finnish Population, the TAMRISK Study. *Medicine (Baltimore)*. 2015;94(32):e1221.

247. Duffy VB. Associations between oral sensation, dietary behaviors and risk of cardiovascular disease (CVD). *Appetite*. 2004;43(1):5-9.

248. van Setten J, Isgum I, Pechlivanis S, Tragante V, de Jong PA, Smolonska J, et al. Serum lipid levels, body mass index, and their role in coronary artery calcification: a polygenic analysis. *Circ Cardiovasc Genet*. 2015;8(2):327-33.

249. Jansen H, Loley C, Lieb W, Pencina MJ, Nelson CP, Kathiresan S, et al. Genetic variants primarily associated with type 2 diabetes are related to coronary artery disease risk. *Atherosclerosis*. 2015;241(2):419-26.

250. 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.

251. Quertermous T, Ingelsson E. Coronary Artery Disease and Its Risk Factors: Leveraging Shared Genetics to Discover Novel Biology. *Circ Res*. 2016;118(1):14-6.

252. Ross S, Gerstein HC, Eikelboom J, Anand SS, Yusuf S, Pare G. Mendelian randomization analysis supports the causal role of dysglycaemia and diabetes in the risk of coronary artery disease. *Eur Heart J*. 2015;36(23):1454-62.

253. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet*. 2012;380(9841):572-80.

254. Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, Shah T, et al. Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. *BMJ*. 2011;342:d548.

255. Vaarhorst AA, Lu Y, Heijmans BT, Dolle ME, Bohringer S, Putter H, et al. Literature-based genetic risk scores for coronary heart disease: the Cardiovascular Registry Maastricht (CAREMA) prospective cohort study. *Circ Cardiovasc Genet*. 2012;5(2):202-9.

256. Webb TR, Erdmann J, Stirrups KE, Stitzel NO, Masca NG, Jansen H, et al. Systematic Evaluation of Pleiotropy Identifies 6 Further Loci Associated With Coronary Artery Disease. *J Am Coll Cardiol*. 2017;69(7):823-36.

257. Dube JB, Boffa MB, Hegele RA, Koschinsky ML. Lipoprotein(a): more interesting than ever after 50 years. *Curr Opin Lipidol*. 2012;23(2):133-40.

258. Lippi G, Guidi G. Lipoprotein(a): an emerging cardiovascular risk factor. *Crit Rev Clin Lab Sci*.

2003;40(1):1-42.

259. van Iperen EP, Sivapalaratnam S, Holmes MV, Hovingh GK, Zwinderman AH, Asselbergs FW. Genetic analysis of emerging risk factors in coronary artery disease. *Atherosclerosis*. 2016;254:35-41.
260. Willeit P, Kiechl S, Kronenberg F, Witztum JL, Santer P, Mayr M, et al. Discrimination and net reclassification of cardiovascular risk with lipoprotein(a): prospective 15-year outcomes in the Bruneck Study. *J Am Coll Cardiol*. 2014;64(9):851-60.
261. Tada H, Melander O, Louie JZ, Catanese JJ, Rowland CM, Devlin JJ, et al. Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. *Eur Heart J*. 2016;37(6):561-7.
262. Antiochos P, Marques-Vidal P, McDaid A, Waeber G, Vollenweider P. Association between parental history and genetic risk scores for coronary heart disease prediction: The population-based CoLaus study. *Atherosclerosis*. 2016;244:59-65.
263. Christiansen MK, Nyegaard M, Pedersen LN, Larsen SB, Wurtz M, Hjort J, et al. A 45-SNP genetic risk score is increased in early-onset coronary artery disease but independent of familial disease clustering. *Atherosclerosis*. 2017;257:172-8.
264. Hughes MF, Saarela O, Stritzke J, Kee F, Silander K, Klopp N, et al. Genetic markers enhance coronary risk prediction in men: the MORGAM prospective cohorts. *PLoS One*. 2012;7(7):e40922.
265. Iribarren C, Lu M, Jorgenson E, Martinez M, Lluís-Ganella C, Subirana I, et al. Clinical Utility of Multimarker Genetic Risk Scores for Prediction of Incident Coronary Heart Disease: A Cohort Study Among Over 51 Thousand Individuals of European Ancestry. *Circ Cardiovasc Genet*. 2016;9(6):531-40.
266. van der Net JB, Janssens AC, Defesche JC, Kastelein JJ, Sijbrands EJ, Steyerberg EW. Usefulness of genetic polymorphisms and conventional risk factors to predict coronary heart disease in patients with familial hypercholesterolemia. *Am J Cardiol*. 2009;103(3):375-80.
267. Paynter NP, Chasman DI, Pare G, Buring JE, Cook NR, Miletich JP, et al. Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA*. 2010;303(7):631-7.
268. Christiansen MK, Nyegaard M, Larsen SB, Grove EL, Wurtz M, Neergaard-Petersen S, et al. A genetic risk score predicts cardiovascular events in patients with stable coronary artery disease. *Int J Cardiol*. 2017;241:411-6.
269. Labos C, Martinez SC, Leo Wang RH, Lenzini PA, Pilote L, Bogaty P, et al. Utility of a genetic risk score to predict recurrent cardiovascular events 1 year after an acute coronary syndrome: A pooled analysis of the RISCA, PRAXY, and TRIUMPH cohorts. *Atherosclerosis*. 2015;242(1):261-7.
270. Labos C, Wang RH, Pilote L, Bogaty P, Brophy JM, Engert JC, et al. Traditional risk factors and a Genetic Risk Score are associated with age of first acute coronary syndrome. *Heart*. 2014;100(20):1620-4.
271. Weijmans M, de Bakker PI, van der Graaf Y, Asselbergs FW, Algra A, Jan de Borst G, et al. Incremental value of a genetic risk score for the prediction of new vascular events in patients with clinically manifest

vascular disease. *Atherosclerosis*. 2015;239(2):451-8.

272. de Vries PS, Kavousi M, Ligthart S, Uitterlinden AG, Hofman A, Franco OH, et al. Incremental predictive value of 152 single nucleotide polymorphisms in the 10-year risk prediction of incident coronary heart disease: the Rotterdam Study. *Int J Epidemiol*. 2015;44(2):682-8.

273. Mega JL, Stitzel NO, Smith JG, Chasman DI, Caulfield MJ, Devlin JJ, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet*. 2015;385(9984):2264-71.

274. Ridker PM, MacFadyen JG, Fonseca FA, Genest J, Gotto AM, Kastelein JJ, et al. Number needed to treat with rosuvastatin to prevent first cardiovascular events and death among men and women with low low-density lipoprotein cholesterol and elevated high-sensitivity C-reactive protein: justification for the use of statins in prevention: an intervention trial evaluating rosuvastatin (JUPITER). *Circ Cardiovasc Qual Outcomes*. 2009;2(6):616-23.

275. Taylor F, Huffman MD, Macedo AF, Moore TH, Burke M, Davey Smith G, et al. Statins for the primary prevention of cardiovascular disease. *The Cochrane database of systematic reviews*. 2013(1):Cd004816.

276. Heller DJ, Coxson PG, Penko J, Pletcher MJ, Goldman L, Odden MC, et al. Evaluating the Impact and Cost-Effectiveness of Statin Use Guidelines for Primary Prevention of Coronary Heart Disease and Stroke. *Circulation*. 2017.

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

278. Tada H, Kawashiri MA, Yamagishi M. Clinical Perspectives of Genetic Analyses on Dyslipidemia and Coronary Artery Disease. *Journal of atherosclerosis and thrombosis*. 2017;24(5):452-61.

279. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. *Stat Med*. 1990;9(12):1501-15.

280. Rallidis LS, Sakadakis EA, Tympas K, Varounis C, Zolindaki M, Dargès N, et al. The impact of smoking on long-term outcome of patients with premature (≤ 35 years) ST-segment elevation acute myocardial infarction. *Am Heart J*. 2015;169(3):356-62.

281. Diabetes Mellitus: A Major Risk Factor for Cardiovascular Disease. A Joint Editorial Statement by the American Diabetes Association; the National Heart, Lung, and Blood Institute; the Juvenile Diabetes Foundation International; the National Institute of Diabetes and Digestive and Kidney Diseases; and the American Heart Association. *Circulation*. 1999;100(10):1132-3.

282. Stamler J, Wentworth D, Neaton JD. Is relationship between serum cholesterol and risk of premature death from coronary heart disease continuous and graded? Findings in 356,222 primary screenees of the Multiple Risk Factor Intervention Trial (MRFIT). *JAMA*. 1986;256(20):2823-8.

283. Franklin SS, Wong ND. Hypertension and cardiovascular disease: contributions of the framingham heart study. *Global heart*. 2013;8(1):49-57.
284. Mora S, Ridker PM. Justification for the Use of Statins in Primary Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER)- can C-reactive protein be used to target statin therapy in primary prevention? *Am J Cardiol*. 2006;97(2a):33a-41a.
285. Hansson GK. Inflammation, Atherosclerosis, and Coronary Artery Disease. *N Engl J Med*. 2005;352(16):1685-95.
286. Pereira A, Palma dos Reis R, Rodrigues R, Sousa AC, Gomes S, Borges S, et al. Association of ADAMTS7 gene polymorphism with cardiovascular survival in coronary artery disease. *Physiol Genomics* 2016;48(11):810-5.
287. You L, Tan L, Liu L, Shen R, Chaugai S, Wang DW, et al. ADAMTS7 locus confers high cross-race risk for development of coronary atheromatous plaque. *Mol Genet Genomics*. 2016;291(1):121-8.
288. Arroyo AG, Andrés V. ADAMTS7 in cardiovascular disease: from bedside to bench and back again? *Circulation*. 2015;131(13):1156–59.
289. Kessler T, Zhang L, Liu Z, Yin X, Huang Y, Wang Y, et al. ADAMTS-7 inhibits re-endothelialization of injured arteries and promotes vascular remodeling through cleavage of thrombospondin-1. *Circulation*. 2015;131(13):1191-201.
290. Bauer RC, Tohyama J, Cui J, Cheng L, Yang J, Zhang X, et al. Knockout of Adamts7, a novel coronary artery disease locus in humans, reduces atherosclerosis in mice. *Circulation*. 2015;131(13):1202-13.
291. Bayoglu B, Arslan C, Tel C, Ulutin T, Dirican A, Deser SB, et al. Genetic variants rs1994016 and rs3825807 in ADAMTS7 affect its mRNA expression in atherosclerotic occlusive peripheral arterial disease. *J Clin Lab Anal*. 2017;15.
292. Saleheen D, Zhao W, Young R, Nelson CP, Ho W, Ferguson JF, et al. Loss of Cardioprotective Effects at the ADAMTS7 Locus as a Result of Gene-Smoking Interactions. *Circulation*. 2017;135(24):2336-53.
293. Wirtwein M, Melander O, Sjogren M, Hoffmann M, Narkiewicz K, Gruchala M, et al. Relationship between selected DNA polymorphisms and coronary artery disease complications. *Int J Cardiol*. 2017;228:814-20.
294. Iakoubova OA, Tong CH, Rowland CM, Kirchgessner TG, Young BA, Arellano AR, et al. Association of the Trp719Arg polymorphism in kinesin-like protein 6 with myocardial infarction and coronary heart disease in 2 prospective trials: the CARE and WOSCOPS trials. *J Am Coll Cardiol*. 2008;51(4):435-43.
295. Shiffman D, Chasman DI, Zee RY, Iakoubova OA, Louie JZ, Devlin JJ, et al. A kinesin family member 6 variant is associated with coronary heart disease in the Women's Health Study. *J Am Coll Cardiol*. 2008;51(4):444-8.
296. Shiffman D, O'Meara ES, Bare LA, Rowland CM, Louie JZ, Arellano AR, et al. Association of gene variants with incident myocardial infarction in the Cardiovascular Health Study. *Arterioscler Thromb Vasc*

Biol. 2008;28(1):173-9.

297. Assimes TL, Holm H, Kathiresan S, Reilly MP, Thorleifsson G, Voight BF, et al. Lack of association between the Trp719Arg polymorphism in kinesin-like protein-6 and coronary artery disease in 19 case-control studies. *J Am Coll Cardiol*. 2010;56(19):1552-63.

298. Ward H, Mitrou PN, Bowman R, Luben R, Wareham NJ, Khaw KT, et al. APOE genotype, lipids, and coronary heart disease risk: a prospective population study. *Arch Intern Med*. 2009;169(15):1424-9.

299. Brautbar A, Pompeii LA, Dehghan A, Ngwa JS, Nambi V, Virani SS, et al. A genetic risk score based on direct associations with coronary heart disease improves coronary heart disease risk prediction in the Atherosclerosis Risk in Communities (ARIC), but not in the Rotterdam and Framingham Offspring, Studies. *Atherosclerosis*. 2012;223(2):421-6.

300. Lluís-Ganella C, Subirana I, Lucas G, Tomas M, Munoz D, Senti M, et al. Assessment of the value of a genetic risk score in improving the estimation of coronary risk. *Atherosclerosis*. 2012;222(2):456-63.

301. Krarup NT, Borglykke A, Allin KH, Sandholt CH, Justesen JM, Andersson EA, et al. A genetic risk score of 45 coronary artery disease risk variants associates with increased risk of myocardial infarction in 6041 Danish individuals. *Atherosclerosis*. 2015;240(2):305-10.

302. Ellis KL, Frampton CM, Pilbrow AP, Troughton RW, Doughty RN, Whalley GA, 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. *Circ Cardiovasc Genet*. 2011;4(6):636-46.

303. Bressler J, Folsom AR, Couper DJ, Volcik KA, Boerwinkle E. Genetic variants identified in a European genome-wide association study that were found to predict incident coronary heart disease in the atherosclerosis risk in communities study. *Am J Epidemiol*. 2010;171(1):14-23.

304. Kathiresan S, Melander O, Anevski D, Guiducci C, Burt NP, Roos C, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. *N Engl J Med*. 2008;358(12):1240-9.

305. Abraham G, Havulinna AS, Bhalala OG, Byars SG, De Livera AM, Yetukuri L, et al. Genomic prediction of coronary heart disease. *Eur Heart J*. 2016;37(43):3267-78.

ARTIGOS PUBLICADOS

Associação independente da variante rs1333049, no locus 9p21, com a doença coronária, numa população portuguesa [45]

ISABEL MENDONÇA, ROBERTO PALMA DOS REIS, ANDREIA PEREIRA, HUGO CAFÉ, MARCO SERRÃO, ANA CÉLIA SOUSA, ANA ISABEL FREITAS, GRAÇA GUERRA, SÓNIA FREITAS, CAROLINA FREITAS, ILÍDIO ORNELAS, ANTÓNIO BREHM, JOSÉ JORGE ARAÚJO

Unidade de Investigação e Serviço de Cardiologia do Hospital Nélio de Mendonça (SESARAM), Funchal, Portugal
Laboratório de Genética Humana da Universidade da Madeira, Funchal, Portugal
Faculdade de Ciências Médicas da Universidade Nova de Lisboa, Lisboa, Portugal

Rev Port Cardiol 2011; 30 (06): 575-591

RESUMO

Introdução: Estudos recentes de associação genómica em larga escala (GWAS) identificaram vários polimorfismos de um único nucleótido (SNP), localizados no locus 9p21, associados com doença arterial coronária (DAC). De entre eles o SNP rs1333049 demonstrou uma associação consistente com a DAC tendo sido reproduzida, com sucesso, em várias populações.

Objectivo: Investigar se a nova variante rs1333049, no cromossoma 9p21, é um factor de risco independente para DAC, na população Portuguesa.

Material e métodos: Estudo caso-controlo, que incluiu 1406 indivíduos, 723 doentes coronários internados consecutivamente (idade média de 53,7±8,9 anos 79,9% do sexo masculino) e 683 controlos sem doença coronária (idade média de 53,3±10,5 anos, 73,9 % do sexo masculino) seleccionados para não serem significativamente diferentes quanto ao sexo e idade. Estudou-se o SNP rs1333049, em todos os indivíduos, com recurso à técnica convencional de PCR combinada com a técnica TaqMan

Independent association of the variant rs1333049 at the 9p21 locus and coronary heart disease

ABSTRACT

Introduction: Recent genome-wide association studies have identified single-nucleotide polymorphisms (SNPs) at the 9p21 locus as risk factors for coronary artery disease (CAD). Among them, the SNP rs1333049 has demonstrated a consistent association with CAD, which has been successfully replicated in several populations.

Aim: To investigate whether the SNP rs1333049 located on the 9p21 chromosome is an independent risk factor for CAD in a Portuguese population.

Methods: We performed a case-control study which included 1406 individuals, 723 consecutive coronary patients (mean age 53.7±8.9 years, 79.9% male) and 683 controls without coronary disease (mean age 53.3±10.5 years, 73.9% male). Cases and controls were selected so as not to be significantly different in terms of gender and age. We studied the SNP

(Applied Biosystems). Determinou-se a distribuição alélica e genotípica (C/G), *odds ratio* e respectivo intervalo de confiança para risco de DAC.

Foi construído um modelo de regressão logística *forward wald* ajustado para a idade, sexo, factores de risco convencionais, marcadores bioquímicos e genótipos em estudo, afim de avaliar quais as variáveis associadas de forma significativa e independente com DAC.

Resultados: 60% dos doentes coronários e 53% dos controlos apresentaram o alelo C (OR=1,33; p=0,0002), 35,7% dos doentes e 29,3% dos controlos tinham o genótipo homocigoto CC (OR=1,34;p=0,010).

O heterocigoto CG estava presente em 48,1% dos doentes e 47% nos controlos, não atingindo significância estatística, para risco vascular (OR=1,05;p=0,670). Após análise multivariada de regressão logística o genótipo CC do cromossoma 9p21 ficou na equação com um OR=1,7, p=0,018 e o genótipo heterocigoto CG com um OR=1,5, p=0,048.

Conclusão: Com o presente trabalho replicou-se, numa população portuguesa, o risco coronário ligado à nova variante rs1333049 do cromossoma 9p21.

A robustez deste genótipo, tanto em homo como em heterocigotia, tem sido consistente e relevante na estratificação de risco para a DAC, mesmo em contextos populacionais muito diversos. Nestas circunstâncias, a utilização do genótipo CC ou CG poderá vir a revelar-se útil para a previsão do risco de DAC na nossa população.

Palavras-chave:

Factores de risco; Polimorfismos genéticos;
Doença arterial coronária; Locus 9p21;

rs1333049 at the 9p21 locus in all individuals, using standard PCR combined with the TaqMan technique (Applied Biosystems). The allelic and genotype distribution (C/G), odds ratios and corresponding confidence intervals for CAD risk were determined. A forward Wald logistic regression analysis model was constructed, adjusted for age, gender, conventional risk factors, biochemical markers and the genotypes under study, in order to determine which variables were linked significantly and independently with CAD.

Results: The C allele was found in 60% of the CAD patients and 53% of the controls, with OR=1.33; p=0.0002. The CC genotype appeared in 35.7% of CAD patients, with OR=1.34, p=0.010. The heterozygous CG genotype was present in 48.1% of the CAD patients and 47% of the controls, and did not present vascular risk (OR=1.05, p=0.670). After logistic regression analysis, the CC genotype remained in the equation with OR=1.7; p=0.018 and CG with OR=1.5, p=0.048.

Conclusion: In the present study we replicated the coronary risk linked to the recently discovered variant rs1333049 on the 9p21 chromosome in a Portuguese population. Although the mechanism underlying the risk is still unknown, the robustness of this risk allele in risk stratification for CAD has been consistent, even in very different populations. The presence of the CC or CG genotype may thus prove to be useful for predicting the risk of developing CAD in the Portuguese population.

Key words:

Risk factors; Genetic polymorphisms;
Coronary artery disease; Locus 9p21.

INTRODUÇÃO

A doença das artérias coronárias (DAC) é actualmente um dos maiores problemas de saúde no mundo e também no nosso país, devido à sua enorme prevalência e elevada mortalidade. Como foi demonstrado em estudos efectuados em gémeos e famílias, o desenvolvimento da DAC tem um fundo genético consistente^(1,2).

Recentemente a utilização de tecnologias de alta precisão na análise genética e os estudos de associação genómica em larga escala (GWAS), simultaneamente, identificaram um locus no cromossoma 9p21.3, que apresentou uma forte associação com a DAC^(3,4). Esta associação estende-se ao longo de uma região superior a 100 kilobases (Kb) na qual foram identificados vários polimorfismos de um único nucleótido (SNPs) sendo o rs1333049 o que apresentou a associação mais forte com a DAC. Esta associação é mais intensa a alguma distância dos genes mais próximos que codificam proteínas (CDKN2A, CDKN2B, e MTAP). A zona de interesse era atravessada, quase completamente, por um gene recentemente identificado, “antisense” e não codificante, denominado ANRIL.

O sinal de associação estende-se ao longo de um segmento onde não existem genes codificadores e o SNP com a associação mais intensa é o rs1333049 que se encontra aproximadamente a 100 Kb dos genes codificadores mais próximos, o CDKN2A que codifica a proteína p16^{INK4a} e CDKN2B que codifica a p15^{INK4B}, sendo a P14^{ARF} codificada por exões alternativos do gene CDKN2A e do gene MTAP, situado também nesta zona.

Estes genes, CDKN2A e CDKN2B, são considerados supressores tumorais e codificam proteínas inibidoras do ciclo celular, as quinases ciclina-dependentes (CDKs), que desempenham um papel importante na regulação do ciclo celular, podendo as mutações destes genes estar implicadas na patogénese da arteriosclerose e de certos tumores malignos como o melanoma^(5,6).

A demonstração de associação entre os nucleótidos de um único polimorfismo (SNPs) situados nos loci de risco coronário e a doença, em outras amostras independentes, é de interesse primordial para a validação dos resultados dos estudos.

INTRODUCTION

Coronary artery disease (CAD) is one of the greatest health problems worldwide and in Portugal, with extremely high prevalence and mortality. Twin and family studies have demonstrated that the development of CAD has a strong genetic component^(1,2).

High-precision genetic analysis techniques, together with genome-wide association studies, have recently identified a locus on chromosome 9p21.3 that is strongly associated with CAD^(3,4). This association extends across a region of over 100 kilobases (kb) on which several single-nucleotide polymorphisms (SNPs) have been identified, of which rs1333049 has the most consistent link with CAD. The association is strongest some distance from the closest protein-coding genes (*CDKN2A*, *CDKN2B*, and *MTAP*). The region of interest is almost completely spanned by a recently identified non-coding antisense gene denominated *ANRIL*.

The signal extends across a segment where there are no protein-coding genes; rs1333049 is almost 100 kb from the nearest genes, *CDKN2A* (encoding p16^{INK4a}) and *CDKN2B* (p15^{INK4b}); p14^{ARF} is encoded by alternative exons of *CDKN2A* and by *MTAP*, which is located in the same region. *CDKN2A* and *CDKN2B* code for cyclin-dependent kinase inhibitors that play an important role in regulation of the cell cycle; mutations in these genes have been implicated in the pathogenesis of arteriosclerosis and certain malignant tumors, including melanoma^(5,6).

In order to validate these studies, it is essential to replicate the association between SNPs in regions of coronary risk and disease in other populations. Lack of reproducibility is one of the main problems with genetic association studies, and results must be replicated in various populations in order to clarify the contribution to the pathogenic process of the factors under study^(7,8). Populations differ not only in genetic profile but also in their environments and lifestyles.

This study aims to investigate the association of the SNP rs1333049 located on the

A falta de reprodutibilidade é um dos grandes problemas nos estudos de associação e a replicação dos resultados em várias populações torna-se indispensável para clarificar a contribuição dos factores investigados no processo patogénico^(7, 8). As populações diferem não somente no perfil genético, mas também no meio ambiente que partilham e no estilo de vida. Deste modo, justifica-se replicar estes estudos de base genética em populações diferentes, com perfis genéticos e comportamentais diferentes.

Neste estudo propomo-nos investigar a associação do SNP rs1333049, localizado no locus do cromossoma 9p 21.3, com a DAC numa população Portuguesa.

MATERIAL E MÉTODOS

Estudo caso-controlo, que incluiu 1406 indivíduos, 723 doentes coronários consecutivos (idade média 53,7±8,9 anos, 79,9% sexo masculino) e 683 controlos sem doença coronária conhecida (idade média 53,3±10,5 anos, 73,9% do sexo masculino). Os doentes foram seleccionados consecutivamente das altas hospitalares (2002-2006) entre os admitidos por enfarte do miocárdio (EM) confirmado usando os critérios da *Joint European Society of Cardiology/American College of Cardiology Committee*⁽⁹⁾, ou doença coronária confirmada por coronariografia com pelo menos 75% de obstrução de uma artéria coronária principal. O grupo controlo foi constituído por voluntários saudáveis seleccionados das listas eleitorais, sem história sugestiva de DC. Para cada controlo efectuava-se uma observação clínica com referência aos factores de risco tradicionais, um ECG e em casos duvidosos, um teste de esforço e um eco-doppler cardíaco. Casos e controlos foram seleccionados de forma a não serem significativamente diferentes em termos de sexo e idade.

O uso das listas eleitorais para seleccionar os controlos tinha a finalidade de assegurar que não diferiam, de forma significativa, dos casos em termos de sexo e idade. Todos os indivíduos deram o seu consentimento informado. O protocolo do estudo teve a aprovação da Comissão de Ética do nosso Hospital.

9p21 chromosome with CAD in a Portuguese population.

METHODS

We performed a case-control study which included 1406 individuals, 723 consecutive coronary patients (mean age 53.7±8.9 years, 79.9% male) and 683 controls without coronary disease (mean age 53.3±10.5 years, 73.9% male). The cases were selected from patients discharged between 2002 and 2006 after being admitted for myocardial infarction diagnosed according to the criteria of the Joint European Society of Cardiology/American College of Cardiology Committee⁽⁹⁾, or with CAD confirmed by coronary angiography with at least 75% occlusion of a major coronary artery. The control group consisted of healthy volunteers selected from electoral rolls, with no history of CAD. All controls underwent clinical assessment of conventional cardiovascular risk factors, an ECG, and, in doubtful cases, exercise testing and cardiac Doppler echocardiography. Cases and controls were selected so as not to be significantly different in terms of gender and age.

Electoral rolls were used for selection of controls in order to ensure that they did not differ significantly from cases in terms of gender and age. Informed consent was obtained from all participants and the study protocol was approved by the hospital's ethics committee.

The data, which were stored in a standardized database, included demographic and clinical characteristics and conventional risk factors (gender, age, body mass index [BMI], diabetes, hypertension, family history, smoking, alcohol consumption and physical activity). Conventional risk factors were defined on the basis of previously published standard criteria⁽¹⁰⁻¹³⁾. All participants underwent biochemical tests and genetic analysis for the SNP rs1333049.

Genetic analysis

Genomic DNA was extracted from peripheral blood leukocytes using a standard salting-

Os dados foram guardados numa base “standard”, que compreendia dados demográficos, características clínicas e factores de risco clássicos (género, idade, índice de massa corporal, diabetes, hipertensão arterial, história familiar, hábito de fumar, de beber e tipo de actividade física). A definição dos factores de risco convencionais foi baseada nos critérios “standard” descritos previamente⁽¹⁰⁻¹³⁾. Foram efectuadas, em todos os participantes, análises bioquímicas e genéticas (SNP rs 1333049).

Análise genética

O ADN genómico foi extraído das células leucocitárias do sangue periférico usando o método standard de “salting-out”. A análise genotípica foi efectuada usando sonda oligonucleotídica marcada com fluorescência específica para cada um dos alelos num ensaio que combina a técnica convencional de PCR e a técnica TaqMan (Applied Biosystems). Os “primers” e as sondas eram as pré-estabelecidas pelo fornecedor (*TaqMan SNP Genotyping Assays, Applied Biosystems*) para o polimorfismo a analisar (SNP rs1333049). Os oligonucleotídeos foram sintetizados e os marcadores fluorogénicos FAM e VIC acoplados às extremidades 5’ das sondas de modo a alcançar a discriminação alélica. A reacção de polimerização em 2 passos consistiu, de uma forma geral, em 40 ciclos de desnaturação a 92°C durante 15 segundos e de “primer annealing” e extensão a 60°C durante 1 minuto e foi efectuada num aparelho 7300 Real-Time PCR System (Applied Biosystems). Os genótipos foram determinados recorrendo ao Software (*Applied Biosystems*) sem qualquer conhecimento prévio dos dados clínicos individuais (Fig. 2).

out method. For the genetic analysis, oligonucleotide probes with fluorescent tags specific to each allele were used in an assay that combines conventional PCR and the TaqMan technique (Applied Biosystems). The primers and probes used were as specified by the manufacturer (TaqMan SNP Genotyping Assays, Applied Biosystems) for the SNP under analysis. The oligonucleotides were synthesized and the FAM and VIC tags were coupled to the 5’ end of the probe in order to discriminate the allele. A two-stage polymerase reaction was used, consisting of 40 cycles of denaturation at 92 °C for 15s and primer annealing and extension at 60 °C for 1 min in a 7300 Real-Time PCR System (Applied Biosystems). The genotypes were determined using Applied Biosystems software with no previous knowledge of the participants’ clinical data (*Figure 2*).

