

**ANDRÉ GUSTAVO PIRES DE SOUSA**

**Associação entre *TCF7L2* e outras variantes  
genéticas de risco para Diabetes Mellitus Tipo 2 e  
doença cardiovascular**

Tese apresentada à Faculdade de Medicina da  
Universidade de São Paulo para obtenção do  
título de Doutor em Ciências

Programa de Cardiologia

Orientador: Prof. Dr. José Eduardo Krieger

**São Paulo**

**2010**

**Dados Internacionais de Catalogação na Publicação (CIP)**

Preparada pela Biblioteca da  
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Sousa, André Gustavo Pires de  
Associação entre *TCF7L2* e outras variantes genéticas de risco para Diabetes Mellitus Tipo 2 e doença cardiovascular / André Gustavo Pires de Sousa. -- São Paulo, 2010.  
Tese(doutorado)--Faculdade de Medicina da Universidade de São Paulo.  
Programa de Cardiologia.  
Orientador: José Eduardo Krieger.

Descritores: 1.Diabetes mellitus 2.Doença da artéria coronariana 3.Doenças cardiovasculares 4.*TCF7L2* 5.Fatores de risco 6.Prevalência 7.Genética

USP/FM/DBD-524/10

Sousa, A.G.P. **Associação entre *TCF7L2* e outras variantes genéticas de risco para Diabetes Mellitus Tipo 2 e doença cardiovascular.** Tese apresentada à Faculdade de Medicina da Universidade de São Paulo para obtenção do título de Doutor em Ciências, Programa de Cardiologia.

Aprovado em:

Banca Examinadora

Prof. Dr.: \_\_\_\_\_ Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_ Assinatura: \_\_\_\_\_

Prof. Dr.: \_\_\_\_\_ Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_ Assinatura: \_\_\_\_\_

Prof. Dr.: \_\_\_\_\_ Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_ Assinatura: \_\_\_\_\_

Prof. Dr.: \_\_\_\_\_ Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_ Assinatura: \_\_\_\_\_

Prof. Dr.: \_\_\_\_\_ Instituição: \_\_\_\_\_

Julgamento: \_\_\_\_\_ Assinatura: \_\_\_\_\_

**Dedicatória**

---

**À Virna, minha esposa, companheira nos bons  
e maus momentos, especial em todos**

**A Valdeci e Graça, meus pais, que me ensinaram  
as mais importantes coisas da vida: os valores**

## **Agradecimientos**

---

Em um momento de significância ímpar para mim, julgo importante agradecer e homenagear pessoas e instituições que, de diferentes formas, muito contribuíram para o sucesso desse trabalho. De antemão, peço desculpas a quem, por ventura, eu me esqueci de mencionar ou a quem minhas seguintes palavras não alcançaram a real magnitude do meu reconhecimento e gratidão.

Ao Prof. Dr. José Eduardo Krieger, pelo privilégio de executar minha pós-graduação em seu laboratório e pelo exemplo de professor, pesquisador e gestor. A sua visão estratégica, ao mesmo tempo precavida e inovadora, o seu relacionamento com alunos, funcionários e colegas e, é claro, o seu desempenho científico, são verdadeiros estímulos para os iniciantes na vida acadêmica.

Ao Dr. Alexandre da Costa Pereira, pelas orientações, pelos conselhos, por todo o aprendizado. Agradeço por ter acreditado em mim mesmo quando nem eu mesmo acreditava. Sua confiança é uma honra para mim.

Aos meus irmãos, por compartilharem comigo nossas infâncias, adolescências e nosso presente, e certamente nosso futuro. Nossos laços fraternos são muito mais robustos agora.

A Marcos Rocha e Edione Rocha, meus sogros, e Vinicius Rocha, meu cunhado, pelo apoio e aconchego familiar necessários à plena execução das atividades profissionais.

Aos Profs. Josivan Lima, Lucia Helena Nóbrega e Maria Lúcia Nóbrega, por terem me despertado para a ciência e pelos incontáveis estímulos para que eu tente alcançar o máximo do meu potencial.

Aos Profs. José Carlos Nicolau, Carlos Serrano Júnior e Pedro Lemos, membros da minha banca de qualificação, pela preciosa participação na melhoria da qualidade da versão final desse trabalho.

Ao amigo Guilherme Marquezine, pela colaboração mútua e pelas orientações básicas nos primórdios do processo de doutoramento.

Aos Profs Marcos Oliveira, Fátima Baracho e José Brandão Neto, professores da Disciplina de Endocrinologia da UFRN, pelos ensinamentos e partilha das suas experiências docentes, bem como o apoio dedicado a mim durante as minhas ausências para a conclusão dessa tese.

À Profa. Adriana Nunes, pela valiosa contribuição na revisão desta tese com o intuito de melhorar a sua qualidade.



Aos amigos Profs. Ricardo Oliveira e Zilma Souza, por terem se esforçado em dobro, em meu lugar, na árdua tarefa de gestão do Internato de Clínica Médica do Curso de Medicina da UFRN.

Aos amigos Júlio Vieira e Cinthia Bastia, pela amizade, compreensão e paciência durante a árdua empreitada de dividir um teto em São Paulo.

Aos amigos Lidiane Moura, Liana Viana, Lissandro Barros e Gustavo Marques, meus sócios, por partilharem os difíceis momentos do início de uma carreira profissional.

Aos Drs. Adriana Angelucci, Eduardo Lorena, Gustavo Akerman, Luciana Audi, Luciana Seoane, Marcos Tadashi e Patrícia Pugliese, amigos residentes de Endocrinologia da FMUSP, pelo companheirismo e amizade durante os fantásticos anos de residência. Vocês foram minha família paulista. Não posso esquecer ainda dos Drs. Tatiana Goldbaum, Daniel Laguna, Daniel Vicária, Fernando Sebastianes, Leandro Michelin, Michelle Vidinha, Patrícia Zanoni, Roberta Lordelo, Sharon Admoni, Ana Priscila Soggia, Daniela Iguchi, Eduardo Quadros, Isabel Carvalho, Marcelo Shimizu, Márcio Aurélio Pinto, Rafael Pergher e Tiago Cavalcanti. Vocês tornaram a vida em São Paulo bastante agradável e feliz.

Aos amigos do laboratório de Genética e Cardiologia Molecular do InCor: Noely Ferreira, Alessandro Betito, Rafael Alvim, Paulo Caleb, Diogo Biaggi, Ruth, Alessandra Medeiros, Lívia Marques, Lívia Selvatici, Cléber René, Isolmar Schettert, Amanda Brasil, André Vaquero, Júlia Marsiglia, Marina França, Luciana Turolla, Ana Carolina Moron, Flávia Credidio, Ashila Almeida, Lea Bozzini, Silvana Campos, Maria de Lourdes (Maúde) Junqueira, Luciana Soares, os recém-chegados, os que foram e os que virão e os que eu não lembrei, pelo auxílio e apoio. Fico feliz com o sucesso de vocês.

A Neusa, Juliana, Eva, da secretaria da pós-graduação do InCor, pela simpatia e eficiência com que sempre trataram os problemas e necessidades dos alunos de pós-graduação.

A Faculdade de Medicina da Universidade de São Paulo, uma instituição a qual aprendi a respeitar, admirar e honrar, não apenas pela tradição, mas pelo que representa para a comunidade científica brasileira.

A Universidade Federal do Rio Grande do Norte, pela graduação, pós-graduação e pela oportunidade de oferecer a outras gerações o meu amor e devoção pela Medicina.

A Deus, por todas as graças alcançadas.

A todos os meus professores, oficiais ou eventuais, pelos ensinamentos sobre a arte e ciência da Medicina. Em um momento em que se poderia imaginar que estou mais perto de lhes alcançar, sinto, na verdade, que estou mais preparado para lhes dar o merecido reconhecimento.

## Normalização Adotada

Esta tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: estilo Vancouver

Universidade de São Paulo. Faculdade de Medicina. Serviço de Biblioteca e Documentação. Guia de apresentação de dissertações, teses e monografias. Elaborado por Annelise Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 2ª ed. São Paulo: Serviço de Biblioteca e Documentação; 2005.

Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

## **Sumário**

---

## SUMÁRIO

1	Introdução.....	1
1.1	Genética do Diabetes Mellitus Tipo 2.....	4
1.1.1	Marcadores Genéticos Estudados.....	6
1.2	O Importante Papel do <i>TCF7L2</i> .....	18
1.3	Efeitos da Combinação da Informação de Diferentes Variantes Genéticas.....	20
1.4	Racional e justificativa do presente estudo .....	21
2	Objetivos.....	25
3	Métodos.....	27
3.1	Objetivo Específico 1 .....	28
3.1.1	População 1 - Indivíduos referenciados para cateterismo cardíaco para diagnóstico de DAC .....	29
3.1.2	População 2 – Indivíduos com Doença Arterial Coronariana Multiarterial do Estudo MASS II .....	30
3.1.3	Construção de um Escore de Gravidade da DAC .....	37
3.1.4	Determinação do genótipo do polimorfismo rs7903146 do gene <i>TCF7L2</i> .....	38
3.1.5	Análise Estatística .....	41
3.2	Objetivo Específico 2 .....	43
3.2.1	Seleção de SNPs, alelos de risco e genotipagem .....	44
3.2.2	Construção dos Escores Genotípicos .....	47
3.2.3	Análise Estatística .....	48
4	Resultados.....	52
4.1	Objetivo Específico 1 .....	53
4.1.1	Avaliação da Prevalência e Gravidade das Lesões Coronarianas na População Sumetida a Cateterismo Cardíaco.....	53
4.1.2	Incidência de Eventos Cardiovasculares na População do Estudo MASS II .....	59
4.1.3	Resultados do Objetivo Secundário 1 .....	63
4.2	Objetivo Específico 2 .....	67

5	Discussão.....	82
5.1	Relação entre hiperglicemia e Doença Cardiovascular.....	83
5.2	Hiperglicemia como um fator de risco cardiovascular em indivíduos não-diabéticos.....	86
5.3	Relação entre <i>TCF7L2</i> e Doença Arterial Coronariana.....	91
5.4	Análise crítica e limitações dos resultados do Objetivo Específico 1.....	95
5.5	Combinação da informação de variantes genéticas de risco para diabetes: prováveis implicações na fisiopatologia do DM2 ....	97
5.6	Efeitos preditivos dos Escores Genotípicos na avaliação de risco cardiovascular em portadores de doença arterial coronariana crônica.....	100
5.7	Análise crítica e limitações dos resultados do objetivo específico 2.....	104
5.8	Associação entre as variantes genéticas e o diagnóstico de DM2 .....	108
5.9	Considerações sobre os escores genéticos.....	112
5.10	Possíveis aplicações clínicas dos fatores de risco genético associados ao DM2.....	113
5.11	Perspectivas na avaliação do risco genético do DM2 .....	117
6	Conclusões.....	120
7	Referências .....	122

Apêndices

## **Lista de Figuras**

---



<b>Figura 1</b> - Papel do canal de potássio ATP-dependente na secreção de insulina.....	9
<b>Figura 2</b> - Estimativas do tamanho do efeito do risco de DM2.....	21
<b>Figura 3</b> - Esquema de determinação do genótipo por meio da técnica de RFLP.....	40
<b>Figura 4</b> - Curva de Kaplan-Meier para eventos cardiovasculares combinados em não-diabéticos (A) e diabéticos (B) após 5 anos de seguimento, de acordo com os grupos de genótipos do <i>TCF7L2</i> .....	63
<b>Figura 5</b> - Curvas ROC para predição de DM2 dos dois modelos em Vitória.....	65
<b>Figura 6</b> - Curvas de Kaplan-Meier dos escores genotípicos com e sem informação do <i>TCF7L2</i> e desfechos cardiovasculares compostos em indivíduos não-diabéticos .....	74
<b>Figura 7</b> - Curvas de Kaplan-Meier dos escores genotípicos e defechos cardiovasculares compostos para indivíduos diabéticos e não-diabéticos separados em grupos, de acordo com os tercís dos escores genéticos, após 5 anos de seguimento .....	76
<b>Figura 8</b> - Curvas ROC para o diagnóstico de DM2 para o Escore Clínico (contendo fatores de risco para DM2) e a combinação dos escores genotípicos (Escore 1, 2 ou 3) e o Escore Clínico.....	81

## **Lista de Tabelas**

---

<b>Tabela 1</b> -	Descrição do sistema de pontuação empregado na construção do escore de gravidade das lesões coronarianas .....	37
<b>Tabela 2</b> -	Variantes genéticas associadas ao DM2: SNPs, localização cromossômica e alelos de risco .....	44
<b>Tabela 3</b> -	Variáveis de risco, análise de regressão logística múltipla e pontuação atribuídas às variantes genéticas para o Escore 3 e Escore 3 sem <i>TCF7L2</i> .....	50
<b>Tabela 4</b> -	Variáveis de risco, análise de regressão logística múltipla e pontuação atribuídas às variáveis para construção do Escore Clínico e escores derivados .....	51
<b>Tabela 5</b> -	Características demográficas e laboratoriais de indivíduos não-diabéticos e diabéticos da população submetida á cateterismo cardíaco, de acordo com o genótipo do <i>TCF7L2</i> .....	56
<b>Tabela 6</b> -	Prevalência e estimativa de risco de lesões coronarianas na população submetida a cateterismo cardíaco, de acordo com o genótipo do <i>TCF7L2</i> .....	57
<b>Tabela 7</b> -	Distribuição de lesões coronárias em pacientes não-diabéticos com DAC de acordo com os grupos genotípicos, na população submetida a cateterismo cardíaco .....	58
<b>Tabela 8</b> -	Distribuição dos tercís do escore de gravidade das lesões coronarianas em pacientes não-diabéticos, de acordo com os grupos de genótipos do <i>TCF7L2</i> , na população submetida a cateterismo cardíaco .....	59
<b>Tabela 9</b> -	Características demográficas e laboratoriais da população do estudo MASS-II, de acordo com diagnóstico de diabetes e grupos de genótipos do <i>TCF7L2</i> .....	61

<b>Tabela 10</b> - Performance do modelo clínico e do modelo clínico com informação do <i>TCF7L2</i> na população geral de Vitória – ES....	66
<b>Tabela 11</b> - Características clínicas dos indivíduos diabéticos e não-diabéticos genotipados para os 10 <i>loci</i> de risco na população do projeto MASS II .....	68
<b>Tabela 12</b> - Fatores Genéticos Preditivos de Diabetes Mellitus Tipo 2 na população do estudo MASS-II .....	70
<b>Tabela 13</b> - Regressão logística multivariada para o diagnóstico de DM2 na população do projeto MASS-II.....	71
<b>Tabela 14</b> - Proporção de indivíduos com DM2, segundo grupo de risco genético (baseado nos tercis dos escores genotípicos) .....	72
<b>Tabela 15</b> - Distribuição dos escores genéticos classificados, de acordo com os tercis das somas dos valores da pontuação de cada SNP .....	73
<b>Tabela 16</b> - Número de eventos cardiovasculares compostos e Hazard cumulativo em indivíduos diabéticos e não diabéticos, de acordo com os tercis dos escores genotípicos .....	77
<b>Tabela 17</b> - Riscos relativos e intervalos de confiança para eventos cardiovasculares compostos e mortalidade em função do aumento de 1 ponto em cada escore genético em indivíduos não-diabéticos .....	78
<b>Tabela 18</b> - Riscos relativos e intervalos de confiança para eventos cardiovasculares compostos e mortalidade em função do aumento de 1 ponto em cada escore genético, em indivíduos não-diabéticos, ajustados para glicemia de jejum ..	80

## **Lista de Abreviaturas e Siglas**

---

A	Adenina
ABCC8	<i>ATP-binding cassette, sub-family C member 8</i>
Acum	Acumulada
Acil-CoA	Acil Coenzima A
ADA	<i>American Diabetes Association</i>
ADP	Difosfato de adenosina
Ala	Alanina
ARIC	<i>Atherosclerosis Risk in Communities</i>
ATC	Angioplastia Transluminal Coronária
ATP	Trifosfato de adenosina
AUC	<i>Area under the curve</i> – área sob a curva
C	Citosina
°C	Grau Célsius
CD	<i>Cluster of differentiation</i> - grupamento de diferenciação
CDC	<i>Centers for Disease Control</i>
CDK5	<i>Cyclin-dependent kinase 5</i> – quinase ciclina-dependente
CDK5RAP1	<i>CDK5 regulatory subunit associated protein 1</i>
CDKAL-1	<i>CDK5 regulatory subunit-associated protein 1-like 1</i>
CDKN2	<i>Cyclin-dependent kinase inhibitor 2</i>
cM	Centimorgan
CPK	Creatinofosfoquinase
Cum	Cumulative – Cumulativa
DAC	Doença Arterial Coronária

DAP	Doença Arterial Periférica
DCCT	<i>Diabetes Control and Complications Trial</i>
DCV	Doença Cardiovascular
DECODE	<i>Diabetes Epidemiology Collaborative Analyses of Diagnosis Criteria in Europe</i>
DIAGRAMA	<i>Diabetes Genetics Replication And Meta-analysis Consortium</i>
dL	Decilitro
DM2	Diabetes melitus tipo 2
DNA	Ácido desoxirribonucléico
dNTP	Desoxirribonucleotídeos fosfatados
E	Ácido glutâmico
ECA	Enzima conversora de angiotensina
ECG	Eletrocardiograma
EDIC	<i>Epidemiology of Diabetes Intervention and Complications</i>
EDTA	Ácido etilenodiamino tetra-acético
<i>ENPP1</i>	<i>Ectonucleotide pyrophosphatase/phosphodiesterase 1</i>
FAR <sub>AF</sub>	Frequência do alelo de risco em indivíduos afetados
Fe(II)	Íon ferroso
FEVE	Fração de Ejeção do Ventrículo Esquerdo
<i>FTO</i>	<i>Fat mass and obesity associated</i>
G	Guanina
GH	<i>Growth hormone</i> – hormônio do crescimento
GIP	<i>Glucose-dependent insulinotropic peptide</i>
GLP-1	<i>Glucagon-like peptide</i>
GLUT	<i>Glucose transporter</i> – transportador de glicose

GWA	<i>Genome wide association studies</i>
HAS	Hipertensão Arterial Sistêmica
HC-FMUSP	Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo
HCl	Ácido Clorídrico
HDL	<i>High density lipoprotein</i> – lipoproteína de alta densidade
<i>HHEX</i>	<i>Haematopoietically expressed homeobox</i>
HMG-CoA	Hidroximetilglutaril Coenzima A
HNF	<i>Hepatic nuclear factor</i> – fator nuclear hepatocitário
HOMA	<i>Homeostatic model assessment</i> - modelo de avaliação da homeostase
HR	<i>Hazard ratio</i>
IAM	Infarto Agudo do Miocárdio
IC	Intervalo de confiança
ICC	Insuficiência Cardíaca Congestiva
IDE	<i>Insulin degrading enzyme</i>
IGF1	<i>Insulin-like growth factor 1</i>
IGFBP	<i>Insulin-like growth factor binding protein</i> – proteínas ligantes do IGF
IMC	Índice de Massa Corpórea
InCor	Instituto do Coração
IRC	Insuficiência Renal Crônica
K	Lisina
KATP	Canal de potássio ATP-dependente
Kbp	Quilobase
KCl	Cloreto de potássio



<i>KCNJ11</i>	<i>Potassium inwardly rectifying channel, subfamily J, member 11</i>
Kg	Quilograma
LDL	<i>Low density lipoprotein</i> – lipoproteína de baixa densidade
M	Molar
m <sup>2</sup>	Metro quadrado
MASS	<i>Medicine, Angioplasty, Surgery Study</i>
Mbp	Megabase
MET	Equivalente metabólico
mg	Miligrama
Mg <sup>2+</sup>	Íon magnésio bivalente
MgCl <sub>2</sub>	Cloreto de magnésio
mL	Mililitro
mm	Milímetro
mM	Milimolar
mmHg	Milímetro de mercúrio
MODY	<i>Mature onset diabetes of young</i>
MONICA	<i>Monitoring of Trends and determinants in Cardiovascular Disease</i>
MRFIT	<i>Multiple Risk Factor Intervention Trial</i>
ms	Milissegundo
NaCl	Cloreto de Sódio
NADPH	<i>Nicotinamide Adenine Dinucleotide Phosphate</i>
NF-κB	<i>Nuclear factor kappa B</i> – fator nuclear kappa B
Ng	Nanograma
NH <sub>4</sub> Cl	Cloreto de amônio

NH <sub>4</sub> HCO <sub>3</sub>	Bicarbonato de amônio
Nm	Nanômetro
NYHA	<i>New York Heart Association</i>
OR	<i>Odds ratio</i> – razão de chances
PA	Pressão Arterial
Pb	Pares de base
PCR	<i>Polimerase Chain Reaction</i> – reação de Polimerase em cadeia
pH	Potencial hidrogeniônico
Pro	Prolina
PIP <sub>2</sub>	<i>Phosphatidyl inositol bisphosphate</i>
PPARG	<i>Peroxisome proliferator-activated receptor-g</i>
R	Arginina
RFLP	<i>Restriction Fragment Length Polymorphism</i>
RNA	Ácido ribonucleico
RNA <sub>m</sub>	RNA mensageiro
ROC	<i>Receiver Operating Characteristic</i>
Rpm	Rotações por minuto
RR	Risco relativo
RVM	Revascularização Cirúrgica do Miocárdio
SDS	<i>Sodium dodecyl sulfate</i>
SLC30A8	<i>Solute carrier family 30 (zinc transporter), member 8</i>
SLC	<i>Solute carriers</i>
SNP	<i>Single nucleotide polymorphism</i> – polimorfismo de base única
Sobrev	Sobrevida

SUR	<i>Sulfonylurea receptor</i> - receptor de sulfonilureia
<i>TCF7L2</i>	<i>Transcription factor 7 like 2 (T-cell specific, HMG box).</i>
TE	Teste ergométrico
TNF $\alpha$	<i>Tumor Necrosis Factor alpha</i> – Fator de Necrose Tumoral alfa
Tris-HCl	<i>Cloridrato tris-hidroximetil-aminometano</i>
$\mu$ l	Microlitro
W	Triptofano
<i>WFS1</i>	<i>Wolfram syndrome gene</i>
WHO	World Health Organization
UKPDS	<i>United Kingdom Prospective Diabetes Study</i>
VPN	Valor preditivo negativo
VPP	Valor preditivo positivo
ZnT	<i>Zinc transporter</i> – transportador de zinco
Zn <sup>2+</sup>	Íon zinco bivalente

**Resumo**

---

Sousa AGP. *Associação entre TCF7L2 e outras variantes genéticas de risco para diabetes mellitus tipo 2 e doença cardiovascular* [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2010. 137p.

**Introdução:** Estudos prévios têm confirmado genes de susceptibilidade para Diabetes Mellitus tipo 2 em diferentes populações. Nesse contexto, os polimorfismos do gene *TCF7L2* são considerados as mais importantes variantes genéticas de risco para o diabetes tipo 2, que, por sua vez, é um dos principais fatores de risco cardiovascular conhecidos, especialmente para doença arterial coronariana. Contudo, a hipótese de que variantes de risco do *TCF7L2* ou a informação combinada de marcadores genéticos associados ao diabetes também podem prever eventos cardiovasculares ainda não foi testada. **Objetivos:** este estudo objetivou avaliar a associação entre o polimorfismo rs7903146 do gene *TCF7L2* e doença arterial coronariana em indivíduos diabéticos e não-diabéticos e determinar se a informação combinada de variantes genéticas de risco para diabetes também está associada com eventos cardiovasculares incidentes e diabetes. **Métodos:** Oitocentos e oitenta e nove indivíduos, que foram referenciados para cateterismo cardíaco, foram avaliados em um formato transversal para lesões coronárias (carga aterosclerótica) e 559 indivíduos do estudo MASS II, um estudo randomizado em pacientes portadores de doença coronariana multiarterial, foram seguidos prospectivamente e avaliados para a incidência de eventos cardiovasculares. Em outra análise, 425 indivíduos, também do estudo MASS II, foram genotipados para 10 polimorfismos associados com diabetes tipo 2. Escores genotípicos foram criados se utilizando da informação combinada destes polimorfismos. **Resultados:** Os indivíduos não-diabéticos portadores do alelo T do *TCF7L2* apresentaram uma frequência significativamente maior de lesões coronarianas que os não-diabéticos que não carregavam o alelo de risco (OR = 2,32 IC95% 1,27-4,24, p = 0,006). O mesmo não foi encontrado para indivíduos diabéticos. Da mesma forma, na análise prospectiva, os indivíduos não-diabéticos portadores dos genótipos CT/TT apresentaram significativamente maior número de eventos cardiovasculares compostos que o genótipo CC (p = 0,049), principalmente devido à mortalidade (p = 0,004). A informação genética combinada também foi capaz de prever a incidência de eventos cardiovasculares e mortalidade global em indivíduos não-diabéticos em 5 anos. Os indivíduos não-diabéticos com alto risco genético tiveram uma incidência de eventos similar aos sujeitos diabéticos. A adição da informação do *TCF7L2*, isoladamente, a um modelo de predição clínica de risco de diabetes não foi capaz de aumentar a sua precisão, mas a adição de informação genética combinada à informação clínica melhorou significativamente a capacidade diagnóstica para diabetes. **Conclusões:** o alelo de risco T do polimorfismo rs7903146 está associado

com diabetes e, em indivíduos não-diabéticos, com maior prevalência e gravidade da doença arterial coronária, bem como eventos cardiovasculares combinados. A informação de variantes genéticas de risco para diabetes está associada à incidência de eventos cardiovasculares em indivíduos não-diabéticos com doença arterial coronária.

**Descritores:** 1.Diabetes mellitus 2.Doença da artéria coronariana  
3.Doenças cardiovasculares 4.*TCF7L2* 5.Fatores de risco 6.Prevalência  
7.Genética

## Summary

---

Sousa AGP. *Association between both TCF7L2 and genetic risk variants for type 2 diabetes mellitus and cardiovascular disease* [thesis]. São Paulo: “Faculdade de Medicina, Universidade de São Paulo”; 2010. 137p.

**Background:** studies have confirmed and have identified novel susceptibility genes for type 2 Diabetes Mellitus in different populations. In this context, polymorphisms in *TCF7L2* gene are considered the main genetic risk variants for type 2 diabetes, which is a major risk factor for cardiovascular disease, especially for coronary artery disease. Nevertheless, the hypothesis that either *TCF7L2* variants or combined information from genetic markers associated with diabetes can also predict major cardiovascular events has not been tested. **Objectives:** this study aimed to evaluate the association between *TCF7L2* polymorphism rs7903146 and coronary artery disease in diabetic and non-diabetic subjects and to determine whether information from combined genetic risk variants for diabetes is also associated with cardiovascular events incidence and with diagnosis of diabetes. **Methods:** eight-hundred and eighty nine subjects who were referred for cardiac catheterization were cross-sectionally evaluated for coronary lesions (atherosclerotic burden) and 559 subjects from the MASS II Study population, a randomized study in patients with multi-vessel coronary artery disease, were prospectively followed-up for 5 years and assessed for major cardiovascular events incidence. In another analysis, 425 subjects from the MASS II Study were also genotyped for single-nucleotide polymorphisms at 10 *loci* associated with type 2 diabetes. Genotype scores using information from genetic markers were created. **Results:** non-diabetic individuals carrying the *TCF7L2* T allele were associated with a significantly higher frequency of coronary lesions than non-diabetic non-carriers of the risk allele (adjusted OR = 2.32 95%CI 1.27-4.24, p = 0.006). This result was not found for diabetic subjects. Similarly, from the prospective sample analysis, non-diabetic individuals carrying the CT/TT genotypes had significantly more composite cardiovascular end-points than CC genotype (p = 0.049), mainly due to death (p = 0.004). Combined genetic information was also able to predict 5-year incidence of major cardiovascular events and overall-mortality in non-diabetic individuals. Non-diabetic individuals with high genetic risk had a similar incidence of events than diabetic individuals. Addition of single polymorphism rs7903146 of *TCF7L2* gene to an established clinical risk prediction score did not increase model accuracy, but the addition of combined genetic information to clinical information significantly improved the prediction of diabetes. **Conclusions:** rs7903146 T allele is associated with diabetes and, in non-diabetic individuals, with a higher prevalence and severity of coronary artery disease and cardiovascular events. Combined information of genetic variants for diabetes risk is associated to major cardiovascular events incidence in non-diabetic individuals with coronary artery disease.