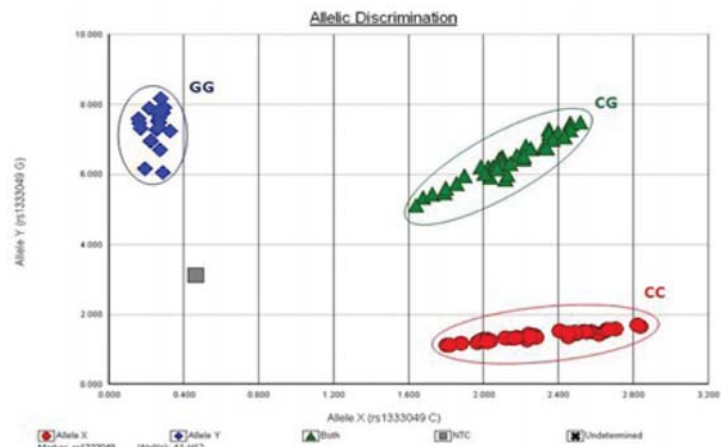
Statistical analysis

Data are expressed as percentages or means \pm standard deviation, except in the case of outliers in continuous variables, when they are presented as medians or a logarithmic transformation was performed.

Allele and genotype frequencies are calculated by gene counts and compared between cases and controls by the chi-square test. Pearson’s chi-square test was used to test deviation from Hardy-Weinberg equilibrium, based on observed and expected genotype frequencies^(14, 15). The Student’s t test was used for continuous variables to compare means between groups; when these had a non-normal

Figura 2. Discriminação genotípica do SNP rs1333049. Aparelho 7300 Real Time PCR System (Applied Biosystems)

Figure 2. Genotyping of the SNP rs1333049 in a 7300 Real-Time PCR System (Applied Biosystems).



Análise Estatística

Os dados são expressos em percentagem (%) ou média \pm desvio padrão, mas quando as variáveis contínuas apresentavam valores extremos (*outliers*), os dados são apresentados pela mediana ou com transformação logarítmica.

As frequências alélicas e genótípicas são calculadas pela contagem dos genes e comparadas, entre casos e controlos, pelo teste do χ^2 quadrado. Para testar o desvio do Equilíbrio de Hardy-Weinberg (EHW) foi efectuado o teste do χ^2 de Pearson, usando as frequências genótípicas observadas e as esperadas⁽¹⁴⁻¹⁵⁾. Para comparar as médias entre grupos foi usado o teste *t de Student*, para as variáveis contínuas, e sempre que as mesmas não apresentavam uma distribuição normal, foi-lhes aplicada uma transformação logarítmica ou usados testes não paramétricos (*Kruskal Wallis*). Os resíduos, são definidos como os valores observados menos os valores preditos, baseados em factores de “confundimento”, sendo usados na análise de associação genótipo-fenótipo (análise de variância one way ANOVA). Para testar a associação, expressa em coeficiente β por alelo C, entre o alelo C do polimorfismo rs1222049 e as características basais (expressas como variáveis contínuas e independentes), pressão arterial sistólica e diastólica, diabetes, IMC, alcoolemia, colesterol total, LDL e HDL, fibrinogénio, Apo B, Lp(a), PCR (as) e VOP, foi usado um modelo de regressão linear.

A força da associação foi expressa pelo *odds ratio* (OR) assim como pelos intervalos de confiança (IC) a 95%. Um valor de $p < 0,05$, foi considerado estatisticamente significativo. Para determinar, em análise multivariada, os factores associados com DC, foi construído um modelo de regressão logística pelo método *forward wald*, após ajuste para idade, sexo e factores de risco clássicos. Foi utilizado o software estatístico SPSS versão 14.0.

RESULTADOS

O Quadro I apresenta diferenças significativas entre os dois grupos, pois as percentagens de diabetes tipo2, hipertensão arterial (HTA),

distribuição a logarítmica transformação was performed or a non-parametric test (*Kruskal-Wallis*) was applied. Residuals were defined as observed minus expected values based on confounding factors and were used in genotype-phenotype association analyses (one-way ANOVA). A linear regression model was used to test the association (expressed as the β coefficient for the C allele) between the C allele of the rs1222049 polymorphism and baseline characteristics (expressed as continuous and independent variables): systolic and diastolic blood pressure (BP), diabetes, BMI, blood alcohol, total, LDL and HDL cholesterol, fibrinogen, apolipoprotein B, lipoprotein(a), high-sensitivity C-reactive protein (CRP) and pulse wave pressure.

The strength of the association was expressed as odds ratios (OR) and 95% confidence intervals (CI). A value of $p < 0.05$ was considered statistically significant. A forward Wald logistic regression model was constructed, adjusted for age, gender and conventional risk factors, to determine the factors associated with CAD in multivariate analysis. The statistical software used was SPSS version 14.0.

RESULTS

Table I shows that there were significant differences between the groups, the prevalences of diabetes, hypertension, obesity, dyslipidemia and smoking being significantly higher in cases than in controls. Blood glucose, blood alcohol, triglycerides and CRP were also significantly higher in the coronary patients, while plasma HDL cholesterol was higher in controls. Systolic and diastolic BP were significantly higher in cases, and fibrinogen was also slightly higher, but without statistical significance.

The frequency of the CC genotype of the SNP rs1333049 was 35.7% in cases and 29.3% in controls, with OR 1.34; 95% CI 1.06-1.69; $p = 0.010$, while the GG (unmutated) genotype had a significant protective effect (OR 0.62; 95% CI 0.47-0.82; $p = 0.0004$). The presence of the C allele, in 60% of cases and

obesidade, dislipidemia, tabagismo e consumo de álcool são mais elevadas, de forma significativa, nos doentes em relação aos controlos. Os níveis plasmáticos de glicemia, triglicéridos e PCR (as) são mais elevados, com significância estatística, nos doentes coronários e os níveis plasmáticos de colesterol HDL, são mais elevados nos controlos. Os valores de pressão arterial tanto sistólica como diastólica, são mais elevados nos doentes, de forma significativa, mas o fibrinogénio foi tendencialmente mais elevado nos doentes, sem, no entanto, apresentar significância estatística.

53% of controls, was significantly associated with CAD (OR 1.33; 95% CI 1.14-1.55, p=0.0002).

Application of Pearson's chi-square test to the observed and expected genotype frequencies gave a value of 0.022 for the coronary patients and 2.22 for controls. As the 5% level of significance for one degree of freedom is 3.84 and the value obtained was less than this figure in both cases and controls, the null hypothesis (H_0) was rejected and hence the population is in Hardy-Weinberg equilibrium.

One-way ANOVA showed no significant

Quadro I. Características basais da população

Características	Total (n=1406)	Casos (n=723)	Controlos (n=683)	Valor-p
Idade (anos)	53,5 ± 9,7	53,7 ± 8,9	53,3 ± 10,5	0,387
Sexo M (%)	1083(77,0%)	578 (79,9%)	505 (73,9%)	0,007
HTA (%)	704 (50,1%)	416 (57,5%)	288 (42,2%)	<0,0001
Diabetes (%)	324 (23%)	240 (33,2%)	324 (23%)	<0,0001
Obesidade (%)	397(28,2%)	231(32,0%)	166 (24,3%)	<0,001
Tabagismo (%)	484 (34,4%)	325 (45,0%)	159 (23,3%)	<0,0001
Consumo de álcool (gr/dia)	0,8	1,6	0,0	<0,0001
Dislipidemia (%)	1246 (88,6%)	694 (96,0%)	552 (80,8%)	<0,0001
Glicemia (mg/dl)	102,0	107,0	98,0	<0,0001
Triglicéridos (mg/dl)	132,0	152,0	115,0	<0,0001
HDL (mg/dl)	44,6 ± 14,6	38,5 ± 10,3	51,2 ± 15,7	<0,0001
PA sistólica (mm Hg)	135,5 ± 20,4	136,4 ± 20,2	134,4 ± 20,6	0,040
PA diastólica (mm Hg)	81,2 ± 11,3	81,7 ± 11,2	80,8 ± 11,5	0,206
Fibrinogénio (mm Hg)	348,9±94,1	344,9±96,6	341,1±96,7	0,397
Lp(a) (mg/dl)	16,5	20,8	15,2	0,010
PCR(as) mg/dl	2,0	3,0	1,6	<0,0001
VOP (m/seg.)	10,3 ± 2,7	10,5 ± 2,3	10,1 ± 3,1	0,007

HTA- Hipertensão Arterial, Lp (a)-Lipoproteína (a), HDL- High Density Lipoprotein, PCR(as)-Proteína C Reactiva de alta sensibilidade, VOP- Velocidade da Onda de Pulso.

Table I. Demographic and clinical characteristics of the population

	Total (n=1406)	Cases (n=723)	Controls (n=683)	p
Age (years)	53.5±9.7	53.7±8.9	53.3±10.5	0.387
Male (%)	1083(77.0%)	578 (79.9%)	505 (73.9%)	0.007
Hypertension (%)	704 (50.1%)	416 (57.5%)	288 (42.2%)	<0.0001
Diabetes (%)	324 (23%)	240 (33.2%)	324 (23%)	<0.0001
Obesity (%)	397(28.2%)	231(32.0%)	166 (24.3%)	<0.001
Smoking (%)	484 (34.4%)	325 (45.0%)	159 (23.3%)	<0.0001
Alcohol consumption (g/day)	0.8	1.6	0.0	<0.0001
Dyslipidemia (%)	1246 (88.6%)	694 (96.0%)	552 (80.8%)	<0.0001
Blood glucose (mg/dl)	102.0	107.0	98.0	<0.0001
Triglycerides (mg/dl)	132.0	152.0	115.0	<0.0001
HDL cholesterol (mg/dl)	44.6±14.6	38.5±10.3	51.2±15.7	<0.0001
Sistolic BP (mmHg)	135.5±20.4	136.4±20.2	134.4±20.6	0.040
Diastolic BP (mmHg)	81.2±11.3	81.7±11.2	80.8±11.5	0.206
Fibrinogen (mg/dl)	348.9±94.1	344.9±96.6	341.1±96.7	0.397
Lipoprotein(a) (mg/dl)	16.5	20.8	15.2	0.010
hs-CRP (mg/dl)	2.0	3.0	1.6	<0.0001
PWV (m/s)	10.3±2.7	10.5±2.3	10.1±3.1	0.007

BP: blood pressure; HDL: high-density lipoprotein; hs-CRP: high-sensitivity C-reactive protein; PWV: pulse wave velocity

O genótipo CC do SNP rs1333049 apresentou uma frequência de 35,7% nos casos e de 29,3% nos controlos com um *odds ratio* (OR) de 1,34; IC 1,06-1,69; p=0,010 e o GG, não mutado, foi protector, de forma significativa, (OR 0,62; IC 0,47-0,42; P=0,0004). A presença do alelo C, com 60% nos casos e 53% nos controlos, associou-se de forma significativa com a DC (*Odds ratio*=1,33; IC 1,14-1,55, p=0,0002).

Após ter sido efectuado o teste do χ^2 de Pearson usando as frequências genóticas observadas e as esperadas, o valor foi de 0,022 para os doentes coronários e 2,22 para os controlos. Sendo o nível de significância de 5% para 1 grau de liberdade = 3,84 e como o valor obtido é inferior a 3,84, tanto nos casos como nos controlos, a hipótese nula (H₀) foi rejeitada logo a população está em equilíbrio de Hardy-Weinberg (Quadro II).

association between the CC, CG and GG genotypes of the SNP rs1333049 and demographic variables, phenotypic characteristics and biochemical markers of our population (Table III).

Linear regression analysis revealed no significant association between the C allele and risk factors or biochemical markers (Table IV).

After multivariate logistic regression analysis, the CC genotype remained in the equation with OR 1.711; 95% CI 1.097-2.668; p=0.018, and the heterozygous genotype CG with OR 1.530; 95% CI: 1.005-2.330; p=0.048, demonstrating that both are independent risk factors for CAD (Table V).

DISCUSSION

The risk allele (C) of the SNP rs1333049, which is very common, has been investigated

Quadro II. Frequência genotípica, alélica associação do SNP rs1333049 com DC, na população portuguesa. Equilíbrio de Hardy Weinberg

SNP rs1333049	Casos (n=723)	Controlos (n=683)	OR (95% IC)	Valor-p
CC	258 (35,7%)	200 (29,3%)	1,34 (1,06 – 1,69)	0,010
CG	348 (48,1%)	321 (47,0%)	1,05 (1,84 – 1,30)	0,670
GG	117 (16,2%)	162 (23,7%)	0,62 (0,47 – 0,82)	0,0004
Frequência alélica C	0,60	0,53	1,33 (1,14 – 1,55)	0,0002
Frequência alélica G	0,40	0,37		
HWE (χ^2)	0,022	2,228	1,504 (toda a população)	

HWE- Equilíbrio de Hardy- Weinberg.

Table II. Genotype frequency, allele association of the SNP rs1333049 with CAD, and Hardy-Weinberg equilibrium

SNP rs1333049	Cases (n=723)	Controls (n=683)	OR (95% CI)	p
CC	258 (35.7%)	200 (29.3%)	1.34 (1.06 – 1.69)	0.010
CG	348 (48.1%)	321 (47.0%)	1.05 (1.84 – 1.30)	0.670
GG	117 (16.2%)	162 (23.7%)	0.62 (0.47 – 0.82)	0.0004
C allele frequency	0.60	0.53	1.33 (1.14 – 1.55)	0.0002
G allele frequency	0.40	0.37		
HWE (chi-square test)	0.022	2.228	1.504 (total population)	

HWE: Hardy-Weinberg equilibrium

Após análise do Quadro III (análise de variância *one way ANOVA*), não houve associação significativa entre os genótipos CC, CG e GG do SNPI333049 e as variáveis demográficas, características fenotípicas e marcadores bioquímicos da nossa população (Quadro III).

in a white population in a genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls published in *Nature* in 2007⁽¹⁶⁾. This study investigated other SNPs on chromosome 9p21, with the strongest signal for association with CAD

Quadro III. Características fenotípicas da população e sua associação com os genótipos (CC, CG e GG) do SNP1333049.

Características Fenotípicas	SNP1333049			Valor p
	CC	CG	GG	
Idade (Anos)	53,9 ± 9,9	53,4 ± 9,7	54,0 ± 9,3	0,710
Sexo M %	354 (77,3%)	515 (77,0%)	214 (76,7%)	0,982
Diabetes %	104 (22,7%)	153 (22,9%)	67 (24,0%)	0,910
HTA %	236 (51,5%)	325 (48,6%)	143 (51,3%)	0,565
Dislipidemia %	411 (89,7)	595 (88,9%)	240 (86,0%)	0,286
Hábito fumar %	158 (34,5%)	236 (35,3%)	90 (32,3%)	0,672
Obesidade %	134 (29,3%)	188 (28,1%)	75 (26,9%)	0,781
Álcool gr/dia	22,5 ± 56,1	26,4 ± 64,9	21,9 ± 55,5	0,784
Log Colest. Total	2,29 ± 0,1	2,30 ± 0,1	2,31 ± 0,1	0,605
LDL mg/dl	117,4 ± 39,4	117,7 ± 38,7	119,9 ± 40,8	0,708
HDL mg/dl	44,0 ± 13,8	44,2 ± 14,3	46,7 ± 16,5	0,101
Triglicéridos mg/dl	164,9 ± 124,3	164,5 ± 123,7	170,5 ± 148,2	0,933
Glicemia mg/dl	115,2 ± 45,4	115,4 ± 44,5	117,3 ± 47,6	0,951
Fibrinogénio mg/dl	348,9 ± 94,1	344,9 ± 96,6	341,1 ± 96,7	0,397
PCR (as) mg/dl	3,7 ± 6,1	3,8 ± 5,3	3,7 ± 5,6	0,488
√Apo B	10,3 ± 1,4	10,3 ± 1,3	10,3 ± 10,4	0,274
Lp(a) mg/dl	31,3 ± 37,0	31,43 ± 31,4	27,4 ± 28,8	0,314

Apo B100-Lipoproteína B 100, HTA-Hipertensão Arterial, Lp(a) -Lipoproteína (a), HDL- High Density Lipoprotein, LDL-Low Density Lipoprotein, PCR(as)-Proteína C Reactiva de alta sensibilidade.

Table III. Demographic variables, phenotypic characteristics and biochemical markers of the study population and their association with the CC, CG and GG genotypes of the SNP rs1333049

	SNP rs1333049			p
	CC	CG	GG	
Age (years)	53.9±9.9	53.4±9.7	54.0±9.3	0.710
Male (%)	354 (77.3%)	515 (77.0%)	214 (76.7%)	0.982
Diabetes %	104 (22.7%)	153 (22.9%)	67 (24.0%)	0.910
Hypertension (%)	236 (51.5%)	325 (48.6%)	143 (51.3%)	0.565
Dyslipidemia (%)	411 (89.7)	595 (88.9%)	240 (86.0%)	0.286
Smoking (%)	158 (34.5%)	236 (35.3%)	90 (32.3%)	0.672
Obesity %	134 (29.3%)	188 (28.1%)	75 (26.9%)	0.781
Alcohol consumption (g/day)	22.5±56.1	26.4±64.9	21.9±55.5	0.784
Log total cholesterol	2.29±0.1	2.30±0.1	2.31±0.1	0.605
HDL cholesterol (mg/dl)	117.4±39.4	117.7±38.7	119.9±40.8	0.708
HDL cholesterol (mg/dl)	44.0±13.8	44.2±14.3	46.7±16.5	0.101
Triglycerides (mg/dl)	164.9±124.3	164.5±123.7	170.5±148.2	0.933
Blood glucose (mg/dl)	115.2±45.4	115.4±44.5	117.3±47.6	0.951
Fibrinogen (mg/dl)	348.9±94.1	344.9±96.6	341.1±96.7	0.397
hs-CRP (mg/dl)	3.7±6.1	3.8±5.3	3.7±5.6	0.488
√Apo B	10.3±1.4	10.3±1.3	10.3±10.4	0.274
Lipoprotein(a) (mg/dl)	31.3±37.0	31.43±31.4	27.4±28.8	0.314

Apo B: apolipoprotein B-100; HDL: high-density lipoprotein; hs-CRP: high-sensitivity C-reactive protein; LDL: low-density lipoprotein

Quadro IV. Associação do alelo C com as variáveis contínuas (expressa em coeficiente β por alelo C)

Variáveis	Coefficiente-β	Valor p
PAS	-1,366	0,261
PAD	-0,501	0,457
IMC	-0,017	0,960
Glicemia	0,168	0,951
Álcool	3,912	0,294
Log Colesterol Total	0,001	0,868
LDL	0,261	0,912
HDL	0,228	0,790
Triglicéridos	-0,380	0,960
Fibrinogénio	-0,476	0,496
Lp(a)	0,120	0,963
PCR (as)	0,113	0,794
√Apo B	0,069	0,508
VOP	0,135	0,418

PCR (as)-Proteína C reactiva de alta sensibilidade; Lp(a)- Lipoproteína a; Apo B – Raiz Quadrada da Apolipoproteína B100; VOP- Velocidade da Onda de Pulso; Log- Logaritmo; PAS- Pressão Arterial Sistólica; PAD- Pressão Arterial diastólica; IMC- Índice de Massa Corporal.

Table IV. Association of the C allele with continuous variables (expressed as β coefficient for the C allele)

	β coefficient	p
SBP	-1.366	0.261
DBP	-0.501	0.457
BMI	-0.017	0.960
Blood glucose	0.168	0.951
Alcohol consumption	3.912	0.294
Log total cholesterol	0.001	0.868
LDL cholesterol	0.261	0.912
HDL cholesterol	0.228	0.790
Triglycerides	-0.380	0.960
Fibrinogen	-0.476	0.496
Lipoprotein(a)	0.120	0.963
hs-CRP	0.113	0.794
√Apo B	0.069	0.508
PWW	0.135	0.418

Apo B: apolipoprotein B-100; BMI: body mass index; DBP: diastolic blood pressure; HDL: high-density lipoprotein; hs-CRP: high-sensitivity C-reactive protein; LDL: low-density lipoprotein; PWW: pulse wave velocity; SBP: systolic blood pressure

Após análise do Quadro IV (Regressão linear), não se verificou associação significativa entre o alelo C do SNP rs1333049 e as variáveis fenotípicas, consideradas factores de risco, ou os marcadores bioquímicos.

being seen at rs1333049 ($p=1.8 \times 10^{-14}$). As stated above, an association was found across a region of over 100 kb that had not been highlighted in previous studies and that contains the coding sequences of two genes,

Quadro V. Variáveis associadas de forma independente com DC

Variáveis	B	S.E.	Wald	df	Odds Ratio (IC 95%)	Valor-p
Sexo	0,329	0,191	3,140	1	1,404 (0,965-2,042)	0,076
Diabetes	0,922	0,204	21,343	1	2,513 (1,684-3,751)	<0,0001
Hábito fumar	0,857	0,175	23,989	1	2,356 (1,672-3,319)	<0,0001
Dislipidemia	1,639	0,323	25,802	1	5,151 (2,737-9,695)	<0,0001
PCR(as)	0,047	0,017	7,802	1	1,048 (1,014-1,084)	0,005
Lp (a)	0,007	0,002	9,728	1	1,008 (1,003-1,012)	0,002
Fibrinogénio	0,04	0,001	12,284	1	1,004 (1,002-1,006)	<0,0001
Rs1333049			5,865	2		0,040
CC	0,537	0,227	5,612	1	1,711 (1,097-2,668)	0,018
CG	0,456	0,226	4,069	1	1,530 (1,005-2,330)	0,048
Constante	-5,238	0,647	65,535	1		<0,0001

Entraram na equação: idade, sexo, HTA, dislipidemia, diabetes, hábito de fumar, obesidade, PCR(as), fibrinogénio, Lp(a), SNP rs133349. B – Coeficiente estimado, com erro padrão S.E.; df – Graus de liberdade; Sig. – Significância estatística ($p < 0,05$).

Table V. Variables independently associated with CAD

	β	SE	Wald	df	Odds Ratio (95% CI)	p
Gender	0.329	0.191	3.140	1	1.404 (0.965-2.042)	0.076
Diabetes	0.922	0.204	21.343	1	2.513 (1.684-3.751)	<0.0001
Smoking	0.857	0.175	23.989	1	2.356 (1.672-3.319)	<0.0001
Dyslipidemia	1.639	0.323	25.802	1	5.151 (2.737-9.695)	<0.0001
hs-CRP	0.047	0.017	7.802	1	1.048 (1.014-1.084)	0.005
Lipoprotein(a)	0.007	0.002	9.728	1	1.008 (1.003-1.012)	0.002
Fibrinogen	0.04	0.001	12.284	1	1.004 (1.002-1.006)	<0.0001
rs1333049			5.865	2		0.040
CC	0.537	0.227	5.612	1	1.711 (1.097-2.668)	0.018
CG	0.456	0.226	4.069	1	1.530 (1.005-2.330)	0.048
Constant	-5.238	0.647	65.535	1		<0.0001

The following variables were included in the equation: age, gender, hypertension, dyslipidemia, diabetes, smoking, obesity, hs-CRP, fibrinogen, lipoprotein(a), and rs133349. β : estimated coefficient; SE: standard error; df: degrees of freedom

Após análise multivariada de regressão logística (Quadro V), o genótipo CC ficou na equação com um OR de 1,711; IC-1,097-2,668; $p=0,018$, e o genótipo heterozigoto CG com um OR de 1,530; IC: 1,005-2,330; $p=0,048$, mostrando serem ambos factores de risco independentes para DAC.

DISCUSSÃO

O alelo de risco (C) do SNP rs1333049, extremamente frequente, foi investigado na população Caucasiana, no Genome-Wide Association Study (Study of 14,000 cases of seven common diseases and 3,000 shared con-

CDKN2A and CDKN2B, which code for two cyclin-dependent kinase inhibitors that inhibit the cell cycle. Both genes have multiple isoforms, have an important role in the regulation of the cell cycle, and are widely expressed in many tissues. The normal progression of the cell cycle is mediated by the activity of the cyclin-dependent protein kinases CDK4 and CDK6, which are inhibited by the proteins P16^{INK4a} and p15^{INK4b}. Transforming growth factor β (TGF- β) induces the expression of p15, which inhibits cell proliferation, and thereby induces the G1 phase of the cell cycle (arrest)^(17, 18). Deletion of regions harboring genes that code for proteins that suppress the cell cycle may increase expression of genes

trols. Wellcome Trust Case Control Consortium) e publicado na revista *Nature* em 2007⁽¹⁶⁾. Neste estudo foram investigados outros SNPs na área do Cromossoma 9p21 tendo o rs1333049 mostrado o sinal mais forte de associação com DC. Como já foi dito, foi encontrada associação ao longo de uma região, que ainda não tinha sido assinalada, de mais de 100 kilobases (Kb) que continha sequências de genes codificantes, inibidores do ciclo celular, os inibidores dependentes das duas ciclinas cinases, *CDKN2A* e *CDNK2B*. Ambos os genes são supressores do ciclo celular, têm múltiplas isoformas e apresentam um papel importante na regulação deste ciclo, expressando-se em muitos tecidos. Como sabemos, a progressão normal do ciclo celular é promovida pela actividade das ciclinas dependentes das proteína kinases CDK4 e CDK6, que são inibidas pelas proteínas P16^{INK4a} e p15^{INK4b}. O factor de transformação do crescimento β (TGF β) ao induzir a expressão da proteína p15 que por sua vez inibe a proliferação celular induz a fase-G1 (paragem do ciclo celular)⁽¹⁷⁻¹⁸⁾. A deleção de zonas portadoras dos genes que codificam estas proteínas supressoras do ciclo celular, podem originar expressão aumentada dos genes que promovem a proliferação celular e assim o aparecimento de aterosclerose e cancro. Além do *CDKN2A* e do *CDKN2B*, o único gene conhecido nas proximidades é o *MTAP* que codifica a metil-tioadenosina fosforilase (*MTAP*), uma enzima que contribui para o metabolismo da poliamina e se pode expressar em muitos tecidos humanos inclusivamente no sistema cardiovascular^(19, 20).

O locus 9p21 é uma região genética particular, pois é bem conhecida dos especialistas em cancro genético por ser um dos locais que apresenta maior número de deleções, que conduzem a formas hereditárias de melanoma maligno⁽²¹⁾ e os investigadores que se têm dedicado ao estudo deste locus de risco, têm procurado uma explicação fisiopatológica lógica para a sua forte associação com a DAC.

Grande esforço tem sido efectuado para identificar as variantes genéticas codificadoras com repercussão funcional, nesse local.

that promote cell proliferation and thus atherosclerosis and cancer. Besides *CDKN2A* and *CDKN2B*, the only other known gene nearby is *MTAP*, which encodes methylthioadenosine phosphorylase, an enzyme that contributes to polyamine metabolism and is expressed in many human tissues, including the cardiovascular system^(19, 20).

The 9p21 locus is well known to cancer specialists as presenting a particularly large number of deletions that lead to hereditary forms of malignant melanoma⁽²¹⁾. Those investigating this risk locus have also sought a pathophysiological explanation for its strong association with CAD, and considerable effort has gone into identifying genetic variants at this site that have functional repercussions. Studies by Horne et al. suggest that the primary effect is on the vessel wall, and again focus attention on *CDKN2A* and *CDNK2B*, which code for two cyclin-dependent kinase inhibitors that play an important role in regulation of the cell cycle and may be implicated in the pathogenesis of atherosclerosis⁽²²⁾.

At the *INK4* locus there is a large antisense gene for non-coding RNA, 20 exons long, that spans the 58-kb region of coronary risk, with multiple transcripts and unknown functions, designated *ANRIL*. Studies by Holdt et al.⁽²³⁾ demonstrated that *ANRIL* is expressed in various tissues and cells affected by atherosclerosis. Transcript-specific primers yield short (DQ485454) or long (DQ485453) transcripts; expression of the short transcripts was significantly increased in carriers homozygous for the C allele, while the same allele was associated with a decrease of expression of the long *ANRIL* variant. Increased expression of the long variant was associated with increased expression of genes that regulate cell proliferation, particularly *CDKN2A* (p14 and p16) and *CDKN2B* (p15).

According to Holdt et al., alleles in the risk region of chromosome 9p21 may affect the expression of *ANRIL* but may also alter the expression of *CDKN2A* and *CDKN2B* through an effect mediated by *ANRIL* itself. Recent studies by Jarinova et al.⁽²⁴⁾ also showed that the risk allele (C) of the rs1333049 polymor-

Os estudos de Horne et al.⁽²²⁾ sugerem um efeito primário na parede do vaso e focam novamente a atenção sobre os genes inibidores dependentes das ciclinas cinase, *CDKN2A* e *CDKN2B*, que desempenham um papel importante na regulação do ciclo celular e podem estar implicados na patogénia da aterosclerose.

Existe também nesta zona, no locus *INK4*, um gene *antisense*, não codificante RNA, longo com 20 exões e que ultrapassa os 58 kb do intervalo de risco coronário, que é transcrito em múltiplas cópias e com funções desconhecidas, designado por *ANRIL*. Estudos de Lesca Holdt e col.⁽²³⁾ demonstraram que o *ANRIL* se expressa em vários tecidos e células afectadas pela aterosclerose, mas baseando-se nos *primers* utilizados, podiam obter uma expressão de transcritos de *ANRIL* com formas curtas (DQ4854549) ou formas longas (DQ485453). O nível de expressão de transcritos curtos estava aumentado de forma significativa nos portadores do alelo de risco C em homozigotia e o mesmo alelo associava-se com a redução da expressão da variante longa do *ANRIL*. Assim o aumento de expressão da variante longa associava-se com níveis elevados de expressão dos genes que controlam a proliferação celular, nomeadamente o *CDKN2A* (p14 e p16) e o *CDKN2B* (p15).

Segundo estes autores, os alelos, no intervalo de risco do cromossoma 9p21, poderiam alterar a expressão do *ANRIL*, mas também poderiam alterar a expressão dos genes *CDKN2A* e *CDKN2B*, através de um efeito mediado pelo próprio gene *ANRIL*. Estudos recentes de Olga Jarinova et al.⁽²⁴⁾ também demonstraram que o alelo de risco (C) do polimorfismo 1333349 se associava com aumento da expressão das variantes curtas do gene *ANRIL* e reduzia a expressão da variante longa. A expressão no sangue total da variante longa do *ANRIL* está positivamente relacionada com a expressão do *CDKN2B* (supressor do ciclo celular) e o alelo de risco C está associado com uma expressão baixa de *CDKN2B* e *upregulation* dos grupos de genes envolvidos na proliferação celular.

Estes SNPs de risco cardiovascular situados dentro dos 58Kb desta região, podem

phism was associated with increased expression of the short *ANRIL* transcripts and decreased expression of the long form. Whole blood expression of the long variant of *ANRIL* was positively correlated with that of *CDKN2B*, an inhibitor of the cell cycle, and the C (risk) allele was associated with low expression of *CDKN2B* and upregulation of gene sets modulating cellular proliferation.

The SNPs associated with cardiovascular risk within this 58-kb region may orchestrate a series of molecular events that promote a proliferative response. They are associated with changes in expression of *ANRIL* transcription and of genes involved in regulation of the cell cycle.

There is evidence of an association with type 2 diabetes in the same region, although different SNPs are involved⁽²⁵⁾.

The link with CAD has been found in genome-wide association studies by other investigators^(26, 27).

In the present study, on a sample of over 1000 participants, the SNP rs1333049 at the 9p21 locus was found in homozygosity (CC) in around 30% of healthy controls. The C allele presented an increased risk for CAD with an odds ratio of 1.7 in homozygosity and 1.5 in heterozygosity. These results are generally in agreement with studies in the literature on various populations.

According to McPherson et al.⁽²⁸⁾, the C allele of rs1333049 is extremely common, since over 25% of white individuals are homozygous for the risk allele (CC) and have double the risk of early-onset CAD compared to those without this allele. This effect is independent of established risk factors including dyslipidemia, hypertension, diabetes, obesity and inflammatory markers and may indicate a new and important biological mechanism in the pathophysiology of CAD.

Samani et al. studied rs1333049 in the Wellcome Trust Case Control Consortium (WTCCC) dataset and found a frequency of homozygous CC carriers of 30.5% in cases and 23.0% in controls; for heterozygous CG individuals the figures were 49.9% in coronary patients and 48.7% in controls (OR 1.9

orquestrar uma série de eventos moleculares que promovem a resposta proliferativa e estão associados com alterações na expressão individual da transcrição do ANRIL e mudanças na expressão dos genes reguladores do ciclo celular.

Na mesma região existem evidências de associação com a diabetes tipo 2 mas são outros grupos de SNPs que estão envolvidos⁽²⁵⁾.

A associação com a DAC foi também encontrada através de GWAS, por outros grupos de investigadores^(26, 27).

No presente estudo, recorrendo a uma amostra de mais de 1000 participantes, a variante rs13333049 no locus 9p21 apresentou-se em homozigotia CC em cerca de 30% dos controlos saudáveis.

Evidenciou um risco acrescido de aparecimento de DAC (odds ratio de 1,7 em homozigotia ou de 1,5 em heterozigotia).

Estes resultados, de um modo geral estão de acordo com o descrito na literatura, em várias populações estudadas.

Segundo Mac Pherson et al⁽²⁸⁾, o rs1333049 é um polimorfismo extremamente comum, pois mais de 25% da população Caucasiana transporta os dois alelos de risco (homozigotia CC), significando um risco de início precoce de DAC duas vezes superior ao risco dos indivíduos que não transportam esses alelos.

Este efeito, no risco, é independente dos conhecidos factores de risco incluindo dislipidémia, hipertensão, diabetes, obesidade e marcadores de inflamação podendo estar implicado um novo mecanismo biológico muito relevante, na fisiopatologia da DAC.

Samani e al identificaram o locus cromossómico 9p21, fortemente associado com DAC, no Wellcome Trust Case Control Consortium (WTCCC) study encontrando para o *SNP rs1333049* uma frequência populacional em homozigotia CC de 30,5% nos casos e 23,0% nos controlos. Quando em heterozigotia CG a frequência populacional era 49,9% nos doentes coronários e 48,7% nos controlos, com um OR em homozigotia de 1,90 e em heterozigotia de 1,47 ($p=1.1 \times 10^{-14}$)^(29, 30).

No estudo do grupo do *Cardiogenics Consortium* liderado por Heribert Schunkert,

for homozygosity and 1.47 for heterozygosity) ($p=1.1 \times 10^{-14}$)^(29, 30).

In a study by the Cardiogenics Consortium group headed by Heribert Schunkert, after adjustment for covariates, rs1333049 was associated with CAD with OR 1.32, 95% CI 1.24-1.42. In a 2008 meta-analysis by the same group, of seven case-control studies with a total of 4654 cases and 5177 controls, the SNP rs1333049 was genotyped; the risk allele was associated with CAD in all seven studies, with OR for heterozygosity of 1.3 and for homozygosity of 1.5⁽³¹⁾.

In 2008, this association was reproduced in Japanese and Korean populations. In the Japanese population, the CC genotype was associated with CAD, with OR 1.44 (95% CI 1.15-1.80; $p=0.0013$) and the C allele with OR 1.30 (95% CI 1.13-1.49; $p=0.00027$), while in the Korean population, the CC genotype presented an OR of 1.28 (95% CI 1.00-1.63; $p=0.049$) and the C allele an OR of 1.19 (95% CI 1.02-1.38; $p=0.025$)⁽³²⁾.

In the Portuguese population investigated in the present study, the risk allele was very common, with a frequency in homozygosity of 35.7% in cases and 29.3% in controls, higher than the figures reported by various previous studies. It can thus be assumed that this genotype is at least as important in the Portuguese population as in other studied populations.

Over 60% of the controls had one or two copies of the C allele, which means that this risk marker is extremely prevalent in this population.

As this was a selected population resident on an island (Madeira) with little genetic variety, these results cannot of course be extrapolated to other populations, even that of mainland Portugal. It would therefore be of considerable interest to investigate the prevalence and repercussions of this genetic variant in other Portuguese populations.

Multivariate logistic regression analysis showed that the risk allele presented an odds ratio of 1.7 in homozygosity and 1.5 in heterozygosity, meaning that individuals with two copies of the risk C allele have a 70% higher risk of CAD, and 50% higher with one copy,

após ajuste para as variáveis susceptíveis de confusão, a associação do SNP rs1333049 com DAC apresentou um odds ratio (OR) de 1.32 com um IC a 95% entre 1,24 e 1,42.

Numa meta-análise do mesmo grupo, efectuada em 2008 e envolvendo 7 estudos caso-controlo com um total de 4 645 doentes e 5177 controlos, foi genotipado o SNP rs1333049. O alelo de risco associou-se à DC em todos os 7 estudos incluídos, apresentando no total um OR em heterozigotia de 1,3 e em homozigotia de 1,5.⁽³¹⁾

Em 2008, esta associação foi reproduzida na população Japonesa e Coreana. Neste estudo, na população Japonesa o genótipo CC associou-se com a DC apresentando um OR de 1,44 (IC1,15-1,80;p=0.0013) e o alelo C um OR de 1,30 (IC1,13-1,49;p=0,00027) e na população Coreana o genótipo CC apresentou um OR de 1,28 (IC1,00-1,63; p=0.049) e o alelo C um OR de 1,19 (IC1,02-1,38;p=0.025)⁽³²⁾.

Na população Portuguesa, investigada no presente estudo, o alelo de risco foi portanto extremamente comum, com uma frequência em homozigotia de 35,7% nos doentes coronários e 29,3% nos controlos, valores acima dos citados pelos vários trabalhos internacionais anteriores. Assim, podemos dizer que este genótipo será tanto ou mais importante nesta população portuguesa como o é em outras populações já estudadas.

A população controlo, com uma ou duas cópias do alelo de risco C, apresentou uma frequência superior a 60%, o que torna este marcador de risco extremamente prevalente nesta População.

Naturalmente, que sendo uma população seleccionada, residente numa ilha e com pouca mistura genética com o exterior, teria o máximo interesse conhecer a prevalência e agressividades desta variante genética noutras populações Portuguesas.

Após análise multivariada de regressão logística, este alelo de risco, apresentou em homozigotia um *odds ratio* de 1,7 e 1,5 em heterozigotia, significando que os indivíduos que possuem duas cópias do alelo de risco C têm um risco de DAC 70% superior ao risco daqueles que não têm nenhuma ou 50% supe-

than those with none. This result, which holds after controlling for other risk factors including gender and age, confirms that the significant difference in gender between cases and controls in the study did not affect its central finding: the risk associated with chromosome 9p21.

In our study, as in all the others cited, this risk variant was not linked to the phenotypic characteristics studied, including hypertension, diabetes, obesity, dyslipidemia, behavior (alcohol consumption) or biochemical risk markers (*Table III*). The risk is independent of all established risk factors, which may indicate a new biological pathway in the pathogenesis of CAD. The finding by McPherson et al. that the risk allele is more prevalent in those with severe and early-onset atherosclerosis than in the elderly suggests that this allele promotes CAD by accelerating atherosclerosis.

An interesting recent letter by Visel et al. published in *Nature* in March 2010⁽³³⁾ sheds some light on this process. The authors deleted the 70-kb non-coding interval on mouse chromosome 4 that is orthologous to the 9p21 risk interval in humans, and observed reduced expression of *CDKN2A* and *CDKN2B* in the mutant mice. They also saw a doubling of proliferation in aortic smooth muscle cells in culture, which is relevant to the appearance of atherosclerosis. Expression of genes that suppress cell proliferation was reduced and that of genes promoting proliferation was enhanced. These results suggest that the risk interval represented by the deleted region has a pivotal role in CAD progression by altering the dynamics of vascular cell proliferation.

In the present study, as expected, significant differences were found between cases and controls with regard to gender, BMI, hypertension, diabetes, smoking, alcohol consumption and biochemical markers. We aimed to minimize possible confounding due to these baseline differences by using multivariate analysis that included several confounding factors. Logistic regression showed that the rs1333049 variant at the 9p21 locus maintained or even increased the significance of

rior quando só apresentam um alelo de risco. Estes resultados, após controlo para os outros factores de risco incluindo sexo e idade, confirmam que a diferença significativa existente no sexo entre casos e controlos, não afecta o seu resultado central- o risco associado com o cromossoma 9 p21.

Também no nosso estudo, assim como em todos os outros já apresentados, esta variante de risco não se associou a qualquer dos fenótipos estudados, como hipertensão arterial, diabetes, obesidade, dislipidémia, comportamentos (hábito de beber) ou outros marcadores bioquímicos de risco (*Quadro III*). Este efeito no risco é independente de todos os conhecidos factores de risco, o que pode implicar um novo caminho biológico relevante na patogenia da DAC. O achado, descrito por Mc Pearson, de que a prevalência do alelo de risco era mais elevada nos doentes com aterosclerose severa e prematura em relação aos idosos, sugere que este alelo promove a DAC acelerando a aterosclerose.

Um artigo interessante de *Visel e al*, muito recente, publicado em Março de 2010 na revista *Nature*, esclarece um pouco mais este processo. Estes autores procederam à deleção de 70 Kb da região não codificada do cromossoma 4 do rato, o qual é ortólogo ao intervalo de risco do 9 p21 humano e observaram uma redução da expressão do CDKN2A e CDKN2B, no rato mutante. Foi também observada o dobro da capacidade proliferativa das células musculares lisas da aorta mutante, na cultura celular, facto relevante para o aparecimento de aterosclerose. Ao mesmo tempo que houve redução da expressão dos genes supressores da proliferação celular, houve aumento da expressão dos genes implicados na proliferação celular. Estes resultados sugerem que o intervalo de risco representado pela região que sofreu a deleção, tem um papel importantíssimo na progressão da DAC ao alterar a dinâmica da proliferação das células vasculares, podendo assim representar um novo mecanismo fisiopatológico para o aparecimento da DAC⁽³³⁾.

No presente estudo, como era de prever, foram encontradas diferenças significativas

the associated risk. This result suggests that the mechanism of this variant is independent of conventional risk factors.

In case-control studies such as this, large cohorts are needed to confirm that genetic variants on chromosome 21 predispose to CAD. However, a considerable body of evidence has accumulated in recent years from studies in different populations, and it is likely that this strong association will also be confirmed in the Portuguese population.

CONCLUSION

In the present study we replicated the close link between the new variant rs1333049 on the 9p21 chromosome and coronary artery disease in a Portuguese population.

The risk allele of this genetic variant is extremely common in the study population, occurring in homozygosity in 30% and in heterozygosity in 50%.

Homozygosity for the risk allele increases susceptibility to CAD by over 70%, and by over 50% in heterozygosity, which is unusual for an isolated genetic variant.

This variant was not associated with any of the conventional phenotypic risk factors, including hypertension, diabetes, obesity, smoking, alcohol consumption, or biochemical risk markers, which suggests that the mechanism of this variant is independent of such factors.

Inclusion of the CC genotype in genetic risk scores could help to identify individuals predisposed to CAD, and may be useful in preventive medicine by enabling health plans to be individualized for the prevention of this deadly and widespread disease.

ACKNOWLEDGEMENTS

This study was supported by the European Regional Development Fund's Operational Programme for the Enhancement of Economic Potential and Territorial Cohesion for the Autonomous Region of Madeira (INTERVIR+).

entre casos e controlos em relação ao sexo, índice de massa corporal, hipertensão arterial, diabetes, hábitos de fumar e beber e outros marcadores bioquímicos entre o grupo doente e o grupo de controlo. Tentámos minimizar os efeitos potenciais de confusão causados por estas diferenças basais, com uma análise multivariada, de regressão logística, na qual vários factores de confusão foram incluídos. No entanto, na regressão logística, a variante rs 1333049 do locus 9p21 manteve ou até acentuou o risco e a significância do mesmo. Este resultado sugere a independência do mecanismo desta variante em relação aos factores de risco tradicionais.

Em estudos caso-controlo, como o presente, serão necessários grandes *cohorts* populacionais para confirmar a predisposição para o aparecimento de DAC associado a estas variantes genéticas no cromossoma 9p21. No entanto, nos últimos tempos, têm-se acumulado evidências sólidas, através de estudos efectuados em diferentes populações, sendo provável que esta forte associação se confirme na população Portuguesa.

CONCLUSÃO

Com o presente trabalho replicou-se numa população portuguesa a forte associação da nova variante rs1333049 do cromossoma 9p21 com a doença coronária.

Esta variante genética é extremamente comum na presente população, ocorrendo em homozigotia em cerca de 30% e em heterozigotia em cerca de 50% da população.

A homozigotia para este alelo de risco confere, na susceptibilidade para a DAC, um aumento de risco superior a 70%, e em heterozigotia um aumento de risco superior a 50% o que é raro numa variante genética isolada.

Esta variante não se associou a nenhum dos fenótipos de risco tradicionais nomeadamente hipertensão arterial, diabetes, obesidade, hábito de fumar, álcool ou outros marcadores bioquímicos de risco o que sugere a independência do mecanismo desta variante em relação aos factores de risco tradicionais.

A inclusão do genótipo CC, em *scores* de risco genético, pode permitir identificar indivíduos predispostos para DAC e revelar-se útil em medicina preventiva individualizando planos de saúde para a prevenção desta doença tão grave e prevalente.

AGRADECIMENTOS

Este trabalho foi apoiado pelo Fundo Europeu de Desenvolvimento Regional, Programa Operacional de Valorização do Potencial Económico e Coesão Territorial da RAM (INTERVIR +).

Pedido de Separatas
Address for reprints:

Maria Isabel Mendonça
Unidade de Investigação do
Hospital Dr. Nélio Mendonça (SESARAM)
Avenida Luís de Camões nº57
9004-514 Funchal
Madeira, Portugal
Tel: +351 291 744312
Fax: +351 291 744312
e-mail: dep.card@srs.pt

BIBLIOGRAFIA / REFERENCES

1. Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med* 1999; 330:1041-6.
2. Murabito JM, Pencina MJ, Nam BH, et al. Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. *JAMA* 2005; 294:3117-23.
3. Jonasdottir A, Jonasdottir A, Sigurdsson A et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 2007; 316:1491-3.
4. Larson MG, Atwood LD, Benjamin EJ, et al. Framingham Heart Study 100K project: genome-wide associations for cardiovascular disease outcomes. *BMC Med Genet* 2007; 8(Suppl. 1):S5.
5. Anderson JL, Horne BD. The 9p21 Locus and Coronary Heart Disease: Initiator, Promoter, or Precipitator? *J Am Coll Cardiol* 2010; 56: 487-489
6. Ake B, Sandberg T, Kestin Nilsson, et al. High Frequency of Multiple Melanomas and Breast and Pancreas Carcinomas in CDKN2A Mutation-Positive Melanoma Families. *JNCI J Natl Cancer Inst* (2000); 92 (15): 1260-1266
7. Lon R, Cardon and Bell JI. Association study designs for complex disease. *Nature Reviews Genetics* 2001;2:91-99
8. Morgan TM, Krumholz HM, Lifton RP, Spertus JA. Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *JAMA* 2007; 297:1551-1561
9. Joint European Society of Cardiology/American College of Cardiology Committee. Myocardial infarction redefined. A consensus document of the Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *Eur Heart J* 2000; 21: 1502-1513
10. Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults. Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 2001, 285:2486-2497
11. Chobanian AV et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure (JNC 7) *JAMA* 2003; 289:2560-2572.
12. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 2001;24(11):2017-8.
13. Pi-Sunyer FX. National Institutes of Health, National Heart, Lung, and Blood Institute North American Association for the Study of Obesity. The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. Obesity Education Initiative, NIH Publication Number 00-4084, October 2000
14. Guo S W, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 1992; 48:361-372.
15. Wigginton J E, Cutler D J, Agbecasis G R. A note on exact tests of Hardy-Weinberg equilibrium. *Am. J. Hum. Gen* 2005; 76: 887-893.
16. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls The Wellcome Trust Case Control Consortium. *Nature* 2007; 447:661-678.
17. Hannon GJ, Beach D. p15INK4B is a potential effector of TGF- β -induced cell cycle arrest. *Nature* 1994; 371:257-261.
18. Kalinina N, et al. Smad Expression in Human Atherosclerotic Lesions: Evidence for Impaired TGF- β /Smad Signaling in Smooth Muscle Cells of Fibrofatty Lesions. *Arterioscler Thromb Vasc Biol* 2004;24: 1391-1396
19. Hustinx SR, Leoni LM, Yeo CJ, et al. Concordant loss of MTAP and p16/CDKN2A expression in pancreatic intraepithelial neoplasia: evidence of homozygous deletion in a noninvasive precursor lesion. *Modern Pathology* 2005; 18:959-963
20. Baudhuin LM. Genetics of coronary artery disease: focus on genome wide association studies. *Am J Transl Res* 2009;1(3):221-234
21. Hayward NK. Genetics of melanoma predisposition. *Oncogene*, 2003; 22:3053-3062.
22. Horne BD, Carlquist JF, Muhlestein JB, Bair TL, Anderson JL. Association of variation in the chromosome 9p21 locus with myocardial infarction versus chronic coronary artery disease. *Circ Cardiovasc Genet*.2008;1:85-92
23. Holdt LM; Beutner F; Scholz M et al. ANRIL Expression Is Associated With Atherosclerosis Risk at Chromosome 9p21. *Arterioscler Thromb and Vasc Biol* 2010; 30:620-627.
24. Jarinova O, Stewart AF, Roberts R et al. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler Thromb Vasc Biol* 2009; 29(10) :1671-7
25. Doria A, Wojcik J, Xu R et al. Interaction between poor glycemic control and 9p21 locus on risk of coronary artery disease in type 2 diabetes. *JAMA* 2008;300(20):2389-97
26. Ye Shu, Willeit J, Kronenberg F, Xu Q, Kiechl S. Association of Genetic Variation on Chromosome 9p21 With Susceptibility and Progression of Atherosclerosis. *J Am Coll Cardiol* 2008;52:378-84.
27. Helgadottir A, Thorleifsson G, Manolescu A et al. Common Variant on Chromosome 9p21 Affects the Risk of Myocardial Infarction. *Science* 2007;316(5846):1491-3
28. McPherson R, Pertsemlidis A, Kavaslar N et al. A Common Allele on Chromosome 9 Associated with Coronary Heart Disease. *Science* 2007;316 (5830):1488-1491
29. Samani NJ, Erdmann J, Hall AS et al for the WTCCC and the Cardiogenics Consortium. Genomewide Association Analysis of Coronary Artery Disease. *N Engl J Med* 2007; 357:443-53.
30. Samani NJ and Schunkert H. Chromosome 9p21 and Cardiovascular Disease: The Story Unfolds. *Circ Cardiovasc Genet* 2008; 1:81-84.
31. Schunkert H, Götz A, Braund P et al. The Cardiogenics Consortium. Repeated Replication and a Prospective Meta-Analysis of the Association between Chromosome 9p21.3 and Coronary Artery Disease. *Circulation* 2008; 117(13): 1675-1684.

CALL FOR PAPERS | *Systems Biology and Polygenic Traits*

Association of ADAMTS7 gene polymorphism with cardiovascular survival in coronary artery disease

A. Pereira,¹ R. Palma dos Reis,² R. Rodrigues,¹ A. C. Sousa,¹ S. Gomes,¹ S. Borges,¹ I. Ornelas,¹ A. I. Freitas,³ G. Guerra,³ E. Henriques,¹ M. Rodrigues,¹ S. Freitas,¹ C. Freitas,¹ A. Brehm,³ D. Pereira,¹ and M. I. Mendonça¹

¹Funchal Hospital Center, Research Unit and Cardiology Department, Funchal, Madeira, Portugal; ²Faculty of Medical Sciences, New University of Lisbon, Lisbon, Portugal; and ³Laboratory of Human Genetics, Madeira University, Campus da Penteadá, Funchal, Madeira, Portugal

Submitted 16 May 2016; accepted in final form 8 September 2016

Pereira A, Palma dos Reis RP, Rodrigues R, Sousa AC, Gomes S, Borges S, Ornelas I, Freitas AI, Guerra G, Henriques E, Rodrigues M, Freitas S, Freitas C, Brehm A, Pereira D, Mendonça MI. Association of ADAMTS7 gene polymorphism with cardiovascular survival in coronary artery disease. *Physiol Genomics* 48: 810–815, 2016. First published September 9, 2016; doi:10.1152/physiolgenomics.00059.2016.—Recent genetic studies have revealed an association between polymorphisms at the ADAMTS7 gene locus and coronary artery disease (CAD) risk. Functional studies have shown that a CAD-associated polymorphism (rs3825807) affects ADAMTS7 maturation and vascular smooth muscle cell (VSMC) migration. Here, we tested whether ADAMTS7 (A/G) SNP is associated with cardiovascular (CV) survival in patients with established CAD. A cohort of 1,128 patients with angiographic proven CAD, who were followed up prospectively for a mean follow-up period of 63 (range 6–182) mo, were genotyped for rs3825807 A/G. Survival statistics (Cox regression) compared heterozygous (AG) and wild-type (AA) with the reference homozygous GG. Kaplan-Meier (K-M) survival curves were performed according to ADAMTS7 genotypes for CV mortality. Results showed that 47.3% of patients were heterozygous (AG), 36.5% were homozygous for the wild-type allele (AA) and only 16.2% were homozygous for the GG genotype. During the follow-up period, 109 (9.7%) patients died, 77 (6.8%) of CV causes. Survival analysis showed that AA genotype was an independent risk factor for CV mortality compared with reference genotype GG (HR = 2.7, $P = 0.025$). At the end of follow-up, the estimated survival probability (K-M) was 89.8% for GG genotype, 82.2% for AG and 72.3% for AA genotype ($P = 0.039$). Carriage of the mutant G allele of the ADAMTS7 gene was associated with improved CV survival in patients with documented CAD. The native overfunctional ADAMTS7 allele (A) may accelerate VSMC migration and lead to neointimal thickening, atherosclerosis progression and acute plaque events. ADAMTS7 gene should be further explored in CAD for risk prediction, mechanistic and therapeutic goals.

ADAMTS7 gene; coronary artery disease; single nucleotide polymorphism (SNP); cardiovascular survival

ADAMTS7 IS A MEMBER of a disintegrin and metalloprotease with thrombospondin type 1 repeats (ADAMTS) family with proteolytic activity against extracellular substrates, mainly extracellular matrix proteins (1). The importance of ADAMTS

proteases is implicated in the establishment of tissue architecture during development and in tissue degradation during disease states such as cancer, Alzheimer's and chronic inflammatory conditions (20).

A recent study revealed the first firm evidence that ADAMTS7 plays a proatherogenic role likely through the promotion of vascular smooth muscle cell (VSMC) migration by degrading its primary substrate, the cartilage oligomeric matrix protein (COMP), and thus promoted neointima formation following vascular mechanical injury (18). Moreover, in vitro studies have shown that VSMCs produce COMP, which inhibits VSMC migration (18). Consistently, genetic deletion of ADAMTS7 protects mice from atherosclerosis by reducing VSMC migration (4). Since VSMC migration is a very relevant process for plaque formation and stability, it is likely that ADAMTS7 may play a role in the development of atherosclerosis, the major cause of coronary artery disease (CAD).

In fact, three independent genome-wide association studies have recently identified the ADAMTS7 in the 15q24.2 chromosome, a novel locus for the development of coronary atherosclerosis (6, 14, 16). The lead CAD-associated SNP was rs3825807, a nonsynonymous polymorphism, with an adenine (A) to guanine (G) change, resulting in a serine-to-proline substitution in the prodomain of ADAMTS7 (16). Previous results suggest that this polymorphism has an effect on ADAMTS7 maturation, COMP degradation, and VSMC migration and thus is associated with atherosclerosis (13).

We aimed to examine whether the carriage of the ADAMTS7 rs3825807 A/G polymorphism plays a role in the cardiovascular (CV) survival, in a cohort of patients with significant angiographic coronary artery disease.

MATERIALS AND METHODS

Study population. A prospective study was performed with a cohort of 1,128 individuals (mean age 53.3 ± 7.9 yr, 79.6% men) admitted for coronary angiography due to angina and clinical suspicion of ischemic heart disease from either sex that gave written informed consent. The study was approved by the Funchal Hospital Ethics Committee.

Subjects were primary local residents, and this population is genetically a Caucasian south European sample. Nonresidents were not considered for this study.

Address for reprint requests and other correspondence: A. M. Pereira, Funchal Hospital Center, Research Unit and Cardiology Dept., Avenida Luís de Camões, n° 57, 9004-514 Funchal, Madeira, Portugal (e-mail: dep.card@sesaram.pt).

Table 1.

Variables	Total (n = 1,128)	Survivors (n = 1,019)	Deaths (n = 109)	Cardiovascular Deaths (n = 77)	P1	P2
Male (%)	898 (79.6)	808 (79.3)	90 (82.6)	64 (83.1)	0.420	0.429
Age, yr	53.3 ± 7.9	53.1 ± 8.0	55.5 ± 7.6	55.5 ± 8.2	0.003	0.015
Sedentarism (%)	718 (63.7)	638 (62.6)	80 (73.4)	55 (71.4)	0.026	0.142
Smoking (%)	511 (45.3)	463 (45.4)	48 (44.0)	34 (44.2)	0.780	0.834
Hypertension (%)	819 (72.6)	745 (73.1)	74 (67.9)	49 (63.6)	0.245	0.067
PWV, m/s	8.7 ± 2.0	8.6 ± 1.9	9.0 ± 2.1	8.9 ± 2.3	0.058	0.261
Diabetes (%)	387 (34.3)	336 (33.0)	51 (46.8)	35 (45.5)	0.004	0.033
Dyslipidemia (%)	1,057 (93.7)	952 (93.4)	105 (96.3)	74 (96.1)	0.235	0.369
BMI, kg/m ²	28.6 ± 4.3	28.6 ± 4.2	28.6 ± 4.7	28.7 ± 4.6	0.994	0.827
Heart rate, bpm	69.4 ± 12.4	68.9 ± 12.1	73.7 ± 14.6	74.7 ± 13.9	0.001	<0.0001

P1 compares patients dying from any cause with survivors, and P2 compares patients dying from cardiovascular causes with others. PWV, pulse wave velocity; BMI, body mass index. Continuous variables are presented as means ± SD; Statistically significant for $P < 0.05$.

Determination of presence and severity of CAD was made by the hemodynamic laboratory staff that included a dedicated intervention cardiologist. Angiographically proven CAD was considered significant if ≥ 1 coronary lesions of $\geq 70\%$ stenosis in ≥ 1 major coronary artery or its primary branch. Absent or nonsignificant CAD was excluded from this study. Index angiography occurred between 2008 and 2013.

This cohort was followed until death or December 2014 [mean follow-up period of 63 mo (range 6–182 mo)]. Each subject was followed up at the outpatient clinic according to the angiography date, and data concerning new hospital admissions or angiograms were collected from the clinical files of patients.

In the case of extrahospital death, death cause was collected consulting the public death report and by family telephone survey. The causes of CV death were determined according to the International Classification of Disease, 10th revision, clinical modification diagnostic criteria. Cardiovascular death included CAD and cerebrovascular disease.

Follow-up using the database and direct contact with patients or families allowed for 100% assessment of survival within this cohort.