**Descriptors:** 1. Diabetes mellitus 2. Coronary artery disease 3. Cardiovascular diseases 4. *TCF7L2* 5. Risk factors 6. Prevalence 7. Genetics



# 1 Introdução

---

O Diabetes Mellitus tipo 2 (DM2) é uma doença comum com incidência e prevalência crescentes em todo o mundo, constituindo-se um verdadeiro problema de saúde pública. Nos Estados Unidos, entre os anos de 1980 a 2005, o número de adultos com idade entre 18 e 79 anos com casos recém-diagnosticados de diabetes quase triplicou, passando de 493.000 em 1980 para 1,4 milhões em 2005; e apenas no ano de 2007, cerca de 1,6 milhões de novos casos de diabetes foram diagnosticados em pessoas com mais de 20 anos (Cdc, 2008). No Brasil, a prevalência de DM2, em indivíduos com 30 a 69 anos, foi de 7,6% em 1990 (Malerbi e Franco, 1992). Contudo, atualmente esse índice provavelmente deve ser mais elevado, em virtude da crescente prevalência da obesidade, inadequados padrões alimentares e envelhecimento da população.

Na fisiopatologia do DM2, uma doença caracterizada pela hiperglicemia, duas condições são essenciais: resistência à insulina e disfunção da célula beta pancreática. Na realidade, a hiperglicemia é resultante da incapacidade da célula beta em aumentar a produção de insulina, em uma situação de reduzida sensibilidade a esse hormônio. Acredita-se que, na história natural do DM2, exista uma fase pré-clínica de aumento progressivo na resistência à insulina que se acompanha do concomitante e compensatório aumento dos níveis séricos de insulina. Com a evolução, o nível de resistência à insulina atinge um platô.

Todavia, caso a célula beta pancreática não mais consiga aumentar a secreção de insulina necessária para atingir a perfeita homeostase glicêmica, a hiperglicemia ocorrerá e o paciente apresentará critérios clínicos de DM2 (Wajchenberg, 2007a).

Adicionalmente, as incretinas, notadamente o *glucagon-like peptide* (GLP-1) e o *glucose-dependent insulintropic peptide* (GIP), vêm sendo consideradas de grande importância na fisiopatologia do DM2. As incretinas são hormônios produzidos pelo trato gastrintestinal e influenciam a secreção de insulina e glucagon em nível pancreático, o esvaziamento gástrico e o apetite. A desregulação da produção hepática de glicose, devido à neoglicogênese e/ou glicogenólise e, conseqüentemente, provocando hiperglicemia, representa outro mecanismo implicado na fisiopatologia do DM2. A produção hepática de glicose é controlada por hormônios: enquanto a insulina suprime essa produção, o glucagon (e catecolaminas) estimula a gliconeogênese e glicogenólise. Portanto, é concebível que, além de secreção de insulina prejudicada e reduzida sensibilidade hepática à insulina, o aumento da secreção de glucagon e/ou aumento da sensibilidade hepática ao glucagon contribuem para a desregulação da produção hepática de glicose (Burcelin, Knauf *et al.*, 2008).

A divulgada preocupação com as crescentes prevalência e incidência de diabetes se justifica pelas complicações crônicas relacionadas a essa condição. Classicamente, essas complicações são qualificadas em microvasculares (retinopatia diabética, neuropatia diabética e nefropatia diabética) e macrovasculares (acidente vascular encefálico, síndromes

coronarianas agudas e crônicas e doença arterial obstrutiva periférica), que, por sua vez, estão cada vez mais presentes quanto maior o tempo de diabetes. Contudo, alguns estudos têm mostrado que mesmo indivíduos diabéticos recém-diagnosticados têm uma ou mais complicações micro ou macrovasculares no momento do diagnóstico, incluindo a doença cardiovascular (DCV), que responde pela maioria dos óbitos nesses indivíduos (Milicevic, Raz *et al.*, 2008). Assim, mesmo indivíduos que não preenchem os critérios para DM2 podem apresentar tais complicações. Portanto, o estabelecimento do diagnóstico precoce de diabetes, bem como a estratificação de indivíduos de alto risco de desenvolver DM2 permitem intervenções em tempo hábil, com a possibilidade de se minimizar tais complicações, o que seria de grande valia.

## **1.1 Genética do Diabetes Mellitus Tipo 2**

A etiologia do DM2 é multifatorial, embora comumente se aceite a ideia de o envelhecimento da população além de fatores ambientais, que favorecem o desenvolvimento da obesidade (como dietas hipercalóricas e sedentarismo), contribuam fortemente para o desencadeamento dessa doença em indivíduos geneticamente susceptíveis. Realmente, o excesso de peso representa um dos principais fatores de risco para o DM2 (Hossain, Kavar *et al.*, 2007). Contudo, cerca de 10% dos portadores de DM2 possuem peso normal e, por outro lado, muitos dos indivíduos obesos nunca

desenvolverão diabetes, indicando que outros fatores são importantes, incluindo os fatores genéticos (Staiger, Machicao *et al.*, 2009).

Resultados de vários estudos confirmam a importância de fatores genéticos no desenvolvimento de DM2: certas populações com baixas taxas de miscigenação, como os índios Pima na América do Norte, apresentam uma elevada prevalência de DM2, em torno de 21% (Knowler, Bennett *et al.*, 1978); filhos de um dos pais com DM2 apresentam um risco 3,5 vezes maior de desenvolvimento da doença, quando comparados a população geral, e de 6,1 vezes, quando ambos os pais possuem DM2 (Meigs, Cupples *et al.*, 2000); finalmente, existe significativa concordância em estudos com gêmeos (Poulsen, Kyvik *et al.*, 1999). Ademais, o DM2 não segue herança Mendeliana e, dessa forma, é considerada uma doença poligênica. De acordo com a hipótese geralmente aceita, a hipótese da variante comum (Staiger, Machicao *et al.*, 2009), as doenças complexas, como diabetes tipo 2, são causadas pela ocorrência simultânea de variantes comuns (alterações da sequência do DNA, com frequências alélicas maiores que 5%) em vários genes. Supõe-se que cada uma destas variantes do DNA exerce apenas efeitos moderados sobre a função e/ou expressão dos genes afetados, mas, na sua soma, essas variações conferem uma aumentada susceptibilidade. Polimorfismos de um único nucleotídeo (SNPs), trocas de simples pares de base, respondem por cerca de 90% da variação na sequência do genoma humano (SNP Ficha de *the human Genome Project*, disponível em [http://www.ornl.gov/sci/techresources/Human\\_Genome/faq/snps.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/faq/snps.shtml)) e são, portanto, considerados como os grandes determinantes da predisposição

individual a doenças complexas. No entanto, inserções, deleções, duplicações, inversões ou variações de números de cópia de segmentos do DNA (devido à deleção e/ou duplicação de segmentos de DNA) podem, também, desempenhar um papel no risco genético de DM2 (Staiger, Machicao *et al.*, 2009). Estas variações em vários *loci* genéticos conferem uma maior susceptibilidade à doença seja diretamente seja por meio da interação com fatores ambientais.

### 1.1.1 Marcadores Genéticos Estudados

Estudos de associação, avaliação de genes candidatos e, mais recentemente, estudos de *genome wide association* (GWA) identificaram alguns SNPs que conferem aos seus portadores um maior risco de desenvolver DM2. Os estudos de GWA, nesta última década, significaram uma revolução na genética das doenças complexas, pois são capazes de avaliar simultaneamente centenas de milhares de SNPs por meio de matrizes de SNPs de alta densidade. Atualmente, as matrizes frequentemente usadas avaliam desde 500.000 a 1.000.000 de SNPs simultaneamente, o que significa uma cobertura de, pelo menos, 65% dos SNPs do genoma humano com  $r^2 > 0,8$  (Staiger, Machicao *et al.*, 2009). Nos estudos de GWA, a frequência desses SNPs genotipados são comparadas entre casos e controles, e os alelos significativamente mais frequentes em casos que controles (comumente se utiliza um nível de significância  $p < 5.10^{-8}$ ) são considerados alelos de risco.

Em 2007, uma metanálise de estudos que usou a tecnologia GWA, com dados de mais de 32.000 indivíduos, confirmou a associação de polimorfismos nos genes *PPARG*, *HHEX*, *TCF7L2* e *KCNJ11* e o risco de desenvolver DM2. Além disso, associações entre DM2 e outros cinco genes também foram significantes, a saber: *SLC30A8*, *IGFBP-2*, *CDKAL1*, *CDKN2A/B* e *FTO* (Scott, Mohlke *et al.*, 2007). Inclusive, exceto para o gene do *PPARG*, não havia dados da associação desses genes com o risco de DM2 na população brasileira.

Posteriormente, foram realizadas outras meta-análises de estudos de GWA que incluíram mais indivíduos e, desta forma, ganharam poder estatístico para encontrar *loci* adicionais associados ao DM2. Contudo, as estimativas de risco da associação (*odds ratio*) encontradas foram mais baixas (variando entre 1.09 e 1,13) que aquelas dos marcadores genéticos avaliados no presente estudo (Lee, Kang *et al.*, 2008; Unoki, Takahashi *et al.*, 2008; Zeggini, Scott *et al.*, 2008; Prokopenko, Langenberg *et al.*, 2009) e não foram incluídos no trabalho. Outros genes são considerados de risco potencial, embora ainda não exista uma confirmação sobre seus verdadeiros papéis no risco de estabelecimento de DM2 (Staiger, Machicao *et al.*, 2009).

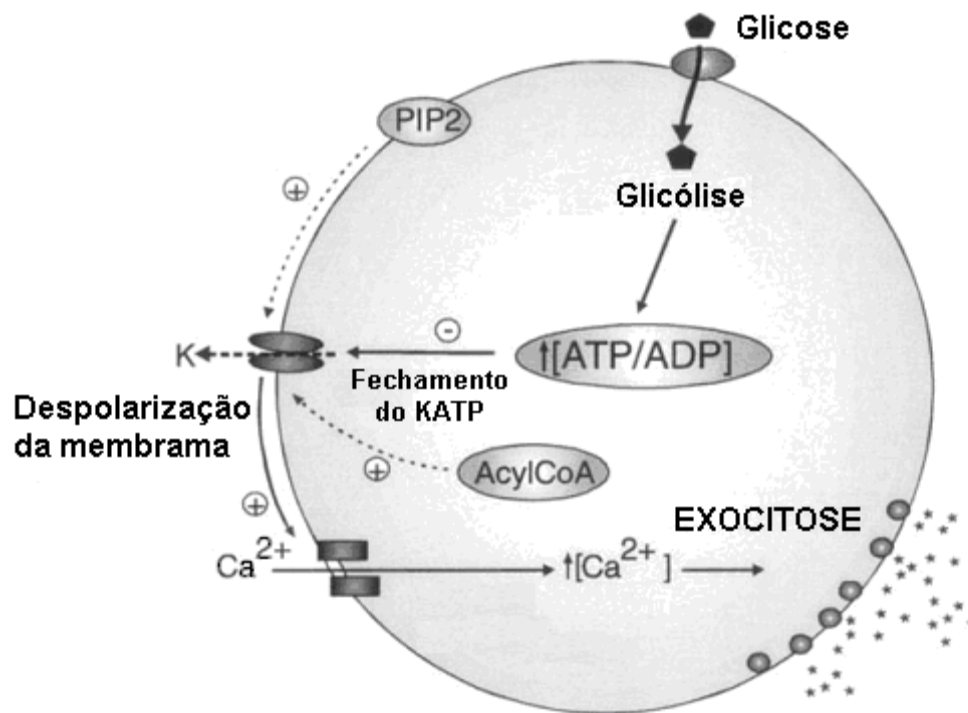
Considerando a fisiopatologia do DM2, a maioria desses *loci* está implicada supostamente com a disfunção da célula beta pancreática. Os genes *KCNJ11*, *CDKAL1*, *SCL30A8*, *TCF7L2* e *HHEX* estão, de alguma forma, associados à secreção de insulina pela célula beta

pancreática. Os genes *PPARG* e *FTO* estão relacionados com a sensibilidade à insulina e adiposidade. Por fim, as funções dos genes do *IGFBP-2* e *CDKN2A/B*, a princípio, não apresentariam relação direta com a patogênese do DM2. Contudo, interações com outros genes, com funções mais conhecidas, poderiam explicar os resultados de associação destes genes e DM2.

### ***KCNJ11***

A secreção de insulina pela célula beta pancreática é glicose-dependente. A glicose é transportada para dentro da célula beta através de uma proteína transmembrana transportadora de glicose específica, a GLUT-2. Após a entrada da glicose na célula por meio da GLUT-2, no citoplasma, ocorre metabolização da glicose (por meio da glicólise e ciclo de Krebs) e o resultado final é o aumento na relação ATP/ADP intracelular. A elevação do ATP, no ambiente intracelular, acarreta o fechamento dos canais de potássio ATP-dependentes (KATP) e consequente despolarização da membrana plasmática da célula beta pancreática. Por sua vez, isto resulta na ativação de canais de cálcio voltagem-dependentes, aumento do conteúdo de cálcio no ambiente intracelular (por influxo de cálcio a partir do extracelular e mobilização dos estoques intracelulares) e, conseqüentemente, na fusão dos grânulos secretores, que contêm insulina, com a membrana plasmática, levando a secreção da insulina.





**Figura 1** - Papel do canal de potássio ATP-dependente na secreção de insulina. Elevações na glicemia resultam, na célula beta pancreática, em aumento da relação ATP/ADP, que promovem o fechamento dos canais de potássio ATP-dependentes e, em via final, secreção de insulina. Alguns efetores, como o PIP2 e AcilCoA, modulam a sensibilidade deste canal. Retirado de Koster, J. C., M. A. Permutt, *et al.* (Diabetes, v.54, n.11, Nov, p.3065-72. 2005)

Decerto, as vias elétricas de secreção de insulina são moduladas pelos KATP. Estes canais consistem no receptor de sulfonilureia 1 (SUR1), de localização periférica, codificado pelo gene *ABCC8*, e a subunidade Kir6.2, localizada interiormente, codificada pelo gene *KCNJ11*. Metabólitos de nutrientes, incretinas e drogas, como as próprias sulfonilureias, afetam o funcionamento desses canais (Koster, Permutt *et al.*, 2005). Alterações nos sinais metabólicos, na sensibilidade dos KATP a esses metabólitos ou no número de canais ativos poderiam alterar a sinalização elétrica e, conseqüentemente, levar a distúrbios na secreção de insulina. Realmente,

estudos demonstraram que a atividade reduzida ou mesmo ausente dos KATP é a causa do Hiperinsulinismo Congênito em humanos (Huopio, Shyng *et al.*, 2002) e que mutações ativadoras em heterozigose no gene *KCNJ11* são responsáveis por alguns casos de diabetes neonatal em humanos (Koster, Permutt *et al.*, 2005).

Um estudo *in vitro* demonstrou que a variante comum K, na posição 23 do gene *KCNJ11* (rs5219), reduz a sensibilidade do KATP ao ATP, inibindo o fechamento do canal e, dessa forma, prejudicando a secreção de insulina (Schwanstecher, Meyer *et al.*, 2002). Apesar de os estudos iniciais não observarem relação entre polimorfismos da subunidade Kir6.2 e DM2 (Sakura, Wat *et al.*, 1996), vários estudos caso-controle e metanálises posteriores encontraram associações significantes entre polimorfismos que aumentam a atividade dos KATP e o DM2 surgido na vida adulta (Van Dam, Hoebee *et al.*, 2005; Scott, Mohlke *et al.*, 2007), com *odds ratios* (OR) variando entre 1,12 (IC 95% 1,01-1,23) e 1,14 (IC 95% 1,10-1,19) nessas metanálises para EK *versus* EE; e 1,44 (IC 95% 1,17-1,78) para KK *versus* EE. Contudo, deve-se salientar que o polimorfismo EK é bem mais prevalente nas populações estudadas (EK = 47% vs KK = 12%), de modo que o risco atribuível à população é ainda mais significante. Por fim, foi observado que a variante E23K prejudica a supressão da secreção de glucagon glicose-dependente, causando níveis plasmáticos elevados de glucagon durante *clamp* hiperglicêmico (Tschritter, Stumvoll *et al.*, 2002). Este achado está em concordância com a presença de KATP nas células alfa-pancreáticas e do papel desses canais na supressão do glucagon de forma glicose-dependente (Macdonald, De Marinis *et al.*, 2007).

## **PPARG**

Um dos principais fatores de transcrição para diferenciação de adipócitos, susceptibilidade à obesidade e para sensibilidade à insulina é o *peroxisome proliferator-activated receptor* gama (PPAR $\gamma$ ). Enquanto o alelo Ala12 do gene *PPARG* está associado com redução de risco de DM2 (Tonjes e Stumvoll, 2007), o polimorfismo comum Pro12Ala é um dos mais replicados fatores genéticos de risco para DM2, conferindo um significativo OR de cerca de 1,25 (Altshuler, Hirschhorn *et al.*, 2000). Scott *et al.* confirmaram esta associação com OR de 1,14 (CI 95% 1,08-1,20) (Scott, Mohlke *et al.*, 2007). Contudo, devido à grande frequência do alelo Pro12, este polimorfismo confere um importante risco atribuível à população da ordem de 25% (Tonjes e Stumvoll, 2007). Pro12Ala é causado por uma mutação missense CCA para GCA no éxon B, na isoforma adipócito-específica  $\gamma$ 2. Como o PPAR $\gamma$ 2 é expresso apenas no tecido adiposo, o mecanismo pelo qual o polimorfismo Pro12Ala influencia no risco de DM2, possivelmente, se relaciona ao adipócito. Estudos têm demonstrado que este polimorfismo está implicado com maior *clearance* de insulina e menores níveis de ácidos graxos livres durante hiperinsulinemia (Tschritter, Fritsche *et al.*, 2003). Possivelmente, o menor aporte hepático de ácidos graxos livres reduz a sensibilidade à insulina no fígado e, conseqüentemente, gera uma maior depuração da insulina da circulação. Adipocitocinas, como a adiponectina e mediadores inflamatórios, gerados por macrófagos, que também são regulados pelo PPAR $\gamma$ , podem desempenhar papéis adicionais neste mecanismo (Tonjes e Stumvoll, 2007). A explicação molecular,

subjacente a este polimorfismo, é uma moderada redução da atividade ligante-dependente do PPAR $\gamma$  (Deeb, Fajas *et al.*, 1998), reduzindo, desta forma, a função do PPAR $\gamma$ .

Paradoxalmente, um estudo sugeriu que, uma vez que os indivíduos desenvolvem DM2, a função da célula beta é mais comprometida em carreadores do alelo de proteção Ala12 (Mori, Ikegami *et al.*, 2001). Em um trabalho com a população brasileira, a frequência do alelo Ala12 foi similar entre indivíduos diabéticos e não-diabéticos (0,09 e 0,06, respectivamente) e o alelo Ala12 não foi relacionado com a proteção contra DM2 (Tavares, Hirata *et al.*, 2005). Contudo, os autores desse estudo encontraram níveis mais baixos de insulina de jejum e melhor sensibilidade à insulina entre os indivíduos diabéticos portadores do alelo Ala12 e concluem ainda que outros estudos com maior número de pacientes são necessários para se obter outras conclusões.

### **SLC30A8**

Na célula beta pancreática, acredita-se que a insulina é armazenada dentro das vesículas secretoras na forma de hexâmeros, ligados a duas moléculas do íon zinco (Zn $^{2+}$ ) para cada hexâmero. Especificamente no pâncreas, as proteínas transportadoras responsáveis pelo efluxo de zinco para o meio extracelular e para dentro das vesículas intracelulares são as *zinc transporters proteins* do tipo 8 (ZnT-8). Estas proteínas pertencem à subfamília SLC30 (*solute carriers*) da grande família dos CDF (*cation diffusion facilitator*) e são codificadas pelo gene *SLC30A8* (Chimienti,

Devergnas *et al.*, 2004). Estudos que utilizaram o princípio do GWA recentemente identificaram outros *loci* cujas variantes também conferem risco para DM2. E um desses *loci* foi o do gene *SLC30A8* (Scott, Mohlke *et al.*, 2007; Sladek R, Rocheleau *et al.*, 2007). O SNP não-sinônimo rs13266634 (R325W), no éxon final do gene *SLC30A8*, foi implicado em um maior risco de DM2, com OR entre 1,12 (IC95% 1,07-1,16) e 1,18 para heterozigotos e 1,53 para homozigotos. Estes resultados também foram confirmados em um estudo caso-controle na população japonesa (Omori, Tanaka *et al.*, 2007). Staiger *et al.* encontraram ainda que este polimorfismo está associado com uma menor liberação de insulina em indivíduos alemães não-diabéticos submetidos à sobrecarga intravenosa de glicose (Staiger, Machicao *et al.*, 2007).

### ***CDKAL1***

Os estudos que utilizaram a tecnologia GWA também encontraram associação entre DM2 e genes, cujos produtos, até então, não estavam diretamente relacionados com a fisiopatologia do diabetes. Um desses *locus* está situado próximo ao gene *CDKAL1* (*cyclin-dependent kinase 5 regulatory subunit associated protein 1-like 1*). A função do produto deste gene é desconhecida. Contudo, esta proteína é similar a outra proteína, a proteína 1 associada à subunidade regulatória do CDK5 (codificada pelo *CDK5RAP1*) (Steinthorsdottir, Thorleifsson *et al.*, 2007). A proteína CDK5RAP1 inibe a quinase ciclina-dependente 5 (CDK-5) por se ligar a subunidade reguladora p35 (Ching, Pang *et al.*, 2002). Nas células beta pancreáticas, estudos têm

demonstrado que o CDK5 desempenha um papel na disfunção da célula beta sob condições de glicotoxicidade (Wei, Nagashima *et al.*, 2005). Dessa forma, uma expressão reduzida do gene *CDKAL1*, ou uma redução na sua função inibitória, poderia levar a resposta deficiente na secreção de insulina em condições de hiperglicemia. O polimorfismo rs7756992 está localizado no íntron 5 do gene *CDKAL1*, no 6p22.3. Steinthorsdottir *et al.* encontraram um significativo risco de DM2 em portadores homocigotos desse polimorfismo, mas não em heterocigotos (Steinthorsdottir, Thorleifsson *et al.*, 2007). Em sua metanálise, Scott *et al.* observaram um OR de 1,12 (IC95% 1,08-1,16).

### ***HHEX***

Também, a partir dos trabalhos de Sladek *et al.* (Sladek R, Rocheleau *et al.*, 2007) usando GWA, outro *loci* considerado de risco para o DM2 foi identificado. Trata-se do *loci* localizado no cromossomo 10q, que contém os genes *HHEX* (*hematopoietically expressed homeobox*) e *IDE* (*insulin-degrading enzyme*). O primeiro gene possui importante função no desenvolvimento embrionário do pâncreas. Camundongos *knockout* para o *HHEX* apresentam distúrbios na formação da região ventral do pâncreas (Habener, Kemp *et al.*, 2005). Adicionalmente, o *HHE* também codifica um fator transcricional envolvido na sinalização do sistema Wnt, assim como o gene *TCF7L2* (Foley e Mercola, 2005). O gene *IDE* codifica uma enzima implicada na função da célula beta. Sladek *et al.* encontraram dois SNP associados ao DM2, rs7923837 e rs1111875. Outros estudos de larga escala em diferentes populações foram capazes de replicar esta associação

(Horikoshi, Hara *et al.*, 2007; Scott, Mohlke *et al.*, 2007; Furukawa, Shimada *et al.*, 2008; Van Vliet-Ostaptchouk, Onland-Moret *et al.*, 2008). Staiger *et al.*, por sua vez, encontraram que o SNP rs7923837 está associado com uma alteração da secreção de insulina avaliada após sobrecarga venosa de glicose, porém o polimorfismo rs1111875 não (Staiger, Stancakova *et al.*, 2008). Por sua vez, outro estudo encontrou que secreção de insulina era significativamente menor nos portadores do polimorfismo rs1111875 (Pascoe, Tura *et al.*, 2007).

### **IGFBP-2**

O locus onde se encontra o gene *IGFBP-2* (*Insulin-like Growth Factor Binding Protein 2*) também foi associado com aumento no risco de DM2. Estudos têm demonstrado a relação entre o eixo GH (*Growth Hormone*), IGF-1 (*Insulin-like Growth Factor*) e o metabolismo da glicose, apesar de que o papel preciso de cada um desses hormônios na homeostase glicêmica ainda é desconhecido. O GH apresenta algumas ações que antagonizam as ações da insulina, ao contrário do IGF-1, que apresenta alguns efeitos semelhantes aos da insulina, opostos aos do GH. A administração de GH ou IGF-1, ou condições clínicas em que os níveis ou funções desses hormônios estejam alterados, sabidamente levam a alterações no metabolismo de carboidratos e na resistência à insulina. As moléculas de IGF-1 circulam ligadas a proteínas ligadoras, as IGFBP, que participam na regulação dos efeitos do IGF-1. A principal proteína ligadora é a IGFBP-3, que reduz os níveis circulantes do IGF-1 e, dessa forma, antagoniza suas funções (Holt, Simpson *et al.*, 2003).

Estudos com GWA, que foram replicados em outras populações (Scott, Mohlke *et al.*, 2007; Zeggini, Weedon *et al.*, 2007), revelaram que um polimorfismo no segundo íntron do gene *IGFBP-2* (rs4402960) está relacionado com risco aumentado de DM2, com OR igual a 1,14 (IC95% 1,11-1,18). A proteína resultante do *IGFBP-2* liga-se à região não traduzida do RNAm do IGF-2 e, dessa forma, controla sua tradução. Por sua vez, o IGF2 está envolvido no desenvolvimento, crescimento e estimulação da ação da insulina (Nielsen, Christiansen *et al.*, 1999). Apesar de Hoeflich *et al.* terem observado que camundongos transgênicos, que superexpressam o gene *IGFBP-2*, apresentam menores níveis de glicose e insulina plasmáticas (Hoeflich, Wu *et al.*, 1999), o mecanismo, pelo qual polimorfismos nesse gene aumentaria o risco de desenvolver DM2, ainda permanece incerto.

### ***CDKN2A/B***

Outro gene implicado no risco de se desenvolver DM2 em estudos de GWA e também replicado em diferentes populações é um que, a princípio, não apresenta nenhuma relação com o metabolismo de carboidratos ou secreção de insulina, o *CDKN2*. Este gene codifica uma proteína (p16) também inibidora de quinase ciclina-dependente (Serrano, Hannon *et al.*, 1993) e foi o primeiro gene relacionado na patogênese do melanoma maligno (Debniak, Scott *et al.*, 2005). O SNP rs10811661, próximo ao genes *CDKN2A* e *CDKN2B*, mostrou evidência de associação com DM2, (OR igual a 1,2; IC95% 1,14-1,25) (Scott, Mohlke *et al.*, 2007). Uma possível explicação é que o *CDKN2A/B* inibe a atividade do *CD4* e *CD6*.