Data collection. Data were collected from all subjects in a standardized file comprising demographic and clinical characteristics and traditional risk factors (sex, age, sedentary lifestyle, smoking habits, arterial hypertension, diabetes, dyslipidemia, body mass index, heart rate and pulse wave velocity). The definition of conventional CV risk

was based on the standard criteria, as previously reported (3, 5, 6a, 7, 12).

Biochemical analysis. Blood samples were extracted after 14–16 h fasting. Biochemical analyses were performed in the central laboratory of the hospital, according to the usual techniques.

To determine total cholesterol and triglycerides, blood samples were placed in dry tubes and centrifuged a half-hour later at 3,500 g and subsequently quantified by an enzymatic technique with a Hitachi 911 auto analyzer. Biochemical markers such as lipoprotein-a, apolipoprotein B (Apo B), and high sensitivity C-reactive protein (hs-CRP) were quantified by nephelometry on a Behring BN 100 automatic system. Homocysteine was measured by fluorescence polarized immunoassay using an Abbot IMX automatic device. To measure fibrinogen, samples placed in a tube containing sodium citrate, and measurements were taken with a Behring BCS automatic analyzer.

Genotype analyses. Genomic DNA was extracted from 80 ml of peripheral blood using a standard phenol-chloroform method (Qiagen, Hilden, Germany).

To identify the ADAMTS7 rs3825807 A/G genotypes, a TaqMan allelic discrimination assay (Applied Biosystems) was performed using labeled probes and primers pre-established by the supplier (TaqMan SNP Genotyping Assays, Applied Biosystems).

The genotyping reaction was amplified and detected on a StepOnePlus Real-Time PCR, and genotypes were determined by

Table 2.

Variables	Total (n = 1,128)	Survivors (n = 1,019)	Deaths (n = 109)	Cardiovascular Deaths (n = 77)	P1	P2
Hemoglobin, g/dl	14.6 (8.6–18.2)	14.6 (8.6–18.2)	14.8 (10.9–18.0)	14.9 (11.0–17.6)	0.889	0.915
Platelets, 10 ³ /μl	214.5 (61.0–893.0)	214.0 (61.0–893.0)	224.0 (116.0–533.0)	227.0 (116.0–533.0)	0.106	0.048
Leukocytes, 10 ³ /μl	7.0 (3.4–24.6)	7.0 (3.4–24.6)	7.4 (4.1–13.3)	7.8 (4.4–12.4)	0.022	0.006
Fibrinogen, mmHg	383.0 (91.0–832.0)	382.5 (91.0–715.0)	399.4 (153.3–832.0)	383.0 (153.3–832.0)	0.096	0.449
Homocysteine, μmol/l	12.4 (4.6–220.2)	12.3 (4.6–220.2)	12.8 (5.9–48.7)	12.7 (5.9–48.7)	0.017	0.044
Lp(a), mg/dl	19.9 (1.9–241.0)	19.3 (1.9–241.0)	26.0 (1.9–219.0)	23.4 (1.9–191.5)	0.065	0.489
Apo B, mg/dl	96.4 (4.9–256.9)	95.9 (4.9–256.9)	101.5 (20.1–181.0)	102.0 (20.1–173.0)	0.007	0.012
Triglycerides, mg/dl	143.0 (31.0–2,500.0)	143.0 (31.0–2500)	149.0 (44.0–608.0)	153.0 (44.0–571.0)	0.123	0.110
Glucose, mg/dl	106.0 (53.0–458.0)	106.0 (53.0–458.0)	116.5 (67.0–393.0)	118.0 (67.0–393.0)	<0.0001	<0.0001
hs-CRP, mg/dl	0.24 (0.01–24.8)	0.23 (0.01–24.8)	0.36 (0.02–15.6)	0.3 (0.02–7.9)	0.005	0.083

P1 compares patients dying from any cause with survivors, and P2 compares patients dying from cardiovascular causes with others. Lp(a), lipoprotein (a); Apo B, apolipoprotein B; hs-CRP, high sensitivity C reactive protein. Values are medians (minimum - maximum); statistically significant for $P < 0.05$.

Table 3.

Group	Wild Type (AA)	Heterozygote (AG)	Homozygote (GG)	WT (A) Allele	Variant (G) Allele
<i>A: All patients/All deaths</i>					
Survivors (1,019)	364 (35.7%)	484 (47.5%)	171 (16.8%)	1,212 (59.5%)	826 (40.5%)
Deaths (109)	48 (44.0%)	49 (45.0%)	12 (11.0%)	145 (66.5%)	73 (33.5%)
Total (1,128)	412 (36.5%)	533 (47.3%)	183 (16.2%)	1,357 (60.2%)	899 (39.8%)
<i>B: CV deaths/CV survivors</i>					
No CV death (1,051)	374 (35.6%)	500 (47.6%)	177 (16.8%)	1,248 (59.4%)	854 (40.6%)
CV deaths (77)	38 (49.4%)	33 (42.9%)	6 (7.8%)	109 (70.8%)	45 (29.2%)
Total (1,128)	412 (36.5%)	533 (47.3%)	183 (16.2%)	1,357 (60.2%)	899 (39.8%)

For A, death by genotype contingency tables gives $P = 0.134$ (χ^2). For B, CV death by genotype gives $P = 0.021$ (χ^2), $P = 0.017$ (likelihood ratio) or $P = 0.006$ (linear-by-linear association). CV, cardiovascular; WT, wild type.

using the 7300 System SDS Software without any prior knowledge of individual clinical data (Applied Biosystems, Foster City, CA).

Statistical analyses. Deviation from Hardy-Weinberg equilibrium for genotypes at individual locus was assessed using the χ^2 -test. Comparisons of characteristics of CV death and survivors were analyzed by χ^2 for categorical variables, and Student's *t*- or Mann-Whitney test was performed for continuous variables as appropriate.

Genotypic frequencies were determined from observed counts and compared by χ^2 analysis. Heterozygous (AG) or homozygous wild-type (AA) carriers were compared with noncarriers (GG) with survival statistics.

Multivariable Cox proportional-hazard regression analyses were used to examine the CV death between genotypes of ADAMTS7 variants, using GG as reference class (genotype model). Hazard ratios and the multiple variable predictive values were determined, conditioned on seven other major CAD risk factors: age, sex, smoking status, diabetes, hypertension, dyslipidemia, and renal failure. The critical value for entering and excluding variables in the model was set at $P = 0.10$. Multivariable Cox proportional-hazard regression analyses also addressed the all-cause mortality including the same set of covariates.

Kaplan Meyer survival function with Log-rank statistics was tested to predict the value of the ADAMTS7 genotypes for CV mortality.

Statistical analyses were performed using Statistical Package for the Social Sciences Software version 19.0 (IBM, Armonk, NY).

RESULTS

A total of 1,128 patients (mean age 53.3 ± 7.9 yr, 79.6% men) with documented CAD entered this study, during a mean follow-up period of 63 mo (range 6–182 mo). Baseline and biochemical characteristics of the studied population are shown in Table 1 and Table 2, respectively.

Of entered patients, 109 (9.7%; mean age 55.5 ± 7.6 yr, 82.6% men) died during the follow-up period and 77 (6.8%, mean age 55.5 ± 8.2 yr, 83.1% men) of CV causes. Nonsurvivors were older at baseline, more sedentary, had more diabetes and higher heart rates than survivors, with statistical significance ($P < 0.05$) (Table 1). Additionally, leukocytes,

homocysteine, apo B levels, glucose and hs-CRP were higher in patients who died during the follow-up than in survivors (Table 2).

ADAMTS7 genotypic distributions are shown in Table 3 for all patients, namely, survivors, those dying of any cause, and those dying of a CV cause. The distribution of alleles was in Hardy-Weinberg equilibrium ($P = 0.890$). Genotypic analysis revealed that 47.3% of patients were heterozygous (AG), 36.5% were homozygous for the wild-type allele (AA), and only 16.2% were homozygous for the GG genotype. The proportion of alleles present in the cohort was 60.2% for the wild-type allele (A) and 39.8% for the G allele. CV deaths occurred in 49.4% of the carriers of the wild-type ADAMTS7 AA, while 7.8% were individuals with the polymorphic GG genotype (Table 3).

In multiple variable Cox regression analysis (Table 4), including age, sex, smoking status, diabetes, history of hypertension and of dyslipidemia, renal failure, and genetic variant (ADAMTS7), only age [hazard ratio (HR) = 1.051, $P = 0.001$], hypertension (HR = 0.634, $P = 0.063$), and ADAMTS7 AA genotype carriers (HR = 2.681, $P = 0.025$) were selected as independent predictors of CV mortality (Table 4).

Kaplan-Meier survival-time plot for CV death by ADAMTS7 genotype is shown in Fig. 1. A significant difference in survival by ADAMTS7 genotype was observed (log-rank statistic 6.48, $P = 0.039$). At the end of follow-up period, the survival probability was 89.8% for the GG genotype carriers, 82.2% for heterozygote individuals (AG) and 72.3% for the wild-type (AA). Within each ADAMTS7 genotype, 3.3% (6/183) of events (CV deaths) occurred for patients carrying the polymorphic genotype (GG), 6.2% (33/533) among the heterozygote (AG) carriers, and 9.2% (38/412) occurred for patients who had the wild-type genotype (AA) as shown in Fig. 1.

Table 4.

Variables	B	SE	Wald	df	Hazard Ratio (CI 95%)	P Value
Age	0.049	0.016	10.131	1	1.051 (1.019–1.083)	0.001
AHT	−0.455	0.245	3.462	1	0.634 (0.393–1.025)	0.063
ADAMTS7			5.481	2		0.065
AG	0.690	0.444	2.412	1	1.993 (0.835–4.760)	0.120
AA	0.986	0.440	5.016	1	2.681 (1.131–6.353)	0.025

Using forward Wald conditional Cox regression method (SPSS v 19.0), entering age (yr), male sex, diabetic status, history of hypertension, history of dyslipidemia, renal failure, smoking and genetic variant ADAMTS7. P value to exclude variables was 0.10. B, beta coefficient; CI, 95% confidence interval.

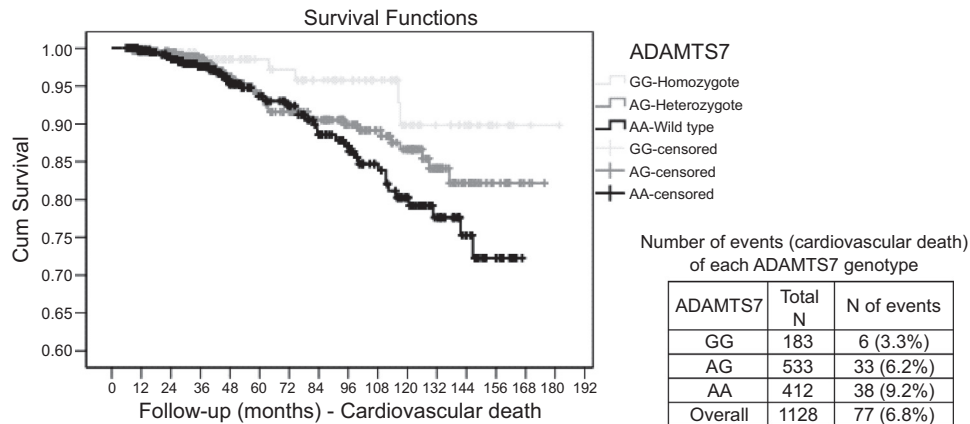


Fig. 1. The log-rank statistic was 6.480, $P = 0.039$. There are 6 events (3.3%) among 183 patients carrying the homozygote, 33 events (6.2%) among the patients who had the heterozygote type, and 38 events (9.2%) among the patients who had the wild type.

A second Cox regression model, adjusted for the same covariates, was performed for predicting all-cause mortality, as shown in Table 5. During the follow-up period, the only independent predictors of all-cause mortality were age (HR 1.05 $P < 0.0001$), Diabetes (HR = 1.469, $P = 0.05$), and smoking habit (HR = 1.58, $P = 0.026$). Neither carriage of homozygous (AA) nor heterozygous (AG) genotypes were independent predictors of all-cause mortality (HR = 1.424, $P = 0.276$, HR = 1.604, $P = 0.147$ respectively) (Table 5).

DISCUSSION

In our study, patients with angiographically documented CAD who were carriers of the polymorphic G variant of ADAMTS7 gene had increased survival from CV causes. For instance, at the end of the follow-up period, patients with ADAMTS7 GG showed approximately three times fewer events of CV mortality in relation to AA carriers.

We consider that the carriers of the G allele had reduced migratory ability with lower atherosclerosis progression and severity, reduced proteinases secretion to influence the stability of the complex plaque, and therefore better prognostic than patients carrying the wild-type A of the ADAMT7 gene.

Similarly, Pu and colleagues (13), observed an inverse association between the ADAMTS7 rs3825807 GG genotype and atherosclerosis, this being the genotype associated with lower carotid atherosclerosis prevalence and severity (score). The same study demonstrated that VSMCs of the protective GG genotype had decreased migratory activity compared with those with the AA genotype, and conditioned media of VSMCs of the GG contained less of the cleaved form of COMP, an ADAMTS7 substrate that had been shown to be produced by VSMCs and inhibit its migration (13). Also, You et al. (19) recently demonstrated a strong and direct association between the proportion of CAD patients with three-vessel disease and increasing ADAMTS7 rs3825807 gene dosage of the risk

variant (A). This condition was investigated in the cohort of our study, but no significant association was found between AA genotype and most severe atherosclerosis (three-vessel disease), although there was a strong trend (data not shown).

However, to date, an association between the CAD prognosis and this polymorphism has not yet been found. As far as we know, this study is the first attempt to establish this association in a cohort of south European descent.

Previously, investigation of the physiological roles of ADAMTS7 has largely focused on its association with the pathogenesis of arthritis and disk diseases (6a). ADAMTS structure contains the thrombospondin type-1 repeats and acts as a negative regulator of endplate chondrocyte differentiation because it downregulates the expression of collagen type II, collagen type X, and *SOX9*, which are early and late marker genes for chondrogenesis (11). Furthermore, suppression of ADAMTS7 expression in human chondrocytes markedly prevents COMP degradation (9). ADAMTS7 mRNA is found in normal human bone, cartilage, synovia, tendons, spleen, heart, and brain, as well as in the meniscus, skeletal muscle, and fat at lower levels (10). However, the physiological implications of this widespread tissue expression of ADAMTS7 remain to be determined. Furthermore, cumulative evidence has demonstrated that ADAMTS7 facilitates VSMC migration and neointima formation through degradation of vascular COMP (18). Consistently, genetic deletion of ADAMTS7 reduces neointimal thickening after wire injuries (4).

In our study, it's interesting to note that besides ADAMTS7 AA genotype, the hypertension status and age were independent predictors of CV mortality; being elderly and lower tensional status both correlated with CV death. The independence of CV death predictors from other traditional risk factors like dyslipidemia, diabetes, and renal failure emphasizes the ADAMT7 gene's mechanistic role being associated with vascular cleavage and migration of VSMCs not currently covered

Table 5.

Variables	B	SE	Wald	df	Hazard ratio (CI 95%)	P Value
Age	0.050	0.014	13.405	1	1.051 (1.023–1.080)	0.000
Diabetic status	0.384	0.197	3.826	1	1.469 (0.999–2.159)	0.050
Smoking	0.456	0.204	4.973	1	1.578 (1.057–2.355)	0.026

Using forward Wald conditional Cox regression method (SPSS v 19.0), entering age (yr), male sex, diabetic status, history of hypertension, history of hyperlipidemia, renal failure, smoking, and genetic variant ADAMTS7. P value to exclude variables was 0.10. B, beta coefficient; CI, 95% confidence interval.

by updated therapies including statins, ACE inhibitors, angiotensin-aldosterone-antagonists, B-blockers, and antidiabetic medication. Monoclonal antibodies targeting several ADAMTS gene family members are on the way in several clinical scenarios, and the ADAMTS7 gene seems a highly favorable candidate for use against for CAD because of the gene's already established association with the development of CAD. Future genetic studies in CAD disease should focus on disease progression and prognosis rather than disease development. So far, established risk factors of atherosclerosis are not able to fully account for CV death due to different pathophysiological mechanisms involved into plaque formation and plaque events. CV death might be associated with the critical mechanisms of plaque stabilization/thrombosis and arrhythmia in CAD patients and gene associated with vascular migration/stabilization are promising targets (14).

Nevertheless, in our study the AA genotype was found to be unassociated with all-cause mortality and associated only with CV mortality, which strengthens the hypothesis that ADAMTS7's being a regulator of the gene maturation, COMP cleavage, and VSMC migration would translate into fewer CV events, including CV death in patients with a reduced function of the ADAMTS gene (G) variant.

Further research in this area could lead to personalized medicine based on identification of ADAMTS7 genetic variants in patients with atherosclerosis who could benefit from strategies targeting this proteolytic pathway (2).

Study Strengths and Limitations

Very few studies have yet addressed common gene variants and outcomes and long-term prognosis, and even the 9p21 gene failed to demonstrate consistent association with long-term myocardial infarction and CV mortality (17). No report on the specific association of ADAMTS7 gene variants with CV mortality in CAD patients has been published.

However, this study is only moderate in size, and the number of events in the follow-up is limited, so that the confidence intervals for CV survival are quite broad.

A significant limitation of the study was the lack of ejection fraction data at admission for determining the relative protective effect of the variant as a function of the ejection fraction.

Patients deceased during the coronary angiogram were also excluded from our study. Those patients probably had severe atherosclerotic coronary disease, and their DNA sample would be useful to enlarge our knowledge of this variant role in CV death. Expression studies of these ADAMTS7 gene variants on the arterial wall of the atherosclerotic plaque or thrombus would also further information from animal studies already published.

In conclusion: Our study emerges as a novel finding in a south European population and provides new evidence implicating ADAMTS7 with CAD prognosis. Its spectrum of interest as a therapeutic target may grow once its association with outcomes and CV mortality is clarified. Our results, although promising, should be amplified in larger and long-term studies that focus on CV prognosis rather than disease development. If validated, they suggest that ADAMTS7 genotyping may provide useful information for prognosis and survival in CAD patients.

ACKNOWLEDGMENTS

We are very grateful to Elsa Sousa who performed the phone calls to the subjects and all the administrative procedures.

GRANTS

This study was supported by the European Regional Development Fund's Operational Programme for the Enhancement of Economic Potential and Territorial Cohesion for the Autonomous Region of Madeira (INTERVIR+).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

A.P., R.P.d.R., and M.I.M. conception and design of research; A.P., R.R., A.C.S., S.G., I.O., A.I.F., G.G., C.F., and M.I.M. performed experiments; A.P., S.B., E.H., M.R., S.F., and M.I.M. analyzed data; A.P. and M.I.M. interpreted results of experiments; A.P., S.B., E.H., M.R., S.F., and M.I.M. drafted manuscript; A.P., R.P.d.R., S.B., A.B., D.P., and M.I.M. edited and revised manuscript; A.P., R.P.d.R., and M.I.M. approved final version of manuscript; E.H., M.R., and S.F. prepared figures.


REFERENCES

- Arner EC. Aggrecanase-mediated cartilage degradation. *Curr Opin Pharmacol* 2: 322–329, 2002.
- Arroyo AG, Andrés V. ADAMTS7 in cardiovascular disease: from bedside to bench and back again? *Circulation* 131: 1156–1159, 2015.
- Asmar R, Benetos A, Topouchian J, Laurent P, Pannier B, Brisac AM, Target R, Levy BI. Assessment of arterial distensibility by automatic pulse wave velocity measurement: validation and clinical application studies. *Hypertension* 26: 485–490, 1995.
- Bauer RC, Tohyama J, Cui J, Cheng L, Yang J, Zhang X, Ou K, Paschos GK, Zheng XL, Parmacek MS, Rader DJ, Reilly MP. Knock-out of Adamts7, a novel coronary artery disease locus in humans, reduces atherosclerosis in mice. *Circ* 131: 1202–1213, 2015.
- Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL Jr, Jones DW, Materson BJ, Oparil S, Wright JT Jr, Roccella EJ; National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure; National High Blood Pressure Education Program Coordinating Committee. The seventh report of the Joint National Committee on Prevention, Detection, Evaluation and treatment of High Blood Pressure: the JNC7 report. *JAMA* 289: 2560–2572, 2003.
- Coronary Artery Disease (CAD) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet* 43: 339–344, 2011.
- Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care* 26, Suppl 1: S5–S20, 2003.
- Expert Panel on Detection Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 285: 2486–2497, 2001.
- Gopalakrishnan K, Kumarasamy S, Abdul-Majeed S, Kalinoski AL, Morgan EE, Gohara AF, Nauli SM, Filipiak WE, Saunders TL, Joe B. Targeted disruption of Adamts16 gene in a rat genetic model of hypertension. *Proc Natl Assoc Sci USA* 109: 20555–20559, 2012.
- Guo NH, Krutzsch HC, Inman JK, Roberts DD. Thrombospondin 1 and type I repeat peptides of thrombospondin 1 specifically induce apoptosis of endothelial cells. *Cancer Res* 57: 1735–1742, 2012.
- Hanby HA, Zheng XL. Biochemistry and physiological functions of ADAMTS7 metalloprotease. *Adv Biochem* 1: 43–50, 2013.
- Kuno K, Matsushima K. ADAMTS-1 protein anchors at the extracellular matrix through the thrombospondin type I motifs and its spacing region. *J Biol Chem* 273: 13912–13917, 1998.
- National Institute of Health, National Heart, Lung, and Blood Institute North American Association for the Study of Obesity. *The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults*. NHLBI Obesity Education Initiative, NIH Publication Number 00–4084, 2000.

13. Pu X, Xiao Q, Kiechl S, Chan K, Ng FL, Gor S, Poston RN, Fang C, Patel A, Senver EC, Shaw-Hawkins S, Willeit J, Liu C, Zhu J, Tucker AT, Xu Q, Caulfield MJ, Ye S. ADAMTS7 cleavage and vascular smooth muscle cell migration is affected by a coronary-artery-disease-associated variant. *Am J Hum Genet* 92: 366–374, 2013.
14. Reilly MP, Li M, He J, Ferguson JF, Stylianou IM, Mehta NN, Burnett MS, Devaney JM, Knouff CW, Thompson JR, Horne BD, Stewart AF, Assimes TL, Wild PS, Allayee H, Nitschke PL, Patel RS; Myocardial Infarction Genetics Consortium; Wellcome Trust Case Control Consortium; Martinelli N, Girelli D, Quyyumi AA, Anderson JL, Erdmann J, Hall AS, Schunkert H, Quertermous T, Blankenberg S, Hazen SL, Roberts R, Kathiresan S, Samani NJ, Epstein SE, Rader DJ; Myocardial Infarction Genetics Consortium; Wellcome Trust Case Control Consortium. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet* 377: 383–392, 2011.
16. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buyschaert I; Cardiogenics Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nothen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgerirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Wittmann JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J; CARDI Consortium; GRAM, Samani NJ. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* 43: 333–338, 2011.
17. Virani SS, Brautbar A, Lee VV, MacArthur E, Morrison AC, Grove ML, Nambi V, Frazier L, Wilson JM, Willerson JT, Boerwinkle E, Ballantyne CM. Chromosome 9p21 single nucleotide polymorphisms are not associated with recurrent myocardial infarction in patients with established coronary artery disease. *Circ J* 76: 950–956, 2012.
18. Wang L, Zheng J, Bai X, Liu B, Liu CJ, Xu Q, Zhu Y, Wang N, Kong W, Wang X. ADAMTS-7 mediates vascular smooth muscle cell migration and neointima formation in balloon-injured rat arteries. *Circ Res* 104: 688–698, 2009.
19. You L, Tan L, Liu L, Shen R, Chaugai S, Wang DW, Cui W. ADAMTS7 locus confers high cross-race risk for development of coronary atheromatous plaque. *Mol Genet Genomics* 242: 351–356, 2015.
20. Zhang Y, Lin J, Wei F. The function and roles of ADAMTS-7 in inflammatory diseases. *Mediators Inflamm* 2015: 801546, 2015.

ORIGINAL PAPER

Genetic risk score and cardiovascular mortality in a southern european population with coronary artery disease

Andreia Pereira^{1,3}  | Maria Isabel Mendonca¹ | Ana Célia Sousa¹ | Sofia Borges¹ |
Sónia Freitas¹ | Eva Henriques¹ | Mariana Rodrigues¹ | Ana Isabel Freitas² |
Graça Guerra² | Ilídio Ornelas¹ | Décio Pereira¹ | António Brehm² |
Roberto Palma Dos Reis³

¹Research Unit and Cardiology
Department, Funchal Hospital Center, Avenida
Luís de Camões, Funchal, Portugal

²Laboratório de Genética
Humana, Universidade da Madeira, Campus
da Penteada, Funchal, Portugal

³Faculdade de Ciências Médicas, Universidade
Nova de Lisboa, Campo dos Mártires da
Pátria, Lisboa, Portugal

Correspondence

Andreia Micaela Pereira, Funchal Hospital
Center, Research Unit and Cardiology
Department, Avenida Luís de Camões,
Funchal, Madeira, Portugal.
Email: dep.card@sesaram.pt

Funding information

This study was supported by the European
Regional Development Fund Operational
Programme for the Enhancement of Economic
Potential and Territorial Cohesion for the
Autonomous Region of Madeira (INTERVIR+)

Summary

Several genetic risk scores (GRS) have been associated with cardiovascular disease; their role, however, in survival from proven coronary artery disease (CAD) have yielded conflicting results.

Objective: The objective of this study was to evaluate long-term cardiovascular mortality according to the genetic risk score in a Southern European population with CAD.

Methods: A cohort of 1464 CAD patients with angiographic proven CAD were followed up prospectively for up to 58.3 (interquartile range: 25.8-88.1) months. Genotyping of 32 single-nucleotide polymorphisms previously associated with CAD was performed using oligonucleotides probes marked with fluorescence for each allele. GRS was constructed according to the additive model assuming codominance and categorised using the median (=26). Cox Regression analysis was performed to determine independent multivariate predictors of cardiovascular mortality. Kaplan-Meier survival curves compared high vs low GRS using log-rank test. C-index was done for our population, as a measure of discrimination in survival analysis model.

Results: During a mean follow-up of 58.3 months, 156 patients (10.7%) died, 107 (7.3%) of CV causes. High GRS (≥ 26) was associated with reduced cardiovascular survival. Survival analysis with Cox regression model adjusted for 8 variables showed that high GRS, dyslipidemia, diabetes and 3-vessel disease were independent risk factors for cardiovascular mortality (HR=1.53, $P=.037$; HR=3.64, $P=.012$; HR=1.75, $P=.004$; HR=2.97, $P<.0001$, respectively). At the end of follow-up, the estimated survival probability was 70.8% for high GRS and 80.8% for low GRS (Log-rank test 5.6; $P=.018$). C-Index of 0.71 was found when GRS was added to a multivariate survival model of diabetes, dyslipidemia, smoking, hypertension and 3 vessel disease, stable angina and dual antiplatelet therapy.

Conclusions: Besides the classical risk factors management, this work highlights the relevance of the genetic profile in survival from CAD. It is expected that new therapies will be directed to gene targets with proven value in cardiovascular survival.

1 | INTRODUCTION

Cardiovascular disease (CVD) is the leading cause of death and a major cause of morbidity worldwide. It is estimated that almost 30% of all-causes of mortality is attributed to CVD.¹ Traditional risk factors are insufficient to identify the asymptomatic high-risk individuals. Epidemiology and family studies have long documented that approximately 50% of the susceptibility for heart disease is genetic.² Comprehensive prevention and treatment will be possible only if we know the genetic predisposition.

Although much of the genetic basis of coronary disease remains still to be discovered, some progress is occurring using both candidate gene and genome-wide association studies (GWAS).³ Until now, the risk attributable to any individual variant has been modest. However, discovering and combining multiple *loci* with modest effects into a global genetic risk score (GRS) could improve the identification of high-risk populations and improve individual risk assessment.

Recently, attempts have been made to use genetic information in the assessment of whether cumulative genetic risk conferred across multiple *loci* is a more robust tool for examining disease risk than associations of SNPs.⁴

While the focus of research in cardiovascular genetics has been in identifying predictors of incident or prevalent cardiovascular disease, little work has been done regarding patients with established cardiovascular disease and genetic susceptibility to the onset of recurrent early or late events. Genetically predisposed individuals for acute coronary syndromes (ACS) may, in fact, be at high risk for early recurrent events given that current medical treatment may be ineffective in reducing the genetic risk post-ACS which led to the initial CV event. Furthermore, a GRS that may further improve the prediction of recurrent events in patients with established cardiovascular disease may have therapeutic implications in the management of these patients. Not much recent and relevant work has been done concerning that subject.

2 | OBJECTIVE

The purpose of this work was to evaluate long-term cardiovascular mortality according to the GRS in a Southern European population with CAD.

3 | METHODS

3.1 | Study population

We performed a prospective study with a cohort of 1464 individuals with coronary disease (mean age 53.3±8.0 years; 79.7% male) admitted for coronary angiography due to clinical suspicion of ischaemic heart disease from either gender. The study was approved by Hospital Ethics Committee. The patients gave written informed consent to participate in the study and they were all informed about blood collection for genetic and relevant clinic data. Blood samples for genetic purposes are kept in our hospital research bio bank according to articles

What's known

- Cardiovascular mortality remains high even after secondary prevention measures and therapeutic intervention.
- At least 50 genetic variants have risen from Genome Wide Association Studies with irrefutable risk association with Coronary Artery Disease, in multiple population sets. Short-term studies have associated genetic profiles to some major adverse cardiovascular events.

What's new

- In this study, in a population with proven CAD disease, we found that individual Genetic Risk Score (GRS), diabetes, dyslipidemia and three-vessel disease were independent predictors for long-term cardiovascular mortality. GRS that may further improve the prediction of recurrent events in patients with established cardiovascular disease may have therapeutic implications in the management of these patients.

6, 15 and 16 from the UNESCO Universal Declaration on Bioethics and Human Rights.

Subjects were primarily local residents and this population is genetically a Caucasian Southern European sample. Population stratification analysis was performed in our population set to account for possible genetic admixture and no significant genetic outliers (<5%) were identified with Principal Component Analysis.⁵

Determination of presence and severity of CAD was made by Hemodynamic Laboratory Staff that included a dedicated Intervention Cardiologist. Angiographically proven CAD was considered significant if ≥1 coronary lesions of ≥70% stenosis in ≥1 major coronary artery or its primary branch. Absent or non-significant CAD (patients with some degree of stenosis but not qualifying as significant stenosis (>70%) were excluded from this study. Index Angiography occurred between 2008 and 2014.

This cohort was followed until death or December 2015 (mean follow-up period of 58.3 months (interquartile range: 25.8-88.1 months). Each subject was followed up at the outpatient clinic according to the angiography date and data concerning new hospital admissions or angiograms that were collected from the clinical files of patients.

In the case of extra-hospital death, cause of death was collected consulting the Public Death Report and by a family telephone survey. The causes of cardiovascular death were determined according to the International Classification of Disease, 10th Revision, Clinical Modification (ICD-10 CM) diagnostic criteria. Cardiovascular death included CAD and cerebrovascular disease. Adjudication of the events was done by two independent certified cardiologists totally blinded to patient level study data.

Follow-up using the database and direct contact with patients or families allowed for 100% assessment of survival within this cohort.

3.2 | Data collection

Data were collected from all subjects in a standardised file comprising demographic, clinical characteristics and traditional risk factors (TRF) (gender, age, sedentary lifestyle, family history of CAD, smoking status, arterial hypertension, diabetes, dyslipidemia, body mass index (BMI), heart rate and pulse wave velocity (PWV)). The definition of conventional CV risk factors was based on the standard criteria, as previously reported.⁶⁻¹¹

Essential Hypertension was considered when patients, at the entry into this study, were already diagnosed and/or had been on antihypertensive medication for more than 3 months or newly diagnosed hypertensives with systolic blood pressure (SBP)/diastolic blood pressure (DBP) $\geq 140/90$ mm Hg measured on at least 3 occasions.⁶

Subjects with a LDL >100 mg/dL, HDL <40 mg/dL for men and <45 mg/dL for women, non-HDL (Total cholesterol-HDL) >130 mg/dL, triglycerides >150 mg/dL or Apo B >100 mg/were classified as having dyslipidemia, as previously described in "Guidelines on management of dyslipidemia 2016" as a primary goal in high-risk patients with known cardiovascular disease.⁷

Subjects were considered to have a family history of premature CVD if the father or brother had been diagnosed with CVD under the age of 55 or the mother or sister under the age of 65.⁸

'Smoking status' refers to current smokers or subjects with less than 5 years of smoking cessation.⁹

Subjects were classified as having diabetes if they were taking oral anti-diabetic medication or insulin or if their fasting plasma glucose was higher than 7.0 mmol/L or 126 mg/dL.¹⁰

3.3 | Biochemical analysis

Blood samples were extracted after 14-16 hours' fasting. Biochemical analyses were performed in the Central Laboratory of the Hospital, according to the usual techniques.

To determine total cholesterol and triglycerides, blood samples were placed in dry tubes and centrifuged half an hour later at 3500 g and subsequently, quantified by an enzymatic technique using a Hitachi 911 auto analyser. Biochemical markers such as lipoprotein-a (Lp(a)), apolipoprotein B (Apo B) and high sensitivity C-Reactive Protein (hs-CRP) were quantified by Immunturbidimetry (AU5400, Beckman Coulter Analyzer, Brea, CA, USA). The sample collected for homocysteine measurement was kept at 2-8°C, at least 30 minutes (no more than 60 minutes). After clot retraction, serum was divided into aliquots and kept at -20°C until measurement by Kinetic method (AU 5400 Beckman Coulter, Brea, CA, USA). To measure fibrinogen, samples were placed in a tube containing sodium citrate, and measurements were taken with a Behring BCS automatic analyser.

3.4 | SNP selection

SNPs were selected either from GWAS or candidate gene association studies.^{12,13}

Entering criteria included genes with an Odds Ratio (OR) for CAD ≥ 1.1 and, simultaneously, with a Minor Allele Frequency (MAF)

$>5\%$. Genes with low Hardy-Weinberg disequilibrium $P < .0016$ (after Bonferroni correction) were automatically excluded.

According to their possible CAD-related function, we have included 32 genes associated to cell cycle, cellular migration and inflammation (rs1333049 (9p21.3), rs4977574 (CDKN2B), rs618675 (GJA4), rs17228212 (SMAD3), rs17465637 (MIA3), rs12190287 (TCF21), rs3825807 (ADAMTS7), rs11556924 (ZC3HC1), rs12526453 (PHACTR1)); genes involved in pro-oxidative status (rs1801133 (MTHFR 677), rs1801131 (MTHFR 1298), rs705379 (PON 1), rs662 (PON192), rs854560 (PON 55), rs6922269 (MTHFD1L)); genes associated with modifiable risk factors such as lipids metabolism, hypertension and diabetes/obesity (rs2114580 (PCSK9), rs20455 (KIF6), rs7412/rs429358 (APOE), rs964184 (ZNF259), rs599839 (PSRC1), rs5186 (AT1R), rs699 (AGT), rs4340 (ACE), rs4402960 (IGF2BP2), rs1326634 (SLC30A8), rs266729 (ADIPOQ), rs7903146 (TCF7L2), rs17782313 (MC4R), rs1801282 (PPARG), rs1884613 (HNF4A), rs8050136 (FTO) and rs1376251 (TAS2R 50).

Details of SNPs used in GRS model are presented in TableS1.

3.5 | Genotype analysis

Genomic DNA was extracted from 80 μ L of peripheral blood using a standard phenol-chloroform method. To identify the 32 genotypes, a TaqMan allelic discrimination assay was performed using labelled probes and primers pre-established by the supplier (TaqMan SNP Genotyping Assays, Applied Biosystems).

The genotyping reaction was amplified and detected on a Applied Biosystems 7300 Real Time PCR System and genotypes were determined using the 7300 System SDS Software without any prior knowledge of individual clinical data (Applied Biosystems, Foster City, CA, USA).

3.6 | GRS Calculation

A GRS with 32 SNPs was constructed as an unweighted score, calculated by adding the number of risk alleles across each SNP. A value of 0 was given for the homozygous wild-type genotype, 1 for heterozygous and 2 for the homozygous mutated genotype (risk). The maximum number of risk alleles possessed by one individual was 37 and the minimum was 14. The average number of risk alleles possessed by the participants was 25.9 ± 3.7 (mean \pm SD).

3.7 | Statistical analysis

Deviation from Hardy-Weinberg equilibrium for genotypes at individual locus was assessed using the Chi-square test.

Comparison of baseline data and biochemical data according to cardiovascular death outcome were analysed by Chi-square for categorical variables. Student's *t* or Mann-Whitney test were performed for continuous variables as appropriate.

GRS median and respective terciles were determined from observed counts and compared by Chi-square analysis. A median of 26 was obtained and used to categorise the GRS variable as high GRS (higher than median) and low GRS (lower than median).

Kaplan-Meier (KM) survival function with Log-rank statistics was tested to address differences in CV mortality according to high/low GRS. Cumulative hazard statistics were also obtained for superior and inferior tertiles of GRS.

Hazard ratios (HR) and the respective confidence intervals values were determined for additive GRS and 7 other major CAD mortality risk factors: smoking status, diabetes, hypertension, dyslipidemia, 3-vessel disease, stable angina and double antiplatelet drug treatment. The critical value for entering the variables in the model was set at $P=.05$. Multivariable Cox proportional-hazard regression analyses also addressed the all-cause mortality including the same set of covariates.¹⁴

C-index was calculated according to Pencina et al.¹⁵ and Harrell et al.,¹⁶ using R software *survcomp* package (Boston, USA). The C-index may take values ranging from 0.5 (no discrimination) to 1.0 (perfect discrimination) and to be statistically significant the 95% confidence interval may not include 0.5.

Other Statistical analyses were performed using Statistical Package for the Social Sciences Software version 19.0 (IBM, Armonk, NY, USA) and R (version 3.2.0). All P -values were two sided, statistically significant for $P<.05$.

4 | RESULTS

4.1 | Baseline and biochemical characteristics of the population

A total of 1464 patients (mean age 53.3 ± 8.0 years, 79.7% male) with documented CAD entered this study, during a mean follow-up period of 58.3 (interquartile range 25.8-88.1) months. Of these patients, 156 patients (10.7%) (mean age 55.7 ± 7.2 years, 80.1% male) died, 107 (7.3%) (mean age 55.5 ± 7.6 years; 79.4% male) of CV causes.

Baseline and biochemical characteristics of the studied population are shown in Table 1 and Table 2, respectively.

Patient population analysis according to the end-point of cardiovascular death found that patients with the adverse outcome were more elderly, physical inactive and had more dyslipidemia and diabetes ($P<.05$) (Table 1). Higher heart rates and 3-vessel disease were also more frequent in non-survivors ($P<.0001$) (Table 1). High additive GRS (higher than median) was found in 65.4% of the patients who died from CV causes as compared to 54.2% of the rest of population ($P=.024$) (Table 1).

In addition, biochemical data (Table 2) showed a pro-oxidative and atherogenic profile in patients who suffer cardiovascular mortality, showing significantly ($P<.05$) increased platelets, leucocytes and serum levels of homocysteine.

4.2 | Multivariate analysis: GRS and traditional risk factors

After dividing the GRS into tertiles, the distribution in each subgroup was as follows: in the first tertile ($n=511$) the GRS mean was 22.0 ± 1.8 , in the second tertile ($n=458$) was 26.1 ± 0.8 and in the third tertile ($n=495$) the mean was 30.0 ± 1.8 .

Survival analysis with Cox regression model (Table 3) adjusted for eight variables showed that high GRS (HR=1.53, $P=.037$), three-vessel disease (HR=2.97, $P<.0001$), dyslipidemia (HR=3.64, $P=.012$) and diabetes (HR=1.75, $P=.004$) were independent risk factors for cardiovascular mortality, with statistical significance ($P<.05$). Patients with high GRS had an augmented probability of cardiovascular mortality of 53% in relation to those with low GRS (Table 3).

4.3 | Higher GRS reduces CV survival

Kaplan-Meier Survival-time plot for cardiovascular death by GRS is shown in Figure 1. A significant difference in survival between high and low GRS was observed (Log-rank test 5.6; $P=.018$). Specifically, the estimated survival probability was 70.8% for high GRS and 80.8% for low GRS, at the end of follow-up period.

Visual analysis of cumulative mortality curves (Figure 2) showed strikingly better survival of the lower GRS tertile. At the end of follow-up, 29.4% of the patients in the highest tertile died of CV causes whereas only 18.7% of the individuals in the lowest tertile suffered CV mortality.

4.4 | GRS improves risk prediction of CV mortality in a model with traditional risk factors

The additive GRS improved the risk prediction when incorporated in the mortality model including dyslipidemia, smoking status, diabetes, hypertension and three-vessel disease (Likelihood Ratio Test: 5.06, $P=.024$). The C-index for the final mortality model in Table 3 was 0.71, reflecting an acceptable ability to discriminate which of two randomly chosen patients had a worst prognostic. The proportional hazards assumption was verified (Harrell Test: 4.9, $P=.56$) and the model showed a good adjustment quality since the estimated survival curves are closer to the KM estimates and it is possible to distinguish between high, medium and low prognostic index groups (Figure 3).

5 | DISCUSSION

Reports relating genetic scores and cardiovascular mortality in patients with established CAD are scarce. Genetic profile does not change overtime and it is expected that higher GRS will translate into reduced survival and incident cardiovascular events. Tada et al.¹⁷ reported that two different GRS with 23 and 50 SNPs were associated with incident CAD in a robust and long-term survival analysis.¹⁷ Nevertheless, Labos et al.¹⁸ have recently reported a 30 SNP CAD/MI GRS unassociated with all-cause mortality, recurrent ACS or cardiac re-hospitalisation. But, follow-up was targeted 1 year after initial admission and we expect new data to be published.¹⁸

Also, Weijmans et al.¹⁹ reported that, in a population with clinical manifest vascular disease but without CAD, GRS did not improve the C-Statistics prediction of 10-year risk of new cardiovascular events, and that the Net Reclassification Index (NRI) was just marginally increased in the stable atherosclerosis group.¹⁹

TABLE 1 Baseline descriptive statistics according to coronary heart disease event status

Variables	Cardiovascular mortality (n=107)	Non-cardiovascular mortality (n=1357)	P-value
Male sex, n (%)	82 (79.4)	1082 (79.7)	.942
Age, years, mean (SD)	55.5±7.6	53.1±8	.002
Dyslipidemia ^a , n (%)	103 (96.3)	1175 (86.6)	.004
Total cholesterol, mg/dL	190 (104-321)	180 (77-437)	.058
HDL, mg/dL	38 (20-70)	41 (18.2-115.8)	.004
LDL, mg/dL	109.4 (23.4-245)	104.6 (15.6-298)	.214
Lipoprotein (a), mg/dL	25.9 (1.9-197.5)	20.4 (1.5-241)	.352
Apolipoprotein B, mg/dL	99.3 (6.8-173)	94.2 (4.9-256.9)	.003
Triglycerides, mg/dL	155 (44-571)	141 (31-2500)	.027
Level of exercise ^b , n (%)	29 (27.1)	510 (37.6)	.030
Smoking status ^c , n (%)	47 (43.9)	633 (46.6)	.587
Hypertension, n (%)	77 (72)	969 (71.4)	.903
SBP, mm Hg, mean (SD)	135.5±20.5	138.1±20.7	.210
DBP, mm Hg, mean (SD)	80.6±11.2	82.7±11.8	.071
PWV, m/s, mean (SD)	8.8±2.1	8.6±1.9	.351
Diabetes, n (%)	54 (50.5)	447 (32.9)	<.0001
Fasting glucose, mg/dL	120 (67-393)	106 (53-458)	<.0001
BMI, kg/m ² , mean (SD)	29±4.4	28.6±4.2	.323
Heart rate, bpm, mean (SD)	75.5±13.6	68.5±12.3	<.0001
3-vessel disease, n (%)	45 (42.1)	282 (20.8)	<.0001
Stable angina, n (%)	25 (23.4)	290 (21.4)	.629
Additive GRS, ≥median, n (%)	70 (65.4)	735 (54.2)	.024
ACEI/ARB, n (%)	57 (53.3)	600 (44.4)	.075
Lipid lowering, n (%)	21 (77.8)	394 (70.6)	.423
Double AntiPlatelet, n (%)	26 (24.3)	565 (41.8)	<.0001

HDL, high-density lipoprotein; LDL, low-density lipoprotein; SBP, systolic blood pressure; DBP, diastolic blood pressure; PWV, pulse wave velocity; BMI, body mass index; bpm, beat per minute; GRS, genetic risk score; ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; Statistically significant for $P < .05$. ^aLDL > 100, HDL < 40 for men and < 45 for women, Triglycerides > 150, Apo B > 100 and non-HDL (Total cholesterol-HDL) > 130; ^bMore than 40 minutes/wk; ^cCurrent smokers or < 5 years of cessation. Biochemical variables are presented by median (minimum–maximum).

TABLE 2 Biochemical characteristics of the population

Variables	Cardiovascular mortality (n=107)	Non-cardiovascular mortality (n=1357)	P-value
Haemoglobin, g/dL	14.6 (9.1-17.6)	14.6 (8.6-18.2)	.696
Platelets, 10 ³ /μL	226 (116-533)	215 (44-893)	.034
Leucocytes, 10 ³ /μL	7.5 (4.4-12.4)	7.1 (3.4-24.6)	.003
Fibrinogen, mg/dL	388 (153.3-832)	384 (91-761)	.378
Homocysteine, μmol/L	13 (4.6-48.7)	12.4 (4.4-220.2)	.005
Hs-CRP, mg/dL	0.24 (0.02-7.9)	0.24 (0.01-24.8)	.203

Variables are presented by median (minimum–maximum). Hs-CRP, high sensitivity C-reactive protein; Statistically significant for $P < .05$.

Disease progression and adverse outcomes after established CAD are mostly related to plaque vulnerability. It has been hard to differentiate true culprit markers associated with CAD itself from risk factors associated with major adverse cardiovascular outcomes and more importantly cardiovascular mortality.

A huge burden of information concerning genetic variants has risen from GWAS and their association with risk for CAD is irrefutable. Some genetic risk variants, namely 9p21 have been associated with MACE in patients with established CAD including incident fatal/non-fatal MI, Cardiovascular mortality and all-cause mortality.²⁰⁻²²

Variables	B	S.E.	Wald	df	Hazard ratio (95% CI)	P-value
Additive GRS	0.427	0.205	4.356	1	1.532 (1.026-2.288)	.037
3-vessel disease	1.087	0.197	30.338	1	2.965 (2.014-4.366)	<.0001
Dyslipidemia	1.292	0.512	6.373	1	3.640 (1.335-9.926)	.012
Diabetes	0.562	0.194	8.365	1	1.754 (1.199-2.567)	.004

TABLE 3 Hazard ratios by Cox Regression Model^a for cardiovascular mortality

^aUsing forward wald cox regression method (SPSS v. 19.0) adjusted for covariates: hypertension; smoking status; stable angina and double antiplatelets. B, beta coefficient; S.E., standard error; df, degrees of freedom; CI, confidence interval; Statistically significant for $P < .05$.

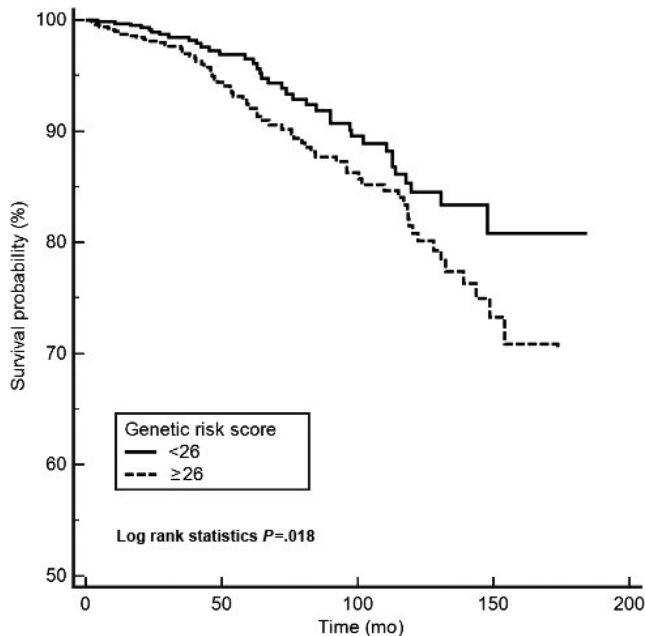


FIGURE 1 Kaplan-Meier survival time plot for cardiovascular mortality by genetic risk score (higher than median vs lower than median). There were 70 events (8.7%) among 805 patients in the higher category of the median and 37 events (5.6%) among 659 patients who were in the lower category. At the end of follow-up the survival probability in patients with high GRS was 80.8% and 70.8% in patients with low GRS. Logrank statistics was 5.6 with statistical significance ($P = .018$)

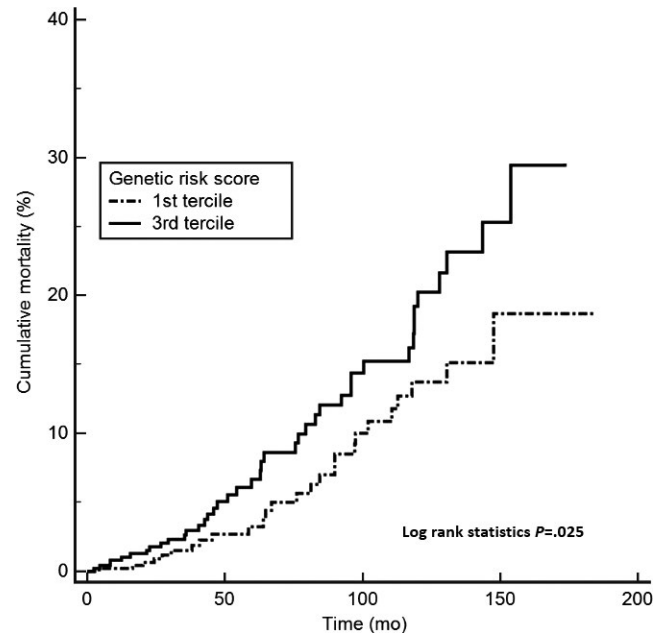


FIGURE 2 Cumulative incidence of cardiovascular mortality according to genetic risk score (3rd tercile vs 1st tercile). 26 patients (5.6%) among the 1st tercile had a cardiovascular death and 43 patients (8.7%) died with a cardiovascular cause between the 3rd tercile. At the end of follow-up, the probability of cardiovascular mortality was 29.4% in patients on the 3rd tercile and 18.7% in the 1st tercile. Log-rank statistics was 5.1 with statistical significance ($P = .025$)

Controversial results have been presented in these cohorts with total population numbers ranging from 478 to 3788 individuals emphasising that being associated with risk for CAD is not certain to influence outcomes. Hoppmann et al.²³ reported in an angiographical controlled cohort that the 40 studied polymorphisms were not significantly related with clinical outcomes at 3 years, including death, myocardial infarction, repeat revascularisation ($P \geq .05$) and the respective composite end-point of adverse events.²³ Recently, our research group has published a study performed in the Madeira Island population with a cohort of 1128 individuals, showing a significant association between a genetic variant of the ADAMTS7 gene and outcome.²⁴ Authors conclude that patients with documented CAD carry a common variant of the ADAMTS7 gene (G allele) associated with improved CV survival when compared to those carrying the A allele.²⁴ The native overfunctional ADAMTS7 allele (A) may accelerate VSMC migration and lead to neointimal thickening and atherosclerosis progression and acute

plaque events. These findings, although promising, should be amplified in larger and long-term studies and should be further explored in the outcome and therapeutic goals.

In this study with proven CAD disease patients, multivariate logistic regression analysis showed that besides the GRS, dyslipidemia, diabetes and three-vessel disease remained associated with cardiovascular mortality. There are several potential hypotheses that might explain these findings. The inadequate control for hyperglycaemic status is a key factor for CAD progression. We found higher glycaemic levels expressed by HBA1C when comparing the diabetic subpopulation with cardiovascular mortality vs the diabetic population without that adverse outcome. Cox et al.⁴ have reported, in a diabetic population, the association of a weighted GRS with 30 variants with CVD mortality (HR: 1.46 (1.08-1.96); $P = .01$). The same group also reported a trend for increased all-cause mortality (HR: 1.28 (1.05-1.56); $P = .01$), emphasising the power of genetic predisposition towards adverse outcomes.⁴

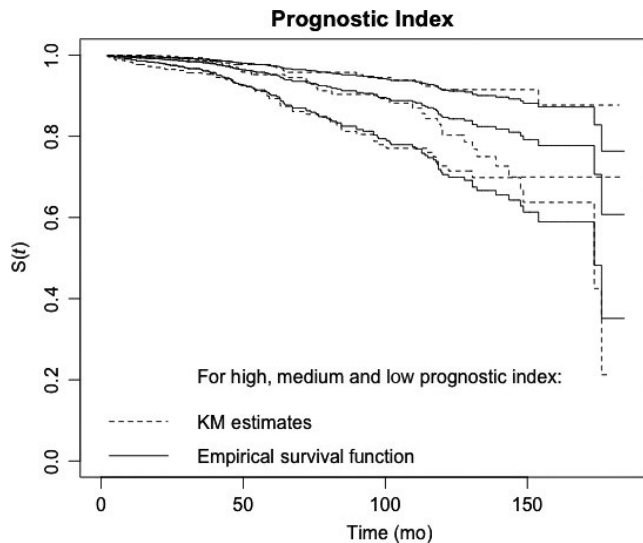


FIGURE 3 Estimated survival function obtained from the Cox model and Kaplan-Meier estimates for high, medium and low prognostic index. The predicted mean survival curves ($S(t)$) for each three risk groups are compared with the respective Kaplan-Meier survival curves, showing a good calibration since the pairs of predicted survival curves nearly coincide (with the exception of medium risk group after survival time of 100 months). KM, Kaplan-Meier

Dyslipidemia was more prevalent in patients with adverse outcome in univariate analysis; and it remained an independent predictor for cardiovascular mortality after multivariate analysis. All patients in this study had received a secondary preventive therapy. We have compared the drug treatment prescribed in the last follow-up visit and found no differences concerning ACE inhibitors ($P=.075$) and lipid lowering ($P=.448$) in CV mortality (53.3% and 77.8%, respectively) vs non-CV mortality (44.4% and 70.6%). Beta Blockers treatment was higher in non-cardiovascular mortality (61.3%) when compared with people who died with CV causes (51.4%) and this difference was significant ($P=.044$). We theorise that even with statin widespread coverage and lower lipid levels in most patients, some residual lipid-associated risk still has impact on cardiovascular disease progression. Antiplatelet drug treatment was superior in non-CV mortality group but apparently it did not influence the outcome in study after adjustment as covariate in Cox regression model.

On the other hand, three-vessel disease was more common among CV mortality group and this was an independent prognostic factor in multivariate logistic regression. Percutaneous transluminal coronary angioplasty was offered at the physician's discretion, but the proportion of diabetic patients being higher in that group might have affected the decision of PTCA vs CABG.

Overall, our findings demonstrated a prospective association of our additive GRS with incident cardiovascular mortality over a mean follow-up period of 58.3 months. This prospective association was not altered by lifestyle intervention and proper instituted therapy.

Our GRS is polygenic and also includes genetic predisposition for some known risk factors for CAD disease. Although we have

introduced secondary preventive strict strategies for cholesterol, and hypertension control, smoking cessation and exercise programs, we can conclude that genetic burden is not yet under total control and that, based on our data, still remains correlated with reduced survival and higher long-term cardiovascular mortality. We theorise that new pathways/targets are still to be found and additional therapies like PCSK9 inhibitors will still arise in a near future.

5.1 | Strengths and limitations

The majority of the studies focused on these aspects included thousands of individuals allocated into groups with heterogeneous populations which can constitute a limitation thereof, like RISCA, GENESIS-PRAXI and TRIUMPH Cohorts.¹⁸ This methodology to compare populations can be criticised because allocation is performed from different population groups which may influence the outcome of a genetic study. Recent results speak in favour of the feasibility and advantages of these studies performed with participants enrolled from homogeneous populations.²⁵

Our work includes solely participants born and living in the Madeira Archipelago. It consists of a Southern European Caucasian population considered homogeneous by its isolation for many years, with no genetic admixtures.

A total of 1464 patients (mean age 53.3 ± 7.9 years, 79.6% male) with documented CAD entered this study. This study had excellent retention rates, over a mean follow-up period of 58.3 months (interquartile range: 25.8–88.1 months) and cardiovascular outcomes were adjudicated by cardiology specialists.

On the other hand, our report has included a limited number of genetic variants associated with CAD disease. The expected P -value of a polygenic score is shown as a function of the inclusion threshold; therefore, including more SNPs into our score could increase the predictive power for CAD and other outcomes.

Furthermore, some non-flow was not an exclusion criterion and we agree that it may be a potential confounding variable. Markers like calcium score have arisen to account for this gap. Nevertheless, high calcium score is perhaps most helpful as a marker of disease burden, rather than for predicting likelihood of an event originating from an individual plaque.

ACKNOWLEDGEMENTS

We are very grateful to Elsa Sousa who made all the administrative procedures and to Rita Freitas, PhD (CIDEHUS, Évora University, Portugal) who reviewed the paper for statistical analysis.

AUTHOR CONTRIBUTIONS

A. Pereira concept/design the study; A. Pereira, MI Mendonça and R. Palma dos Reis performed data analysis/interpretation and drafting/review the article; AC. Sousa contributed to the acquisition of data; S. Borges drafted/reviewed the article; S. Freitas, E. Henriques and M. Rodrigues performed all the statistics analysis of the article; AI.

Freitas, G. Guerra and I. Ornelas performed the biochemical and genetic analysis; D. Pereira and A. Brehm performed a critical revision of the article.

DISCLOSURES

The authors declare that they have no conflict of interest.