Em camundongos, a atividade da proteína Cdk4 parece influenciar a proliferação e a massa de células beta pancreáticas (Rane, Dubus *et al.*, 1999; Tsutsui, Hesabi *et al.*, 1999). Além do mais, *CDKN2A* é altamente expresso em ilhotas, adipócitos, cérebro e pâncreas (Scott, Mohlke *et al.*, 2007) e como supressores tumorais, o *CDKN2A* e *CDKN2B* podem desempenhar um papel no envelhecimento (Kim e Sharpless, 2006).

### ***FTO***

A obesidade é um dos principais fatores de risco para DM2. Dessa forma, variações genéticas, que afetem a adiposidade, poderiam influenciar o risco de diabetes. Vários estudos têm encontrado uma associação entre o gene *FTO* (*fat mass and obesity associated*) e o risco de obesidade. Este gene codifica uma proteína beta-hélica dupla-ociosa, homóloga a Fe(II) e 2-oxoglutarato oxigenase (Gerken, Girard *et al.*, 2007). As oxigenases estão envolvidas com diversos processos biológicos, como reparo do DNA, metabolismo dos ácidos graxos e modificações proteicas pós-traducionais, incluindo hidroxilação e demetilação de aminoácidos. Elas utilizam Fe(II) como co-fator e geralmente usam oxigênio. Alguns autores acreditam que alguns polimorfismos promovam diferenças na demetilação de aminoácidos, principalmente no hipotálamo, e que isso é o responsável pelo ganho de massa adiposa, possivelmente por alteração no balanço energético (Gerken, Girard *et al.*, 2007). O *FTO* sofre significativa *upregulation* após a deprivação de comida, confirmando a importância desse gene na homeostase energética (Fredriksson, Hagglund *et al.*, 2008). Além da relação com obesidade, estudos

de GWA têm considerado polimorfismos no gene *FTO* como de risco para DM2. O polimorfismo rs8050136 gera uma das mais fortes evidências de associação além do *TCF7L2* (OR de 1,27; IC95% 1,16-1,37), com significativa replicação em outras populações (OR de 1,22; IC95% 1,12-1,32) (Zeggini, Weedon *et al.*, 2007). Contudo, este efeito parece ser mediado pelo efeito primário da adiposidade, e o ajuste para o IMC (índice de massa corpórea) abole a associação com DM2 (Frayling, Timpson *et al.*, 2007).

## 1.2 O Importante Papel do *TCF7L2*

Grant *et al.* (Grant, Thorleifsson *et al.*, 2006) descreveram uma associação entre variantes genéticas comuns do gene *TCF7L2* (*transcription factor 7-like 2 T-cell-specific, HMG-box*) e o risco de DM2, dado que vem sendo repetidamente confirmado em diferentes populações (Lyssenko, Lupi *et al.*, 2007; Schafer, Tschritter *et al.*, 2007; Zeggini e McCarthy, 2007). Por isso, o *TCF7L2* é considerado o fator de risco genético mais importante do DM2. Foi observado que dois SNPs, localizados no íntron do *TCF7L2* (rs7903146 e rs12255372), afetam a resposta da célula beta pancreática ao GLP-1, durante *clamp* hiperglicêmico (método que permite a capacidade de secreção individual de insulina em resposta à glicose) combinado com a infusão de GLP-1 (Schafer, Tschritter *et al.*, 2007). Estes dados foram corroborados pela comparação dos efeitos do SNP rs7903146 na secreção de insulina, após infusão intravenosa de glicose comparada com

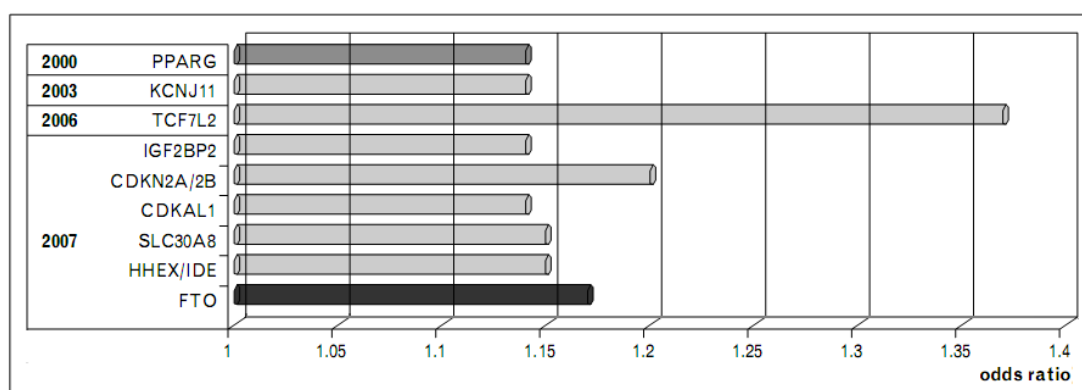
administração oral (Lyssenko, Lupi *et al.*, 2007). Portanto, a possível associação entre polimorfismos do gene *TCF7L2* e diabetes se relaciona a diferentes efeitos de ações dos hormônios incretínicos.

O *TCF7L2* codifica um componente do fator de transcrição do complexo beta-catenina-fator de transcrição *7-like 2*, envolvido na via de sinalização do Wnt (Jin e Liu, 2008). Estudos que utilizaram técnicas de *knockdown*, por interferências no RNA e superexpressão por transfecção demonstraram que, em ilhotas humanas e murinas, o *TCF7L2* é necessário para a sobrevivência e proliferação da célula beta-pancreática, bem como para a secreção de insulina estimulada pela glicose e incretinas (Shu, Sauter *et al.*, 2008). Por fim, a expressão de GLP-1 e GIP, nas ilhotas humanas, dependem da presença do fator de transcrição *7-like 2*, o que explica a relação entre incretinas e o *TCF7L2* (Shu, Matveyenko *et al.*, 2009).

Como citado anteriormente, polimorfismos no gene *TCF7L2* têm sido consistentemente associados com DM2 em diferentes populações (Cauchi, El Achhab *et al.*, 2007), incluindo a população brasileira (Marquezine, Pereira *et al.*, 2008). Além disso, foram encontradas associações entre variantes do *TCF7L2* e outras condições como câncer de próstata, câncer de mama e câncer de cólon (Burwinkel, Shanmugam *et al.*, 2006; Agalliu, Suuriniemi *et al.*, 2008; Folsom, Pankow *et al.*, 2008). Embora a associação entre *TCF7L2* e DCV em indivíduos sem doença coronariana já tenha sido investigada (Bielinski, Pankow *et al.*, 2008), nenhum outro grupo havia estudado essa mesma associação em pacientes sabidamente portadores de Doença Arterial Coronariana (DAC).

### 1.3 Efeitos da Combinação da Informação de Diferentes Variantes Genéticas

A maioria dos polimorfismos descritos anteriormente tem, individualmente, apenas pequenos efeitos sobre o risco de DM2. Para se ter uma ideia, os portadores de variantes de risco do *TCF7L2*, que apresentam os maiores efeitos sobre o risco de DM2 entre os *loci* atualmente conhecidos, possuem um *odds ratio* de 1,3-1,4 por alelo (Figura 2). Considerando a hipótese que a suscetibilidade genética para doenças complexas (como o DM2) é causada pela combinação de múltiplas variantes genéticas, algoritmos de predição de risco construídos com base em informações de um único polimorfismo genético deverão ter valor limitado (Lango, Palmer *et al.*, 2008). No entanto, alguns estudos encontraram que o valor preditivo para o diagnóstico do DM2 pode ser melhorado, através da combinação de múltiplas variantes comuns de risco (Weedon, McCarthy *et al.*, 2006; Cauchi, Meyre *et al.*, 2008; Lango, Palmer *et al.*, 2008; Lyssenko, Jonsson *et al.*, 2008; Meigs, Shrader *et al.*, 2008; Van Hoek, Dehghan *et al.*, 2008; Miyake, Yang *et al.*, 2009), em vez de se utilizar essas variantes isoladamente.



**Figura 2** - Estimativas do tamanho do efeito do risco de DM2, possível mecanismo de risco e ano de descoberta dos polimorfismos estudados. O eixo Y corresponde ao ano de publicação quando cada *loci* genético alcançou nível de significância em estudos de GWA. O eixo X, ao OR estimado. *CDKAL1* - *CDK5 regulatory subunit-associated protein 1-like 1*; *CDKN2* - *cyclin-dependent kinase inhibitor 2A*; *FTO* - *fat mass and obesity associated*; *HHEX* - *haematopoietically expressed homeobox*; *IDE*, *insulin-degrading enzyme*; *IGF2BP2* - *insulin-like growth factor 2mRNA-binding protein 2*; *KCNJ11* - *potassium inwardly rectifying channel, subfamily J, member 11*; *PPARG* - *peroxisome proliferator-activated receptor-g gene*; *SLC30A8* - *solute carrier family 30 (zinc transporter), member 8*; *TCF7L2* - *transcription factor 7 like 2 (T-cell specific, HMG box)*. Barras em cinza claro = implicações na disfunção da célula beta; barra em cinza médio = relação com adiposidade; barra em cinza escuro = relação com peso e IMC

#### 1.4 Racional e justificativa do presente estudo

Uma vez estabelecida a existência e importância relativa de fatores de risco genético na determinação do risco de DM2, surge a necessidade de que estes fatores de risco sejam corretamente validados na população brasileira. Além disso, os fatores de risco genéticos, identificados através de estudos de associação genômicos, provavelmente, estarão associados a

riscos moderados ou mesmo modestos (assim como é, de fato, o caso da grande maioria dos fatores de risco genéticos já caracterizados), ao contrário de valores de risco extremamente altos encontrados nas doenças monogênicas (Pereira, 2008). No entanto, se esse fator de risco for relativamente frequente numa dada população, ele pode responder por uma parcela significativa do risco populacional para aquele fenótipo e ser de relevância no desenvolvimento de algoritmos diagnósticos para essa doença nessa população. Além disso, a hipótese da combinação de pequenos efeitos de múltiplas variantes genéticas de risco, respondendo pelo risco genético das doenças complexas, também se aplica ao DM2, precisa, portanto, ser avaliada. É importante salientar ainda que a utilidade clínica das novas variantes identificadas dependerá de sua capacidade de melhorar a acurácia de algoritmos diagnósticos, baseados em variáveis de risco já conhecidas. Por fim, o entendimento das bases genéticas do desenvolvimento do DM2 facilitará não apenas o diagnóstico precoce da doença ou de suscetibilidade genética ao diabetes, mas, também, pode propiciar abordagens terapêuticas moldadas à suscetibilidade genética em determinado indivíduo, bem como poderá levar à identificação de novas vias para intervenção ou alvos terapêuticos.

Como citado anteriormente, a importância de se prever o risco de um indivíduo de apresentar diabetes está relacionada com a possibilidade do diagnóstico precoce e, dessa forma, prevenção das potenciais complicações relacionadas ao diabetes, em especial, das doenças cardiovasculares. Nesse contexto, o estabelecimento e a confirmação de determinantes

genéticos de risco têm importância potencial, pois explicam uma parte do risco individual de DM2, estão presentes desde o nascimento e são constantes durante toda a vida do indivíduo, isto é, podem ser identificadas em idades mais precoces que a maioria dos fatores de risco ambientais como a obesidade e o sedentarismo. Além disso, as associações entre os *loci* encontrados por meio de estudos que utilizaram o método de GWA têm, por característica básica, a possibilidade de encontrar resultados falso-positivos (Scott, Mohlke *et al.*, 2007) e, conseqüentemente, existe a necessidade de estudos de replicação para a confirmação de resultados em diferentes populações em todo o mundo. Contudo, há uma escassez de estudos em populações sul-americanas. E considerando que essas variantes genéticas de risco para DM2 foram identificadas originalmente em populações caucasianas, não é totalmente adequado extrapolar esses resultados para a miscigenada população brasileira, sendo necessários estudos confirmatórios. Nosso grupo, recentemente, confirmou a associação entre um polimorfismo do *TCF7L2* e DM2 em uma população brasileira do estudo MASS II (Marquezine, Pereira *et al.*, 2008). Isto posto, discute-se repetidamente sobre qual seria a utilidade clínica das novas variantes identificadas e a primeira ideia seria que esses polimorfismos poderiam melhorar a predição diagnóstica para o DM2.

Considerando, ainda, a estreita relação entre DM2 e doenças cardiovasculares, a pergunta que surgiu, em seguida, foi se a informação de SNPs associados ao DM2, em indivíduos com doença cardiovascular estabelecida, pode ser útil, principalmente com relação à predição de eventos

cardiovasculares. Dessa forma, surgiu a ideia de se estudar a relação entre o SNP rs7903146 do *TCF7L2*, a mais importante variante genética associada ao DM2, e doença aterosclerótica coronariana. Em seguida, de acordo com o tópico anterior sobre a combinação das informações genéticas, foi pensada uma análise semelhante, desta vez se estudando vários polimorfismos em conjunto. Para tanto, a utilização de populações, que apresentavam um alto risco cardiovascular, se constituía, a priori, em uma opção adequada na tentativa de responder esses questionamentos. Portanto, foram escolhidas duas populações com essas características: uma população que foi referenciada para cineangiocoronariografia com suspeita de DAC (estudo transversal, denominada população do Laboratório de Hemodinâmica) e outra, oriunda do estudo MASS II, realizado nesta Instituição, que se constituía de indivíduos com DAC multiarterial confirmada com função ventricular preservada e que foram seguidos para eventos cardiovasculares incidentes por um período de 5 anos (longitudinal).

Portanto, a comparação das informações genéticas em diferentes populações, com dados clínicos e laboratoriais disponíveis, bem como com o diagnóstico de diabetes ou doença aterosclerótica, pode permitir profunda análise do impacto destas variantes genéticas na nossa população, além de poder reforçar a noção fisiopatológica disponível na literatura ou até propiciar o surgimento de novos campos de estudos futuros sobre a relação entre variantes genéticas e o processo de aterosclerose.



## **2 Objetivos**

---

O projeto de pesquisa, de que trata esta Tese, tem como objetivo a avaliação da associação entre marcadores de risco genético associados ao Diabetes Mellitus tipo 2 e doença arterial coronariana.

Especificamente, trata-se de:

1. Avaliar a associação entre o gene *TCF7L2* (rs7903146) e tanto prevalência de doença arterial coronariana (em uma população submetida a cateterismo cardíaco) quanto incidência de novos eventos cardiovasculares (em uma população de portadores de DAC multiarterial seguida prospectivamente), em indivíduos diabéticos e em não-diabéticos;
  - 1.1 - Secundariamente, também visa a estudar se a informação do *TCF7L2* é capaz de melhorar uma ferramenta diagnóstica de DM2, baseada em variáveis clínicas;
2. Investigar a associação entre a combinação das informações de variantes genéticas associadas ao DM2 (*PPARG*, *HHEX*, *TCF7L2*, *KCNJ11*, *SLC30A8*, *IGFBP-2*, *CDKAL1*, *CDKN2A/B* e *FTO*) e a frequência de eventos cardiovasculares em pacientes portadores de doença arterial coronária multiarterial durante seguimento de 5 anos, em indivíduos diabéticos e não-diabéticos;
  - 2.1 - Assim como para o *TCF7L2* isoladamente, secundariamente objetiva avaliar se a informação clínica, advinda dos marcadores moleculares estudados em combinação, pode ser utilizada para melhorar algoritmos diagnósticos para diabetes.

## **3 Métodos**

---

O presente estudo utilizou diferentes populações, ferramentas epidemiológicas e métodos de análise para avaliar cada um dos objetivos específicos mencionados anteriormente. Dessa forma, foi optado por descrever a metodologia de cada objetivo específico separadamente.

### **3.1 Objetivo Específico 1**

Avaliar a associação entre o gene TCF7L2 (rs7903146) e tanto prevalência de doença arterial coronariana (em uma população submetida a cateterismo cardíaco) quanto incidência de novos eventos cardiovasculares (em uma população de portadores de DAC multiarterial seguida prospectivamente), em indivíduos diabéticos e em não-diabéticos

Na avaliação do Objetivo 1, foram estudadas duas populações distintas: uma população de indivíduos que foram submetidos à cineangiografia coronariana (cateterismo cardíaco) e outra de indivíduos com DAC multiarterial, oriundos do estudo MASS II (*Medicine, Angioplasty, or Surgery Study*). Para ambas as populações, o diagnóstico de diabetes mellitus foi baseado no tratamento prévio para o diabetes ou de acordo com os critérios da *American Diabetes Association* (ADA) (*Diagnosis and classification of diabetes mellitus, 2008*). Se necessário, foi realizada uma reclassificação para os critérios diagnósticos atuais.

### **3.1.1 População 1 - Indivíduos referenciados para cateterismo cardíaco para diagnóstico de DAC**

Um estudo transversal, que avaliou a gravidade de DAC, foi realizado no Laboratório de Hemodinâmica do Instituto do Coração (InCor) da Faculdade de Medicina da Universidade de São Paulo, Brasil. Todos os pacientes tinham o diagnóstico clínico de angina (suspeita clínica) e nenhum dos pacientes estudados se encontrava em um quadro de Síndrome Coronariana Aguda.

#### *Critérios de Exclusão:*

Os seguintes aspectos foram considerados como critérios de exclusão no estudo:

- Pacientes com eventos cardiovasculares prévios e/ou cateterismo cardíaco prévio, seja diagnóstico ou terapêutico;
- Eventos isquêmicos agudos;
- Insuficiência Cardíaca Classe Funcional III ou IV da NYHA (*New York Heart Association*)
- Hepatopatia crônica
- Hipercolesterolemia familiar;
- Pacientes submetidos a transplantes cardíaco ou renal previamente;
- Pacientes em tratamento com terapia antirretroviral.

De um total de 1.451 pacientes que foram estudados consecutivamente, 167 foram excluídos porque haviam sido submetidos à cirurgia de revascularização do miocárdio (RVM), 140 foram excluídos porque já haviam sido implantados *stents* coronarianos, 78 já haviam recebido angioplastia com balão e 163 por cateterismo cardíaco diagnóstico prévio. Finalmente, 7 indivíduos foram excluídos da análise devido à perda dos dados genotípicos.

As lesões coronarianas foram avaliadas por uma equipe de especialistas que analisou as imagens das angiografias, segundo protocolos padronizados. Os pacientes responderam um questionário clínico a respeito da história médica pregressa, história familiar de DAC, sedentarismo, tabagismo, hipertensão arterial sistêmica (HAS), obesidade, dislipidemia, diabetes e tratamento medicamentoso no momento do procedimento. Durante a entrevista, foram coletados dados antropométricos (peso e altura) e durante o cateterismo cardíaco, uma amostra de sangue periférico foi coletada para extração do DNA genômico.

### **3.1.2 População 2 – Indivíduos com Doença Arterial Coronariana Multiarterial do Estudo MASS II**

O projeto MASS II consiste em um estudo unicêntrico, aleatorizado, prospectivo e multidisciplinar, que compara diferentes opções terapêuticas (tratamento clínico, revascularização cirúrgica do miocárdio e angioplastia transluminal coronária) em pacientes com doença arterial coronária de característica multiarterial, apresentando angina estável e função ventricular

esquerda preservada. Esse projeto vem sendo desenvolvido no Instituto do Coração (InCor) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP), envolvendo a Unidade Clínica de Coronariopatia Crônica, o Serviço de Métodos Gráficos, o Laboratório de Pesquisas Clínicas, o Serviço de Hemodinâmica, a Divisão de Cirurgia e o Laboratório de Genética e Cardiologia Molecular, com aprovação da Comissão de Ética do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo sob o no. 264/94/11.

A população considerada para a inclusão no projeto MASS II compreende pacientes com DAC multiarterial, portadores de angina estável e/ou isquemia miocárdica esforço-induzida, com diagnóstico através de estudo cinecoronariográfico revelador de estenoses coronárias significativas (>70%) em dois ou mais ramos arteriais. As cinecoronariografias foram realizadas em até dois meses antes do procedimento, sendo analisadas por um cardiologista clínico, um cardiologista intervencionista e um cirurgião cardiovascular. Era necessário para a randomização uma tríplice concordância para a indicação de intervenção de revascularização do miocárdio isquêmico, seja por angioplastia transluminal coronária (ATC) seja por RVM (Pereira, 2008).

*Critérios de exclusão:*

Os critérios de exclusão, no estudo MASS II, são citados abaixo:

- Idade menor que 35 ou maior que 75 anos;
- Doenças cardíacas associadas (congenita, valvar, miocardiopatia não-isquêmica, pericárdica);

- Cinecoronariografia de definição insatisfatória;
- FEVE menor que 35% (calculada pela fórmula de Dodge);
- Aneurisma de ventrículo esquerdo;
- Revascularização prévia do miocárdio (RVM ou ATC);
- Angina instável ou Infarto Agudo do Miocárdio (IAM) nas últimas quatro semanas;
- Lesão do tronco da artéria coronária esquerda maior que 50%;
- Extensa calcificação de aorta ascendente ou espasmo coronário primário;
- Acidente Vascular Encefálico ou Ataque Isquêmico Transitório nos últimos seis meses;
- Anatomia coronária inacessível para ATC ou RVM (de acordo com o consenso dos cardiologistas clínico e intervencionista e do cirurgião);
- Expectativa de vida limitada devido à doença grave associada (neoplasia, insuficiência renal ou hepática, doenças hematológicas ou discrasias sanguíneas);
- Gravidez (suspeita, diagnosticada ou planejada);
- Outras morbidades que contraindicassem ATC ou RVM.

### **Definição de critérios e procedimentos de investigação**

Os pacientes, participantes do projeto MASS II, foram aleatorizados para RVM ou ATC e submetidos aos respectivos procedimentos de revascularização do miocárdio isquêmico, com intervalo, quando possível,



de até quatro semanas após a aleatorização. Assim, foram selecionados pacientes de maneira equivalente e consecutiva, de acordo com o número de artérias envolvidas, com a localização das lesões coronárias e ainda indicações semelhantes para ambos os procedimentos de revascularização do miocárdio. A definição do tipo de angina foi considerada de acordo com a classificação funcional da *Canadian Cardiovascular Society* (Campeau, 1976), entendendo-se como portadores de angina refratária aqueles pacientes tratados com terapia anti-isquêmica tripla (betabloqueador e/ou bloqueador de canais de cálcio, nitrato e aspirina) que não obtiveram sucesso terapêutico. O infarto agudo do miocárdio foi definido conforme os seguintes critérios: 1) surgimento de novas ondas Q em, pelo menos, duas derivações eletrocardiográficas contíguas; ou 2) sintomas compatíveis com infarto agudo do miocárdio, associados à elevação da fração MB da enzima creatinofosfoquinase (CPK) maior que três vezes o limite superior ao normal, não sendo considerado diagnóstico para o grupo ATC quando esta elevação ocorria dentro de 96 horas após o procedimento de revascularização (Chaitman, Rosen *et al.*, 1997).

A revascularização do miocárdio anatomicamente completa foi definida, angiograficamente, como a ausência de lesão coronária residual importante (>70%) em todos os ramos epicárdicos considerados passíveis de receber o tratamento através da ATC ou de acordo com o número de anastomoses distais realizadas em todos os vasos considerados doentes na mesma condição (Bourassa, Yeh *et al.*, 1998).

Os eventos cardíacos passaram a ser analisados após a respectiva alta hospitalar dos pacientes sobreviventes aos procedimentos de revascularização do miocárdio. A avaliação clínica foi realizada trimestralmente, durante os cinco primeiros anos após a intervenção, enquanto que a avaliação funcional foi realizada através de teste ergométrico (TE) em esteira, de acordo com o protocolo de Bruce modificado em condições basais, ao final de cada ano de seguimento, até se completar o quinto ano após a intervenção. O teste ergométrico foi considerado ineficaz quando o paciente não conseguiu atingir 5 METs (equivalente metabólico) ou frequência cardíaca máxima de, pelo menos, 80% da frequência esperada. Exceto quando contraindicados, todos os pacientes receberam aspirina, nitratos, inibidores da HMG-CoA redutase (estatinas), inibidores da ECA, betabloqueadores e/ou bloqueadores do canal de cálcio. Os exames de rotina incluíram ECG de repouso, ecocardiograma e rotina laboratorial a cada seis meses. Os pacientes do grupo ATC foram submetidos à angiografia coronária, de rotina, no sexto mês pós-procedimento.

Todos os pacientes foram submetidos à cinecoronariografia em aparelho Philips, modelo Integris 3000, incluindo coronariografia e ventriculografia esquerda pela técnica de Sones ou Seldinger, com registro de, pelo menos, quatro projeções ortogonais para a avaliação da artéria coronária esquerda e duas projeções para a artéria coronária direita, bem como para os enxertos arteriais ou venosos, quando presentes no grupo RVM submetido à cineangiocoronariografia de controle (Sones, 1962).

A fração de ejeção foi calculada pela fórmula proposta por Dodge *et al.* (Dodge, Sandler *et al.*, 1966). Dois observadores experientes quantificaram visualmente as obstruções (considerando significativas aquelas com estenoses de, pelo menos, 70%), bem como a presença de circulação colateral. A classificação das lesões coronárias foi baseada na sua localização, gravidade, extensão e angulação, conforme proposto por Ryan *et al.* (Ryan, Bauman *et al.*, 1993). A doença aterosclerótica coronária multiarterial foi caracterizada como estenose de, pelo menos, 70% em mais de uma artéria coronária ou um de seus ramos, considerando-se apenas vasos com diâmetro luminal superior a 1,5 mm.

A RVM foi realizada, utilizando técnica cirúrgica padronizada, sob hipotermia e uso de solução cardioplégica (Kirklin, Frye *et al.*, 1991). Dessa forma, a estratégia cirúrgica foi determinada pelo cirurgião, com o objetivo de obter revascularização miocárdica completa, utilizando enxertos venosos ou condutos arteriais, sempre que possível. Os pacientes foram submetidos ao procedimento após a definição da estratégia terapêutica e foram medicados com betabloqueadores e medicamentos anti-plaquetários, quando necessário em todo o período perioperatório.

A ATC foi realizada conforme a técnica convencional, incluindo a administração oral de aspirina ou ticlopidina ou clopidogrel nas 24 horas precedentes (Safian, Feldman *et al.*, 2001). Heparina endovenosa (10000 UI) e nitroglicerina foram infundidas imediatamente antes e durante o procedimento. Os cardiologistas intervencionistas foram estimulados a intervir em todas as artérias que pudessem contribuir para a isquemia

miocárdica e/ou demonstrassem obstrução de, pelo menos, 70%. Todas as opções técnicas baseadas em cateteres estavam disponíveis: cateter-balão, *stent*, “laser”, aterectomia direcional e/ou rotacional, de tal modo que a estratégia foi planejada para permitir a revascularização funcional completa, iniciando pelo tratamento da lesão principal, seguida do tratamento dos outros vasos. Quando necessária, a ATC foi planejada e realizada em mais de uma sessão. No caso de oclusão aguda do vaso tratado durante a ATC, o cardiologista intervencionista e o cardiologista clínico tinham a opção de indicar a cirurgia de revascularização do miocárdio de emergência. Em todos os procedimentos percutâneos, havia uma equipe de cirurgia cardiovascular disponível para os casos de emergência. O resultado foi considerado satisfatório quando a obstrução residual foi inferior a 50% ou foi reduzida em, pelo menos, 20% do valor inicial.

Dados clínicos e demográficos (como idade, sexo, história de eventos coronários prévios, de hipertensão, diabetes mellitus, grau de gravidade da angina, tabagismo, peso e altura), laboratoriais (dosagens de glicose sérica em jejum, de colesterol total, HDL, LDL e de triglicérides) e imagenológicos (resultados de angiografias/cateterismo cardíaco) foram coletados antes da aleatorização, bem como amostra de DNA genômico.

Finalmente, 2.076 candidatos, que foram indicados para revascularização do miocárdio, foram avaliados desde maio de 1995 até maio de 2000. Destes, 611 pacientes foram elegíveis e cumpriram todos os critérios de inclusão e foram aleatorizados para um dos três grupos terapêuticos (Hueb, Soares *et al.*, 1999).

### 3.1.3 Construção de um Escore de Gravidade da DAC

A fim de se avaliar a gravidade da DAC, foi criado um escore, em que vinte segmentos coronarianos foram analisados para sua construção. Cada vaso epicárdico ou seu ramo principal foi dividido em 3 segmentos (proximal, medial e distal), exceto ramos secundários da artéria coronária direita, que foram divididos em 2 segmentos: proximal e distal. Esse escore foi definido como a soma da pontuação das lesões coronarianas de cada segmento coronariano analisado, conforme descrito na Tabela 1. Um exemplo da utilização e características desse escore foi publicado previamente (Lanz, Pereira *et al.*, 2005).

**Tabela 1** - Descrição do sistema de pontuação empregado na construção do escore de gravidade das lesões coronarianas

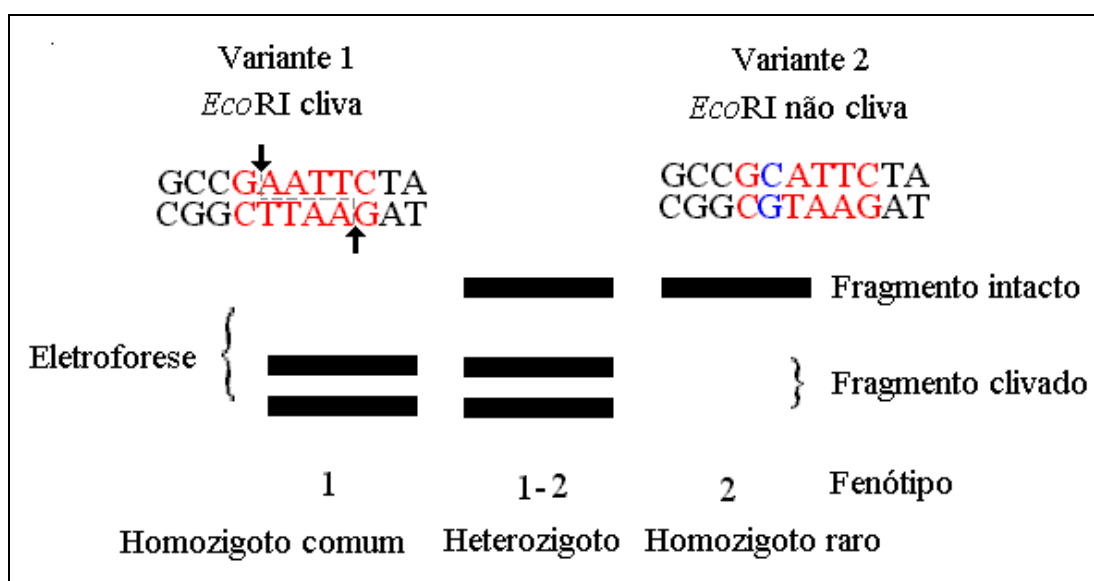
<b>Característica da Lesão</b>	<b>Pontuação</b>
Nenhuma lesão ou irregularidades	0 pontos
Lesão com estenose menor que 50%	0,3 pontos
Lesão com estenose entre 50%-70%	0,6 pontos
Lesão com estenose entre 70%-90%	0,8 pontos
Lesão com estenose entre 90%-100%	0,95

### 3.1.4 Determinação do genótipo do polimorfismo rs7903146 do gene *TCF7L2*

O DNA genômico foi extraído por meio da técnica de *salting-out* dos leucócitos do sangue periférico, colhido em tubo contendo EDTA. Esta técnica consistia em retirar cerca de 8 ml de sangue, que eram transferidos para um tubo de 50ml e o volume completado para 30ml com tampão A (1mM  $\text{NH}_4\text{HCO}_3$  + 144mM  $\text{NH}_4\text{Cl}$ ). A mistura era agitada em vórtex por 20 segundos e incubada a 4°C, por 10 minutos. Em seguida, o material era centrifugado a 4°C por 10 minutos, a 3000 rpm. O sobrenadante era descartado e o *pellet* submetido ao mesmo procedimento descrito acima. O sedimento leucocitário era, finalmente, ressuspensão em 3 mL de tampão B (10 mM Tris-HCl pH 8 + 400 mM NaCl + 2 mM  $\text{Na}_2\text{EDTA}$  pH 8) + 200 µl de SDS 10% + 500 µl de tampão C (50 µl de SDS 10% + 2µl de  $\text{Na}_2\text{EDTA}$  0,5 M pH 8 + 488 mL de água destilada) com proteinase K (2 µl de proteinase K 20 mg/mL diluída em 5 mL de tampão C) e incubado a 37°C por, aproximadamente, 12-18 horas. Posteriormente, ocorria a adição de 1 mL da solução D ( $\text{NaCl}$  6 M), vigorosamente agitada em vórtex por 1 minuto e, em seguida, centrifugada a 4°C por 20 minutos, a 3000 rpm. O sobrenadante era transferido para um tubo de 15 mL e a ele era adicionado 10 mL de etanol absoluto gelado. Após leve agitação, notava-se a precipitação do DNA, que podia ser resgatado e transferido para um eppendorf de 1,5 mL de capacidade, contendo 1 mL de etanol 70% gelado. O eppendorf era, então, centrifugado a 4°C, por 15 minutos, a 13500 rpm.

O etanol era, então, descartado e o sedimento (DNA) posto em temperatura ambiente até ficar totalmente seco. Finalmente, este *pellet* era ressuspenso em 1 ml de TE (10 mM Tris-HCl pH 8 + 1 mM EDTA pH 8). Após a completa diluição, a determinação da concentração do DNA era realizada, diluindo-se 20 µl do material em 980 µl de Tris-EDTA e, em seguida, a mistura era submetida à leitura em espectrofotômetro no comprimento de onda de 260 nm. A concentração final foi corrigida para 100 ng/µl.

Para a reação de polimerização em cadeia (em inglês *Polymerase Chain Reaction* – PCR), foram usados os seguintes *primers*: 5' ACA ATT AGA GAG CTA AGC ACT TTT TAG GTA 3' e 5' GTG AAG TGC CCA AGC TTC TC 3'. O polimorfismo em estudo foi detectado pelo método de PCR-RFLP (em inglês *polymerase chain reaction-restriction fragment length polymorphism assay*). Resumidamente, o racional desta técnica é a seguinte: endonucleases de restrição clivam o DNA em locais de reconhecimento específico. Quando ocorre uma alteração em um único par de bases, dentro de um local de reconhecimento pela enzima, a sequência do DNA já não é mais clivada pela endonuclease de restrição, enquanto que um sítio intacto em outro cromossomo pode ser clivado (Figura 3). Em outras palavras, um alelo tem um local de reconhecimento funcional para uma endonuclease de restrição particular e o outro não. Como consequência desta diferença alélica, fragmentos de DNA de tamanhos diferentes são produzidos, quando o DNA é digerido por uma endonuclease de restrição específica (Pasternack, 2005).



**Figura 3** - Esquema de determinação do genótipo por meio da técnica de RFLP usando, como exemplo, a enzima *EcoRI*. SNPs podem criar ou abolir regiões de reconhecimento pelas endonucleases de restrição, alterando a quantidade e o tamanho dos fragmentos de DNA resultantes da digestão pela enzima

Esta técnica consistia em realizar uma PCR de 30 ciclos, em uma máquina de PTC-DNA Tetrad2, usando uma solução reagente de 10 µL, contendo 10mM de Tris-HCl (pH 9.0), 50mM de KCl, 2.5 mM de MgCl<sub>2</sub>, 100µM de cada dNTP, 0.3U de *Easy Taq* DNA Polymerase, 5 pmol de cada *primer* e 1µL de amostra do DNA genômico. Os produtos da PCR eram digeridos por 1U da enzima de restrição *RsaI* e visualizados por eletroforese em gel de agarose a 3%. O controle de qualidade desses ensaios era realizado pela avaliação randômica de 40 amostras, selecionadas e reagentipadas por um técnico independente. A taxa de concordância, observada entre as genotipagens, foi de 100%.



### 3.1.5 Análise Estatística

A análise de equilíbrio de Hardy-Weinberg para distribuição dos genótipos foi estimada por teste de qui-quadrado. Testes de qui-quadrado, testes *t* e análise de variância foram usados para comparações das características basais. O teste de Kolmogorov-Smirnov foi realizado previamente para cada variável numérica, para se determinar se tal variável apresentava distribuição normal, e, então, foi realizado o teste estatístico adequado. Os *odds ratios* para os diferentes modelos de associação foram calculados com intervalo de confiança (IC) de 95% por regressão logística múltipla pelo método condicional retrógrado para um nível de significância menor que 0,5. Comparações entre curvas prospectivas de sobrevivência livre de eventos foram realizadas tanto para todos os indivíduos quanto para subgrupos de pacientes diabéticos e não-diabéticos em relação ao genótipo do *TCF7L2*. Análise de regressão logística foi usada para estimar a associação transversal dessas variáveis com cada desfecho do estudo MASS II (morte, IAM e isquemia recorrente que requereu revascularização), como também com o desfecho combinado. As curvas de sobrevida livre de eventos foram construídas pelo método de Kaplan-Meier e as diferenças entre as curvas foram calculadas pelo método de *Log-rank*. Na análise prospectiva, a relação entre as variáveis na linha de base e o desfecho composto foi avaliada pelo modelo de sobrevivência proporcional de Cox.

Para se responder ao questionamento levantado pelo objetivo secundário deste tópico, foi utilizado um modelo construído e validado

por nosso grupo em populações urbanas brasileiras (Pires De Sousa, Pereira *et al.*, 2009), utilizando as diretrizes da OMS, projeto MONICA (Tunstall-Pedoe, Kuulasmaa *et al.*, 1994). Nesse caso, fatores de risco clínicos, sabidamente associados com diabetes (idade, sexo, diagnóstico de hipertensão arterial e atividade física), foram selecionados para a construção de um escore diagnóstico. As análises laboratoriais incluíram níveis de colesterol total, de LDL, HDL, triglicérides e ácido úrico. Análises univariadas foram realizadas por meio de regressão logística, assumindo o diagnóstico de DM2 como a variável dependente. Na análise de regressão logística múltipla, foram usadas variáveis que preenchiam o critério de inclusão (valores de  $p$  menores que 0,1). Essas variáveis foram incluídas usando um modelo de eliminação do tipo *stepwise backward*, novamente assumindo DM2 como variável dependente. Os coeficientes de regressão logística foram utilizados para estimar *odds ratio* para cada uma das variáveis independentes no modelo. As variáveis contínuas foram categorizadas em faixas. Para a construção do modelo final, apenas as variáveis com um valor de  $p$  menor que 0,05 foram considerados. Para cada uma dessas variáveis significativas na análise de regressão logística múltipla, foi calculada uma pontuação, multiplicando-se o coeficiente de regressão por 10 e arredondando-se para o número inteiro mais próximo. Finalmente, o escore final para cada participante foi calculado pela somatória da pontuação para cada variável significativa no modelo de risco. Para a criação do escore com informações genéticas, o procedimento foi o mesmo. A fim de se avaliar o desempenho do modelo, uma curva ROC (*receiver operating characteristic*) foi construída e a área sob a curva ROC (AUC) foi

utilizada para medir a acurácia para o diagnóstico de DM2. Este procedimento é uma forma útil para se avaliar o desempenho dos sistemas de classificação em que há uma variável com duas categorias, pelas quais os sujeitos são classificados. A estimativa da área sob a curva ROC foi calculada ou não-parametricamente ou parametricamente, utilizando um modelo exponencial binegativo. As curvas ROC foram plotadas para estimativa do risco de diabetes pelo escore, em que a sensibilidade foi plotada no eixo-y e a taxa de falsos positivos (1-especificidade) no eixo-x. Quanto mais discriminativo for o modelo em classificar corretamente o indivíduo, maior a inclinação da porção ascendente da curva ROC e maior a área sob a curva (AUC). As áreas sob as curvas ROC (AUC) para o escore clínico e os escores genotípicos adicionados à pontuação clínica foram comparados usando o *software* ROCKIT ([www.radiology.uchicago.edu/kr1/KRL\\_ROC/software\\_index6.htm](http://www.radiology.uchicago.edu/kr1/KRL_ROC/software_index6.htm)). Um valor de p menor que 0,05 foi considerado significativo para comparações. O programa SPSS versão 13 para windows foi usado para análise estatística.

### **3.2 Objetivo Específico 2**

*Investigar a associação entre a combinação das informações de variantes genéticas associadas ao DM2 (PPARG, HHEX, TCF7L2, KCNJ11, SLC30A8, IGFBP-2, CDKAL1, CDKN2A/B e FTO) e a frequência de eventos cardiovasculares em pacientes portadores de doença arterial coronária multiarterial durante seguimento de 5 anos, em indivíduos diabéticos e não-diabéticos*

Para fins de se avaliar se existe associação entre eventos cardiovasculares e a combinação das informações das variantes de risco, foi estudada a população do projeto MASS II. Detalhes das características do estudo MASS II já foram descritos na seção 3.1.2. Para esses indivíduos, o diagnóstico de diabetes mellitus também foi baseado no tratamento prévio para o diabetes ou de acordo com os critérios atuais da ADA (Diagnosis and classification of diabetes mellitus, 2008).

### 3.2.1 Seleção de SNPs, alelos de risco e genotipagem

Uma recente metanálise (Scott, Mohlke *et al.*, 2007) confirmou consistentemente a associação entre DM2 e 10 SNPs, que foram selecionados para este estudo. A descrição dos SNPs, bem como o cromossomo em que estão localizados e o alelo de risco podem ser observados na tabela abaixo.

**Tabela 2** - Variantes genéticas associadas ao DM2: SNPs, localização cromossômica e alelos de risco

SNP	Posição	Gene	Cromossomo	Alelo Risco / Alelo sem risco
rs5219	17.366.148	<i>KCNJ11</i>	11	T/C
rs1801282	12.368.125	<i>PPARG</i>	3	C/G
rs7903146	114.748.339	<i>TCF7L2</i>	10	T/C
rs13266634	118.253.964	<i>SLC30A8</i>	8	C/T
rs1111875	94.452.862	<i>HHEX</i>	10	G/A
rs7754840	20.769.229	<i>CDKAL1</i>	6	C/G
rs4402960	186.994.389	<i>IGF2BP2</i>	3	T/G
rs10811661	22.124.094	<i>CDKN2A/B</i>	9	T/C
rs9300039	41.871.942	<i>CDKN2A/B</i>	11	C/A
rs8050136	52.373.776	<i>FTO</i>	16	A/C

O DNA genômico foi extraído de leucócitos do sangue periférico por meio da técnica de *salting-out* padrão, conforme descrito no item 3.1.4. As genotipagens do polimorfismo rs7903146 do *TCF7L2* foram realizadas pela técnica de PCR-RFLP, também descritas nessa mesma seção. Para os demais polimorfismos, as genotipagens foram executadas como ensaios de PCR baseados na técnica de *Array Tape*. Diferentemente de outras técnicas, este método não utiliza placas de microtitulação. Dessa forma, são capazes de oferecer maior rendimento nas genotipagens. *Array Tape* (fita matriz) é uma fita plástica contínua que é usada em conjunto com uma configuração flexível de dosificação, pipetagem, impermeabilização e módulos de detecção. Isto permite a transformação automática contínua de centenas de equivalentes de microplacas, sem a exigência de robôs ou equipamentos complexos de manipulação de placas. Além da capacidade de transferência, uma outra vantagem desse método é a economia de custos de materiais, reagentes e mão-de-obra. Dessa forma, esse método é especialmente adequado para processos baseados em PCR. A ciclagem térmica é feita em fita matriz, selada em imersão aquosa no termociclador. Isso permite o processamento simultâneo de centenas de equivalentes de microplaca em uma hora, e fornece ciclagem rápida de temperatura em comparação a métodos baseados em placas. A fita matriz é produzida em polipropileno 0,3 mm de espessura, e a transferência de calor, através deste material, é muito mais rápida que através de microplacas comuns. Durante cada ciclo, as amostras em banho alcançam a temperatura ambiente mais rapidamente

e permanecem uma porcentagem maior de tempo na temperatura-alvo do termociclador. Isto permite tempos de ciclos mais curtos e devem ajudar na redução da ligação não específica. Foi utilizada uma técnica de PCR alelo-específica modificada, conforme descrito por Myakishev *et al.* (Myakishev, Khripin *et al.*, 2001). O protocolo TaqMan é basicamente semelhante ao protocolo de PCR alelo-específico modificado e seguiu os seguintes procedimentos: 5 ng da amostra de DNA molde são pipetados na placa e se espera secar. Então, adicionam-se 800 nL de uma solução de reação, contendo 1 tampão de PCR, 100 mM de dNTPs, 0,05 unidades/ $\mu$ L de Taq Platinum®, 1,5 mM  $Mg^{+2}$ , solução corante de referência de Rox a 1% (Invitrogen) e 0,15 primers e sondas TaqMan (primers: 0,134 mM, sondas: 0,030 mM) ao DNA seco na fita. As amostras são amplificadas em um banho de água no termociclador, de acordo com o seguinte: 95°C por 3 minutos; em seguida, 35 ciclos de 95°C por 20 segundos, 60°C por 1 minuto, com uma incubação final a 60°C por 6 minutos. As amostras são, então, passadas em módulos de secagem, onde todo o líquido é evaporado, a partir das amostras de DNA. Após a sequência de ciclos térmicos, as matrizes são transferidas para um módulo de detecção e a leitura é realizada e os dados processados por um software específico. Os dados dos genótipos foram obtidos em mais de 90% das amostras de DNA utilizadas e amostras idênticas de controle (32 amostras) foram genotipados com 100% de concordância.

### 3.2.2 Construção dos Escores Genotípicos

A fim de se avaliar a utilização das informações de todos os dez SNPs combinados, foram criados três escores genéticos ligeiramente diferentes. Para o primeiro, foram considerados o número de genes para os quais o indivíduo era homozigoto para o alelo de risco (Escore 1), isto é, considerando uma distribuição genotípica hipotético de XX (homozigoto para o alelo de proteção), XY (heterozigoto) e YY (homozigoto para o alelo de risco), os indivíduos ganhavam um ponto, quando eram YY, em relação a cada polimorfismo. O Escore 2 corresponde ao número total de alelos de risco. Para cada polimorfismo estudado, se o indivíduo fosse heterozigoto para o alelo de risco (XY), ele recebia um ponto; se fosse homozigoto (YY), dois pontos. No modelo aditivo, os valores de *odds ratio* são apresentados para cada alelo de risco. Dessa forma, caso o indivíduo seja homozigoto para um determinado alelo de risco, o *odds ratio* deve ser, por dedução matemática, duplicado. Finalmente, para criar o terceiro escore, um modelo de regressão logística foi criado utilizando-se o diagnóstico de DM2 como variável dependente. Cada SNP no seu modelo de melhor desempenho na análise univariada (modelos aditivo, dominante ou recessivo) foi incluído como variável independente. O modelo dominante foi definido como YY + XY *versus* XX, e o modelo recessivo como YY *versus* XX + XY. Os valores dos coeficientes de regressão para cada SNP foram multiplicados por 10 e arredondados para o número inteiro mais próximo (Tabela 3). O Escore 3 foi calculado como a soma destes valores. Foi testado um modelo dominante,

cujos resultados foram inferiores aos demais e, portanto, não foi incluído. Um escore clínico, que incorporou fatores de risco conhecidos como sexo, idade, índice de massa corporal (IMC), hipertensão arterial e tabagismo, foi construído separadamente, à semelhança do Escore 3 (Tabela 4) e o valor incremental da adição da informação genética combinada ao escore clínico de fatores de risco também foi avaliada.

### 3.2.3 Análise Estatística

As análises estatísticas foram realizadas com o uso do *software* SPSS, versão 13.0. O teste de Kolmogorov-Smirnov foi realizado previamente para cada variável numérica para se determinar se tal variável apresentava distribuição normal, e, então, foi realizado o teste estatístico adequado. Testes de qui-quadrado, testes *t* e análise de variância foram utilizados para comparações das características dos grupos no momento da aleatorização (linha de base). Quando foram realizados múltiplos testes, um valor de *p* mais restrito foi considerado, baseado na correção de Bonferroni (divisão de *p* pelo número de testes, isto é,  $p = 0,05$  dividido por 10 testes, o valor de *p* para se considerar a exclusão da hipótese nula seria de 0,005). A fim de se avaliar o desempenho de modelos, curvas ROC foram construídas e as áreas sob as curvas ROC (AUC) foram usadas para se aferir o poder de discriminação do modelo. Comparações de curvas de sobrevivência em longo prazo foram realizadas para o grupo inteiro de estudo e para os subgrupos de pacientes diabéticos e não-diabéticos na



linha de base. Um modelo de regressão de Cox foi utilizado para estimar a associação prospectiva entre as variáveis e cada um dos desfechos finais (morte, infarto do miocárdio e isquemia recorrente necessitando revascularização), bem como o desfecho combinado após 5 anos de seguimento. As curvas de sobrevida foram calculadas com o método de Kaplan-Meier e as diferenças entre as curvas foram comparadas pelo método estatístico *Log-rank*. Nesse estudo, o diagnóstico de DM2 foi utilizado como variável categórica (sim ou não) e as pontuações de cada escore como variáveis independentes para construção da curva. Para fins de análise estatística em que o diagnóstico de DM2 foi usado como variável dependente, foram considerados diabéticos os indivíduos que possuíam o diagnóstico de diabetes no momento da aleatorização (linha de base) e aqueles que se tornaram diabéticos no decorrer do seguimento de 5 anos.

**Tabela 3** - Variáveis de risco, análise de regressão logística múltipla e pontuação atribuídas às variantes genéticas para o Escore 3 e Escore 3 sem *TCF7L2*

Genes	SNPs	Modelo	Variantes	Escore 3		Escore 3 sem <i>TCF7L2</i>	
				Coefficiente $\beta$	Pontuação ( $\beta \times 10$ )	Coefficiente $\beta$	Pontuação ( $\beta \times 10$ )
<i>KCNJ11</i>	rs5219	Recessivo	CT	0,210	0	0,260	0
			TT		2		3
<i>PPARG</i>	rs1801282	Aditivo	GC	0,431	4	0,458	5
			CC		8		10
<i>TCF7L2</i>	rs7903146	Aditivo	CT	0,382*	4	-	-
			TT		8		-
<i>SLC30A8</i>	rs13266634	Recessivo	TC	0,193	0	0,312	0
			CC		2		3
<i>HHEX</i>	rs1111875	Recessivo	AG	0,332	0	0,203	2
			GG		3		2
<i>CDKAL1</i>	rs7754840	Dominante	GC	0,170	2	0,247	2
			CC		2		4
<i>IGF2BP2</i>	rs4402960	Aditivo	GT	0,271	3	0,324	0
			TT		6		3
<i>CDKN2A/B</i>	rs10811661	Recessivo	CT	0,321	0	0,179	0
			TT		3		2
<i>CDKN2A/B</i>	rs9300039	Recessivo	AC	0,148	0	0,118	1
			CC		1		1
<i>FTO</i>	rs8050136	Dominante	CA	0,086	1	0,260	0
			AA		1		3

(\*) o valor de p para esta análise foi 0,034. Para todas as demais p > 0,1.

**Tabela 4** - Variáveis de risco, análise de regressão logística múltipla e pontuação atribuídas às variáveis para construção do Escore Clínico e escores derivados, na população do estudo MASS II

Variável	Grupos	Coefficiente $\beta$	Pontuação ( $\beta \times 10$ )
<b>Para Escores 1 e 2</b>			
Sexo	Masculino	0,066	1
Tabagismo	Tabagismo (+)	0,170	2
Hipertensão	Hipertensão (+)	0,533	5
Obesidade	IMC $\geq 30$ kg/m <sup>2</sup>	0,264	3
Idade	55-65 anos	0,347	3
	$\geq 65$ anos	0,250	3
Escore 1	4	0,344	3
	$\geq 5$	0,824	8
Escore 2	12-13	0,231	3
	$\geq 14$	0,752	8
<b>Para Escore 3 e Escore 3 sem TCF7L2</b>			
Sexo	Masculino	0,046	0
Tabagismo	Tabagismo (+)	0,187	2
Hipertensão	Hipertensão (+)	0,543	5
Obesidade	IMC $\geq 30$ kg/m <sup>2</sup>	0,315	3
Idade	55-65 anos	0,447	4
	$\geq 65$ anos	0,348	4
Escore 3	19-22	0,422	4
	$\geq 23$	1,108	11
Escore 3 sem TCF7L2	17-19	0,696	7
	$> 20$	0,810	8

## **4 Resultados**

---

Da mesma forma que foi feito na seção Casuística e Métodos, os resultados do presente estudo serão apresentados de acordo com os objetivos específicos.

#### **4.1 Objetivo Específico 1**

Avaliar a associação entre o gene *TCF7L2* (rs7903146) e tanto prevalência de doença arterial coronariana (em uma população submetida a cateterismo cardíaco) quanto incidência de novos eventos cardiovasculares (em uma população de portadores de DAC multiarterial seguida prospectivamente), em indivíduos diabéticos e em não-diabéticos

##### **4.1.1 Avaliação da Prevalência e Gravidade das Lesões Coronarianas na População Submetida a Cateterismo Cardíaco**

Inicialmente, foi estudada uma população com suspeita diagnóstica de DAC e que havia sido encaminhada para cateterismo cardíaco no Laboratório de Hemodinâmica do InCor, Brasil. Após critérios de exclusão, 896 pacientes foram selecionados e submetidos, pela primeira vez, à cineangiografia coronariana. Esta população do estudo tinha média de idade de 59,5 anos e 52,4% dos indivíduos eram do sexo masculino. As prevalências de

hipertensão, diabetes, obesidade (IMC > 30,0 kg/m<sup>2</sup>), tabagismo, insuficiência cardíaca (ICC) classe I-II, doença arterial periférica (DAP) e insuficiência renal crônica (IRC) foram de 73%, 32,1%, 28%, 37,1%, 1,8%, 9,2% e 2,7%, respectivamente. Os valores médios de glicemia em jejum, colesterol total, LDL-colesterol, HDL-colesterol e triglicerídeos foram 130,2 mg/dl, 229,0 mg/dl, 145,0 mg/dl, 44,8 mg/dl e 181,4 mg/dl, respectivamente.

Em toda a população, as frequências do polimorfismo rs7903146 foram: CC = 46,8% (n = 419); CT = 42,6% (n = 382); TT = 10,6% (n = 95), caracterizando que tal polimorfismo estava em equilíbrio de Hardy-Weinberg. Para os genótipos CC, CT e TT, respectivamente, a prevalência de diabetes foi de 32,2%, 30,9% e 36,8%. Não houve diferença nas frequências genóticas em relação à idade, ao sexo, ao tabagismo, ao IMC, ao colesterol total, ao HDL e ao LDL-colesterol, aos triglicerídeos e à pressão arterial sistólica e diastólica, DAP, ICC e IRC.

Apesar da suspeita clínica, 267 pacientes não apresentaram lesões identificadas pela angiografia coronariana. Como esperado, os pacientes diabéticos apresentaram maior prevalência de lesões coronarianas obstrutivas (cerca de 80%). Contudo, não foi observada relação entre os genótipos do *TCF7L2* e diabetes ( $p = 0,93$ ).

Embora não houvesse diferença na prevalência de lesões coronárias entre os três genótipos do *TCF7L2* em toda a população, indivíduos não-diabéticos que carregavam o alelo T, isto é, indivíduos não-diabéticos com genótipos CT e TT (que a partir de então foram agrupados nas análises posteriores) apresentaram uma frequência significativamente

maior de lesões coronárias em comparação aos indivíduos não-diabéticos com o genótipo CC. A análise alélica também revelou uma associação significativa entre o alelo T e lesões coronárias neste subgrupo (OR = 1,37 por alelo, IC95% = 1,05-1,78, valor de  $p = 0,025$ ). As características demográficas e laboratoriais da população de indivíduos não-diabéticos e diabéticos, de acordo também com os grupos genotípicos, são apresentadas na Tabela 5.