REFERENCES

- Roger VL, Go AS, Lloyd-Jones DM, Benjamin EJ, Berry JD, Borden WB; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics – 2012 update: a report from the American Heart Association. *Circulation* 2012;125:e2-e220.
- Chan L, Boerwinkle E. Gene-environment interactions and gene therapy in atherosclerosis. *Cardiol Rev*. 1994;2:130-137.
- Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661-678.
- Cox AJ, Hsu FC, Ng MC, et al. Genetic risk score associations with cardiovascular disease and mortality in the Diabetes Heart Study. *Diabetes Care*. 2014;37:1157-1164.
- Hervé Abdi, Williams Lynne J.. Principal component analysis. *WIREs Comp Stat*. 2010;2:433-459.
- Chobanian AV, Bakris GL, Black HR, et al. The seventh report of the joint national committee on prevention, detection, evaluation and treatment of high blood pressure: the JNC 7 report. *JAMA*. 2003;289:2560-2572.
- Catapano AL, Graham I, De Backer G, et al. ESC/EAS Guidelines for the Management of Dyslipidaemias: the Task Force for the Management of Dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS) Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J*. 2016;37:2999-3058.
- Mons U, Müezziner A, Gellert C, et al. Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. *BMJ*. 2015;350:h1551.
Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care*. 2003; 26: 3160-3167.
- National Institute of Health, National Heart, Lung, and Blood Institute North American Association for the Study of Obesity. The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. NHLBI Obesity Education Initiative, NIH Publication Number: 00-4084, 2000.
- Asmar R, Benetos A, Topouchian J, et al. Assessment of arterial distensibility by automatic pulse wave velocity measurement: validation and clinical application studies. *Hypertension*. 1995;26:485-490.
- Coronary Artery Disease Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet*. 2011;43:339-344.
- Schunkert H, König IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011;2011(43):333-338.
- Hosmer D, Lemeshow S. *Applied Survival Analysis*, 3rd edn. New York: John Wiley & Sons; 2008.
- Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med*. 2004;23:2109-2123.
- Harrell FE, Lee KL, Mark DB. Tutorial in Biostatistics: multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361-387.
- Tada H, Melander O, Louie JZ, et al. Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. *Eur Heart J*. 2016;37:561-567.
- Labos C, Martinez SC, Leo Wang RH, et al. Utility of a genetic risk score to predict recurrent cardiovascular events 1 year after an acute coronary syndrome: a pooled analysis of the RISCA, PRAXY, and TRIUMPH cohorts. *Atherosclerosis*. 2015;242:261-267.
- Weijmans M, de Bakker PI, van der Graaf Y, et al. Incremental value of a genetic risk score for the prediction of new vascular events in patients with clinically manifest vascular disease. *Atherosclerosis*. 2015;239:451-458.
- Dutta A, Henley W, Lang IA, et al. The coronary artery disease-associated 9p21 variant and later life 20-year survival to cohort extinction. *Circ Cardiovasc Genet*. 2011;4:542-548.
- Gioli-Pereira L, Santos PC, Ferreira NE, Hueb WA, Krieger JE, Pereira AC. Higher incidence of death in multi-vessel coronary artery disease patients associated with polymorphisms in chromosome 9p21. *BMC Cardiovasc Disord*. 2012;12:61.
- Andreassi MG, Adlerstein D, Carpeggiani C, et al. Individual and summed effects of high-risk genetic polymorphisms on recurrent cardiovascular events following ischemic heart disease. *Atherosclerosis*. 2012;223:409-415.
- Hoppmann P, Erl A, Türk S, et al. No association of chromosome 9p21.3 variation with clinical and angiographic outcomes after placement of drug-eluting stents. *JACC Cardiovasc Interv* 2009;2:1149-1155.
- Pereira A, Palma Dos Reis R, Rodrigues R, et al. Association of ADAMTS7 gene polymorphism with cardiovascular survival in coronary artery disease. *Physiol Genomics*. 2016;48:810-815.
- Jorde LB, Wooding SP. Genetic variation, classification and 'race'. *Nat Genet*. 2004;36:S28-S33.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Pereira A, Mendonca MI, Sousa AC, et al. Genetic risk score and cardiovascular mortality in a southern european population with coronary artery disease. *Int J Clin Pract*. 2017;71:e12956. <https://doi.org/10.1111/ijcp.12956>

Genetic Risk Analysis of Coronary Artery Disease in a Population-based Study in Portugal, Using a Genetic Risk Score of 31 Variants

Andreia Pereira,¹ Maria Isabel Mendonça,¹ Sofia Borges,¹ Sónia Freitas,¹ Eva Henriques,¹ Mariana Rodrigues,¹ Ana Isabel Freitas,² Ana Célia Sousa,¹ António Brehm,² Roberto Palma dos Reis³

Unidade de Investigação, Hospital Dr. Nélio Mendonça,¹ Funchal – Portugal

Laboratório de Genética Humana, Universidade da Madeira,² Funchal – Portugal

Faculdade de Ciências Médicas, Universidade Nova de Lisboa,³ Lisboa – Portugal

Abstract

Background: Genetic risk score can quantify individual's predisposition to coronary artery disease; however, its usefulness as an independent risk predictor remains inconclusive.

Objective: To evaluate the incremental predictive value of a genetic risk score to traditional risk factors associated with coronary disease.

Methods: Thirty-three genetic variants previously associated with coronary disease were analyzed in a case-control population with 2,888 individuals. A multiplicative genetic risk score was calculated and then divided into quartiles, with the 1st quartile as the reference class. Coronary risk was determined by logistic regression analysis. Then, a second logistic regression was performed with traditional risk factors and the last quartile of the genetic risk score. Based on this model, two ROC curves were constructed with and without the genetic score and compared by the Delong test. Statistical significance was considered when *p* values were less than 0.05.

Results: The last quartile of the multiplicative genetic risk score revealed a significant increase in coronary artery disease risk (OR = 2.588; 95% CI: 2.090-3.204; *p* < 0.0001). The ROC curve based on traditional risk factors estimated an AUC of 0.72, which increased to 0.74 when the genetic risk score was added, revealing a better fit of the model (*p* < 0.0001).

Conclusions: In conclusion, a multilocus genetic risk score was associated with an increased risk for coronary disease in our population. The usual model of traditional risk factors can be improved by incorporating genetic data. (Arq Bras Cardiol. 2018; [online].ahead print, PP.0-0)

Keywords: Coronary Artery Disease / history; Coronary Artery Disease / morbidity; Mortality; Polymorphism, Genetic; Epidemiology; Risk Factors.

Introduction

Coronary artery disease (CAD) has become a major health problem worldwide, with increasing prevalence and high morbidity and mortality. Traditional risk factors (TRFs) are insufficient to identify asymptomatic high-risk individuals. Epidemiology and family studies have long documented that approximately 50% of the susceptibility for heart disease is genetic.¹ Knowledge of genetic predisposition to cardiac disease is crucial for its comprehensive prevention and treatment.

Although much of the genetic basis of coronary disease remains to be discovered, some progress has been made using both candidate gene and genome-wide association studies (GWAS).² In fact, a number of genetic variants have been previously identified at several genomic regions associated with CAD.²

Until now, the risk attributable to any individual variant has been modest. However, discovering and combining multiple loci with modest effects into a global genetic risk score (GRS) could improve the identification of high-risk populations and improve individual risk assessment.

Therefore, the purpose of this work was to generate a multilocus GRS based on common variants previously shown to be associated with CAD, and evaluate whether it is independent of TRFs and improves the predictive ability of a model based only on TRFs.

Methods

Study Population

Study population was enrolled from GENEMACOR (GENEs in Madeira Island Population with CORonary artery disease), a population-based ongoing case-control registry of CAD with 2,888 participants, 1,566 cases (mean age 53.3 ± 8.0 years, 79.1% male) and 1322 controls (mean age 52.7 ± 7.8 years, 76.4% male). Cases were selected from patients discharged after being admitted for myocardial infarction/unstable angina diagnosed according to the previously described

Mailing Address: Andreia Pereira •

Avenida Luís de Camões, nº 57. Postal Code 9004-514, Funchal – Portugal

E-mail: andreiapereira21@gmail.com

Manuscript received June 14, 2017, revised manuscript October 12, 2017, accepted February 22, 2018

DOI: 10.5935/abc.20180107

criteria,³ or with CAD confirmed by coronary angiography with ≥ 1 coronary lesions of $\geq 70\%$ stenosis in ≥ 1 major coronary artery or its primary branches. Absent or non-flow limiting atheroma was excluded from the analysis. The control group consisted of healthy volunteers, without symptoms or history of CAD, selected from the same population. All controls underwent clinical assessment of conventional cardiovascular risk factors, an electrocardiogram (ECG), and, in doubtful cases, an exercise stress test, a stress echocardiography or computerized tomography for calcium scoring. Cases and controls were matched for gender and age.

Inclusion criteria comprised an age limit of 65 years and being a permanent resident to avoid genetic admixture. Principal Component Analysis (PCA)⁴ was used for analysis of population stratification for possible genetic admixture and detection of significant genetic outliers ($< 5\%$).⁴

The study was approved by the Hospital ethics committee according to the Declaration of Helsinki and all patients provided written informed consent.

Data collection

Data was collected from all subjects in a standardized file comprising demographic, clinical characteristics and TRFs traditional risk factors (gender, age, level of exercise, smoking status, arterial hypertension, dyslipidemia, diabetes, family history of CAD, body mass index (BMI), heart rate and pulse wave velocity (PWV).

"Smokers" referred to current smokers or subjects with less than 5 years of smoking cessation.⁵

Essential hypertension was considered when patients, at the entry into this study, were already diagnosed and/or had been on antihypertensive medication for more than 3 months or newly diagnosed hypertensives with systolic blood pressure (SBP)/diastolic blood pressure (DBP) $\geq 140/90$ mmHg measured on at least 3 occasions.⁶

Dyslipidemia was defined for control population as low-density lipoprotein (LDL) > 140 mg/dL, high-density lipoprotein (HDL) < 45 mg/dL for women and < 40 mg/dL for men, Triglycerides > 150 mg/dL and apolipoprotein (Apo) B > 100 mg/dL. For patients (at high risk) dyslipidemia was considered when LDL > 100 , HDL < 45 mg/dL for women and < 40 mg/dL for men, triglycerides > 150 mg/dL, Apo B > 100 mg/dL and non-HDL (total cholesterol-HDL) > 130 mg/dL.⁷

Subjects were classified as having diabetes if they were taking oral anti-diabetic medication or insulin or if their fasting plasma glucose was higher than 7.0 mmol/L or 126 mg/dL.⁸

Subjects were considered to have a family history of premature cardiovascular disease (CVD) if the father or brother had been diagnosed with CVD under the age of 55 or mother or sister under the age of 65.

The definition of other TRFs was based on the standard criteria, as previously reported.^{9,10}

Biochemical analysis

Blood samples were extracted after 12 hours' fasting. Biochemical analyses were performed at the Central

Laboratory of the Hospital, according to standard techniques. In order to determine total cholesterol, HDL, LDL, triglycerides and glucose, blood samples were placed in dry tubes, centrifuged half an hour later at 3,500 g and subsequently quantified by an enzymatic technique using an "AU 5400" (Beckman Coulter) autoanalyzer. Biochemical markers such as lipoprotein-a – Lp(a), (Apo B), and high-sensitivity C-reactive protein (hs-CRP) were quantified by immunoturbidimetry also using an "AU 5400" (Beckman Coulter) automatic system.

Single Nucleotide Polymorphisms (SNP) selection

Two parallel approaches were employed to identify SNPs for the GRS. In the first approach, we searched the National Human Genome Research Institute database, which included SNPs identified by means of GWAS and catalogued based on phenotype and/or trait. We searched for the keywords: "coronary artery disease", "coronary disease", "myocardial infarction" and "early myocardial infarction." The second approach included SNPs that were identified through candidate gene approaches, included in a published GRS for CAD.

Including criteria included genes described in previous studies with an Odds Ratio (OR) for CAD ≥ 1.1 and a minor allele frequency (MAF) $> 5\%$. Genes with low Hardy-Weinberg equilibrium ($p < 0.002$) (after Bonferroni correction) were excluded.

In total, 33 SNPs were selected according to their possible CAD-related function: association with cell cycle, cellular migration and inflammation (rs1333049 (9p21.3), rs4977574 (CDKN2B), rs618675 (GJA4), rs17228212 (SMAD3), rs17465637 (MIA3), rs12190287 (TCF21), rs3825807 (ADAMTS7), rs11556924 (ZC3HC1), rs12526453 (PHACTR1)); genes involved in pro-oxidative status (rs1801133 (MTHFR 677), rs1801131 (MTHFR 1298), rs705379 (PON 1), rs662 (PON 192), rs854560 (PON 55), rs6922269 (MTHFD1L)); genes associated with modifiable risk factors such as lipids metabolism, hypertension and diabetes/obesity (rs3798220 (LPA), rs2114580 (PCSK9), rs20455 (KIF6), rs7412/rs429358 (APOE), rs964184 (ZNF259), rs599839 (PSRC1), rs5186 (AT1R), rs699 (AGT), rs4340 (ACE), rs4402960 (IGF2BP2), rs1326634 (SLC30A8), rs266729 (ADIPOQ), rs7903146 (TCF7L2), rs17782313 (MC4R), rs1801282 (PPARG), rs1884613 (HNF4A), rs8050136 (FTO) and rs1376251 (TAS2R 50)) (Supplementary Table 1).

Genetic analyses

Genetic analyses were performed at the Human Genetics Lab of the University of Madeira. Genomic DNA was extracted from 80 μ l of peripheral blood using a standard phenol-chloroform method. A TaqMan allelic discrimination assay for genotyping was performed using labelled probes and primers pre-established by the supplier (TaqMan SNP Genotyping Assays, Applied Biosystems).

All reactions were done on an Applied Biosystems 7300 Real Time PCR System and genotypes were determined using the 7300 System SDS Software (Applied Biosystems, Foster City, USA) without any prior knowledge of individual's clinical data. Quality of genotyping techniques was controlled by the

inclusion of one non-template control (NTC) in each plate of 96 wells. All SNPs TaqMan assays had blind duplicates accounting for 20% of all samples. Some SNP genotypes were randomly confirmed by conventional direct DNA sequencing, as 10-15% of all samples were re-amplified for sequencing. Call rates for SNPs in the GRS were 98%-100% and a minimum 95% call rate was set for quality control.

Computation of the GENETIC RISK SCORE

We have tested several models to construct the GRS using both non-weighted and weighted scores, taking into consideration each pattern of inheritance for each gene locus. An additive score (AGRS) was generated, i.e., for each one of the 31 variants a score of 0, 1, and 2 was defined as there were 0, 1 or 2 risk alleles, by calculating the accumulated sum of the risk alleles in these variants. Each individual could be assigned a GRS of 0-62. Additionally, a multiplicative GRS (MGRS) was calculated by multiplying the relative risk for each genotype.

Validation of the risk score calculation was performed in a random sample of 597 patients (20%).

Statistical analysis

Categorical variables were expressed by frequencies and percentages and compared by the Chi-squared test or Fisher's exact test. Continuous variables were expressed as mean \pm standard deviation (SD) or median (1st quartile – 3rd quartile) and compared by Student's t-test (unpaired) or Mann-Whitney, as appropriate. The Kolmogorov-Smirnov test and the Levene's test were used to test the assumption of normality and the homogeneity of the variables. All analyses were considered significant when p values were less than 0.05.

Binary logistic regression was used to determine the combined and separate effects of the variables on the risk for angiographic CAD. GRS was modeled using as a continuous variable and as quartiles, using the first quartile as the reference category. Multivariate analyses were used to adjust for 7 covariates also reported to be associated with CAD. We plotted receiver operating characteristic (ROC) curves and calculated the area under the curve (AUC) for logistic regression models including TRFs without and with GRS (quartiles). Pairwise comparison of ROC curves was performed using the Delong test.¹¹ The model calibration was tested with Hosmer-Lemeshow goodness-of-fit test. A P-value less than 0.05 was considered statistically significant. Collinearity between the variables was measured by assessment of tolerance and variance inflation factor (VIF).

Associations of SNPs with CAD were considered significant at $p < 0.05$ and in aggregate with GRS models at $p < 0.0015$ applying Bonferroni correction.

For MAF of 30%, the study had 70% power to detect an OR for CAD of 1.3 and $> 90\%$ for $OR \geq 1.35$, for 2-sided alpha of < 0.05 for 2,000 cases and 1,000 controls. Power calculations used *G power Statistical Power Analyses*.

The potential of GRS to improve individual risk stratification then was measured using the net reclassification improvement (NRI) method,¹² defined as the percentage of subjects in

each subgroup changing categories when the new model of GRS (in quartiles) was added. The integrated discrimination improvement (IDI), defined as the incremental improvement prognostic value of GRS, was compared between cases and controls. NRI was computed by categorical and non-categorical (continuous) variables using the PredictABEL package available in R software (version 3.2.0).

Statistical analyses were performed using SPSS version 19.0 (IBM), MedCalc version 13.3.3.0 and R software version 3.1.2.

Results

Baseline characteristics of the population

Table 1 shows the baseline characteristics of our population. As expected, cases and controls showed no significant differences concerning gender and age, since this was a selection criterion. Higher frequency of dyslipidemia, diabetes, hypertension, physical inactivity, smoking habit, alcohol consumption, and family history of premature cardiovascular disease was found in CAD patients when compared to the controls ($p < 0.0001$). Also, PWV, BMI and waist-to-height ratio were higher in cases than in controls, with statistical significance ($p < 0.05$) (Table 1). The other biochemical variables analyzed such as hemoglobin, leucocytes, fibrinogen, homocysteine and hs-CRP > 3 showed significantly higher levels in the coronary patients group when compared to the controls ($p < 0.05$) (Table 1).

Computation and analysis of Genetic Risk Score

Deviation from Hardy-Weinberg equilibrium for the 33 genotypes at individual loci were assessed using the Chi-squared test and $p < 0.002$ with Bonferroni correction for all SNPs included. LPA gene variant was excluded for further analyses due to its low Hardy-Weinberg p-value ($p < 0.002$). Linkage disequilibrium for the mutually adjusted SNPs within the genes was studied. CDKN2B gene was excluded because of the strong linkage disequilibrium with another selected SNP, rs1333049, which resides in the 9p21 region. The remaining 31 SNPs were included for further analysis (Supplementary Table 1).

In this study, the MGRS had the highest AUC value for assessing the risk for CAD disease with a specificity of 62.3% and sensitivity of 54% (data not shown) and therefore this model was computed in the subsequent analyzes (Supplementary Table 2).

The MGRS of 31 SNPs was significantly higher in CAD cases than in controls (0.67 ± 0.73 vs 0.48 ± 0.53 ; $p < 0.0001$), even by quartile and gender discrimination (Table 2).

A normal distribution of risk alleles in the total sample set including cases and controls is shown in Figure 1. While CAD patients exhibited lower GRS values, risk alleles were more prevalent in this group than in controls. In CAD patients, a mean of 27 risk alleles was seen in 52% of the individuals, and a mean of 26 risk alleles was found in 53% of controls (Figure 1).

When analyzed in deciles, GRS showed that the increase in the number of risk alleles was significantly associated with CAD

Table 1 – Baseline characteristics of our study population

Variables	Cases (n = 1566)	Controls (n = 1322)	P value
Age, years	53.3 ± 8.0	52.7 ± 7.8	0.053
Male Gender, n (%)	1238 (79.1%)	1010 (76.4%)	0.087
Dyslipidemia [†] , n (%)	1398 (89.3)	1103 (83.4)	0.0001
Total Cholesterol, mg/dl	180.0 (154.0 – 213.0)	205.0 (181.0 – 234.0)	< 0.0001
LDL, mg/dl	104.6 (82.8 – 128.7)	127.2 (104.7 – 152.3)	< 0.0001
HDL, mg/dl	41.0 (35.0 – 49.0)	48.0 (41.0 – 57.0)	< 0.0001
Triglycerides, mg/dl	141.0 (102.0 – 210.0)	121.0 (89.0 – 174.0)	< 0.0001
Apolipoprotein B, mg/dl	93.9 (75.5 – 113.3)	92.5 (43.0 – 115.8)	< 0.0001
Lipoprotein (a), mg/dl	20.4 (9.2 – 62.0)	12.8 (8.8 – 29.3)	< 0.0001
Diabetes, n (%)	533 (34.0)	175 (13.2)	< 0.0001
Fasting glucose, mg/dl	106.0 (96.0 – 129.0)	99.0 (91.0 – 109.0)	< 0.0001
Hypertension, n (%)	1114 (71.1)	700 (53.0)	< 0.0001
SBP, mmHg	137.9 ± 20.8	136.2 ± 18.1	0.024
DBP, mmHg	82.6 ± 11.8	83.9 ± 11.1	0.002
Heart rate, bpm	68.8 ± 12.5	72.3 ± 11.5	< 0.0001
PWV, m/s	8.6 ± 1.9	8.3 ± 1.7	< 0.0001
Smoking status [*] , n (%)	730 (46.6)	309 (23.4)	< 0.0001
Level of exercise [*] , n (%)	573 (36.6)	761 (57.6)	< 0.0001
Alcohol, g/day	24.7 ± 49.7	18.2 ± 28.2	< 0.0001
BMI, kg/m ²	28.6 ± 4.2	28.1 ± 4.5	0.007
Waist/Height	0.61 ± 0.06	0.59 ± 0.07	< 0.0001
Family history, n (%)	373 (23.8)	167 (12.6)	< 0.0001
Hemoglobin, g/dl	14.6 (13.8 – 15.4)	14.7 (14 – 15.4)	0.001
Leucocytes, 103/ μ l	7.1 (6 – 8.3)	6.6 (5.6 – 7.8)	< 0.0001
Fibrinogen, mg/dl	387 (337 – 444)	361 (315 – 409)	< 0.0001
Homocysteine, μ mol/L	12.2 (10 – 14.9)	11.4 (9.7 – 13.6)	< 0.0001
Hs-CRP, mg/L > 3, n (%)	648 (41.4)	496 (37.5)	0.035

[†] Controls: LDL > 140 mg/dL, HDL < 40 mg/dL for men and < 45 mg/dL for women; triglycerides > 150 mg/dL, APO B > 100 mg/dL. Cases: LDL > 100 mg/dL; triglycerides > 150 mg/dL, HDL < 40 mg/dL for men and < 45 mg/dL for women; APO B > 100 mg/dL, non HDL > 130 mg/dL; *More than 40 min/week; †Current smokers or < 5 years of cessation; HDL: high density lipoprotein; LDL: low density lipoprotein; SBP: systolic blood pressure; DBP: diastolic blood pressure; PWV: pulse wave velocity; BMI: body mass index; Hs-CRP: high sensitivity C-reactive protein. Categorical variables compared by the Chi-square test. Continuous variables expressed as mean ± standard deviation (using Student's t-test) and biochemical variables as median (1st quartile – 3rd quartile) (using Mann-Whitney's test). Statistical significance: p < 0.05.

as shown by inter-deciles p values (1st decile: OR = 0.612 (0.439 – 0.853), p = 0.004; 9th decile: OR = 0.957 (0.1400 – 2.734), p < 0.0001 and last decile: OR = 2.472 (1.755 – 3.482), p < 0.0001) (Figure 2).

A logistic regression analysis was performed with GRS quartiles, using the first as the reference category. Results showed an increase in CAD risk with statistical significance across the 2nd, 3rd and 4th quartiles with respective ORs and CIs of 1.372 (1.114 – 1.689), 1.878 (1.522 – 2.317) and 2.588 (2.090 – 3.204), respectively (data not shown).

A multivariable predictive model for CAD incorporating GRS quartiles and TRFs is presented in Table 3. The 4th GRS quartile has intermediate contribution to CAD phenotype

– OR = 2.727 (2.162 – 3.439), greater than dyslipidemia – OR = 1.298 (1.023 – 1.646) and hypertension – OR = 2.067 (1.744 – 2.450). The reduced contribution of dyslipidemia on CAD risk may be due to standard use of statins in CAD patients. Extended adjustment for confounding variables (gender, age, heart rate, PWV, low exercise level, BMI and family history of CAD) revealed modest increases in the OR for TRFs and the 2nd and 3rd quartiles of GRS.

We used VIF to test for multi-collinearity among the variables included in our GRS adjusted logistic regression model. Tolerance and VIF were respectively > 0.1 and < 10 attesting for no significant collinearity between variables included in the adjustment model.

Table 2 – Distribution of multiplicative genetic risk score (MGRS) for cases and controls by quartiles and gender

Variables	Cases (n = 1566)	Controls (n = 1322)	p value
MGRS	0.67 ± 0.73	0.48 ± 0.53	< 0.0001
1 st Quartile	0.18 ± 0.05	0.17 ± 0.05	
2 nd Quartile	0.33 ± 0.05	0.33 ± 0.05	
3 rd Quartile	0.52 ± 0.07	0.52 ± 0.07	< 0.0001
4 th Quartile	1.35 ± 1.02	1.18 ± 0.88	
MGRS male	0.67 ± 0.77	0.48 ± 0.44	< 0.0001
MGRS female	0.65 ± 0.58	0.51 ± 0.74	0.006

MGRS was expressed as mean ± standard deviation (SD) (using Student's t-test). Statistical significance: $p < 0.05$.

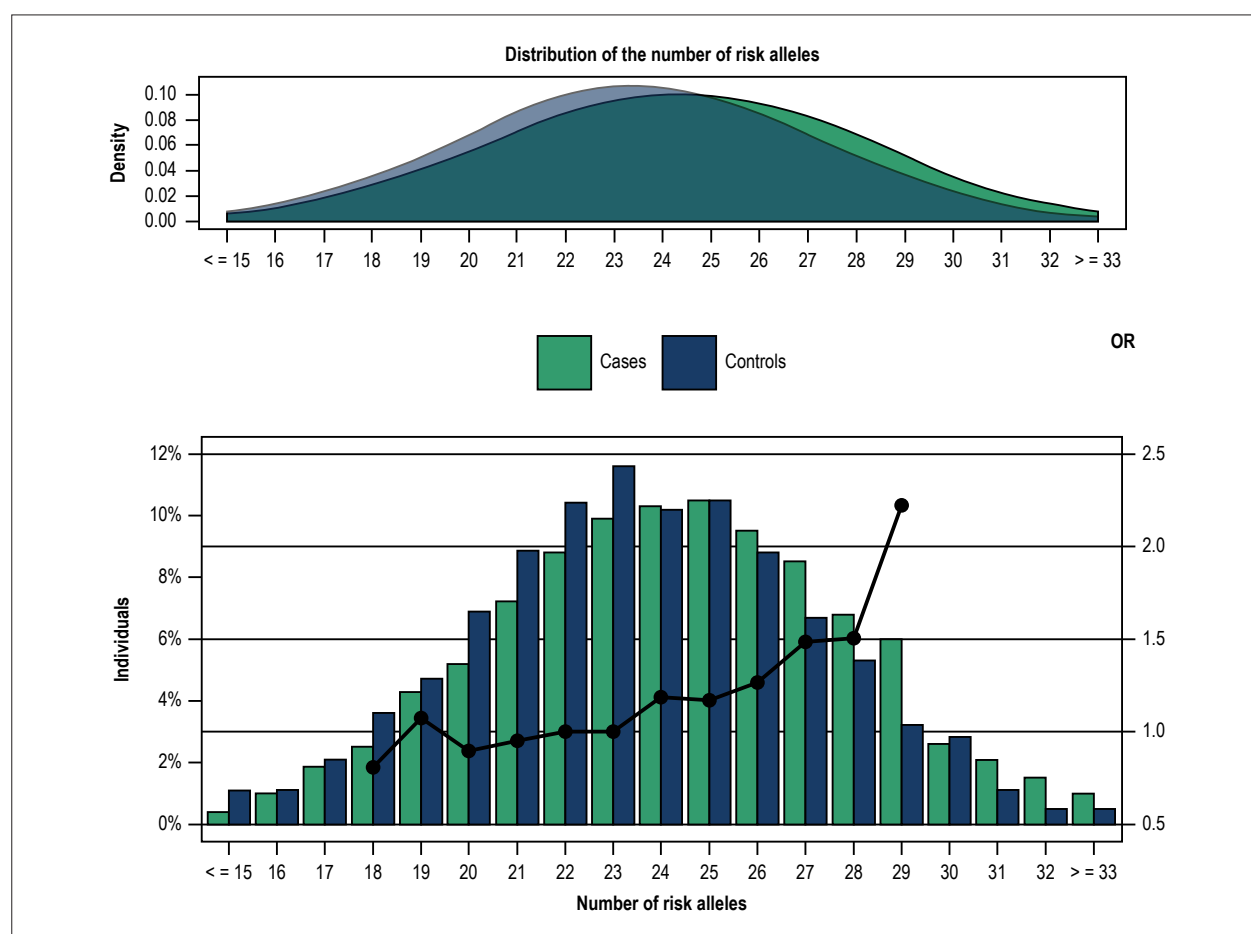


Figure 1 – Distribution of the number of risk alleles by cases and controls. A logistic regression model was used to determine the coronary artery disease risk by the number of risk alleles compared to the number of reference alleles (23 alleles, in relation to the median value of the controls). Dots: regression analysis odds ratio for coronary artery disease.

Two ROC curves were plotted based on the TRFs without and with the GRS (Figure 3). The first ROC curve estimated an AUC of 0.72, which increased to 0.74 when the GRS was added, revealing a better fit of the model ($p < 0.0001$) (Figure 3).

The NRI and its p value were used to make conclusions about improvements in prediction performance gained by

adding a set of biomarkers to an existing risk prediction model. The addition of GRS quartiles to TRF improved the risk classification of the models (Table 4). This new marker provided a continuous NRI of 31% (95% CI: 23.8-38.3%; $p < 0.0001$) with 14.6% reclassification of CAD patients and 16.4% of healthy control population (Table 4).

Table 3 – Multivariate analysis performed with the multiplicative genetic risk score (MGRS) (quartiles) and traditional risk factors

Variables	OR* (95% CI)	p value	OR* (95% CI)	p value
MGRS (Quartiles)	-----	-----	-----	< 0.0001
2 nd	1.355 (1.082 – 1.698)	0.008	1.406 (1.107 – 1.786)	0.005
3 rd	1.934 (1.539 – 2.429)	< 0.0001	2.006 (1.575 – 2.554)	< 0.0001
4 th	2.727 (2.162 – 3.439)	< 0.0001	2.657 (2.083 – 3.389)	< 0.0001
Smoking	3.440 (2.887 – 4.100)	< 0.0001	3.651 (3.030 – 4.401)	< 0.0001
Diabetes	3.138 (2.559 – 3.847)	< 0.0001	3.436 (2.763 – 4.273)	< 0.0001
Hypertension	2.067 (1.744 – 2.450)	< 0.0001	2.187 (1.816 – 2.633)	< 0.0001
Dyslipidemia	1.298 (1.023 – 1.646)	0.032	1.344 (1.044 – 1.731)	0.022
Constant	0.186	< 0.0001		

Using forward Wald method (SPSS vs. 19.0); Dyslipidemia. Controls: LDL > 140 mg/dL, HDL < 40 mg/dL for men and < 45 mg/dL for women; triglycerides > 150 mg/dL, APO B > 100 mg/dL. Cases: LDL > 100 mg/dL; triglycerides > 150 mg/dL, HDL < 40 mg/dL for men and < 45 mg/dL for women; APO B > 100 mg/dL, non HDL > 130 mg/dL; OR*: odds ratio adjusted for age and gender; OR: odds ratio adjusted for gender, age, heart rate, pulse wave velocity, sedentary life style, alcohol, body mass index and family history; CI: confidence interval; Statistically significant for p < 0.05.

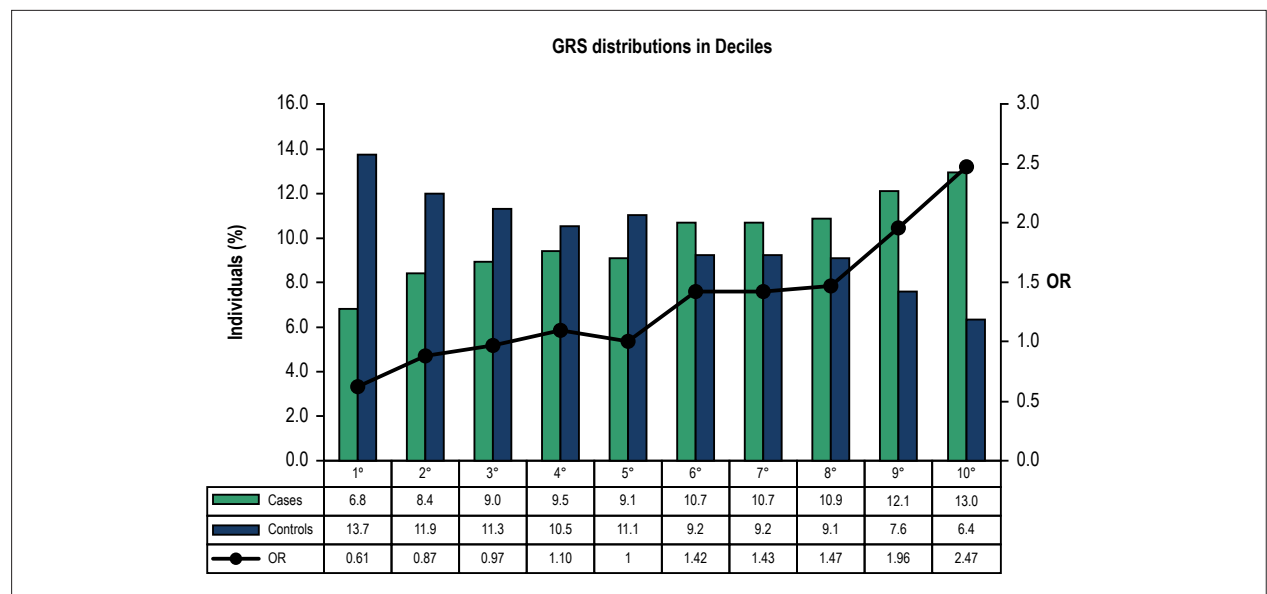


Figure 2 – Distribution of genetic risk score in deciles by cases and controls. A logistic regression model was used with the 5th decile of the controls as the reference class.

NRI was also computed using categorical variables and applied to this case-control study and was defined as the percentage of subjects changing categories in each subgroup when adding the new marker (CAD quartile score). Movement towards a better category (higher in patients than in controls) was calculated to address a potential impact for clinical use. NRI showed higher improvement capacity in reclassifying 19.5% of patients from the 50-75% category to the highest risk (75-100%) category. Likewise, 14.1% of healthy controls were moved down into a lower risk category, from 25-50% risk category to < 25% one (Table 5).

Furthermore, the inclusion of GRS quartiles to TRF also provided an IDI of 2.5% (95%CI: 1.9-3.1%; p < 0.0001) (data not shown).

Discussion

Several years ago, polymorphisms involved in specific biological pathways, relevant to coronary atherosclerosis, were genotyped to determine their association with CAD. This candidate gene approach revealed about 30 high-confidence SNPs loci with significant effects on atherosclerosis.¹³ However, following traditional candidate gene approach has generated many conflicting results or with weak associations; replication studies are necessary for consistent validation of these results.

In 2004, Mendonça et al. first genotyped angiotensin-converting enzyme (ACE) I/D polymorphisms in a Portuguese population yielding similar reports as described in literature.¹⁴

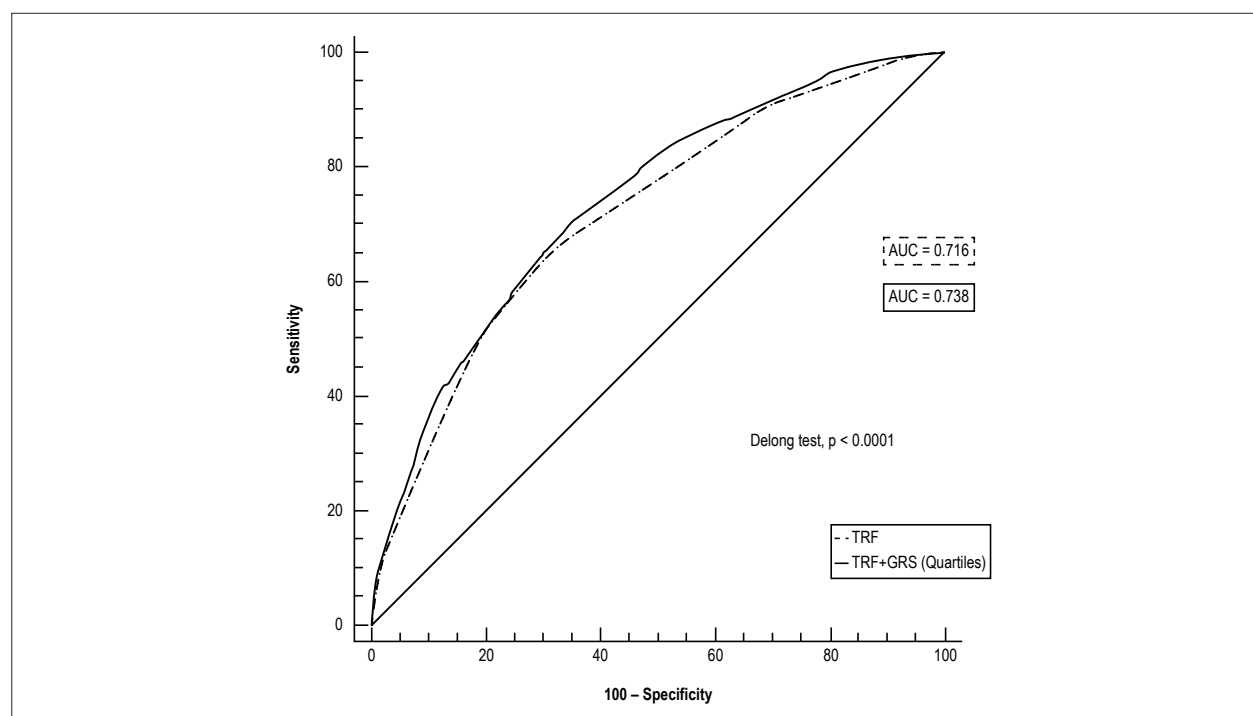


Figure 3 – ROC curves based on the baseline model (traditional risk factors, TRFs) and after adding the genetic risk score (GRS) (quartiles) in predicting the risk for coronary artery disease. The two curves are based on logistic regression models incorporating conventional risk factors (diabetes, dyslipidemia, smoking and hypertension) with and without the GRS. AUC indicates area under curve. The Delong test compares the difference between the two AUCs ($p < 0.0001$).

Table 4 – The category-free net reclassification index (cfNRI) after addition of the GRS quartiles

Group	n	Higher risk n (%)	Lower risk n (%)	p (cfNRI)	cfNRI (%)	cfNRI (95% CI)
CAD patients	1566	897 (57.3%)	669 (42.7%)	< 0.0001	14.6%	(9.7-19.5%)
Healthy controls	1322	553 (41.8%)	769 (58.2%)	< 0.0001	16.4%	(11.2-21.8%)
Total	2888	---	---	< 0.0001	31%	(23.8-38.3%)

GRS: genetic Risk Score; CAD: coronary Artery Disease; CI: confidence Interval; cfNRI: category-free net reclassification index. This analysis uses the function "improveProb" from R software package "Hmisc".

After the development of high capacity arrays in 2008,¹⁵ GWAS examined millions of polymorphisms simultaneously in several ethnical subpopulations with a case-control design. The standardized minimum significance level set at 1×10^{-5} added reliability to cardiovascular genetics and put it into perspective.¹⁶

In 2007, Samani et al.¹⁷ first identified chromosomal loci that were strongly associated with CAD in the Wellcome Trust Case Control Consortium (WTCCC) study (which involved 1,926 case subjects with CAD and 2938 controls) and looked for replication in the German MI (Myocardial Infarction) Family Study.¹⁷

In the following years, a surprisingly large number of gene variants were consistently reported to be associated with CAD. The 9p21 variant was the most frequently gene variant reported across populations. The huge consortium of Wellcome Trust and three other European research groups joined for the CARDIOGRAM project that confirmed, in a

very large sample (> 22,000 cases) of individuals of European ancestry, a 29% increase in risk for MI per copy of the rs1333049 9p21 variant ($p = 2 \times 10^{-20}$).¹⁸

Our research group replicated this 9p21 variant analysis in the Portuguese population and found a CC genotype prevalence of 35.7% in CAD patients, with an adjusted OR of 1.34, $p = 0.010$. The adjusted OR for TRF of CC genotype was 1.7 ($p = 0.018$) and CG genotype of OR = 1.5, $p = 0.048$. The authors concluded that although the mechanism underlying the risk is still unknown, the robustness of this risk allele in risk stratification for CAD has been consistent, even in very different populations. The presence of the CC or CG genotype may thus prove to be useful for predicting the risk of developing CAD in the Portuguese population.¹⁹

The most recent meta-analysis of GWAS for (CAD) identified 46 genome-wide loci with significant association and 104 genome-wide loci potentially associated with increased risk.^{20,21}

Table 5 – Reclassification table comparing predicted coronary artery disease (CAD) risk with and without genetic risk score (GRS) quartiles

Predicted risk (without GRS)	Reclassified predicted risk (with GRS)				% Increase	% / Decrease
	< 25%	25-50%	50-75%	75-100%		
CAD patients (n = 1,566)						
< 25%	6	11	0	0	0,7%	0%
25-50%	44	335	123	0	7,9%	2.8%
50-75%	0	59	471	305	19,5%	3.8%
75-100%	0	0	9	203	0%	0.6%
NRI CAD patients			20.9%			
Healthy controls (n = 1,322)						
< 25%	65	36	0	0	2,7%	0%
25-50%	186	504	88	0	6,7%	14.1%
50-75%	0	60	268	79	6%	4.5%
75-100%	0	0	1	35	0%	0.1%
NRI controls			3.3%			
NRI total			24.2%			

NRI: net reclassification improvement (categorical NRI); CAD: coronary artery disease.

In our study, we found a gradual and continual increase in CAD risk with increasing number of CAD risk alleles carried. Individuals in the bottom decile are naturally protected and subjects in top decile of the GRS had a CAD risk of 2.472 (1.755 – 3.482). Even though the score distribution overlaps between cases and controls, the GRS is significantly associated with CAD risk and can be used to identify subjects at highest risk in terms of lifestyle or therapeutic interventions.

Our results are similar to others reports in Caucasians populations where GRS with 13, 29 or 109 SNPs²²⁻²⁴ were independent and marginally increased the predictive power of TRF conferred either by AUC increases, C-index changes or more modern discriminative statistical methods like reclassification measures or improved discrimination.

We report a higher OR for the 4th quartile of GRS (2.59) compared to 1.66 reported by Ripatti et al. in the highest quintile.²² When comparing the relative weight of the GRS in the multivariate logistic analysis we found slightly lower OR than smoking, hypertension, and dyslipidemia. In Ripatti's²² cohort, a weighted GRS was also an independent predictor even after adjusting for age, sex and TRFs in a Northern European population-based trial. The relative risk of the GRS based on 13-SNP was also lower than that of dyslipidemia and comparable to the effects of hypertension.²²

An increased power to TRFs definition has been given in this study. For instance, we have used a broad dyslipidemia term including Apo B levels as indicated by 2016 lipid guidelines.⁷ Moreover, we have not considered ex-smokers until 5 years of cessation to account for the risk for CV disease events decrease be comparable to a nonsmoker.⁵

Thanassoulis et al.²⁴ demonstrated that adding to a 13 SNP-based GRS, 89 SNPs associated with modifiable risk factors did not increase the power of the GRS reporting a HR of 1.01 (95% CI 0.99 – 1.03; $p = 0.48$). This revealed that the weak association of polymorphisms with CAD risk

factors in GRS analysis could be masked by the relative stronger effect of other polymorphisms. Considering the lack of a significant association of lipid profiles with CAD risk, Jansen et al. reported in 2015 that several SNPs associated with type 2 diabetes mellitus were related with CAD risk.²⁵ Recently, Webb et al. identified 6 new loci associated with CAD at genome-wide significance. The study confirmed a pleiotropy between lipid traits, blood pressure phenotypes, body mass index, diabetes, and smoking behavior.²⁶ Our GRS is an assembly of risk factors and non-risk factors-related SNP, reinforcing the genotype-phenotype interactions.

Limitations of this study

The main clinical utility of the GRS in our population is a modest improvement in risk stratification. GRS seems to be a better indicator of patients at a higher than average risk for DAC as compared with TRF stratification. The number and type of SNPs included is limited in our study and a larger number of GWAS hit SNPs should be included in further studies. Nevertheless, the increasing capability of analyzing multiple SNPs in GRS so far have not been translated into increasing ability of risk prediction.

Finally, this study did not include a gene-gene (G-G) and gene-environment (G-E) analysis. It is expected that, as better statistical significance arises from those interplays, the G-G and G-E incorporation in GRS plus TRF will increase our ability to accurately and individually predict risk.

Conclusions

We conclude that a multilocus GRS based on multiple variants of genetic risk was associated to an increased cardiovascular risk in a Portuguese population. We found that a GRS calculated with the 31 studied SNPs was significantly associated to CAD and that 25% of individuals who carry the greatest risk alleles have, approximately,

2.5 times increased CAD risk when compared to those in the lowest quartile. This GRS has provided a slight improvement of the predictive ability compared to the initial model and can enhance individual risk stratification. These results highlight the potential value of including genetic information in the usual models.

Acknowledgments

We are very grateful to Elsa Sousa who made all the administrative procedures and to Rita Freitas (CIDEHUS, Évora University, Portugal) who reviewed the paper for statistical analysis.

Author contributions

Conception and design of the research and Writing of the manuscript: Pereira A; Acquisition of data: Pereira A, Freitas AI, Sousa AC, Brehm A; Analysis and interpretation of the data and Statistical analysis: Pereira A, Freitas S, Henriques E, Rodrigues M; Critical revision of the manuscript for intellectual content: Pereira A, Mendonça MI, Borges S, Brehm A, Reis RP.

References

1. Lawrence C, Boerwinkle E. Gene-environment interactions and gene therapy in atherosclerosis. *Cardiol Rev*. 1994;2:130-7.
2. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447(7145):661-78.
3. Myocardial infarction redefined. A consensus document of the Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *Eur Heart J*. 2000;21(18):1502-13.
4. Abdi H, Williams LJ. Principal component analysis. *WIREs Computacional Statistics*. 2010;2:433-59.
5. Mons U, Müezziner A, Gellert C, Schöttker B, Abnet CC, Bobak M, et al; CHANCES Consortium. Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. *BMJ*. 2015;350:h1551.
6. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL Jr, et al; National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure; National High Blood Pressure Education Program Coordinating Committee. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure: the JNC 7 report. *JAMA*. 2003;289(19):2560-72. Erratum in: *JAMA*. 2003;290(2):197.
7. Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults. Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA*. 2001;285(19):2486-97.
8. Genuth S, Alberti KG, Bennett P, Buse J, DeFronzo R, Kahn R, et al; Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care*. 2003;26(11):3160-7.
9. National Institute of Health, National Heart, Lung, and Blood Institute North American Association for the Study of Obesity. The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. NHLBI Obesity Education Initiative. Bethesda (MD); 2000. NIH Publication Number: 00-4084.
10. Asmar R, Benetos A, Topouchian J, Laurent P, Pannier B, Brisac AM, et al. Assessment of arterial distensibility by automatic pulse wave velocity measurement: validation and clinical application studies. *Hypertension*. 1995;26(3):485-90.
11. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-45.
12. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30(1):11-21.
13. Humphries SE, Morgan L. Genetic risk factors for stroke and carotid atherosclerosis: insights into pathophysiology from candidate gene approaches. *Lancet Neurol*. 2004;3(4):227-35.
14. Mendonça I, Freitas AI, Sousa AC, Gomes S, Faria P, Drumond A, et al. Polymorphism of the ACE gene is associated with extent and severity of coronary disease. *Rev Port Cardiol* 2004;23(12):1605-11.
15. Keating BJ, Tischfield S, Murray SS, Bhargale T, Price TS, Glessner JT, et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS One*. 2008;3(10):e3583.
16. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9(5):356-69.
17. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al; WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007;357(5):443-53.
18. Preuss M, König IR, Thompson JR, Erdman J, Absher D, Assimes TL, et al; CARDIoGRAM Consortium. Design of the Coronary ARtery Disease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: A Genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ Cardiovasc Genet*. 2010;3(5):475-83.
19. Mendonça I, dos Reis RP, Pereira A, Café H, Serrão M, Sousa AC, et al. Associação independente da variante rs1333049, no locus 9p21, com a doença coronária, numa população portuguesa. *Rev Port Cardiol*. 2011;30(6):575-91.

Potential Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Sources of Funding

There were no external funding sources for this study.

Study Association

This article is part of the thesis of Doctoral submitted by Andreia Pereira, from Universidade Nova de Lisboa.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the SESARAM, EPE under the protocol number 50/2012. All the procedures in this study were in accordance with the 1975 Helsinki Declaration, updated in 2013. Informed consent was obtained from all participants included in the study.

20. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, et al; CARDIoGRAMplusC4D Consortium. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet.* 2013;45(1):25-33.
21. Roberts R. Genetics of coronary artery disease. *Circ Res.* 2014;114(12):1890-903.
22. Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet.* 2010;376(9750):1393-400.
23. Vaarhorst AA, Lu Y, Heijmans BT, Dollé ME, Böhringer S, Putter H, et al. Literature-based genetic risk scores for coronary heart disease: the Cardiovascular Registry Maastricht (CAREMA) prospective cohort study. *Circ Cardiovasc Genet.* 2012;5(2):202-9.
24. Thanassoulis G, Peloso GM, Pencina MJ, Hoffman U, Fox CS, Cupples LA. A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham Heart Study. *Circ Cardiovasc Genet.* 2012;5(1):113-21.
25. Jansen H, Loley C, Lieb W, Pencina MJ, Nelson CP, Kathiresan S, et al; CARDIoGRAM consortium. Genetic variants primarily associated with type 2 diabetes are related to coronary artery disease risk. *Atherosclerosis.* 2015;241(2):419-26.
26. Webb TR, Erdmann J, Stirrups KE, Stitzel NO, Masca NG, Jansen H, et al. Systematic evaluation of pleiotropy identifies 6 further loci associated with coronary artery disease. *J Am Coll Cardiol.* 2017;69(7):823-36.



Supplementary Table 1 – List of the 33 genetic variants previously associated with coronary artery disease risk, used for the development of the genetic risk score in the study population

SNP ID	Nearest gene	Chr	Position	Genotypic OR (95%CI)	p value	Allelic OR (95%CI)	p value	MAF (%)	Potential Mechanism of Action
rs1333049	9p21.3	9	22125504	1.147 (1.036-1.270)*	0.008	1.155 (1.041-1.282)	0.007	45.8	Cellular
rs4977574	CDKN2B	9	22098575	1.161 (1.049-1.286)*	0.004	1.172 (1.056-1.302)	0.003	42.0	Cellular
rs618675	GJA4	1	34922761	1.143 (0.792-1.649)*	0.475	1.046 (0.918-1.191)	0.502	19.6	Cellular
rs17228212	SMAD3	15	65245693	1.202 (0.888-1.629)*	0.234	1.025 (0.910-1.155)	0.684	25.3	Cellular
rs17465637	MIA3	1	222650187	1.088 (0.971-1.220)*	0.148	1.088 (0.971-1.220)	0.147	28.6	Cellular
rs12190287	TCF21	6	134256218	1.230 (1.100-1.375)*	< 0.0001	1.226 (1.098-1.368)	0.0003	32.7	Cellular
rs3825807	ADAMTS7	15	76876166	1.073- (0.967-1.191)*	0.185	1.074 (0.967-1.194)	0.181	41.2	Cellular
rs11556924	ZC3HC1	7	130023656	1.227 (1.058-1.423)*	0.007	1.157 (1.037-1.290)	0.009	34.3	Cellular
rs1332844	PHACTR1	6	12927312	1.113 (1.003-1.235)*	0.044	1.113 (1.003-1.236)	0.043	44.3	Cellular
rs2114580	PCSK9	1	55167236	1.079 (0.821-1.417)*	0.587	0.974 (0.866-1.096)	0.665	26.3	Lipids
rs3798220	LPA	6	160540105	1.484 (1.212-1.816)*	< 0.0001	2.167 (1.452-3.235)	< 0.0001	2.1	Lipids
rs20455	KIF6	6	39357302	1.129 (0.896-1.424)*	0.306	1.060 (0.949-1.184)	0.302	32.8	Lipids
rs7412/ rs429358 ¹	APOE ¹	19	44908822/ 44908684	1.261 (1.062-1.497) [#]	0.008	1.231 (1.056-1.435)	0.008	13.4	Lipids
rs964184	ZNF259	11	116778201	1.131 (0.986-1.298)*	0.078	1.130 (0.986-1.295)	0.079	17.7	Lipids
rs599839	PSRC1	1	109279544	1.059 (0.933-1.203)*	0.375	1.058 (0.933-1.201)	0.379	21.4	Lipids
rs1801133	MTHFR 677	1	11796321	1.178 (1.017-1.365) [#]	0.029	1.114 (0.998-1.243)	0.055	33.5	Oxidation
rs1801131	MTHFR 1298	1	11794419	0.944 (0.816-1.093) [#]	0.443	0.958 (0.854-1.075)	0.465	28.0	Oxidation
rs705379	PON -108	7	96324583	1.135 (0.950-1.355) [#]	0.163	1.068 (0.962-1.184)	0.217	46.4	Oxidation
rs662	PON 192	7	95308134	0.836 (0.652-1.072)*	0.157	0.927 (0.828-1.037)	0.186	30.1	Oxidation
rs854560	PON 55	7	95316772	1.161 (1.044-1.290)*	0.006	1.161 (1.044-1.290)	0.006	40.4	Oxidation
rs6922269	MTHFD1L	6	150931849	1.067 (0.804-1.416)*	0.653	0.996 (0.887-1.118)	0.943	27.3	Oxidation
rs5186	AT1R	3	148742201	1.245 (0.906-1.710)*	0.177	1.062 (0.942-1.198)	0.323	24.7	RAS
rs699	AGT	1	230710048	0.932 (0.798-1.090) [#]	0.380	0.969 (0.873-1.076)	0.552	42.9	RAS
rs4340	ACE	17	61565892	1.165 (1.001-1.355)*	0.048	1.083 (0.973-1.205)	0.143	38.1	RAS
rs4402960	IGF2BP2	3	185793899	1.124 (0.876-1.443)*	0.358	1.020 (0.911-1.141)	0.736	30.8	Diab/Obes
rs1326634	SLC30A8	8	117172544	1.213 (0.914-1.609) [#]	0.181	1.081 (0.961-1.217)	0.195	25.8	Diab/Obes
rs266729	ADIPOQ	3	186841685	1.209 (1.041-1.403) [#]	0.013	1.165 (1.030-1.318)	0.015	23.3	Diab/Obes
rs7903146	TCF7L2	10	112998590	0.961 (0.862-1.072)*	0.480	0.962 (0.863-1.072)	0.482	35.3	Diab/Obes
rs17782313	MC4R	18	60183864	1.314 (0.931-1.855)*	0.120	1.016 (0.896-1.152)	0.806	21.6	Diab/Obes
rs1801282	PPARG	3	12351626	1.427 (0.717-2.843) [#]	0.309	1.164 (0.970-1.396)	0.102	8.8	Diab/Obes
rs1884613	HNF4A	20	44351775	1.159 (0.987-1.360) [#]	0.072	1.106 (0.960-1.273)	0.163	16.2	Diab/Obes
rs8050136	FTO	16	53782363	1.194 (1.026-1.390) [#]	0.022	1.129 (1.016-1.255)	0.025	39.7	Diab/Obes
rs1376251	TAS2R 50	12	11030119	1.556 (0.767-3.155)*	0.217	1.080 (0.920-1.267)	0.349	11.9	Diab/Obes

SNP: Single Nucleotide Polymorphism; Chr: Chromosome; OR: Odds Ratio; CI: Confidence Interval; MAF: Minor Allele Frequency; RAS: Renin-Angiotensin System; Diab/Obes: Diabetes/Obesity; *Additive model; #Recessive model; #Dominant model; ¹Resulting from a Haplotype; Table shows susceptibility loci for coronary artery disease, genotypic and allelic ORs and p values for the lead SNP within each locus reported in genome-wide association studies and candidate gene studies. Genotypic ORs are given for additive, recessive and dominant models. Potential mechanism of action is on the basis of what is already known about the function of the nearby genes. It includes "Cellular" (genes associated to cell cycle, cellular migration and inflammation); "Oxidation" (genes involved in pro-oxidative status) and associated with modifiable risk factors such as "Lipids" metabolism, hypertension ("RAS") and Diabetes/Obesity.

Supplementary Table 2 – Logistic regression with respective ORs and ROC curves with respective AUCs of the GRS models

GRS models	OR (95% CI)	P value ¹	AUC (95% CI)	Sensitivity (%)	Specificity (%)	P value ²
Multiplicative	1.78 (1.52 – 2.10)	< 0.0001	0.61 (0.59 – 0.62)	54.0	62.3	< 0.0001
Additive	1.06 (1.04 – 1.09)	< 0.0001	0.56 (0.54 – 0.58)	58.7	50.5	< 0.0001
Weighted (Best model OR)	1.02 (0.94 – 1.10)	0.660	0.57 (0.55 – 0.59)	41.0	70.3	< 0.0001
Weighted (Beta)	2.23 (1.88 – 2.65)	< 0.0001	0.60 (0.58 – 0.61)	43.0	71.5	< 0.0001
Weighted (Literature OR)	1.35 (1.12 – 1.62)	0.001	0.54 (0.52 – 0.55)	53.4	54.1	0.008
Classic weighted	3.01 (2.32 – 3.89)	< 0.0001	0.59 (0.57 – 0.61)	59.4	54.4	< 0.0001

OR: Odds ratio; ROC: Receiver Operating Characteristic; AUC: Area under curve; GRS: Genetic Risk Score; CI: Confidence interval; P value¹: Obtained by logistic regression to evaluate the significance of the odds ratio; P value²: Obtained by the ROC Curve to verify the significance of the area under the curve; Statistically significant for $p < 0.05$.

29-Jan-2018

2018

Dear Dr. Pereira:

Your manuscript entitled "**ADDITIONAL VALUE OF A COMBINED GENETIC RISK SCORE TO STANDARD CARDIOVASCULAR STRATIFICATION**" submitted for publication in Genetics and Molecular Biology has been evaluated by the assigned Associate Editor who concluded that you have adequately addressed the concerns raised by the reviewers.

Therefore, I am pleased to inform you that your manuscript is now accepted for publication.

As a next step, your manuscript will be checked by one of our copy editors for correction of language errors and technical inconsistencies.

You will receive an edited manuscript by e-mail for appreciation and approval of eventual corrections before starting the layout process.

From now on, all correspondence will be directly with the Editorial Office by e-mail: editor@gmb.org.br. Please remember to refer to your Manuscript ID number whenever you contact us.

You will be contacted by gestao.editorial@scielo.org in order for checking the final PDF proofs.

NOTE: If you are a corresponding author who is a Member-in-good-standing of the Sociedade Brasileira de Genética (before the submission date), you are exempt from publication charges. Authors who were granted a waiver under our policies are exempt as well.

On behalf of the Editors of Genetics and Molecular Biology, I would like to take the opportunity to thank you for submitting your interesting contribution to our journal.