**Tabela 5** - Características demográficas e laboratoriais de indivíduos não-diabéticos e diabéticos da população submetida a cateterismo cardíaco, de acordo com o genótipo do *TCF7L2*

	Indivíduos não-diabéticos			Indivíduos diabéticos		
	CC (n = 284)	CT + TT (n = 324)	Valor de p*	CC (n = 135)	CT+TT (n = 153)	Valor de p*
Idade (anos)	58,4 ± 11,1	59,4 ± 10,4	0,229	61,0 ± 9,6	60,3 ± 9,4	0,507
Sexo Masculino (%)	53,5	57,1	0,376	45,9	46,4	0,935
Tabagismo (%)	35,6	41,7	0,123	29,6	36,6	0,210
Sedentarismo (%)	94,0	90,1	0,214	94,8	93,5	0,303
IMC (kg/m <sup>2</sup> )	27,8 ± 5,3	27,3 ± 4,6	0,087	28,9 ± 5,5	28,6 ± 5,4	0,672
HAS (%)	70,1	65,7	0,254	80,7	86,3	0,205
ICC (%)	10,2	8,6	0,508	6,7	15,7	0,016
IRC (%)	1,8	1,5	0,770	5,2	3,9	0,755
Glicemia (mg/dL)	98,1 ± 9,1	97,8 ± 8,4	0,288	172,1 ± 69,1	177,7 ± 73,2	0,522
Colest. total (mg/dL)	229,1 ± 47,3	230,3 ± 53,3	0,510	231,2 ± 47,4	223,9 ± 47,0	0,248
LDL-c (mg/dL)	145,4 ± 38,5	147,1 ± 43,1	0,698	154,0 ± 42,3	132,4 ± 41,2	0,002
HDL-c (mg/dL)	46,0 ± 13,4	44,7 ± 12,9	0,577	42,9 ± 10,4	44,6 ± 13,0	0,394
Triglicerídeos (mg/dL)	164,7 ± 104,5	163,3 ± 102,5	0,831	223,4 ± 192,5	210,2 ± 188,1	0,602

(\*) Comparação entre os grupos genotípicos para indivíduos diabéticos e não-diabéticos, calculados por qui-quadrado para variáveis categóricas e teste *t* de *student* para variáveis contínuas; IMC – Índice de Massa Corporal; HAS – Hipertensão Arterial Sistêmica; ICC – Insuficiência Cardíaca Congestiva; IRC – Insuficiência Renal Crônica.



Os indivíduos não-diabéticos CT/TT tinham quase 2,5 vezes mais chance de ter lesões obstrutivas coronárias que o grupo CC (Tabela 6). Esta chance aumentada de ter lesões ateroscleróticas ainda estava presente mesmo após ajuste para idade, sexo, IMC, tabagismo, hipertensão, glicemia, colesterol total, LDL, HDL e triglicerídeos e o uso de estatinas, aspirina e betabloqueadores.

**Tabela 6** - Prevalência e estimativa de risco de lesões coronarianas, de acordo com o genótipo do *TCF7L2*, na população submetida a cateterismo cardíaco

<b>Indivíduos Diabéticos</b>				
	<b>Prevalência (n)</b>	<b>OR (IC95%)</b>	<b>OR Ajustado*</b>	<b>Valor de p†</b>
CC	80,7% (109)	-	-	-
CT+TT	79,1% (121)	0,902 (0,51-1,61)	-	0,727
<b>Indivíduos Não-diabéticos</b>				
	<b>Prevalência (n)</b>	<b>OR (IC95%)</b>	<b>OR Ajustado*</b>	<b>Valor de p†</b>
CC	60,6% (172)	-	-	-
CT+TT	70,1% (227)	1,52 (1,09-2,13)	2,32 (1,27-4,24)	0,006
<b>Todos os Indivíduos</b>				
	<b>Prevalência (n)</b>	<b>OR (IC95%)</b>	<b>OR Ajustado*</b>	<b>Valor de p†</b>
CC	67,1% (281)	-	-	-
CT+TT	73,0% (348)	1,325 (0,99-1,76)	1,47 (0,91-2,36)	0,112

(\*) Ajustado para sexo, IMC, tabagismo, hipertensão, glicemia, colesterol total, LDL-c, HDL-c, triglicerídeos. (†) Valor de p para o OR ajustado.

Em relação ao escore de gravidade das lesões coronarianas, os indivíduos não-diabéticos que carregavam o alelo T tinham uma prevalência significativamente maior de DAC triarterial do que pacientes com o genótipo

CC (Tabela 7). Além disso, considerando o escore de gravidade dividido em tercís, os pacientes do grupo CT/TT também tinham uma proporção maior de indivíduos no tercil mais elevado (Tabela 8). No entanto, após ajustes para múltiplos fatores de risco, estas diferenças entre os tercís do escore de gravidade da DAC não alcançaram significância estatística. A análise de regressão linear demonstrou que a presença de um alelo T aumentou o valor da pontuação do escore de gravidade em 1,4 pontos ( $p = 0,038$ ).

**Tabela 7** - Distribuição de lesões coronárias em pacientes não-diabéticos com DAC, de acordo com os grupos genotípicos, na população submetida a cateterismo cardíaco

	CC % (n)	CT + TT % (n)	Valor de p	OR*	OR Ajustado†	Valor de p‡
Normal	43,7 (124)	34,0 (110)	0,015			
Uniarterial	20,8 (59)	27,8 (90)		1,72 (1,13-2,61)	2,09 (1,08-4,04)	0,029
Biarterial	19,4 (55)	16,0 (52)		1,07 (0,67-1,68)	1,47 (0,75-2,90)	0,262
Triarterial	16,2 (46)	22,2 (72)		1,76 (1,12-2,77)	2,06 (1,00-4,25)	0,050

(\*) *Odds Ratio* calculado conforme comparação entre DAC triarterial, biarterial e uniarterial versus coronárias normais por meio de regressão logística múltipla;

(†) Ajustado para sexo, IMC, tabagismo, hipertensão, glicemia, colesterol total, LDL-c, HDL-c, triglicerídeos e uso de estatinas, aspirina e betabloqueadores; (‡) Valor de p para OR ajustado.

**Tabela 8** - Distribuição dos tercís do escore de gravidade das lesões coronarianas em pacientes não-diabéticos, de acordo com os grupos de genótipos do *TCF7L2*, na população submetida a cateterismo cardíaco

Tercil (valor)	CC % (n)	CT + TT % (n)	Valor de p	OR* (IC95%)	OR Ajustado† (IC95%)	Valor de p (‡)
1º Tercil (≤ 0,9)	38,7 (110)	28,4 (92)	0,025			
2º Tercil (0,91-2,2)	29,9 (85)	36,1 (117)		1,65 (1,11-2,44)	1,004 (0,53-1,91)	0,99
3º Tercil (≥ 2,2)	31,3 (89)	35,5 (115)		1,545 (1,04-2,28)	1,31 (0,69-2,47)	0,40

(\*) *Odds Ratio* calculado conforme comparação entre tercís do escore por meio de regressão logística múltipla; (†) Ajustado para sexo, IMC, tabagismo, hipertensão, glicemia, colesterol total, LDL-c, HDL-c, triglicerídeos e uso de estatinas, aspirina e betabloqueadores; (‡) Valor de p para OR ajustado

Portanto, esses dados sugerem que os indivíduos não-diabéticos portadores do alelo de risco T do *TCF7L2*, quando comparados aos não-carreadores, apresentam tanto uma maior prevalência de lesões coronarianas obstrutivas quanto uma maior gravidade dessas lesões (de acordo com a proporção de artérias acometidas e a pontuação do escore de gravidade).

#### 4.1.2 Incidência de Eventos Cardiovasculares na População do Estudo MASS II

A fim de se avaliar se as variantes do gene *TCF7L2* também estão associadas com eventos cardiovasculares ocorridos durante 5 anos após a randomização em uma população de indivíduos com DAC estabelecida, foram analisados os pacientes do estudo MASS II. Esta população de

estudo incluiu 559 indivíduos elegíveis, dos quais 69,2% eram do sexo masculino, com uma média de idade de 59,7 anos (semelhante à população do Laboratório de Hemodinâmica) e IMC médio de 27,1 kg/m<sup>2</sup>. A prevalência de HAS foi de 59,6%, de diabetes 30,9% e de tabagismo 33,9%. O polimorfismo rs7903146 apresentava a seguinte distribuição de genótipos: CC = 159 (28,4%), CT = 336 (60%), e TT = 65 (11,6%). As características dos indivíduos dessa população, de acordo com o diagnóstico de diabetes e subdivididos, do mesmo modo que anteriormente, de acordo com a presença ou não do alelo de risco T, são apresentadas na Tabela 9. Observa-se que o grupo CT/TT apresentava uma maior prevalência de indivíduos com diagnóstico prévio de IAM. Contudo, isso não influenciou os resultados porque foram realizados ajustes para essa variável.

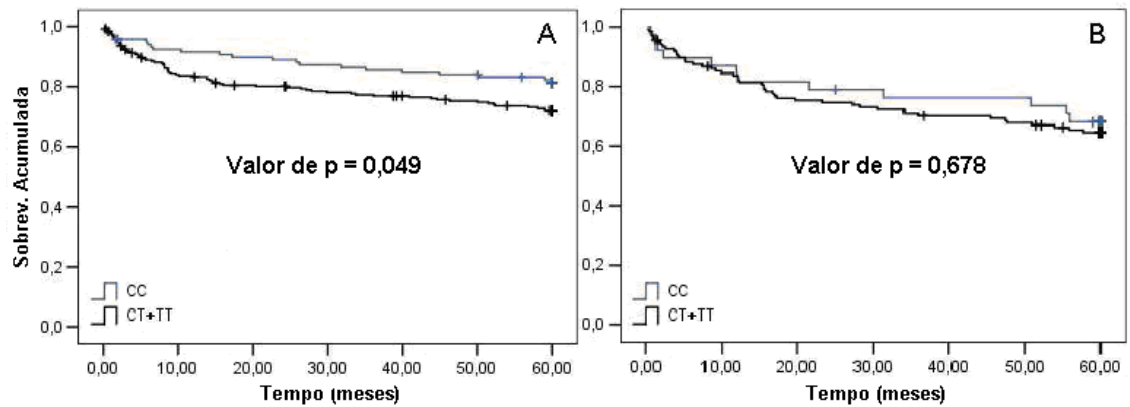
**Tabela 9** - Características demográficas e laboratoriais da população do estudo MASS-II, de acordo com diagnóstico de diabetes e grupos de genótipos do *TCF7L2*

	Indivíduos Não-diabéticos			Indivíduos Diabéticos		
	CC (n = 120)	CT + TT (n = 266)	Valor de p*	CC (n = 38)	CT + TT (n = 135)	Valor de p*
Idade (anos)	59,6 ± 9,2	59,1 ± 9,4	0,836	62,7 ± 6,4	59,8 ± 9,1	0,071
Sexo Masculino (%)	68,3	71,4	0,537	60,5	65,2	0,280
Tabagismo (%)	34,2	36,5	0,663	31,6	25,9	0,479
IMC (kg/m <sup>2</sup> )	27,6 ± 3,9	26,4 ± 4,0	0,966	27,9 ± 4,4	28,0 ± 4,6	0,880
HAS (%)	56,7	53,8	0,595	73,7	66,7	0,673
IAM prévio (%)	39,2	50,0	0,048†	26,3	45,9	0,030
Glicemia (mg/dL)	99,1 ± 9,4	102,4 ± 11,1	0,551	161,3 ± 63,6	183,0 ± 79,7	
Colest. Total (mg/dL)	220,5 ± 44,4	225,9 ± 48,3	0,409	221,0 ± 49,6	219,2 ± 50,6	0,848
LDL (mg/dL)	147,3 ± 38,2	149,9 ± 44,7	0,349	144,7 ± 41,9	141,2 ± 46,8	0,688
HDL (mg/dL)	37,5 ± 9,4	37,0 ± 10,7	0,783	37,2 ± 11,9	38,1 ± 10,6	0,673
Triglicerídeos (mg/dL)	186,2 ± 109,9	185,0 ± 94,6	0,191	213,3 ± 156,2	205,4 ± 141,5	0,765

(\*) Comparação entre os grupos genotípicos para indivíduos diabéticos e não-diabéticos, calculados por qui-quadrado para variáveis categóricas e teste *t* de *student* para variáveis contínuas; IMC – Índice de Massa Corporal; HAS – hipertensão arterial sistêmica.

A Figura 4 corresponde a uma curva de sobrevivência de Kaplan-Meier. Analisando-a, observa-se que o grupo de indivíduos não-diabéticos (assim como na população anterior) com genótipos CT ou TT apresentou uma incidência significativamente maior de eventos cardiovasculares (Figura 4A). A análise da interação entre as variáveis diabetes e grupos genotípicos foi significativa (valor de  $p = 0,032$ ) para eventos cardiovasculares compostos, o que valoriza os resultados nesse subgrupo. Em valores absolutos, o grupo de não-diabéticos CT/TT teve 47 eventos combinados (morte cardíaca, infarto do miocárdio e angina refratária necessitando revascularização ou novo cateterismo cardíaco), em comparação com 12 eventos no grupo CC. No entanto, pôde-se observar que o polimorfismo rs7903146 não foi associado com o desfecho combinado de eventos cardiovasculares em indivíduos diabéticos (Figura 4B). Decompondo os desfechos em indivíduos não-diabéticos, observou-se que a associação entre os genótipos do *TCF7L2* e eventos cardiovasculares ocorreu, principalmente, em virtude do desfecho “mortalidade por todas as causas”. Os indivíduos carreadores do alelo T apresentaram uma maior mortalidade (número de óbitos: CT/TT = 27 *versus* CC = 4, valor de  $p = 0,004$ ). Embora a incidência de infarto do miocárdio, revascularização miocárdica, e angioplastia coronária percutânea (ATC) terem sido maiores no grupo CT/TT nos não-diabéticos, estes valores não alcançaram significância estatística ( $p > 0,10$ ). Após ajuste para vários fatores de risco estabelecidos para DCV listados na tabela 9 (incluindo o diagnóstico de IAM prévio), apenas o grupo de genótipos CT/TT do *TCF7L2* ( $p = 0,047$ ), idade ( $p = 0,003$ ) e tabagismo ( $p = 0,013$ ) foi significativamente e independentemente associado a eventos cardiovasculares compostos.

Portanto, a informação do genótipo do *TCF7L2* foi capaz de estratificar populações brasileiras não apenas com relação à prevalência e à gravidade de DAC, mas também em relação à incidência de eventos cardiovasculares em uma população com DAC estabelecida.



Número sob risco														
CC	120	110	106	103	100	99	93	38	33	31	29	29	29	24
CT+TT	266	218	206	198	192	187	177	135	112	100	97	92	89	81

**Figura 4** - Curva de Kaplan-Meier para eventos cardiovasculares combinados em não-diabéticos (A) e diabéticos (B) após 5 anos de seguimento, de acordo com os grupos de genótipos do *TCF7L2*. Em pacientes não-diabéticos, os indivíduos CT e TT tiveram uma incidência significativamente maior de desfechos cardiovasculares compostos. Valor de p para interação entre diabetes e grupos genotípicos = 0,032. Embora os indivíduos diabéticos tivessem maior incidência de eventos cardiovasculares do que os não-diabéticos, a presença do alelo T não foi associada com desfechos cardiovasculares

#### 4.1.3 Resultados do Objetivo Secundário 1

*Estudar se a informação do *TCF7L2* é capaz de melhorar uma ferramenta diagnóstica de DM2, baseada em variáveis clínicas.*

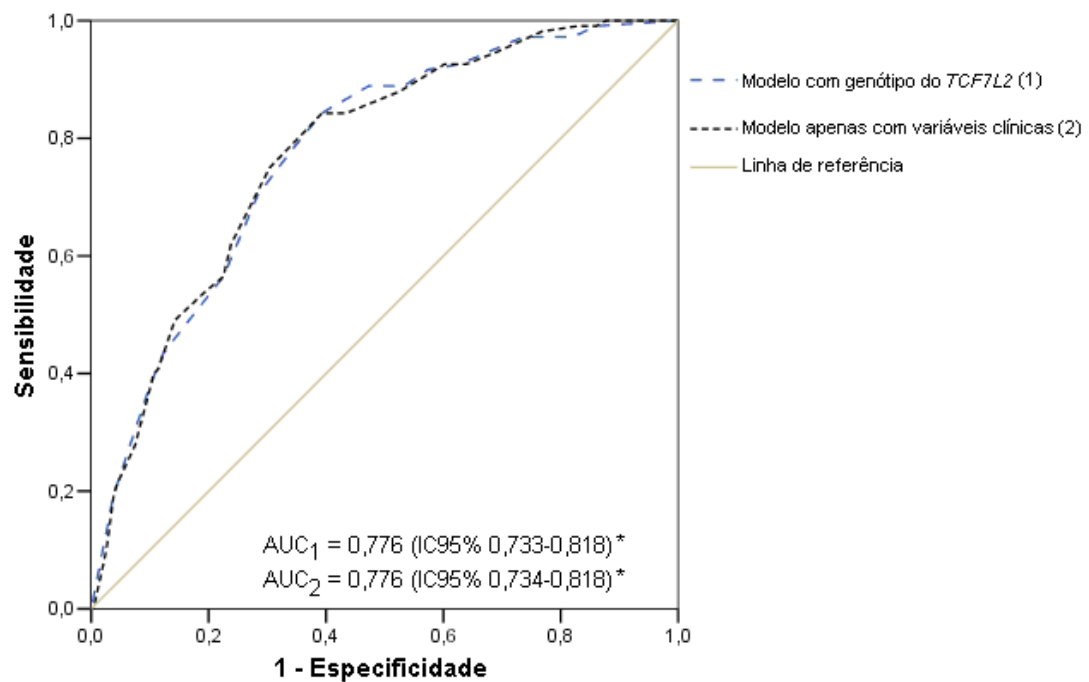
Uma das potenciais aplicações da determinação de variantes genéticas associadas às doenças complexas seria a possível capacidade

de acrescentar informações relevantes a modelos diagnósticos ou de predição dessas doenças, além daquilo que pode ser retirado de variáveis clínicas ou facilmente disponíveis. Por isso, nesse campo, foi avaliada se a adição da informação de apenas uma das variantes genéticas (*TCF7L2*) a um modelo de predição diagnóstica de DM2, validado em uma população geral, melhoraria expressivamente o desempenho desse modelo. Em uma população de Vitória, Espírito Santo, Brasil, 1.440 amostras de DNA foram genotipadas para a variante rs7903146 do *TCF7L2*. Nesta população, destaca-se a prevalência de DM2 de 7,9%. A prevalência do genótipo CC foi 43,4% (n = 614), do genótipo CT igual a 46,6% (n = 660) e a prevalência do genótipo TT do *TCF7L2* foi de 10,0% (n = 142), de forma que as frequências alélicas e genotípicas estavam em equilíbrio de Hardy-Weinberg nessa amostra. A prevalência do genótipo TT foi um pouco maior nos indivíduos diabéticos (12,1%) em comparação com sujeitos não-diabéticos (9,9%), mas essa diferença não foi estatisticamente significativa (p = 0,46). De particular interesse, o número total de indivíduos diabéticos que carregavam o genótipo TT, apesar do tamanho relativamente grande da amostra inicial, caiu para apenas 14 pessoas. Portanto, o poder estatístico para se detectar essa associação na amostra estudada, a partir dessa população geral, foi de apenas 24%.

Considerando o modelo de predição diagnóstica de diabetes nessa mesma população (Pires De Sousa, Pereira *et al.*, 2009; Sousa, Marquezine *et al.*, 2009), baseado apenas em critérios clínicos (idade, presença de hipertensão arterial e IMC), que alcançou um bom desempenho para identificar indivíduos com maior probabilidade de ter diabetes, a inclusão da informação apenas do genótipo TT, em tal modelo, não resultou em melhoria



na área sob a curva ROC (AUC = 0,776 IC95% 0,73-0,82 em ambos os modelos), como se observa na Figura 5. As características de desempenho dos dois modelos são apresentadas na Tabela 10. Houve uma ligeira melhora na especificidade quando o genótipo TT do *TCF7L2* foi incluído, às custas de uma diminuição da sensibilidade. Portanto, a inclusão de apenas uma variante genética, mesmo sendo esta a principal variante associada ao DM2, a um modelo de predição diagnóstica em uma população, cuja prevalência de diabetes foi pequena, não foi capaz de melhorar o desempenho do modelo.



**Figura 5** - Curvas ROC para predição de DM2 dos dois modelos em Vitória. Observa-se que tanto o modelo com apenas variáveis clínicas quanto aquele com adição do genótipo do *TCF7L2* alcançaram capacidade preditiva significativa. Contudo, a adição da informação genética não aumentou significativamente o valor da AUC. (\*) Significância assintótica - valores de  $p < 0,0001$

**Tabela 10** - Desempenho do modelo clínico e do modelo clínico com informação do *TCF7L2* na população geral de Vitória – ES

<b>Modelo 1 - Apenas Variáveis Demográficas</b>								
Valores de corte	Sensibilidade	Especificidade	VPP	VPN	Acurácia	OR (IC95%)	Valor de p	Necessidade de testes adicionais* (%)
16,5	0,8426	0,6071	0,19	0,972	0,63	8,27 (3,86-12,68)	<0,00001	43,7
18	0,7037	0,7162	0,213	0,957	0,71	5,99 (3,38-8,60)	<0,00001	32,51
<b>Modelo 2 – Variáveis Demográficas e Informação Genética do <i>TCF7L2</i></b>								
Valores de corte								
16,5	0,8426	0,5939	0,185	0,972	0,62	7,83 (3,66-12,0)	<0,00001	44,89
18	0,7222	0,7081	0,213	0,959	0,71	6,30 (3,51-9,1)	<0,00001	33,42

VPP = valor preditivo positivo; VPN = valor preditivo negativo; (\*) percentual da população total que atingiria a específica pontuação (de acordo com o valor de corte da pontuação) e que seria necessária investigação adicional para o diagnóstico de DM2.

## 4.2 Objetivo Específico 2

*Investigar a associação entre a combinação das informações de variantes genéticas associadas ao DM2 e eventos cardiovasculares em pacientes portadores de doença arterial coronária multiarterial seguidos por 5 anos, também em indivíduos diabéticos e não-diabéticos.*

Na população do projeto MASS II, um total de 425 indivíduos foram genotipados para todos os dez SNPs estudados (informação genética completa), dos quais 134 tinham diabetes (31,5%). Foi realizada uma comparação entre os pacientes do projeto MASS II incluídos e excluídos nessa análise, com fins de avaliar viés de seleção, mas não houve diferenças significativas nas características clínicas e laboratoriais entre estes grupos. Todos os SNPs estavam em equilíbrio de Hardy-Weinberg (valor de  $p > 0,02$ ). Os participantes eram provenientes do Brasil, pertenciam a vários grupos étnicos e a população não foi estratificada por etnia. As características clínicas dos participantes, no momento da randomização, são mostradas na Tabela 11. Observa-se que, exceto para a prevalência de hipertensão arterial, que foi maior no grupo de indivíduos diabéticos, os sujeitos portadores de DM2 apresentavam características semelhantes aos não-diabéticos.

Antes da análise da associação entre as variantes genéticas e os eventos cardiovasculares, de que se trata esta seção especificamente, foi testada a associação entre os SNPs e o diagnóstico de DM2. As características gerais desses SNPs e suas associações com o diagnóstico

de DM2 nessa população podem ser encontradas na Tabela 12. Analisando-se essa tabela, observa-se que apenas o gene *TCF7L2* isoladamente talvez estava associado ao DM2, se considerado um valor de  $p < 0,05$  para se excluir a hipótese nula que nenhum polimorfismo estava independentemente associado ao DM2. Contudo, utilizando-se um valor de  $p$  mais restrito ( $p < 0,005$  após correção de Bonferroni para múltiplos testes), nenhum dos polimorfismos estava associado ao DM2. As frequências dos alelos de risco foram bastante variáveis entre os genótipos, embora comparáveis a outros estudos (Lango, Palmer *et al.*, 2008; Lyssenko, Jonsson *et al.*, 2008; Meigs, Shrader *et al.*, 2008).

**Tabela 11** - Características clínicas dos indivíduos diabéticos e não-diabéticos genotipados para os 10 *loci* de risco na população do projeto MASS II

	<b>Pacientes com diabetes (n=134)</b>	<b>Pacientes sem diabetes (n=291)</b>	<b>Valor de P</b>
Idade (anos)	60,40 ± 8,82	59,31 ± 9,36	0,256
Sexo Masculino (%)	63,4%	69,1%	0,250
IMC (kg/m <sup>2</sup> )	27,68 ± 4,44	26,98 ± 4,00	0,109
Colest, Total (mg/dL)	219,23 ± 50,56	226,31 ± 50,27	0,181
HDL-c (mg/dL)	38,19 ± 11,27	36,76 ± 9,86	0,200
LDL-c (mg/dL)	143,78 ± 45,67	150,93 ± 44,34	0,141
Triglicerídeos (mg/dL)	204,78 ± 145,07	190,26 ± 105,24	0,247
HAS (%)	67,9%	54,3%	0,008
Tabagismo (%)	28,4%	35,1%	0,173
IAM prévio (%)	42,5%	46,7%	0,419
DAC triarterial (%)	56,7%	54,6%	0,689

A informação combinada dos 10 SNPs estudados foi analisada por meio dos três escores genotípicos: Escore 1 (Escore Recessivo), Escore 2 (Aditivo) e Escore 3 (Ponderado). A fim de se avaliar a associação entre os três escores genotípicos e o diagnóstico de DM2, foi realizada uma regressão logística multivariada, incluindo fatores de risco clínicos para DM2 (como idade, sexo, IMC, hipertensão arterial e tabagismo) e a informação genética combinada (correspondente aos três escores genotípicos) e apenas os três escores e hipertensão arterial foram independentemente associados com DM2 (Tabela 13). Houve um aumento no risco de diabetes tipo 2 por um fator de 1,27 para cada ponto do Escore 1 ( $p = 0,011$ ), 1,18 para um ponto do Escore 2 ( $p = 0,002$ ) e 1,09 por um ponto do Escore 3 ( $p = 0,0003$ ). Baseado em tercis de valores dos escores genotípicos, que classificou os indivíduos em grupos de baixo risco genético (menores valores de escores) e grupos de alto risco genético (maiores valores de escores), como esperado, o percentual de indivíduos com diabetes foi significativamente maior nos grupos de alto risco genético do que nos grupos de baixo risco, independentemente do escore utilizado na análise (Tabela 14).