Sincerely,
Dr. Carlos Menck
Editor in Chief, Genetics and Molecular Biology
editor@gmb.org.br

**ADDITIONAL VALUE OF A COMBINED GENETIC RISK SCORE TO
STANDARD CARDIOVASCULAR STRATIFICATION**

Andreia Pereira^{1,2}, Maria Isabel Mendonca¹, Ana Célia Sousa^{1,2}, Sofia Borges¹, Sónia Freitas¹, Eva Henriques¹, Mariana Rodrigues¹, Ana Isabel Freitas^{1,3}, Graça Guerra^{1,3}, Carolina Freitas¹, Décio Pereira¹, António Brehm³, Roberto Palma Dos Reis²

*corresponding author

Andreia Micaela Pereira, PhD

Email: andreiapereira21@gmail.com

Phone: +351-291-709600

Fax: +351-291-709677

¹Unidade de Investigação, Hospital Dr. Nélio Mendonça

Avenida Luís de Camões, nº 57

9004-514 Funchal-Portugal

²Faculdade de Ciências Médicas, Universidade Nova de Lisboa

Campo dos Mártires da Pátria, n.º 130

1169-056 Lisboa-Portugal

³Laboratório de Genética Humana, Universidade da Madeira

Campus da Penteadá

9020-105 Funchal-Portugal

Running title: GRS and CAD risk prediction

Keywords: Genetic risk score; Framingham risk score; Traditional risk factors; Coronary Artery Disease

Abstract

The utility of genetic risk scores (GRS) as independent risk predictors remains inconclusive. Here, we evaluate the additive value of a multi-locus GRS to the Framingham risk score (FRS) in coronary artery disease (CAD) risk prediction. A total of 2888 individuals (1566 coronary patients and 1322 controls) were divided in three subgroups according to FRS. Multiplicative GRS was determined for 32 genetic variants associated to CAD. Logistic Regression and Area Under the Curve (AUC) were determined first using the TRF for each FRS subgroup and secondly adding GRS. Different models (TRF, TRF+GRS) were used to classify the subjects into risk categories for the FRS 10-year predicted risk. The improvement offered by GRS was expressed as Net Reclassification Index and Integrated Discrimination Improvement.

Multivariate analysis showed that GRS was an independent predictor for CAD (OR=1.87;p<0.0001). Diabetes, arterial hypertension, dyslipidemia and smoking status

were also independent CAD predictors ($p < 0.05$). GRS added predictive value to TRF across all risk subgroups. NRI showed a significant improvement in all categories with In conclusion, GRS provided a better incremental value in intermediate and high-risk subgroups. In these subgroups, inclusion of genotyping may be considered to better stratify cardiovascular risk.

Keywords: Coronary Artery Disease; Genes; Genetic Risk Score; Framingham Score; Risk prediction; Risk factors

Introduction

The most important Traditional Risk factors (TRF) for Coronary Artery Disease (CAD) include dyslipidemia, arterial hypertension, diabetes, obesity, smoking, lack of physical activity and stress.⁴ However; some patients can develop vascular disease without conventional risk factors.

Although the familial nature of CAD has been documented for many years^{15,16,20} and the addition of family history has been shown to improve risk prediction,^{27,28} the genetic variants responsible for the increased familial risk were, until recently, unknown. Genome-Wide Association Studies (GWAS) have uncovered several common genetic variants (single nucleotide polymorphisms, or SNPs) that are robustly associated with CAD and have been replicated in multiple independent samples.³⁴ The identification of these genetic variants provides an opportunity to evaluate whether addition of a genetic risk score (GRS) to risk models may improve predictive power.

There is an increasing interest in the potential use of GRS in cardiovascular disease, because this could enlarge the preventive and therapeutic interventions in individuals and groups with high genetic risk that are not obvious candidates to those interventions using the current standard stratification. Usual cardiovascular risk stratification uses family history, TRF evaluation and is quantified into scores like Framingham risk score and EuroSCORE.²³ However, due to the potential financial and medical costs associated with measuring these new markers like GRS, their ability in improving the prediction of CAD outcomes over existing risk models needs to be rigorously assessed. Effective statistical tools for evaluating the incremental value of the novel markers over the routine clinical risk factors are crucial in the field of outcome prediction.

Here, we evaluate whether a GRS, based on 33 SNPs associated to CAD, is independent of the vascular risk explainable by conventional risk factors and can improve the predictive capacity of TRF in the assessment of CAD risk, according to FRS subgroups (low risk <10%, Intermediate risk 10-20% or high risk >20%). Using the Net Reclassification Index (NRI) and Integrated Discrimination Index (IDI), we investigate the performance of combined stratification including TRF and GRS in CAD risk assessment.

Methods

Study population

A case-control study was performed with 2888 individuals (mean age of 53.0 ± 7.9 years), including 1566 consecutive coronary patients and 1322 controls. These were divided into three subgroups according to FRS: low risk (FRS<10, n=1312, 68.7% male), intermediate risk ($10 \leq \text{FRS} \leq 20$, n=1049, 83.1% male) and high risk (FRS>20, n=527, 90.1% male). Cases and controls were matched for age and gender.

Determination of CAD was adjudicated by a dedicated Intervention Cardiologist. Angiographically proven CAD was considered significant if ≥ 1 coronary lesions of $\geq 70\%$ stenosis in ≥ 1 major coronary artery or its primary branches. Absent or non-flow limiting CAD was excluded from this study. The control group consisted of healthy volunteers selected from the same population with no symptoms or history of CAD. All controls underwent clinical and phenotype assessment of TRF, an electrocardiogram, and, if needed, complementary exercises stress test or a Stress Echocardiography.

Population stratification analysis was performed in our population set to account for possible genetic admixture and no significant genetic outliers (<5%) were identified with Principal Component Analysis (PCA).¹

The study was conducted according to the Declaration of Helsinki, the protocol was reviewed and approved by the Hospital ethics committee, and all patients provided written informed consent.

Data collection

Data was collected from all subjects in a standardized file comprising demographic, clinical characteristics and TRF (gender, age, level of exercise, smoking status, arterial hypertension, dyslipidemia, diabetes, and family history of CAD, body mass index (BMI), heart rate and pulse wave velocity (PWV).

'Smoking status' refers to current smokers or subjects with less than 5 years of smoking cessation.¹⁸

Arterial Hypertension was considered when patients, at the entry into this study, were already diagnosed and/or had been on antihypertensive medication for more than 3 months or newly diagnosed hypertensives with systolic blood pressure (SBP)/diastolic blood pressure (DBP) $\geq 140/90$ mmHg measured on at least 3 occasions.⁵

Subjects with a LDL > 100 mg/dL, HDL <40 mg/dL for men and <45 mg/dL for women, non HDL (Total cholesterol – HDL) > 130 mg/dL, triglycerides > 150 mg/dL or Apo B > 100 mg/dL were classified as having dyslipidemia.⁹

Subjects were classified as having diabetes if taking oral anti-diabetic medication or insulin or if fasting plasma glucose was higher than 7.0 mmol/l or 126 mg/dl.²⁶

Family history of CAD was considered if a female relative (mother or sister) had presumed CAD disease before 65 years or male relative (father or brother) had CAD disease before 55 years old.

The definition of other traditional risk factors was based on standard criteria, as previously reported^{2,21}.

Biochemical analysis

Blood samples were extracted after 14-16 hours' fasting. Biochemical analyses were performed at the Central Laboratory of the Hospital, according to standard techniques. In order to determine total cholesterol, HDL, LDL, triglycerides and glucose, blood samples were placed in dry tubes, centrifuged half an hour later at 3500 g and subsequently quantified by an enzymatic technique using a "AU 5400" (Beckman Coulter) auto analyzer. Biochemical markers such as lipoprotein-a (Lp(a)), apolipoprotein B (Apo B), and high sensitivity C-Reactive Protein (hs-CRP) were quantified by Immunospectrometry using a "AU 5400" (Beckman Coulter) automatic system.

SNP selection

SNPs were selected either from GWAS or candidate gene association studies^{6,30}. Entering criteria included genes described in previous studies with an Odds Ratio (OR) for CAD ≥ 1.1 and, simultaneously, with a Minor Allele Frequency (MAF) $> 5\%$. Genes with low Hardy-Weinberg equilibrium $p < 0.002$ (after Bonferroni correction) were excluded.

According to their possible CAD-related function, we have included 32 genes associated to cell cycle, cellular migration and inflammation (rs1333049 (9p21.3), rs4977574 (CDKN2B), rs618675 (GJA4), rs17228212 (SMAD3), rs17465637 (MIA3), rs12190287 (TCF21), rs3825807 (ADAMTS7), rs11556924 (ZC3HC1), rs12526453 (PHACTR1)); genes involved in pro-oxidative status (rs1801133 (MTHFR 677), rs1801131 (MTHFR 1298), rs705379 (PON 1), rs662 (PON192), rs854560 (PON 55), rs6922269 (MTHFD1L); genes associated with modifiable risk factors such as lipids metabolism, hypertension and diabetes/obesity (rs2114580 (PCSK9), rs20455 (KIF6), rs7412/rs429358 (APOE), rs964184 (ZNF259), rs599839 (PSRC1), rs5186 (AT1R), rs699 (AGT), rs4340 (ACE), rs4402960 (IGF2BP2), rs1326634 (SLC30A8), rs266729 (ADIPOQ), rs7903146 (TCF7L2), rs17782313 (MC4R), rs1801282 (PPARG), rs1884613 (HNF4A), rs8050136 (FTO) and rs1376251 (TAS2R 50).

Details of SNPs used in GRS model are presented in Supplementary Table S1.

Genotype analysis

Genetic analysis was performed at the Human Genetics Lab of the University of Madeira. Genomic DNA was extracted from 80 µl of peripheral blood using a standard phenol-chloroform method. A TaqMan allelic discrimination assay for genotyping was performed using labelled probes and primers pre-established by the supplier (TaqMan SNP Genotyping Assays, Applied Biosystems).

All reactions were done on an Applied Biosystems 7300 Real Time PCR System and genotypes were determined using the 7300 System SDS Software (Applied Biosystems, Foster City, USA) without any prior knowledge of the individual's clinical data. Quality

check of genotyping techniques was maintained by the inclusion of one non-template control (NTC) in each plate of 96 wells. All SNPs TaqMan assays had blind duplicates accounting for 20% of all samples. Some SNP genotypes were randomly confirmed by conventional direct DNA sequencing, as 10-15% of all samples were re-amplified for sequencing.

Statistical analysis

Deviation from Hardy-Weinberg equilibrium for the 33 genotypes at individual locus were assessed using the Chi-square test and $p < 0.002$ with Bonferroni correction for all SNPs included. LPA gene variant was excluded for further analyses due to its low Hardy-Weinberg equilibrium ($p < 0.002$).

Comparisons of baseline characteristics and biochemical data of cases and controls were analyzed by Chi-square for categorical variables and Student's t-test or Mann-Whitney were performed for continuous variables, as appropriate. Genotypic frequencies were determined from observed counts and compared by Chi-square analysis.

To evaluate the impact of genotype frequencies and Odds Ratios (OR) on the overall discriminative accuracy of genetic risk models, we assessed the AUC. SNPs associated with p values less than or equal to 0.05 entered the race-specific GRS with a coding value of 2 for the mutated genotype (risk), 1 for the heterozygote, and 0 for the homozygous wild-type genotype. The GRS was constructed under a multiplicative model (multiplying the odds ratio of each genotype for all the 32 genetic variants). The maximum number of risk alleles possessed by one individual was 36 and the minimum was 13

(mean±standard deviation (SD)=23.96±3.63). The increasing loci were assumed to have an additive effect, and the GRS was normally distributed.

For each risk subgroup, the GRS was included in a ROC model containing the TRF. A new combined risk score was calculated for each individual (TRF + GRS) by multiplying each model coefficient by the associated risk variable and then summing the products.

The AUC equals the probability that classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one.

ROC curve analysis was performed using the MedCalc software. Non-parametric methods developed by DeLong et al.⁷ were used to test for significant differences between ROC curves.

NRI was computed according to the continuous method and applied to the case-control studies.²⁴ NRI was defined as the percentage of subjects changing categories in each subgroup when adding the new markers and IDI as the improvement of the difference in average of predicted probabilities between cases and controls. Both measures were obtained with the PredictABEL package available in R software (version 3.2.0). Discriminative capacity (AUC) was accessed for evaluation of specificity and sensitivity with MedCalc version 13.3.3.0 and other statistics were analyzed by the Statistical Package for the Social Sciences Software version 19.0 (IBM, Armonk, NY, USA). All p-values were two sided, statistically significant for $p < 0.05$.

Results

Table 1 compares baseline characteristics between CAD and control groups. According to the selection criteria, age and gender did not differ significantly between the two groups. In contrast, CAD patients were more likely to have a higher prevalence of modifiable risk factors such as dyslipidemia, tobacco use, hypertension and diabetes ($p<0.05$). Also BMI, PWV, family history of CAD and lower level of exercise were significantly more frequent among cases ($p<0.05$). Lower heart rates among patients potentially reflects a better heart rate control ($p<0.0001$). Across all Framingham score subgroups, the GRS was significantly lower in controls than CAD patients ($p<0.0001$).

A Logistic regression under a Forward Wald method was performed including GRS and TRF (Table 2). In this multivariable analysis, GRS was an independent predictor for CAD (OR=1.87; 95%CI: 1.58 – 2.21), with statistical significance ($p<0.0001$). Even after a Bonferroni Correction, GRS remained significant at a $p<0.01$ ($0.05/5$) level. Furthermore, smoking status (OR=3.44; 95%CI: 2.89 – 4.10); diabetes (OR=3.19; 95%CI: 2.61 – 3.91); arterial hypertension (OR=2.10; 95%CI: 1.77 – 2.48) and dyslipidemia (OR=1.30; 95% CI: 1.03 – 1.65) were independently and significantly associated to CAD risk ($p<0.05$) (Table 2).

The AUC for the TRF + GRS model was then compared with the basal AUC for the TRF model as shown in Table 3. TRF yielded an AUC of 0.72 (95% CI=0.70 – 0.73), predicting CAD moderately well. The model including the GRS was more discriminating than TRF model (AUC=0.74; 95% CI=0.72 – 0.76; Delong test $p<0.0001$).

ROC curves were performed for each subgroup, according to expected 10-year FRS stratification (Low Risk FRS subgroup (<10%), Intermediate FRS subgroup

($10\% \leq \text{FRS} \leq 20\%$), High Risk $>20\%$ FRS and the respective AUC's were compared after the GRS inclusion (Table 3).

For each subgroup, we found that including the GRS to the model with TRF only, the AUC increased significantly (Table 3), showing that GRS adds predictive value to TRF in all the 3 risk subgroups. Specifically, in individuals within low risk subgroup, the AUC was 0.72 for the TRF model and 0.75 for the TRF + GRS model ($p < 0.0001$). The increase in predictive accuracy in this subgroup was 2.5% (95% CI: 1.4 – 3.6%). This difference was small, but significant (DeLong's test for correlated ROC curves $p < 0.0001$). Similarly, in individuals within intermediate FRS subgroup, the AUC was 0.70 for the TRF model and 0.73 for the TRF + GRS model. The increase in the predictive accuracy for the AUC model in this subgroup was 2.6% (95% CI: 1.2 – 3.9%) at a significant p level ($p = 0.0002$). Finally, a surprising significant (DeLong's test for correlated ROC curves $p = 0.003$) addition to the model's predictive accuracy was seen in the higher risk of FRS group with an AUC increasing from 0.68 to 0.72. The increase in predictive accuracy for the AUC model (TRF + GRS model) in this group was 4.4% (95% CI: 1.5 – 7.3%). For all subgroups, the 95% CI of the increase did not include zero, suggesting that the increase in AUC caused by the GRS inclusion is statistically significant.

The addition of GRS to TRF improved the risk classification of the models. The new marker provided a continuous NRI of 32.3% (95%CI: 22.4-42.3%; $p < 0.0001$) in the low risk group; 30.4% (95%CI: 19.0-41.8%; $p < 0.0001$) in the intermediate FRS group and 29.8% (95%CI: 13.1-46.6%; $p = 0.0005$) in the high risk subgroup (Table 4). Furthermore, the inclusion of GRS to TRF also provided an IDI of 2.4% (95%CI: 1.6-3.2%; $p < 0.0001$) in

the low risk group; 2.1% (95%CI: 1.3-2.9%; $p < 0.0001$) in the intermediate risk group and 3.3% (95%CI: 1.9-4.7%; $p < 0.0001$) in the high risk group. Full reclassification results are provided in Table 4.

The value of Nagelkerke R square increased in each FRS subgroup with inclusion of GRS, that is, this new variable improved the proportion of variation explained in the model. Therefore, it can be stated that the disease is explained in 19.1% by the independent variables in the intermediate risk group. The values presented in Table 4 confirm that the selected variables present a significant explanatory degree about the dependent variable.

Discussion

Genetic risk stratification could be of interest in some clinical settings: at a young age when interventions like statins are more effective if started early and continued over a long-time period. In the intermediate's Framingham or Heart-Score calculators risk groups, where classical and standard risk stratification seems to be open to sampling errors.

This case-control study assesses the contribution of genetic risk for CAD in a Southern European's population controlled for age and sex.

We have downgraded our expectations after an unfulfilled enthusiastic time, of expecting genes to account for the hidden heritability in CAD, not meant for generalized use in the actual guidelines on cardiovascular prevention.

1-5% of allelic frequencies and ORs for CAD/MI at GWAS level from 1.1- 1.5 in literature will produce moderate risk improvement if added to Standard risk stratification models like Framingham.^{19,32}

These models have reasonable to good risk prediction curves as modeled with C-statistics ranging from 68-72%.^{10,22} Nevertheless improved accuracy is expected when gene-gene inter-mapping and some epigenetic factors impact are accounted on CAD's risk. Additional stratification strategies for CAD, either biomarkers, imaging modalities or genetic profile, will have to be calculated individually or for the intermediate risk group.

In our study, in a multivariate model, GRS presents a cardiovascular risk independent of the risk conferred by TRF. Thus it can improve the predictive capacity of TRF scores. According to the results presented, standard risk Framingham evaluation states good AUC levels across all risk subgroups with an AUC ranging from 0.676 to 0.765. Apart from the "maths", the intermediate risk subgroup represents the vast majority of patients followed at the outpatient's clinics. The low aged patient with one or two risk factors and low calculated cardiovascular risk is frequently left out of intensive prevention, as statin introduction, active measures for tobacco eviction or inclusion in exercise program.

Here, we report that a GRS based on 32 CAD risk alleles, identified previously in GWAS databases, independently and significantly predicts CAD with an increased risk of 80% (1.8 times). Our observations line up with several reports already addressing the impact of adding GRS to standard risk stratification. Recently, Iribarren et al.¹¹ have reported, in a large scale cohort living in USA, improved predictive capacity and discrimination

index for four GRS starting from just 8 or 12 and up to 36 to 51 genetic variants. Likewise Tada et al.³¹ also reported that two GRS of 27 and 50 SNPs, in a 23595 participants followed-up of 14 years, improved all measures of discrimination and reclassification.

Despite the modest predictive ability, our genetic risk model seems useful in preventive health care and disease prevention to correctly identify individuals at intermediate and high risk groups. We observed a general tendency for both measures of reclassification improvement, the NRI and IDI, to increase after addition of the GRS to the basic risk function. The net reclassification improvement was 30.8% for the entire cohort (95% CI: 24.0%-37.5%). However, reclassification improvement was more noticeable in the intermediate risk group, where it was statistically significant (NRI: 33.7%, 95%CI 25.6%-41.8%; IDI: 1.8%, 95%CI 1.2%-2.3%). It is important to note that the constructed reclassification tables follow clinically relevant strata of 5, and 20 percent 10-year risk, in order to have direct clinical application. But limitations in the interpretation of data apply because the prevalence of CAD in any patient study population is much higher than in a normal population, and thus, frequencies and risks are highly overestimated. The feasibility and clinical application of the GRS approach strategy will depend not only on the predictive capacity of the risk model, but also on the threshold level that is chosen.

We believe that genetic information will improve the preventive strategies probably targeted to individuals at very high genetic risk even though this may be a small subgroup.

Nevertheless, another issue deserves some conclusion remarks, relating genetic information and adverse events and even cardiovascular mortality. It's irrefutable that GRS is associated with worse prognosis and outcomes. Several GRS including SNPs

associated with CHD have been associated with incident events in distinct cohorts.³ These studies have been conducted in both primary and secondary prevention subpopulations. Long-term prospective cohorts have validated GRS as a marker for incident events across all risk categories in primary care.^{14,33} Mega et al.¹⁷ have also analyzed GRS orientated therapy with statins reporting interesting number needed to treat reaching just 20 to 25 patients with high GRS to avoid one incident adverse outcome. Also Voight et al.³³ have already tested GRS in a mendelian randomization design trial and thus GRS has been pointed out as an effective tool for early statin introduction in patients with a concordant genotypic score and phenotypic LDL elevation.³³ Similarly, in a secondary prevention setting, our group has reported, in 1555 patients with documented CAD, significantly higher cardiovascular mortality in patients with higher GRS.²⁵

Quite disparate number of gene variants and different population settings have been validated in these trials, ranging from 13 to a few hundred genes. Not in all of these studies with positive association with adverse outcomes at the long-term follow-up, the NRI was statistically significant,²⁹ not being this fact a justification for its clinical uses.²³ Pencina et al.²⁴ have already addressed this limitation of comparing NRI stating that continuous NRI may not be adequate in case-control studies, but also distinct risk cut-offs and number of risk categories in categorical NRI makes it quite difficult to inter-studies comparison. Nevertheless, in this study we have used the continuous metric formula recognizing its limitation for interpretation because when using continuous NRI it is harder to improve models that already perform well.²⁴

Irrespective of its preventive strategy control role and relative mortality reduction, classical risk factors are quite changeable over time with lifestyle and drug interventions. Also many biochemical and phenotypic biomarkers like hs-CRP and PWV are powerful predictors of disease risk for short-term concern but less accurate in assessing lifetime risk. Screening for risk stratification is an important issue. Genetic print is an attractive measure for risk prediction with a number of advantages over classical risk stratification. Though its true predictive power in different clinical scenarios can be quite disparate, being self “tailored” and highly stable over time makes it desirable over classical measures.

It is expected that financial cost related to genome-sequencing will further plunge in near future and, in the short-term, additional genetic prediction will be available as regular biomarkers determinations and, in this way, can reach larger population coverage. These findings confirm the importance of genetic contributions to CAD disease even adjusting for traditional risk stratification.

Strengths and Limitations

A significant strength was the use of a genetically isolated population. This has been especially valuable for mapping rare recessive disorders, but many researchers believe this could be a solution for more complex disorders as well because of the relatively uniform genetic background of the population. Some culturally and genetically isolated populations have a more similar lifestyle, eating habits and natural environment that reduces environmental variation. Often these populations have been founded by a small number of individuals, followed by a period of genetic isolation, during which genetic

drift might have been seen and population expansion mainly occurred by population growth and not by immigration.¹²

Our study has included a limited number of genetic variants associated with CAD. Although there is no established minimal number of genes to include in a genetic score and this number may be dependent on alleles frequencies specific to the population set in analysis, including more SNPs into our score could increase the predictive power for CAD and other outcomes. Further studies evaluating the clinical utility of adding a GRS in very large samples of individuals are underway.

Continuous NRI may have improved our prediction performance for each category risk. Nevertheless as previously stated some criticism about limitation of using NRI either as a continuous or categorical variable are still under debate and no conclusive remarks are yet accepted.

Conclusions

Overall, our findings demonstrate that GRS is independent of TRF in a single-center cohort with adequate population stratification. GRS improved the predictive risk of FRS across all subgroups in 10-year risk estimates according to Framingham risk prediction function. Marginal increase in reclassifications measures were found with a similar pattern across all subgroups. Further debate is warranted to incorporate genetic data in subgroups known to be subestimated with standard stratification.

Acknowledgments

We are very grateful to Elsa Sousa who made all the administrative procedures and to Rita Freitas (CIDEHUS, Évora University, Portugal) who reviewed the paper for statistical analysis.

References

Abdi Hervé and Williams Lynne J (2010) Principal component analysis. *WIREs Comp Stat* 2: 433-459.

Asmar R, Benetos A, Topouchian J, Laurent P, Pannier B, Brisac AM, Target R, Levy BI (1995) Assessment of arterial distensibility by automatic pulse wave velocity measurement: validation and clinical application studies. *Hypertension* 26:485–490.

CARDIoGRAMplusC4d Consortium (2013) Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet* 45:25-33.

Chan L and Boerwinkle E (1994) Gene-environment interactions and gene therapy in atherosclerosis. *Cardiol Rev* 2:130–137.

Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL Jr, Jones DW, Materson BJ, Oparil S, Wright JT Jr (2003) The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure: the JNC 7 report. *JAMA* 289:2560-2572.

Coronary Artery Disease Genetics Consortium (2011) A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet* 43: 339–344.

DeLong ER, DeLong DM and Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44: 837–845.

Ganna A, Magnusson PKE, Pedersen NL, de Faire U, Reilly M, Arnlöv J, Sundström J, Hamsten A, Ingelsson E (2013) Multilocus genetic risk score for coronary heart disease prediction. *Arterioscler Thromb Vasc Biol* 33:2267-2272.

Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults (2001) Executive Summary of the Third Report of the National Cholesterol Education

Program (NCEPT) Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 285:2486-2497.

Humphries SE, Yiannakouris N and Talmud PJ (2008) Cardiovascular disease risk prediction using genetic information (gene scores): is it really informative? *Curr Opin Lipidol* 19:128–132.

Iribarren C, Lu M, Jorgeson E, Martinez M, Lluís-Ganella C, Subirana I, Salas E, Elosua R (2016) Clinical Utility of Multi-marker Genetic Risk Scores for Prediction of Incident Coronary Heart Disease: A Cohort Study among over 51 Thousand Individuals of European Ancestry *Circ Cardiovasc Genet* Dec; 9(6):531-540.

Jorde LB and Wooding SP (2004) Genetic variation, classification and 'race'. *Nat genet* 36(11Suppl):S28-33.

Krøger NT, Borglykke A, Allin KH, et al (2015) A genetic risk score of 45 coronary artery disease risk variants associates with increased risk of myocardial infarction in 6041 Danish individuals. *Atherosclerosis* 240(2):305-10.

Lloyd-Jones DM, Nam BH, D'Agostino RB, Levy D, Murabito JM, Wang TJ, Wilson PW, O'Donnell CJ (2004) Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA* 291:2204–2211.

Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U (1994) Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med* 330:1041–1046.

Mega JL, Stitzel NO, Smith JG, Chasman DI, Caulfield MJ, Devlin JJ, Nordio F, Hyde CL, Cannon CP, Sacks FM (2015) Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet* 385(9984):2280-7.

Mons U, Müezzinler A, Gellert C, Schöttker B, Abnet CC, Bobak M, de Groot L, Freedman ND, Jansen E, Kee F (2015) Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. *BMJ* 350:h1551.

Morrison AC, Bare LA, Chambless LE, Ellis SG, Malloy M, Kane JP, Pankow JS, Devlin JJ, Willerson JT, Boerwinkle E (2007) Prediction of coronary heart disease risk using a genetic risk score: the Atherosclerosis Risk in Communities Study. *Am J Epidemiol* 166:28–35.

Murabito JM, Pencina MJ, Nam BH, D'Agostino RB Sr, Wang TJ, Lloyd-Jones D, Wilson PW, O'Donnell CJ (2005) Sibling cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults. *JAMA* 294:3117–3123.

National Institute of Health, National Heart, Lung, and Blood Institute North American Association for the Study of Obesity (2000) *The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults*. NHLBI Obesity Education Initiative, NIH Publication Number: 00-4084.

Paynter NP, Chasman DI, Pare G, Buring JE, Cook NR, Miletich JP, Ridker PM (2010) Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA* 303:631–637.

Pencina MJ, D'Agostino RB Sr, Larson MG, Massaro JM, Vasan RS (2009) Predicting the 30-year risk of cardiovascular disease: the Framingham heart study. *Circulation* 119:3078-3084.

Pencina MJ, Steyerberg EW and D'Agostino RB Sr (2011) Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 30(1): 11–21.

Pereira A, Mendonça MI, Sousa AC, Borges S, Freitas S, Henriques E, Rodrigues M, Freitas AI, Guerra G, Ornelas I (2017) Genetic risk Score and Cardiovascular Mortality in a Southern European Population with Coronary Artery Disease. *Int J Clin Pract* doi: 10.1111/ijcp.12956.

Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (2003) *Diabetes Care* 26: 3160-3167.

Ridker PM, Buring JE, Rifai N, Cook NR (2007) Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA* 297:611–619.

Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR (2008) C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation* 118:2243–2251.

Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, Guiducci C, Perola M, Jula A, Sinisalo J, Lokki ML (2010) A multilocus genetic risk score for coronary heart disease: case control and prospective cohort analyses. *Lancet* 376:1393-1400.

Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C (2011) Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* 43:333-338.

Tada H, Melander O, Louie JZ, Catanese JJ, Rowland CM, Devlin JJ, Kathiresan S, Shiffman D (2016) Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. *Eur Heart J* 37(6):561-7.

Van der Net JB, Janssens AC, Defesche JC, Kastelein JJ, Sijbrands EJ, Steyerberg EW (2009) Usefulness of genetic polymorphisms and conventional risk factors to predict

coronary heart disease in patients with familial hypercholesterolemia. *Am J Cardiol* 103:375–380.

Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, Hindy G, Hólm H, Ding EL, Johnson T (2012) Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* 380(9841):572–80.

Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 42(Database issue):D1001-D1006.

Supplementary material – the following online material is available for this article:

Table S1 - List of the 32 genetic variants previously associated with CAD risk, selected to create the GRS in our Population.