**Tabela 12** - Fatores Genéticos Preditivos de Diabetes Mellitus Tipo 2 na população do estudo MASS II

Gene	Cromosomo	SNP	Alelos Risco/ sem Risco	FAR <sub>AF</sub>	Modelo Aditivo		Modelo Alternativo	
					Odds Ratio (IC95%)	Valor p	Odds Ratio (IC95%)	Valor p
<i>KCNJ11</i>	11	rs5219	T/C	0,33	1,003 (0,738-1,364)	0,984	Recessivo 1,398 (0,744-2,627)	0,298
<i>PPARG</i>	3	rs1801282	C/G	0,93	1,464 (0,837-2,560)	0,182	Recessivo 1,444 (0,819-2,548)	0,204
<i>TCF7L2</i>	10	rs7903146	T/C	0,47	1,511 (1,069-2,135)	0,019	Recessivo 2,039 (1,106-3,760)	0,022
<i>SLC30A8</i>	8	rs13266634	C/T	0,76	1,137 (0,818-1,582)	0,444	Recessivo 1,238 (0,816-1,880)	0,315
<i>HHEX</i>	10	rs1111875	G/A	0,68	1,274 (0,946-1,716)	0,111	Recessivo 1,420 (0,939-2,147)	0,096
<i>CDKAL1</i>	6	rs7754840	C/G	0,37	1,198 (0,885-1,622)	0,243	Dominante 1,291 (0,850-1,963)	0,231
<i>IGF2BP2</i>	3	rs4402960	T/G	0,39	1,237 (0,899-1,702)	0,191	Recessivo 1,419 (0,744-2,707)	0,289
<i>CDKN2A/B</i>	9	rs10811661	T/C	0,85	1,307 (0,880-1,941)	0,184	Recessivo 1,406 (0,894-2,212)	0,141
<i>CDKN2A/B</i>	11	rs9300039	C/A	0,90	1,095 (0,684-1,753)	0,705	Recessivo 1,214 (0,718-2,051)	0,470
<i>FTO</i>	16	rs8050136	A/C	0,40	1,010 (0,743-1,374)	0,948	Dominante 1,167 (0,756-1,801)	0,486

Os valores de *odds ratio* para o risco de diabetes tipo 2 foram calculados por regressão logística univariada com ajuste para idade e sexo. Os modelos genéticos primários são aditivos; modelos alternativos estão indicados. IC indica intervalo de confiança; NA: não aplicável, e FARaf frequência do alelo de risco em indivíduos afetados.

**Tabela 13** - Regressão logística multivariada para o diagnóstico de DM2 na população do projeto MASS-II

	<i>Odds Ratio</i> (IC95%)	Valor de p
<b>Escore Recessivo (Escore 1)</b>		
Sexo masculino	1,181 (0,762-1,830)	0,456
Idade, para cada ano	1,018 (0,995-1,042)	0,127
Tabagismo	1,275 (0,806-2,017)	0,299
Hipertensão	<b>1,582 (1,037-2,412)</b>	<b>0,033</b>
IMC, para cada 1 kg/m <sup>2</sup>	1,035 (0,985-1,088)	0,170
Combinação dos alelos de risco do Escore 1	<b>1,226 (1,047-1,435)</b>	<b>0,011</b>
<b>Escore Aditivo (Escore 2)</b>		
Sexo masculino	1,206 (0,777-1,871)	0,404
Idade, para cada ano	1,019 (0,995-1,043)	0,115
Tabagismo	1,245 (0,787-1,970)	0,348
Hipertensão	<b>1,558 (1,020-2,378)</b>	<b>0,040</b>
IMC, para cada 1 kg/m <sup>2</sup>	1,037 (0,987-1,090)	0,154
Combinação dos alelos de risco do Escore 2	<b>1,177 (1,060-1,306)</b>	<b>0,002</b>
<b>Escore Ponderado (Score 3)</b>		
Sexo masculino	1,185 (0,762-1,844)	0,450
Idade, para cada ano	1,021 (0,997-1,045)	0,087
Tabagismo	1,254 (0,791-1,989)	0,336
Hipertensão	<b>1,580 (1,033-2,418)</b>	<b>0,035</b>
IMC, para cada 1 kg/m <sup>2</sup>	1,041 (0,990-1,094)	0,119
Combinação dos alelos de risco do Escore 3	<b>1,092 (1,042-1,146)</b>	<b>0,0003</b>

**Tabela 14** - Proporção de indivíduos com DM2 segundo grupo de risco genético (baseado nos tercís dos escores genotípicos)

	<b>Diabetes (%)</b>	<b>Odds Ratio (IC95%)</b>	<b>Valor de P</b>
<b>Escore 1</b>			
Baixo risco ( $\leq 3$ )	33,9%		
Alto risco ( $\geq 5$ )	47,9%	1,786 (1,126-2,833)	0,014
<b>Escore 2</b>			
Baixo risco ( $\leq 11$ )	27,7%		
Alto risco ( $\geq 14$ )	45,1%	2,146 (1,276-3,610)	0,005
<b>Escore 3</b>			
Baixo risco ( $\leq 12$ )	30,6%		
Alto risco ( $\geq 17$ )	52,2%	2,479 (1,524-4,032)	<0,0001

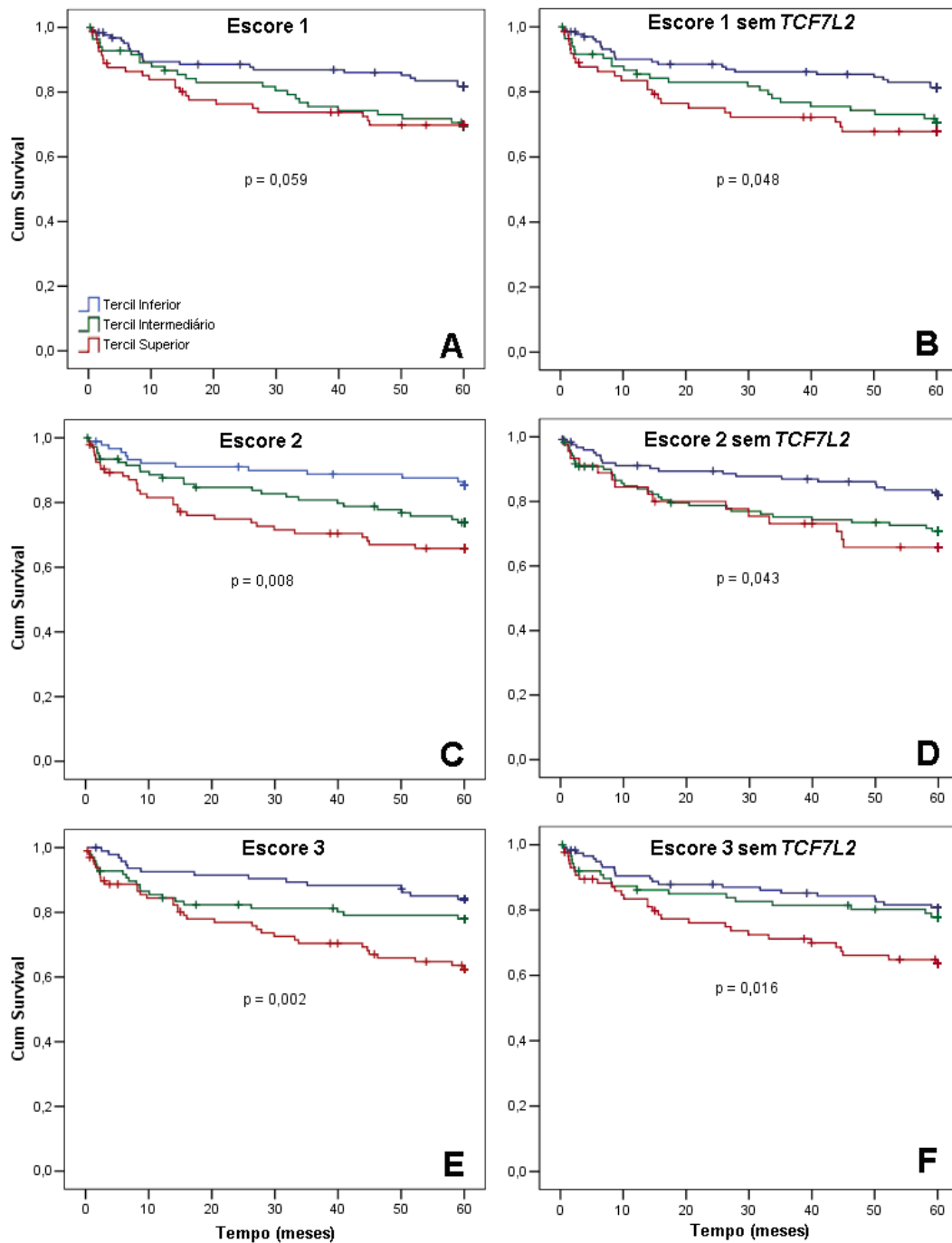
Após a confirmação da associação entre a informação genética combinada (escores) e o diagnóstico de DM2, foi estudada a associação entre os escores genotípicos e eventos cardiovasculares na população do projeto MASS II. Com esse fim, foi realizada uma divisão dos indivíduos em três grupos, baseada nos tercís dos valores dos três escores. Os valores dos escores genotípicos, de acordo com a classificação em tercís, podem ser encontrados na Tabela 15.



**Tabela 15** - Distribuição dos escores genéticos classificados, de acordo com os tercís das somas dos valores da pontuação de cada SNP

	<b>Escore 1</b> <b>(Variação: 1 a 7)</b>	<b>Escore 2</b> <b>(Variação: 4 a 17)</b>	<b>Escore 3</b> <b>(Variação: 4 a 27)</b>
Tercil Inferior (Baixo risco)	$\leq 3$	$\leq 10$	$\leq 17$
Tercil Intermediário (Risco médio)	4	11 – 12	18 – 21
Tercil Superior (Risco elevado)	$\geq 5$	$\geq 13$	$\geq 22$

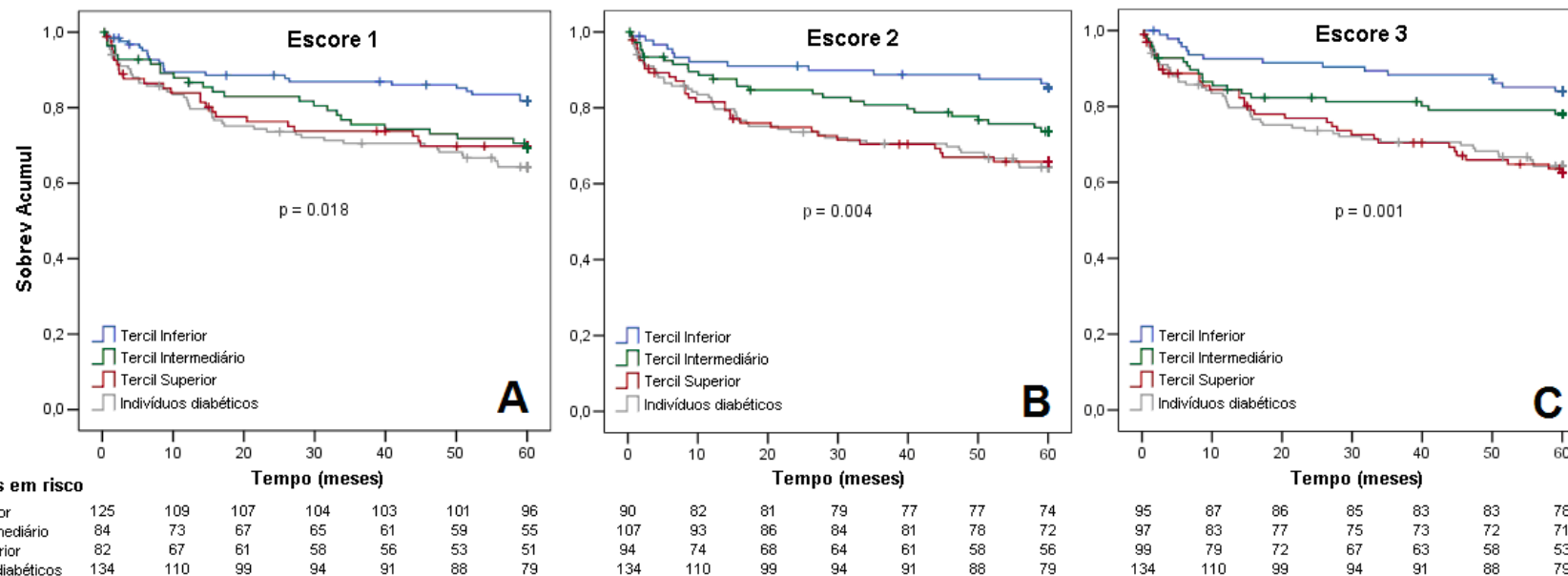
Embora os indivíduos diabéticos tenham apresentado uma maior incidência de eventos cardiovasculares combinados que os não-diabéticos (35,1 *versus* 24,4%,  $p = 0,022$ ), observou-se que não houve associação entre escores genéticos e desfechos cardiovasculares compostos no grupo de indivíduos diabéticos. No entanto, indivíduos não-diabéticos com maiores valores nos Escores 2 e 3 apresentaram uma incidência significativamente maior de eventos cardiovasculares (Figura 6). Considerando a associação positiva entre alelos de risco do gene *TCF7L2* e eventos cardiovasculares nessa mesma população, mostrada na seção anterior, essa mesma análise foi realizada subtraindo-se os dados referentes a esse polimorfismo. Contudo, mesmo quando dados do *TCF7L2* foram excluídos, os Escores genotípicos 2 e 3 permaneceram significativamente associados com o desfecho cardiovascular composto em indivíduos não-diabéticos (Figura 6). Além disso, o Escore 2 e o Escore 2 sem informação do *TCF7L2* também foram significativamente associados com mortalidade por todas as causas (valor  $p = 0,009$  e  $0,036$ , respectivamente).



**Figura 6** - Curvas de Kaplan-Meier dos escores genotípicos com e sem informação do *TCF7L2* e desfechos cardiovasculares compostos em indivíduos não-diabéticos. A figura é mostrada com os grupos separados, de acordo com tercís dos valores de cada escore, após 5 anos de seguimento. As seções A, C e E revelam as curvas de Kaplan-Meier para os escores originais e as seções B, D e F para os escores 1, 2 e 3, respectivamente, após a exclusão da informação do *TCF7L2*, confirmada na seção anterior. Dessa forma, entende-se que a associação entre os desfechos cardiovasculares e os escores não se deveu apenas a maior força do *TCF7L2*.

Interessantemente, como mostrado na Figura 7, não existem diferenças significativas entre as estimativas da probabilidade de sobrevivência de indivíduos diabéticos (independentemente dos escores genéticos) e não-diabéticos com valores altos de escores de risco genético (baseados nos tercís dos valores dos três escores genotípicos). Após 5 anos de seguimento, os indivíduos diabéticos apresentaram uma incidência acumulada de eventos cardiovasculares combinados de 35,1%, comparáveis a de indivíduos não-diabéticos dos grupos de tercís mais elevados, que apresentaram 29,3%, 33,0% e 35,4% para o tercil superior dos escores 1, 2 e 3, respectivamente (Tabela 16). Isto sugere que os indivíduos não-diabéticos com DAC multi-arterial e função ventricular preservada que têm elevados valores de escores de risco genéticos apresentam resultados comparáveis aos indivíduos com diabetes já estabelecido, em relação a eventos cardiovasculares.

Baseado nos escores genotípicos obtidos e suas respectivas estimativas de risco, foi encontrado um poder estatístico *post-hoc* acima de 80% (82 e 83%, respectivamente, para os Escores 2 e 3) para uma associação com a incidência de eventos cardiovasculares com um erro alfa de 0,05. Para Escore 1, foi encontrado um poder estatístico *post hoc* de 46%.



**Figura 7** – Curvas de Kaplan-Meier dos escores genotípicos e desfechos cardiovasculares compostos para indivíduos diabéticos e não-diabéticos separados em grupos, de acordo com os tercís dos escores genéticos, após 5 anos de seguimento. Observe que as curvas de indivíduos diabéticos são similares às curvas dos tercís superiores. Valores de P para as comparações: A) inferior *versus* intermediário:  $p = 0,042$ ; inferior *versus* superior:  $p = 0,035$ ; diabéticos *versus* inferior:  $p = 0,002$ ; intermediário *versus* superior:  $p = 0,909$ ; diabéticos *versus* intermediário:  $p = 0,381$ ; superior *versus* diabéticos:  $p = 0,482$ . B) inferior *versus* intermediário:  $p = 0,052$ ; inferior *versus* superior:  $p = 0,002$ ; diabéticos *versus* inferior:  $p = 0,001$ ; intermediário *versus* superior:  $p = 0,178$ ; diabéticos *versus* intermediário:  $p = 0,101$ ; superior *versus* diabéticos:  $p = 0,870$ . C) inferior *versus* intermediário:  $p = 0,239$ ; inferior *versus* superior:  $p = 0,001$ ; diabéticos *versus* inferior:  $p = 0,001$ ; intermediário *versus* superior:  $p = 0,035$ ; diabéticos *versus* intermediário:  $p = 0,037$ ; superior *versus* diabéticos:  $p = 0,849$ .

**Tabela 16** – Número de eventos cardiovasculares compostos e *Hazard* cumulativo em indivíduos diabéticos e não-diabéticos, de acordo com os tercís dos escores genotípicos

	Número de eventos	Hazard Cumulativo (%)	Valor de p	Valor de p - interação
<b>Indivíduos Não diabéticos</b>				
<b>Escore 1</b>			0,018*	0,053 §
Tercil Inferior	22	17,6		
Tercil Intermediário	25	29,1		
Tercil Superior	24	29,3		
<b>Escore 2</b>			0,004*	0,029 §
Tercil Inferior	13	14,4		
Tercil Intermediário	27	25,2		
Tercil Superior	31	33,0		
<b>Escore 3</b>			0,001*	0,034 §
Tercil Inferior	15	15,8		
Tercil Intermediário	21	21,6		
Tercil Superior	35	35,4		
Indivíduos Diabéticos	47	35,1	0,022‡	-

(\*) Valor de p para comparação entre os tercís dos escores e indivíduos diabéticos; (‡) Valor de p para comparações entre indivíduos diabéticos e indivíduos não-diabéticos; (§) Valor de p para interação entre variável diabetes e cada escore genotípico, com relação aos eventos combinados

Finalmente, após análise ajustada para vários fatores de risco estabelecidos (hipertensão, infarto do miocárdio prévio, idade, sexo, IMC, tabagismo e níveis de colesterol total, LDL-colesterol, HDL-colesterol e triglicérides) em um modelo multivariado baseado na análise de riscos proporcionais de Cox, todos os três escores foram significativamente e independentemente associados com eventos cardiovasculares compostos (valores de p menores que 0,05, Tabela 17). Além disso, os Escores 2 e 3 também foram independentemente associados à mortalidade por todas as causas.

**Tabela 17** – Hazard ratios e intervalos de confiança para eventos cardiovasculares compostos e mortalidade em função do aumento de 1 ponto em cada escore genético em indivíduos não-diabéticos \*

	Indivíduos Não-diabéticos		
	HR	IC95%	Valor p
<b>Eventos Cardiovasculares Compostos</b>			
Escore 1	1,233	1,01-1,50	0,038
Escore 2	1,170	1,03-1,33	0,017
Escore 3	1,074	1,01-1,14	0,019
Escore 2 sem <i>TCF7L2</i>	1,137	0,99-1,30	0,058
Escore 3 sem <i>TCF7L2</i>	1,023	1,00-1,30	0,050
	Indivíduos Não-diabéticos		
	HR	IC95%	Valor p
<b>Mortalidade</b>			
Escore 1	1,167	0,90-1,52	0,248
Escore 2	1,246	1,04-1,49	0,018
Escore 3	1,119	1,03-1,22	0,008
Escore 2 sem <i>TCF7L2</i>	1,181	0,98-1,42	0,078
Escore 3 sem <i>TCF7L2</i>	1,094	0,99-1,21	0,084

(\*) Ajustado para idade, sexo, hipertensão arterial, IMC, colesterol total, HDL, LDL, triglicerídeos, IAM prévio e tabagismo

Poder-se-ia imaginar que a glicemia de jejum (como variável contínua), mesmo em níveis não-diabéticos, pudesse explicar os resultados encontrados até então. Contudo, mesmo se os valores de glicemia de jejum também forem incorporadas no modelo proporcional de Cox, ambos os Escores 2 e 3 permanecem significativamente e independentemente associados com os eventos cardiovasculares compostos e mortalidade por todas as causas (Tabela 18). Da mesma forma, os resultados encontrados

poderiam ser explicados apenas pelo fato desses indivíduos com altos valores dos escores genéticos tenham se tornado diabéticos durante o seguimento. No entanto, também foram reavaliados 250 pacientes que completaram o seguimento e desses, 38 pacientes foram reclassificados como diabéticos, de acordo com critérios baseados na glicemia de jejum, ao fim do período de 5 anos. Analisando apenas os 212 indivíduos restantes, ambos os Escores 2 e 3 ainda permaneceram significativamente associados com desfechos cardiovasculares compostos (valores de p iguais a 0,035 e 0,036, respectivamente).

Esses dados sugerem que a combinação das informações genéticas (analisada pelos escores genotípicos) também é capaz de estratificar uma população de pacientes com DAC em relação a eventos cardiovasculares. Considerando que apenas o gene *TCF7L2* foi previamente associado a esses eventos cardiovasculares e que a exclusão desse mesmo gene da análise ainda resultou em uma associação positiva, pode-se inferir que a informação combinada é melhor que a utilização de cada SNPs isoladamente e que isso não ocorreu devido à importância relativa do *TCF7L2*. Finalmente, mesmo considerando a possibilidade de alguns desses indivíduos terem se tornado diabéticos durante o seguimento, isso não foi capaz de explicar totalmente os resultados encontrados, bem como a glicemia de jejum não foi capaz de estratificar os indivíduos de maneira semelhante aos escores genéticos.

**Tabela 18** – *Hazard ratios* e intervalos de confiança para eventos cardiovasculares compostos e mortalidade em função do aumento de 1 ponto em cada escore genético em indivíduos não-diabéticos, ajustados para glicemia de jejum\*

	Indivíduos Não-diabéticos		
	HR	IC95%	Valor p
<b>Eventos Cardiovasculares Compostos</b>			
Escore 1	1,209	0,99-1,48	0,062
Escore 2	1,163	1,02-1,32	0,022
Escore 3	1,079	1,01-1,15	0,015
Escore 2 sem <i>TCF7L2</i>	1,125	0,99-1-29	0,084
Escore 3 sem <i>TCF7L2</i>	1,076	0,99-1,16	0,064
<b>Mortalidade</b>			
Escore 1	1,117	0,85-1,46	0,419
Escore 2	1,217	1,01-1,46	0,038
Escore 3	1,110	1,02-1,21	0,017
Escore 2 sem <i>TCF7L2</i>	1,145	0,95-1,38	0,163
Escore 3 sem <i>TCF7L2</i>	1,076	0,97-1,19	0,168

(\*) Ajustado para idade, sexo, hipertensão arterial, IMC, colesterol total, HDL, LDL, triglicerídeos, IAM prévio, tabagismo e glicemia de jejum.

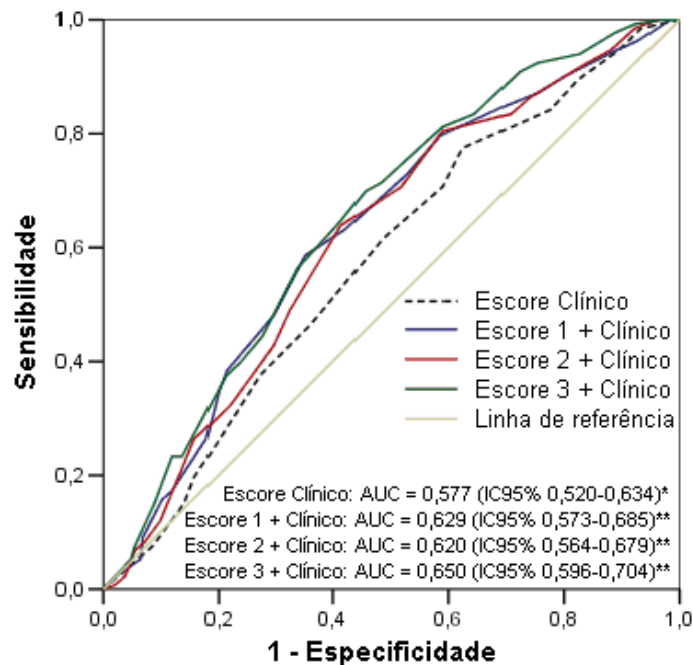
#### 4.2.1 – Resultados do Objetivo Secundário 2

*Avaliar se a informação clínica, advinda dos marcadores moleculares estudados em combinação, pode ser utilizada para melhorar algoritmos diagnósticos para diabetes*

Se apenas as variáveis clínicas (combinação de sexo, idade, obesidade, tabagismo e hipertensão arterial) forem utilizadas para se prever o diagnóstico do DM2 na população do projeto MASS II, o desempenho desse modelo verificada pela área sob a curva ROC (AUC) seria significativa, embora modesto (AUC = 0,58, p = 0,011). A inclusão de outras variáveis, como triglicerídeos e colesterol HDL, nesse escore clínico não melhorou a capacidade preditiva do



escore (AUC = 0,57) e, portanto, não foram utilizadas. Entretanto, a adição da informação dos escores genotípicos (combinação da informação dos 10 polimorfismos estudados) foi capaz de melhorar significativamente o desempenho do modelo clínico (Figura 8), de acordo com as AUC: a adição do Escore 1 eleva a AUC para 0,63 ( $p = 0,024^1$ ), a adição do Escore 2 aumenta a AUC para 0,62 ( $p = 0,029^1$ ) e a adição do Escore 3 incrementa a AUC para 0,65 ( $p = 0,038^1$ ).



**Figura 8** - Curvas ROC para o diagnóstico de DM2 para o Escore Clínico (contendo fatores de risco para DM2) e a combinação dos escores genotípicos (Escore 1, 2 ou 3) e o Escore Clínico. Observa-se que todos os escores genotípicos alcançaram capacidade preditiva significativa. Contudo, a adição da informação genética combinada, correspondente a qualquer um dos escores genéticos, aumenta significativamente o valor da AUC. Os valores de  $p$  correspondem à significância assintótica: (\*) valor de  $p = 0,011$ ; (\*\*) valor de  $p < 0,0001$ . Valores de  $p$  para comparação entre o Escore Clínico e Escores combinados (adição de cada escore genético): Escore 1 – valor de  $p = 0,0243$ ; Escore 2 – valor de  $p = 0,0292$ ; Escore 3 – valor de  $p = 0,0380$ .

<sup>1</sup> Os valores de  $p$ , descritos nessa análise, correspondem à comparação dessas curvas com a do escore clínico de fatores de risco

## **5 Discussão**

---

## 5.1 Relação entre hiperglicemia e Doença Cardiovascular

Sabe-se, há décadas, que o diabetes mellitus aumenta, sobremaneira, o risco de doença cardiovascular, de forma que aproximadamente 65% dos óbitos em pacientes com DM2 estão relacionados à doença coronariana ou acidente vascular encefálico (Morrish, Wang *et al.*, 2001). Apesar dessa clara associação entre doença aterosclerótica e diabetes, os mecanismos subjacentes que relacionam essas duas condições ainda não são completamente entendidos (Milicevic, Raz *et al.*, 2008). Enquanto fatores não-glicêmicos relacionados à Síndrome Metabólica (hipertensão arterial, dislipidemia, estado de hipercoagulabilidade e inflamação crônica) são classicamente implicados, estudos epidemiológicos sugerem um efeito independente da hiperglicemia no risco cardiovascular (Wei, Gaskill *et al.*, 1998; Stratton, Adler *et al.*, 2000). Um exemplo disso vem de dados do estudo MRFIT (*Multiple Risk Factor Intervention Trial*), que mostraram que a mortalidade por doença cardiovascular aumenta com o número de fatores de risco cardiovascular coexistentes, mesmo na população diabética (Stamler, Vaccaro *et al.*, 1993). Contudo, esse risco foi ainda majorado quando havia hiperglicemia concomitante.

Dados epidemiológicos de estudos como o DCCT (*Diabetes Control and Complications Trial*), EDIC (*Epidemiology of Diabetes Interventions and*

*Complications*) e UKPDS (*United Kingdom Prospective Diabetes Study*) têm fornecido evidências significativas para se afirmar que a relação entre o grau de hiperglicemia e o risco de DCV é independente de outros fatores de risco tanto em indivíduos com diabetes tipo 1 quanto com diabetes tipo 2 (Stratton, Adler *et al.*, 2000; Nathan, Cleary *et al.*, 2005). Contudo, ainda não está definido como os vários componentes do perfil de glicemia diária se correlacionam com o risco de DCV (Milicevic, Raz *et al.*, 2008). O *San Luigi Gonzaga Diabetes Study*, que avaliou o efeito da glicemia de jejum e pós-prandial no risco de novos eventos cardiovasculares, em uma população de pacientes com DM2, revelou que apenas a glicemia pós-prandial foi capaz de prever novos eventos (Cavalot, Petrelli *et al.*, 2006). Portanto, acredita-se que o grau de elevação da glicemia pós-prandial seja um melhor preditor de DCV que a glicemia de jejum.

Há cerca de 30 anos, estudos de necropsia e baseados em angiografia têm mostrado que pacientes com diabetes têm doença arterial coronariana de maior gravidade, caracterizada por um maior número de segmentos vasculares envolvidos e por uma distribuição difusa das lesões ateroscleróticas (Vigorita, Moore *et al.*, 1980; Waller, Palumbo *et al.*, 1980). Estes dados são corroborados por nossos resultados, em que os indivíduos diabéticos apresentaram uma prevalência significativamente maior de lesões ateroscleróticas na angiografia coronariana (aproximadamente 80%). Além disso, indivíduos diabéticos apresentam um maior risco de ulceração da placa aterosclerótica e consequente trombose, comparados a sujeitos não-diabéticos (Silva, Escobar *et al.*, 1995). Dessa forma, pode-se concluir que

existem diferenças nas características da doença aterosclerótica entre indivíduos diabéticos e não-diabéticos e essas estão provavelmente relacionadas aos efeitos da hiperglicemia e outras desordens diabetes-específicas (Milicevic, Raz *et al.*, 2008). Realmente, pacientes com DM2 têm caracteristicamente resistência à insulina e frequentemente hiperinsulinemia, hipertensão arterial, dislipidemia, obesidade central, elevação do inibidor do ativador do plasminogênio e inflamação (DeFronzo e Ferrannini, 1991), que contribuem de forma importante para a formação da placa aterosclerótica. Por outro lado, mesmo pequenas alterações da homeostase glicêmica apresentam efeitos diretos no processo aterogênico (Bianchi, Miccoli *et al.*, 2008). Elevações da glicemia (notadamente a pós-prandial) contribuem significativamente para exposição dos tecidos periféricos aos efeitos da hiperglicemia. Além disso, excursões dos níveis de glicemia excessivas e repetidas levam a uma maior variabilidade glicêmica e ambas podem contribuir no desenvolvimento da aterosclerose (Bianchi, Miccoli *et al.*, 2008).

A parede arterial é um conhecido alvo para o insulto provocado pela hiperglicemia: a glicosilação de proteínas causa cruzamento das fibras de colágeno e de outras proteínas da matriz extracelular na parede vascular, aumentando a susceptibilidade à aterosclerose; moléculas de LDL glicosiladas são mais aterogênicas e promovem ativação de macrófagos e, conseqüentemente, formação de células espumosas; a hiperglicemia pode ter efeito direto nas células endoteliais, levando a disfunção endotelial (Bianchi, Miccoli *et al.*, 2008). Além disso, uma elevação aguda na glicemia pode levar a aumento da pressão arterial, aumento nos níveis circulantes de ICAM-1

(*Intracellular Adhesion Molecule 1*), aumento no grau de inflamação (produção de interleucina 6, interleucina 18 e fator de necrose tumoral  $\alpha$  – TNF $\alpha$ ), além de aumento no estresse oxidativo (Ceriello, 2005; Wajchenberg, 2007b). Ademais, os radicais livres podem, também, mediar alguns efeitos associados à hiperglicemia, como vasoconstrição, expressão de moléculas de adesão e ativação da coagulação. A inibição do NADPH por superóxidos prejudica o fluxo da glicólise intracelular e, conseqüentemente, leva ao acúmulo de precursores glicolíticos, que, por sua vez, geram ativação de vias normalmente responsáveis por pequeno efeito no metabolismo de carboidratos, como a via do poliol, via das hexosaminas, a formação de produtos de glicação avançada e, finalmente, a ativação da proteína quinase C (Brownlee, 2001). A proteína quinase C aumenta a secreção de endotelina, colágeno tipo 4 e fibronectina, além de aumentar a expressão de moléculas de adesão e ativar o processo de migração dos macrófagos. Diante disso, pode-se concluir que estas alterações podem desempenhar um papel não apenas no desenvolvimento das complicações microvasculares do DM2, mas também na formação da placa aterosclerótica.

## **5.2 Hiperglicemia como um fator de risco cardiovascular em indivíduos não-diabéticos**

Como descrito anteriormente, a hiperglicemia é considerada um fator de risco cardiovascular independente em pacientes com DM2. Contudo, vários estudos têm demonstrado que os efeitos da glicemia ocorrem mesmo

em indivíduos não-diabéticos. Uma metanálise de dados publicados de 20 estudos totalizando quase 100.000 indivíduos, seguidos por 12,4 anos, encontrou uma relação exponencial entre o risco de eventos cardiovasculares e as glicemias de jejum e pós-prandial (Coutinho, Gerstein *et al.*, 1999). Mais interessante foi que essa associação estava presente desde valores de glicemia abaixo dos limiares diagnósticos de diabetes ou pré-diabetes. Esses dados foram corroborados pela análise de coortes asiáticas, em que se observou uma associação positiva e contínua entre glicemia e o risco cardiovascular (Lawes, Parag *et al.*, 2004). Da mesma forma, essa associação estava presente mesmo em níveis de glicemia abaixo de 88,2 mg/dl, ou seja, níveis considerados normais. Entretanto, um risco cardiovascular ainda maior está relacionado com os níveis de glicemia pós-prandial ou pós-sobrecarga de glicose do que com os níveis de glicemia de jejum. O estudo DECODE (*Diabetes Epidemiology Collaborative Analyses of Diagnosis Criteria in Europe*) avaliou dados de 13 estudos epidemiológicos e encontrou que o risco de mortalidade nos indivíduos com glicemia de jejum alterada foi de 1,2 comparado a 1,5 nos sujeitos com intolerância à glicose (pós-sobrecarga de 75g de glicose anidra) (Glucose tolerance and mortality: comparison of WHO and American Diabetes Association diagnostic criteria. The DECODE study group. European Diabetes Epidemiology Group. *Diabetes Epidemiology: Collaborative analysis Of Diagnostic criteria in Europe*, 1999). Nesse mesmo estudo, após ajustes para vários outros conhecidos fatores de risco para DCV, o risco relativo para mortalidade em função da glicemia de jejum não alcançou significância estatística, embora para valores de glicemia

pós-sobrecarga o risco relativo permaneceu significativo. Similarmente, dados do *Funagata Diabetes Study* (Tominaga, Eguchi *et al.*, 1999) confirmaram que a mortalidade em indivíduos com intolerância à glicose é semelhante a de pacientes com diabetes e muito maior do que em indivíduos com glicemia de jejum alterada. Estes dados sugerem que a hiperglicemia de jejum e, principalmente, pós-prandial, mesmo em níveis não-diabéticos, estão associados com DCV.

Esses dados apoiam a explicação sugerida para os resultados do presente trabalho: indivíduos com maiores valores nos escores genéticos teriam menor capacidade de secreção de insulina, conseqüentemente maiores valores de glicemia pós-prandial e/ou excursões glicêmicas (mesmo em níveis não-diabéticos) e, finalmente, maior prevalência de DAC e incidência de eventos cardiovasculares. A partir do momento em que os indivíduos se tornam diabéticos, quando a hiperglicemia passa a ser frequente e muitas vezes constante, a informação genética combinada perde valor.

Além disso, essa maior excursão glicêmica pós-prandial também pode estar associada com a conhecida resposta metabólica anormal ao estresse. Sabe-se que a hiperglicemia é um fator diretamente associado à mortalidade em doenças graves, como, por exemplo, o infarto agudo do miocárdio, principalmente no que se refere à mortalidade intra-hospitalar (Pesaro, Nicolau *et al.*, 2009). Dados do estudo CARDINAL, que avaliou quase 1500 indivíduos com IAM, confirmaram a associação entre mortalidade em 180 dias e tanto a hiperglicemia no momento da admissão quanto a falha em normalização da glicemia em 24h em, interessantemente, indivíduos não



diabéticos. O mesmo não foi observado em sujeitos diabéticos (Goyal, Mahaffey *et al.*, 2006).

Considerando ainda que a associação entre glicemia e o risco cardiovascular está bem estabelecida, seria bem possível imaginar que a glicemia de jejum (como variável contínua, mesmo em níveis não-diabéticos) pudesse ser responsável por boa parte do risco creditado às variantes genéticas em indivíduos não-diabéticos. Em parte, isso foi verdade porque, após a inclusão desta variável nos modelos de regressão, todos os três escores genotípicos apresentaram uma redução na magnitude dos efeitos. No entanto, a glicemia de jejum, certamente mais prática e barata que a análise genética, não foi capaz de responder por toda a estratificação de risco encontrada nesse estudo. Como observado, ambos os escores 2 e 3 mantiveram-se associados com os desfechos cardiovasculares compostos e mortalidade por todas as causas após ajuste para glicemia de jejum.

É importante ressaltar ainda que outros autores, utilizando dados do *Hoorn Study*, que havia encontrado que o risco relativo de mortalidade cardiovascular se iniciava de um nível de glicemia de jejum a partir de 110 mg/dl, recentemente atualizaram seus resultados e mostraram que a associação entre glicemia de jejum alterada e o risco de mortalidade cardiovascular dependeu da conversão desses pacientes para diabetes (Rijkelijkhuizen, Nijpels *et al.*, 2007), de forma que aqueles indivíduos com glicemia de jejum alterada e que não desenvolveram diabetes não apresentaram maior mortalidade. Contudo, isso não se aplicou ao presente estudo. Mesmo após a exclusão dos indivíduos que se tornaram diabéticos até

o final do período de seguimento, a associação entre eventos cardiovasculares e os Escores 2 e 3 persistiu. Neste contexto, vale a pena lembrar que estes escores genéticos foram associados com a mortalidade e apenas o grupo que completou os cinco anos de seguimento foi avaliado em relação à evolução para o diagnóstico de DM2, o que poderia comprometer a força dessa associação. Mesmo assim, analisando-se apenas o grupo que permaneceu sem diabetes, ainda foi encontrada uma associação positiva e significativa.

Estes resultados vêm ressaltar um dos problemas na análise de estudos epidemiológicos: a categorização de variáveis contínuas. No caso particular da glicemia, os indivíduos são classificados em normais, pré-diabéticos ou diabéticos e os valores de corte, que justificam essa divisão, foram propostos de acordo com os níveis a partir dos quais ocorre uma elevação exponencial da prevalência de complicações microvasculares (como referência foi usada a retinopatia diabética) e isso pode não ser verdadeiro para as complicações macrovasculares. Contrariamente, como demonstrado anteriormente por alguns estudos epidemiológicos, o aumento do risco cardiovascular atribuído à glicemia é melhor explicado utilizando-a como uma variável contínua do que categorizando os indivíduos em normais, pré-diabéticos ou diabéticos. E esses ajustes foram realizados no presente trabalho. Dentro desse mesmo raciocínio, isso também poderia explicar, em parte, porque os indivíduos não-diabéticos com altos valores de escores genéticos, categorizados em tercis de valores, tenham taxas de eventos cardiovasculares parecidos com sujeitos diabéticos. Mas, também, foi encontrada uma associação significativa entre os três escores

genotípicos e os eventos cardiovasculares, quando analisados por regressão proporcional de Cox, em que foram considerados os valores absolutos dos escores, isto é, quando estes foram incluídos no modelo como variável contínua.

### **5.3 Relação entre *TCF7L2* e Doença Arterial Coronariana**

Os resultados do presente estudo mostraram uma associação significativa entre o polimorfismo rs7903146 do gene *TCF7L2* e doença arterial coronariana em indivíduos não-diabéticos, em duas populações independentes. Curiosamente, não se foi capaz de demonstrar essa associação em pacientes diabéticos, embora os indivíduos diabéticos tenham apresentado maior gravidade da DAC e mais eventos cardiovasculares que os não-diabéticos. Inclusive, o valor de p para interação entre diabetes e eventos cardiovasculares foi significativo, o que corrobora os resultados encontrados. Considerando que não havia nenhuma associação entre o polimorfismo rs7903146 e fatores de risco cardiovasculares conhecidos além do próprio diabetes, supõe-se que a associação entre o genótipo do *TCF7L2* o fenótipo de eventos cardiovasculares ocorreu devido a alterações da homeostase da glicemia ainda em um nível não-diabético. No entanto, quando o indivíduo desenvolvia DM2, o risco de doença coronariana já estaria aumentado por si.

O gene *TCF7L2* está localizado no cromossomo 10 e o seu polimorfismo rs7903146 está associado fortemente tanto com a prevalência quanto com a incidência de DM2 em diversas populações independentes (Cauchi, El Achhab *et al.*, 2007). O alelo de risco T está associado com distúrbios da secreção de insulina, bem como alterações nos efeitos dos hormônios incretínicos (Lyssenko, Lupi *et al.*, 2007), com tendência à hiperglicemia pós-prandial. Dados do estudo MONICA/KORA revelaram que o alelo T do rs7903146 está inversamente associado com o Beta HOMA% log-transformado (considerado uma medida da secreção basal de insulina), mas não foram encontradas associações entre esse polimorfismo e resistência à insulina ou síndrome metabólica. Portanto, a explicação mais provável seria a de que os indivíduos não-diabéticos portadores de genótipos CT ou TT teriam maiores elevações na glicemia pós-prandial, ainda em níveis não-diabéticos, do que indivíduos CC e, conseqüentemente, maior gravidade e incidência de eventos cardiovasculares. Como o rs7903146 do *TCF7L2* é o polimorfismo de risco mais significativo para o DM2 em diversos estudos e, conseqüentemente, contribui de forma importante para a capacidade individual de secreção de insulina mediada pelas características genéticas, supõe-se, por isso, que apenas o *TCF7L2* foi isoladamente associado a eventos cardiovasculares, mantendo um padrão de resultados semelhante à combinação de todos os polimorfismos estudados.

Infelizmente, não havia informações sobre hemoglobina glicada ou glicemia pós-prandial em nosso estudo para que se pudesse avaliar essa

correlação. Considerando a glicemia como uma variável contínua, sem limiares associados com o risco de eventos cardiovasculares, foi avaliada a influência da glicemia de jejum nos resultados, mas não foram encontradas diferenças das glicemias entre os genótipos em indivíduos não-diabéticos. Além disso, após o ajuste para vários fatores de risco cardiovasculares estabelecidos, incluindo a glicemia de jejum no momento da randomização (linha de base), os genótipos CT e TT permaneceram significativamente associados à gravidade da DAC e eventos cardiovasculares, principalmente a morte por todas as causas. Portanto, a informação preditiva do genótipo do *TCF7L2* não pôde ser substituída pelo valor da glicemia de jejum.

Contudo, não se pode excluir a possibilidade de um outro mecanismo influenciar nesse resultado. Alguns estudos têm observado que a ativação da transcrição do TCF-4, codificado pelo *TCF7L2*, está relacionada com a via de sinalização do fator nuclear-kB (NF-kB), que regula as vias de sinalização inflamatórias (Wang, Xiao *et al.*, 2002). Dessa forma, existe a hipótese de que o NF-kB poderia estar envolvido em um mecanismo epistático adicional que atuaria na própria parede vascular e isso não pode ser descartado com os dados desse estudo.

O infarto do miocárdio prévio foi significativamente mais frequente em indivíduos não-diabéticos com genótipos CT e TT que em sujeitos CC, na população do projeto MASS II. No entanto, este fator não foi independentemente associado com eventos cardiovasculares compostos, assim como foram o genótipo do *TCF7L2*, o tabagismo e a idade. Portanto, a

maior prevalência de infarto do miocárdio, no momento da randomização, não explica completamente a associação entre *TCF7L2* e eventos cardiovasculares compostos. Na verdade, esses dados até contribuem para a ideia de que os indivíduos não-diabéticos carreadores do alelo T do rs7903146 apresentam DAC mais grave.

Bielinski *et al.* (Bielinski, Pankow *et al.*, 2008) publicaram recentemente um estudo em que os SNPs rs7903146, rs12255372, rs7901695, rs7895340 rs11196205 do gene *TCF7L2* foram avaliados com um objetivo similar, usando mais de 13.000 indivíduos da população do estudo ARIC (*Atherosclerosis Risk in Communities*). No entanto, esses autores não demonstraram uma associação significativa entre qualquer SNP do *TCF7L2* e a incidência de doença coronariana, acidente vascular encefálico isquêmico, prevalência de doença arterial periférica ou mortalidade por todas as causas na coorte completa, nem tampouco quando a população foi estratificada por raça ou diagnóstico de diabetes. Embora os resultados sejam divergentes, há grandes diferenças metodológicas entre os dois estudos: no estudo ARIC, todos os pacientes com doença arterial coronariana prevalente foram excluídos da análise e, conseqüentemente, os pacientes estavam livres de DCV prévias no início do estudo. No entanto, no presente estudo, foram avaliadas duas populações: a primeira contendo pacientes com suspeita de DAC (população de estudo de corte transversal) e a outra com angina estável por DAC multi-arterial (coorte do projeto MASS II). Estas diferenças metodológicas em relação às populações estudadas, bem como no risco

*a priori* de eventos cardiovasculares nessas populações, poderiam explicar os resultados aparentemente contraditórios. Nesse contexto, outro estudo também encontrou uma associação significativa entre eventos macrovasculares (notadamente acidente vascular encefálico) e a combinação de alelos de risco relacionados à vulnerabilidade da placa aterosclerótica e eventos aterotrombóticos. Da mesma forma como observado no presente trabalho, a análise dos alelos individualmente não mostrou associações significativas, mas a informação genética combinada aumentou em 15% por cada alelo de risco a chance de eventos, independentemente dos fatores de risco convencionais (Katakami, Takahara *et al.*). Além disso, como comentado anteriormente, indivíduos diabéticos apresentam um maior risco de ulceração da placa aterosclerótica e consequente trombose (Silva, Escobar *et al.*, 1995) e, portanto, poder-se-ia sugerir que a hiperglicemia teria maior efeito na ulceração do que na formação da placa aterosclerótica.

## **5.4 Análise crítica e limitações dos resultados do Objetivo Específico 1**

Mesmo reconhecendo que as interpretações desse estudo em relação ao gene *TCF7L2* foram derivadas da avaliação de duas populações diferentes e independentes, o que fortalece as conclusões do presente trabalho, pode-se discutir que o subgrupo de indivíduos

diabéticos tinha um menor número de componentes e, dessa forma, não se pode excluir que esse menor número de indivíduos, além da maior prevalência de DAC e eventos cardiovasculares, corresponde a um baixo poder estatístico para se detectar a mesma associação observada entre DAC, eventos cardiovasculares e escores genotípicos em indivíduos não-diabéticos. Entretanto, mesmo com um tamanho de amostra relativamente pequena de indivíduos não-diabéticos, o presente estudo foi capaz de alcançar um resultado significativo e com poder estatístico adequado, devido ao relativamente grande tamanho do efeito encontrado. O maior problema de amostras pequenas seria o erro tipo beta (tipo II), ou seja, resultados falso-negativos, mas a magnitude do efeito encontrado foi capaz de superar o pequeno tamanho amostral. Considerando ainda que se trata de um estudo de associação genética e como tal deve ser interpretado mais como gerador de hipóteses que como comprovador do nexo de causalidade, uma outra limitação é a possibilidade de outras variáveis que não foram estudadas (além dos conhecidos fatores de risco cardiovascular) desempenhassem um papel significativo nos resultados encontrados ou mesmo que a maioria dos SNPs associados ao DM2 não sejam variantes causais. Se isso for verdadeiro, significa que o poder preditivo dos *loci* de susceptibilidade provavelmente estaria subestimado (Lango, Palmer *et al.*, 2008).



## 5.5 Combinação da informação de variantes genéticas de risco para diabetes: prováveis implicações na fisiopatologia do DM2

Há cada vez mais evidências que sugerem que o componente genético relacionado ao DM2 tem maiores implicações na disfunção das células betas pancreáticas que na resistência à insulina. Oito dos 10 polimorfismos estudados têm sido associados com distúrbios da secreção de insulina na população geral: as variantes de risco dos genes *TCF7L2*, *CDKAL1*, *SLC30A8*, *HHEX*, *CDKN2A/B*, *KCNJ11* e *IGF2BP2* predisõem ao diabetes tipo 2 por reduzir a função das células beta e secreção de insulina (Perry e Frayling, 2008). Da mesma forma, a maioria dos outros genes, que não foram avaliados no presente estudo e que são confirmadamente ou potencialmente associados ao risco de diabetes, também estão relacionados à função da célula beta pancreática (Staiger, Machicao *et al.*, 2009). Nesses casos, a avaliação da função pancreática foi realizada por índices substitutos, baseados em cálculos da glicemia e insulinemia em jejum ou foram derivados de testes de tolerância à glicose (por medições da insulina ou peptídeo C) e, dessa maneira, não permite inferências a respeito de aspectos relacionados à disfunção da célula beta, como a maturação da molécula de insulina, sensibilidade da célula beta à glicose ou às incretinas. A partir dessas estimativas da função pancreática, mecanismos fisiopatológicos mostrando como estas variantes genéticas comuns prejudicam a função das células betas pancreáticas só foram propostos para

os poucos genes, após estudos *in vitro* e *in vivo* (Staiger, Machicao *et al.*, 2009). De acordo com a fisiopatologia do DM2, a resistência à insulina faz com que as células betas pancreáticas aumentem a capacidade de secreção de insulina de forma compensatória, para que se mantenha a normoglicemia. Dessa forma, defeitos geneticamente determinados apenas se tornarão evidentes na presença da resistência à insulina. A resistência à insulina é, portanto, considerado um passo inicial e fundamental na patogênese do DM2 (Staiger, Machicao *et al.*, 2009). Embora a relação causa-efeito não seja absoluta, acredita-se que a resistência à insulina seja resultado da obesidade (principalmente visceral) e é predominantemente causada por fatores ambientais, como dieta rica em calorias e/ou sedentarismo, apesar de alguns estudos revelarem que certas variantes genéticas interferem na sensibilidade à insulina, sem influenciar a massa total de gordura.

Isso posto, é importante salientar que o DM2 é uma condição que tem um importante componente ambiental e cada polimorfismo associado ao DM2 contribui apenas com um pequeno efeito no risco de desenvolver a doença, como observado pelos relativamente baixos valores dos ORs encontrados (ao contrário de alguns fatores de risco clínico). Dessa forma, pode-se supor que as condições ambientais aumentam o risco de DM2, principalmente (mas não absolutamente) por prejudicar a sensibilidade à insulina (por meio da relação entre a obesidade visceral-sedentarismo-resistência à insulina). Caso o indivíduo tenha uma capacidade suficiente de secreção de insulina pelas células beta, ele será capaz de aumentar a

secreção de forma compensatória à ação da insulina prejudicada e, conseqüentemente, não desenvolveria hiperglicemia. No entanto, se o indivíduo apresenta uma combinação de alelos de risco (e, conseqüentemente, menor capacidade de secreção insulínica pancreática), ele não será capaz de aumentar a secreção de insulina e hiperglicemia se desenvolverá. No entanto, as variantes genéticas atuais não são absolutamente discriminatórias para a avaliação do risco genético de DM2 de um indivíduo ou de uma família, mas podem explicar uma parcela da herdabilidade do DM2 (Lango, Palmer *et al.*, 2008). Portanto, alguns autores sugerem que os SNPs associados ao DM2 poderiam contribuir significativamente para a previsão do risco populacional, bem como o rastreamento dessas variantes poderia reduzir o número de indivíduos necessários para ser incluídos em programas de prevenção de DM2 (Njolstad, Molven *et al.*, 2009).

Além disso, supõem-se algumas razões pelas quais genes relacionados à resistência à insulina não são identificados como genes de susceptibilidade ao DM2 (Staiger, Machicao *et al.*, 2009): 1) pode realmente existir menos genes de resistência à insulina ou alelos de risco que, no entanto, quando em conjunto, poderia conferir um aumento substancial no risco de diabetes, mas isso ainda precisa ser determinado; 2) os genes de resistência à insulina podem depender estritamente da interação com os fatores ambientais específicos para causar DM2 e estes fatores podem ser ainda desconhecidos e, portanto, podem não ter sido devidamente avaliados em estudos anteriores; 3) os genes de resistência à insulina podem estar

sub-representados nas matrizes de SNPs utilizadas nos estudos que utilizaram a tecnologia de GWA. O desenvolvimento de matrizes que avaliem maiores densidades de SNPs, com uma cobertura quase completa do genoma, em breve, pode superar este problema; 4) Pode haver uma grande quantidade de genes relacionados à resistência à insulina, exercendo individualmente apenas efeitos muito pequenos. Portanto, a detecção desses SNPs exigiria coortes enormes bem definidas, abrangendo centenas de milhares de casos, ou, alternativamente, grandes grupos de vários milhares de indivíduos completamente fenotipados por meio de métodos que avaliem fidedignamente a sensibilidade à insulina (como o *clamp* euglicêmico hiperinsulinêmico), que são onerosos e de difícil execução.

## **5.6 Efeitos preditivos dos Escores Genotípicos na avaliação de risco cardiovascular em portadores de doença arterial coronariana crônica**

Diferente de outros estudos que utilizaram a combinação da informação de diversas variantes genéticas comuns (SNPs) apenas para predição do diagnóstico de DM2, o presente estudo também avaliou os efeitos dos escores genéticos em predizer a incidência de eventos cardiovasculares em uma população de pacientes com DAC estável. Nesse contexto, o projeto MASS II é considerado uma excelente ferramenta para responder esta pergunta.

Os dados do presente estudo sugerem que escores genéticos, criados conforme associação com o diagnóstico de DM2, são também capazes de prever eventos cardiovasculares em indivíduos não-diabéticos. Foi observado que os indivíduos não-diabéticos que tiveram os maiores valores dos Escores 2 e 3 (mas não o escore 1) tinham significativamente mais desfechos cardiovasculares compostos (óbito, infarto do miocárdio e angina refratária necessitando revascularização ou nova ATC) do que aqueles que tinham valores mais baixos (Escore 2: OR = 2,31, valor de  $p = 0,02$ ; Escore 3: OR = 1,94, valor de  $p = 0,035$ ). Nos indivíduos diabéticos, no entanto, não foi encontrada associação entre os desfechos cardiovasculares e os escores genéticos. Essa informação é ainda mais valorizada com o resultado significativo da interação entre essas variáveis. Mesmo após o ajuste para vários fatores de risco estabelecidos para doença cardiovascular (hipertensão arterial, infarto do miocárdio prévio, idade, sexo, tabagismo e níveis de colesterol total, LDL-colesterol, HDL-colesterol e triglicérides) em um modelo de regressão de Cox, ambos os Escores 2 e 3 foram significativamente e independentemente associados aos eventos cardiovasculares (*Hazard ratio* para o aumento de 1 ponto: Escore 2 = 1,16, valor de  $p = 0,022$ ; Escore 3 = 1,08, valor de  $p = 0,015$ ). Mesmo após decomposição dos desfechos combinados, resultados semelhantes foram encontrados em relação à mortalidade por todas as causas em indivíduos não-diabéticos.

Haffner *et al.* (Haffner, Lehto *et al.*, 1998) já demonstraram que a incidência de infarto do miocárdio em sete anos tanto em indivíduos não-

diabéticos com infarto do miocárdio prévio quanto em indivíduos diabéticos sem infarto do miocárdio prévio é semelhante (cerca de 20%). Curiosamente, os indivíduos que apresentaram os maiores valores dos escores genotípicos tiveram taxas de eventos cardiovasculares comparáveis aos que já tinham diabetes. Portanto, a hipótese levantada é que indivíduos não-diabéticos com alto risco genético, de acordo com a combinação de variantes comuns, apresentam comportamento clínico semelhante aos pacientes diabéticos, isto é, apresentam um elevado risco cardiovascular.

Até o momento, não há confirmação da associação entre esses polimorfismos e outros conhecidos fatores de risco para doenças cardiovasculares, além da homeostase glicêmica. Contudo, três diferentes estudos de GWA mostraram uma associação entre DAC ou infarto do miocárdio e SNPs próximos aos genes *CDKN2A* e *CDKN2B* (Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls, 2007; Helgadóttir, Thorleifsson *et al.*, 2007; Mcpherson, Pertsemlidis *et al.*, 2007), que também foram avaliados no presente estudo. Em um estudo de replicação, as associações com DAC e DM2 foram independentes entre si e esses SNPs não foram associados com alterações nos níveis de colesterol, fibrinogênio, albumina, ácido úrico, bilirrubina e homocisteína (Broadbent, Peden *et al.*, 2008). Portanto, novamente se supõe que a associação entre a combinação de alelos de risco, DM2 e doenças cardiovasculares ocorra devido à elevação da glicemia, mesmo em níveis considerados não-diabéticos, embora não pode ser desconsiderada a possibilidade da influência de outros fatores não-glicêmicos.

Como citado anteriormente, o polimorfismo rs7903146 do gene *TCF7L2* é o principal fator de risco genético para o DM2. Esta variante foi consistentemente replicada em diferentes populações em todo o mundo e diferentes estudos encontraram maiores valores de *odds ratio* para rs7903146, com fins de predição do risco de DM2 entre todas as variantes conhecidas (Perry e Frayling, 2008). Como houve uma associação independente entre o polimorfismo rs7903146 do gene *TCF7L2* e DM2 (Marquezine, Pereira *et al.*, 2008) e do mesmo polimorfismo e eventos cardiovasculares na população do projeto MASS II, foi considerado importante avaliar se a associação entre os escores genéticos, DM2 e doenças cardiovasculares ocorreu apenas devido ao peso relativo do *TCF7L2*. Contudo, mesmo após a exclusão das informações do *TCF7L2* da análise, os Escores 1, 2 e 3 sem *TCF7L2* ainda mantinham a capacidade de predizer o diagnóstico de DM2. No entanto, houve uma piora significativa no desempenho dos escores sem os dados do *TCF7L2*. Além disso, os Escores 2 e 3 sem informação do *TCF7L2* ainda permaneceram significativamente associados com eventos cardiovasculares compostos (principalmente a mortalidade). Somente quando vários ajustes para vários outros fatores de risco foram feitos, um nível marginal de significância foi encontrada (tendência de associação). Para um ponto de aumento em cada Escore: HR = 1,12, valor de p = 0,084 para o Escore 2 sem *TCF7L2*; e HR = 1,076, valor de p = 0,064 para o Escore 3 sem *TCF7L2*). Isto denota que o polimorfismo rs7903146 do *TCF7L2* apresenta uma significativa importância relativa, mas não é o único marcador responsável pelos resultados encontrados.

Além disso, a proposta dos escores de risco genético também é capaz de estratificar os indivíduos não-diabéticos em diferentes graus de risco de eventos cardiovasculares, permitindo, assim, identificar os indivíduos não-diabéticos com risco semelhante àquele de indivíduos diabéticos. E esta estratificação foi além dos fatores de risco cardiovascular estabelecidos. Portanto, esses dados sugerem que um subgrupo de indivíduos com risco potencialmente aumentado de se tornarem diabéticos poderia ser alvo de esforços preventivos mais precoces e rigorosos, semelhante ao que é preconizado para indivíduos diabéticos com doença arterial coronariana.

## **5.7 Análise crítica e limitações dos resultados do objetivo específico 2**

Além das limitações anteriormente citadas em relação ao desenho do estudo, que também podem ser aplicadas ao objetivo específico 2, é importante destacar que, atualmente, existem cerca de 25 variantes genéticas de risco associadas com DM2 (Dupuis, Langenberg et al., ; Zeggini, Scott et al., 2008) e apenas 10 polimorfismos foram usados no presente estudo. Na verdade, no momento de preparação do projeto de tese, apenas os 10 SNPs estudados haviam sido replicados em diferentes populações e metanálises. Nos anos posteriores, com o aumento do número de estudos e populações incluídas, essas outras variantes de risco tiveram a



associação com DM2 confirmada. Contudo, a ideia que se surge é que a não-inclusão das informações desses outros marcadores não invalida os resultados encontrados. Pelo contrário, se os resultados descritos no presente estudo foram encontrados utilizando-se apenas 10 SNPs, imagine-se que se forem incluídas as informações de outros polimorfismos comprovadamente associados ao DM2, a capacidade de predição dos escores genéticos seria ainda maior. Por sua vez, como a magnitude da associação (demonstrada pelos odds ratios) entre essas outras variantes de risco e DM2 é menor que as incluídas neste estudo (Perry e Frayling, 2008), estima-se que o aumento na capacidade preditiva desses novos escores genotípicos não seja tão substancial como se poderia supor. Para corroborar esta hipótese, os estudos que utilizaram um maior número de variantes genéticas para a predição de DM2 (mas não de eventos cardiovasculares) em populações diversas (que não a brasileira) encontraram resultados semelhantes ao do presente trabalho (Weedon, McCarthy et al., 2006; Lango, Palmer et al., 2008; Lyssenko, Jonsson et al., 2008; Van Hoek, Dehghan et al., 2008). Houve mudanças nos critérios diagnósticos de DM2, desde a data inicial de inclusão dos pacientes no estudo MASS II, mas como foi realizada uma reclassificação dos pacientes de acordo com os critérios vigentes, a possibilidade de falha de classificação foi minimizada e possivelmente não compromete a validade dos resultados encontrados.

Os indivíduos participantes do estudo MASS II, que foi utilizado para responder o questionamento levantado no objetivo específico 2, foram arrolados em um ensaio clínico randomizado, que objetivou a comparação

entre três modalidades terapêuticas possíveis para pacientes com DAC multiarterial e com função ventricular preservada. No entanto, o presente estudo não utilizou em seu desenho esse processo de randomização, isto é, os indivíduos participantes não foram aleatorizados em grupos semelhantes, de acordo com o tipo de tratamento utilizado, como no estudo original do projeto MASS II. Como os participantes passam, então, a ser seguidos em relação aos eventos cardiovasculares descritos e foram separados de acordo com os grupos genotípicos (mas essa separação não corresponde a uma verdadeira intervenção), dessa forma ocorre uma mudança na característica do desenho do estudo, que passa a ser considerado como um estudo de coorte (longitudinal ou estudo de incidência). Nesse contexto, o presente estudo poderia ser caracterizado como um estudo associação genética entre variantes de risco (genéticos) e incidência de eventos cardiovasculares. Isso posto, como comentado anteriormente, o estabelecimento de umnexo de causalidade é limitado pela impossibilidade de se controlar todas as variáveis que poderiam influenciar nos resultados. Portanto, as inferências que surgem com a análise dos dados não são, sobremaneira, definitivas.

Uma importante limitação desse trabalho foi que a população não foi estratificada pela etnia. Portanto, não é possível excluir que algum(ns) polimorfismo(s) estudado(s) poderia(m), em vez de prever eventos cardiovasculares, prever subpopulações étnicas que apresentam um maior risco de eventos cardiovasculares. Resultados aparentemente contraditórios poderiam ser explicados, em parte, pela utilização de SNPs como marcadores

de risco genético em populações diferentes daquelas em que descoberta a associação entre SNP e o fenótipo ou doença. Contudo, considerando que a população brasileira é altamente miscigenada, composta por ancestralidade europeia, africana e indígena, essa possibilidade de uma variante de risco estar associada com um subgrupo étnico específico fica reduzida. Ainda neste contexto, é importante lembrar que os dados dos polimorfismos estudados têm sido replicados em diversas populações com determinantes genéticos distintos. Mesmo assim, há uma limitação em extrapolar os resultados para outras populações, mesmo para uma população caucasiana em que essas variantes de risco foram descobertas.

Outra limitação desta parte do estudo é com relação às validades interna e externa de suas inferências. Com relação à validade interna, deve-se lembrar que indivíduos do estudo MASS II deveriam necessariamente ser candidatos ou à revascularização cirúrgica, ou percutânea, ou serem candidatos a permanecerem em tratamento clínico (medicamentoso). Desta maneira, o próprio critério de inclusão limita a extrapolação dos resultados para além de indivíduos com doença multiarterial e função ventricular esquerda preservada. Já com relação à validade externa, a limitação é ainda mais importante, visto que não é fidedignamente representativa dos pacientes com DAC multiarterial, uma vez que grande parte dessa população tem função ventricular prejudicada, bem como a extrapolação dos resultados para uma população geral de indivíduos não-diabéticos pode ser inadequada. Considerando ainda que se trata de um estudo de associação genética, para se comprovar a validade e a dimensão das estimativas de

risco encontradas no presente estudo, se fazem necessários estudos de replicação, utilizando as mesmas variantes genéticas estudadas em uma população com características genéticas e ambientais semelhantes (Bracken, 2005). Já para se extrapolar os resultados em outras populações com DAC (por exemplo, uniarterial ou com função ventricular diminuída), a validação dos resultados também se faz necessária.

Finalmente, mesmo excluindo os indivíduos que se tornaram diabéticos na análise final, esta exclusão foi baseada nos valores de glicemia de jejum. Caso tivesse sido realizado um teste oral de tolerância à glicose (ou até hemoglobina glicada) nos indivíduos ao final do estudo, a classificação em indivíduos diabéticos e não-diabéticos ou estados intermediários, como o de intolerância à glicose, poderia ter sido diferente.

## **5.8 Associação entre as variantes genéticas e o diagnóstico de DM2**

Evidências crescentes têm encontrado sucesso na confirmação da associação entre vários polimorfismos e risco de DM2 em diferentes populações. Inicialmente (objetivo secundário 1), foi avaliada a capacidade que a adição de um único polimorfismo (*TCF7L2*) tinha de melhorar a capacidade diagnóstica de um validado modelo composto apenas de variáveis clínicas e foi encontrado que a inclusão da informação do genótipo de risco do *TCF7L2* não melhorou de forma significativa a performance do

modelo. Contudo, alguns autores afirmam que a utilização de uma combinação de várias dessas variantes comuns associadas ao diabetes pode sim ser útil para se prever o risco da doença (Lango, Palmer *et al.*, 2008). Dessa forma, no objetivo secundário 2 do presente estudo, os resultados demonstraram que todos os três escores genéticos foram associados com a chance de diabetes, em uma população de portadores de doença coronariana multiarterial. Mesmo se supondo que apenas o alelo T do gene *TCF7L2* foi isoladamente associado com DM2, este estudo encontrou que a utilização das informações combinadas de vários alelos de risco, ao invés de usar estas informações separadamente, aumentou a capacidade de predição de DM2 que era obtida pelos fatores de risco clínicos. Ambos os Escores 1, 2 e 3 apresentaram uma capacidade discriminatória, avaliada pela área sob a curva ROC (AUC), variando entre 0,592 e 0,616. Estes valores são comparáveis aos resultados da maioria dos outros estudos em diferentes populações, que avaliaram de 3 a 18 polimorfismos, tanto em estudos no formato caso-controle (prevalência de diabetes) quanto em estudos prospectivos (incidência de diabetes), que encontraram AUCs variando entre 0,60 e 0,62 (Weedon, McCarthy *et al.*, 2006; Lango, Palmer *et al.*, 2008; Lyssenko, Jonsson *et al.*, 2008; Van Hoek, Dehghan *et al.*, 2008). No entanto, outros autores encontraram modelos com desempenhos melhores: Miyake *et al.* (Miyake, Yang *et al.*, 2009) encontraram uma AUC equivalente a 0,72 na população japonesa, após avaliação de 18 polimorfismos e Cauchi *et al.* (Cauchi, Meyre *et al.*, 2008) criaram um modelo de previsão de risco baseado em 15 alelos associados

ao DM2, em indivíduos franceses, cujo valor de AUC foi 0,86. Contudo, para análises globais de doenças complexas, alguns autores estimaram que 20 a 25 variantes de risco, com as frequências alélicas superiores a 0,1 e com *odds ratios* de 1,5, são necessárias para uma AUC de cerca de 0,8 (Yang, Khoury *et al.*, 2005).

Embora os escores genotípicos tenham alcançado apenas uma modesta capacidade discriminatória para o diagnóstico de DM2, a utilização da combinação de informações de diversos alelos de risco, juntamente com variáveis clínicas de risco para DM2 (idade, sexo, IMC, presença de hipertensão arterial e tabagismo), teve uma capacidade preditiva significativamente melhor que o uso apenas das variáveis clínicas (AUC variando de 0,620 a 0,637 para modelos com variáveis clínicas e com informações dos escores genéticos comparados com AUC de 0,577 para o modelo apenas com variáveis clínicas, valor de  $p < 0,05$  para todos). Esses resultados estão na mesma direção que outros estudos, nos quais a adição de informações de polimorfismos de risco aumentaram a capacidade preditiva de DM2 (Van Hoek, Dehghan *et al.*, 2008; Miyake, Yang *et al.*, 2009), embora em alguns estudos este aumento não foi significativo (Lango, Palmer *et al.*, 2008; Lyssenko, Jonsson *et al.*, 2008; Meigs, Shrader *et al.*, 2008).

O valor da AUC do modelo clínico, neste estudo, foi comparável ao estudo de Van Hoek *et al.* (Van Hoek, Dehghan *et al.*, 2008) e ao modelo contendo apenas as variáveis clínicas mais simples do estudo por Meigs *et al.* (Meigs, Shrader *et al.*, 2008). Contudo, os modelos clínicos de outros

estudos (Lyssenko, Jonsson *et al.*, 2008; Meigs, Shrader *et al.*, 2008; Miyake, Yang *et al.*, 2009) conseguiram um melhor desempenho (AUC variando entre 0,72 e 0,9), possivelmente porque incluíram outras variáveis, como valores de triglicérides, HDL, glicemia de jejum e história familiar de diabetes. Nesses casos, a inclusão da informação genética proporcionou um aumento discreto na AUC, embora significativo. No presente estudo, a inclusão de variáveis laboratoriais não melhorou o desempenho do modelo baseado em variáveis clínicas, bem como a informação de história familiar de diabetes não estava disponível. Dessa forma, é possível que o desempenho relativamente pior do modelo clínico do presente estudo tenha influenciado o grau de melhora do desempenho preditivo após a inclusão da informação genética combinada.

Por fim, é importante destacar que a interpretação desses resultados deve ser feita em um contexto de probabilidade pré-teste, isto é, a probabilidade de doença antes da aplicação dos resultados de um teste diagnóstico qualquer. O conhecimento das características de um teste diagnóstico é útil apenas se houver o conhecimento das características basais da situação envolvida. Por exemplo: se um determinado teste aumenta probabilidade de doença em 50%, esta informação deve ser entendida de forma completamente diferente se a probabilidade pré-teste for de 10% ou de 60%. E a probabilidade pré-teste é razão direta, no contexto do diagnóstico de DM2 de que trata o presente estudo, da prevalência de DM2 nas populações estudadas. No objetivo secundário 1 (adição de apenas um polimorfismo ao modelo clínico), deve-se interpretar os

resultados se baseando em uma prevalência de DM2 mais baixa (em torno de 8%) e com uma capacidade diagnóstica do modelo clínico moderada. Já para o objetivo secundário 2 (adição da combinação de polimorfismos ao modelo clínico), a interpretação se baseia em uma prevalência de DM2 de mais de 30% além de um modelo clínico de performance modesta. Entretanto, é importante ressaltar que a possibilidade de melhorar a capacidade diagnóstica em um contexto de probabilidade pré-teste maior que 30% não pode ser negligenciada.

## 5.9 Considerações sobre os escores genéticos

A ideia de construir diferentes modelos surgiu com o objetivo de se encontrar o melhor modelo para prever DM2: recessivo, aditivo ou ponderado. Na verdade, os escores genéticos transmitem informações diferentes: o Escore 1 analisou os polimorfismos em um modelo recessivo. De antemão, supunha-se que esse modelo teria resultados mais modestos, visto que a incorporação de polimorfismos em modelos de predição de risco pode até funcionar para um gene específico, mas ao se analisar um conjunto de variantes, inicialmente se pensaria que um modelo aditivo ou outro modelo que atribuiria pesos diferentes às variáveis genéticas seriam mais eficazes, o que realmente aconteceu. O Escore 2 correspondeu ao número de alelos de risco e o Escore 3 conferiu pesos diferentes para cada polimorfismo de risco avaliado para o diagnóstico do DM2, em um modelo complexo. Curiosamente,



estes dois modelos apresentaram resultados semelhantes na avaliação de eventos cardiovasculares compostos em indivíduos não-diabéticos: aqueles com maiores valores desses escores apresentaram maior incidência de eventos cardiovasculares, principalmente a mortalidade. Considerando, então, dois dos três modelos (o Escore 1 não foi tão bom quanto os demais para prever eventos cardiovasculares), supõe-se que a principal explicação para os resultados encontrados é realmente o número de alelos de risco. Na verdade, comparando com o Escore 2, o Escore 3 conferiu valores significativamente superiores aos portadores de alelos de risco dos genes *TCF7L2*, *IGFBP2* e *PPARG* e como o *TCF7L2* foi independentemente associado com eventos cardiovasculares na população do estudo MASS II, isso talvez explique a ligeira preponderância do Escore 3, em relação ao Escore 2, na predição de eventos.

### **5.10 Possíveis aplicações clínicas dos fatores de risco genético associados ao DM2**

Ao contrário dos casos de diabetes monogênico, para os quais o entendimento das bases genéticas do diabetes tem guiado a prática clínica, a maior parte desse conhecimento não é aplicado para o DM2, notadamente uma condição poligênica. Há um grande otimismo que o conhecimento de novos *loci* de risco vai alargar a compreensão clínica do DM2 e direcionar abordagens específicas para pacientes específicos. Na

teoria, os estudos de associação genética podem causar impacto em três diferentes áreas: nosologia da doença, predição de risco e farmacogenética (Majithia e Florez, 2009).

A nosologia corresponde à classificação da doença e, apesar do diagnóstico clínico do diabetes ser simplesmente baseado nos níveis de glicemia de jejum ou após sobrecarga, é naturalmente aceita a ideia de que a hiperglicemia representa um ponto comum na síndrome, mas não implica em uma patogênese única. Portanto, todas as doenças que cursam com hiperglicemia são denominadas diabetes mellitus, mas possivelmente são condições diversas com fisiopatologias subjacentes diferentes. Um exemplo, em que estudos genéticos mudaram a nosologia no diabetes, se refere ao MODY5: após a descoberta de que mutações no gene *HNF1B* correspondiam ao defeito genético único encontrado nesses pacientes, foi observado que pacientes com MODY5 sempre tinha cistos renais, resultando na definição de uma nova síndrome (Majithia e Florez, 2009). Portanto, a descoberta de um genótipo subjacente a um subgrupo de pacientes diabéticos permitiu uma melhoria na fenotipagem clínica. E nesse contexto, estão sendo realizados estudos que visam a avaliar se os marcadores genéticos associados ao DM2 podem refinar o fenótipo de indivíduos diabéticos ou tentar relacioná-los a vias fisiopatológicas específicas.

Como o genótipo está determinado desde o nascimento, é bastante atrativa a utilização da informação de marcadores genéticos de risco para predição de doenças futuras e, conseqüentemente, estabelecer medidas

mais precoces de prevenção ou retardamento do início da doença. Nesta área, encontra-se os modelos de predição clínica, abordados no presente trabalho, em relação a eventos cardiovasculares e DM2, no formato de escores. Quando incorporados aos modelos que incluíam apenas variáveis clínicas, a informação dos genótipos combinados foi capaz de melhorar a estatística verificada pela AUC. E nesses modelos, apenas o sexo e o genótipo (e em alguns modelos, a história familiar) são determinados desde o nascimento e permanecerão certamente estáveis durante toda a vida do indivíduo. Isso apoia a ideia de que fatores genéticos são úteis na detecção precoce de grupos de risco antes de fatores de risco clínico, como o IMC ou glicemia de jejum se manifestar, permitindo que os médicos elaborem medidas eficazes e intervenções em longo prazo, como a modificação do estilo de vida. Além disso, se muitos *loci* adicionais forem descobertos, eles poderão ser incorporados aos escores de risco, com o intuito de melhorar o poder discriminatório da informação genética. Contudo, como citado anteriormente, quanto melhor for o modelo atual (maior os valores prévios de AUC), maiores terão de ser os limiares de informação genética adicional, para que possam contribuir de forma significativa (Majithia e Florez, 2009). Por fim, a informação resultante terá de ser validada em populações específicas, assim como foi realizada no presente estudo.

A farmacogenética corresponde ao estudo das interações entre *loci* genéticos e a terapia farmacológica. Sua aplicação clínica consiste em direcionar terapias específicas para subgrupos de pacientes (determinados geneticamente) com base na eficácia e efeitos colaterais dos medicamentos.

Embora as abordagens farmacogenéticas utilizem normalmente *loci* candidatos e suas respectivas vias farmacológicas, a compreensão efetiva das funções biológicas de *loci* associados ao DM2 ainda não foi totalmente estabelecida. Um exemplo disso se aplica ao gene *TCF7L2*, que desempenha um papel na secreção de insulina, por influenciar a via das incretinas (Shu, Sauter *et al.*, 2008). Contudo, um estudo avaliou a presença do alelo de risco T e a resposta à terapia hipoglicemiante oral: os portadores do genótipo TT apresentaram maiores taxas de falha à utilização de sulfonilureias (que estimulam a secreção pancreática de insulina), enquanto que a resposta à metformina (que atua sobre a resistência à insulina) foi independente do genótipo (Pearson, Donnelly *et al.*, 2007). Esses dados sugerem que a seleção do tratamento do DM2 pode ser baseada na farmacogenômica, bem como existe a possibilidade que novas vias de resposta às medicações sejam descobertas, independentemente do mecanismo molecular do DM2.

Reconhece-se que a aplicação prática, no momento atual dos resultados do presente estudo, é limitada: em virtude dos altos custos relacionados às genotipagens, bem como a existência de outros biomarcadores (fatores de risco clínico ou laboratoriais) de custo mais acessível e que conferem alterações no risco do DM2 de magnitude igual ou superior aos marcadores genéticos, a expectativa da utilização nos próximos anos da genotipagem de variantes relacionadas ao diabetes, com fins de predição de eventos cardiovasculares ou do próprio diagnóstico de DM2, é pequena. Portanto, para agora se entende que os resultados do estudo

contribuem mais para o campo de estudo da fisiopatologia do processo de aterosclerose em indivíduos diabéticos e não-diabéticos. Contudo, com a evolução tecnológica e redução progressiva de custos das genotipagens, acredita-se que os marcadores genéticos serão fatores de risco importantes para a maioria das doenças complexas, como o diabetes. Inclusive, em um ambiente de crescente gastos em saúde, a determinação precoce de fatores de risco para várias doenças pode desencadear estratégias de prevenção em indivíduos sob maior risco e, em vias finais, reduzir os custos em saúde por diminuição dos custos em assistência à saúde.

### **5.11 Perspectivas na avaliação do risco genético do DM2**

A identificação de variantes genéticas que influenciam vias metabólicas ou a doença cardiovascular permanece um grande desafio. Para se encontrar *loci* adicionais, amostras de populações cada vez maiores ou combinação de coortes serão necessárias, e a heterogeneidade entre esses estudos poderá ser um problema significativo (Mohlke, Boehnke *et al.*, 2008). Com relação ao DM2, a metanálise de estudos de GWA, denominada DIAGRAMA (Zeggini, Scott *et al.*, 2008), tinha poder estatístico próximo de 100% para detectar SNPs de risco para DM2, cujo tamanho do efeito (representado pelo *odds ratio*) alcançasse 40% ou mais, o que torna muito improvável que *loci* comuns de grandes efeitos não tenham sido detectados (pelo menos em populações de ascendência europeia). Vale lembrar que,

dos SNPs de risco atualmente conhecidos, apenas o *TCF7L2* atinge tais estimativas de risco. No entanto, o atual catálogo de variantes genéticas humanas permitiu que os estudos de GWA capturassem apenas variantes comuns, isto é, SNPs que têm frequência do alelo raro maior que 5%). Portanto, pode-se afirmar com certeza que não existem SNPs comuns com grandes efeitos sobre a herdabilidade do DM2, mas não se pode dizer o mesmo para *loci* raros e seus respectivos tamanhos de efeitos (Majithia e Florez, 2009). Quanto às variantes comuns, os *loci* identificados pelos estudos de GWA atuais explicam apenas 5% a 10% da base genética do DM2 (Florez, 2008). Portanto, muitos *loci* adicionais, tanto raros quanto comuns, ainda podem ser descobertos e identificar esses alelos pode aumentar a parcela da herdabilidade do DM2 que pode ser explicada (Mohlke, Boehnke *et al.*, 2008). Além disso, algumas dessas falhas podem ser ultrapassadas por: 1) identificação de novos genes de risco através da aplicação de novos métodos e estratégias robustas; e 2) mapeamento fino e meticuloso, por sequenciamento completo de todas as variantes genéticas causais conhecidas que tanto estão em ligação quanto apresentam efeitos de tamanho (*odds ratio*) maiores do que o SNP inicialmente implicado. Além disso, muitos dos SNPs identificados são intrônicos ou localizados nas regiões flangeadoras 5' ou 3' dos genes e de que maneira SNPs não-codificantes influenciam a função dos genes ainda não foi totalmente esclarecido. Supõe-se que ocorra alterações nos locais de ligação de fatores de transcrição ou de *enhancer-binding proteins*, com consequentes mudanças nas taxas de transcrição de alguns genes (Staiger, Machicao *et al.*, 2009).

Outras análises, também, podem revelar conhecimentos biológicos adicionais, como estudos que avaliam fenômenos epigenéticos e interações gene-gene e gene-ambiente. Algumas perturbações ambientais, como baixo peso ao nascimento, obesidade, sedentarismo e envelhecimento, podem levar ao DM2 por afetar a expressão de genes por meio de modificações epigenéticas. Por sua vez, modificações epigenéticas do genoma proporcionam mecanismos que permitem a propagação estável da expressão de alguns genes de uma geração celular para outra, tanto por modificações nas histonas quanto por metilação do DNA (Pinney e Simmons). Por outro lado, estudos de interações gene-gene podem indicar como a combinação de variantes genéticas podem causar impacto sobre o risco de desenvolver o DM2, bem como estudos de interação gene-ambiente, especialmente interações entre fatores dietéticos e relacionados à prática de atividade física, podem futuramente elucidar os mecanismos de variabilidade interindividual (Mohlke, Boehnke *et al.*, 2008).

## **6 Conclusões**

---



1. Existe uma associação importante entre o polimorfismo rs7903146 do gene *TCF7L2* e a gravidade da doença aterosclerótica coronariana em indivíduos não-diabéticos, bem como entre este mesmo polimorfismo e a incidência de eventos cardiovasculares em sujeitos não-diabéticos portadores de doença coronariana crônica estável;
  - 1.1 Secundariamente, a adição da informação do polimorfismo do *TCF7L2* não melhorou significativamente a capacidade diagnóstica de um modelo validado na população brasileira;
2. A informação combinada dos 10 polimorfismos estudados apresentou associação com eventos cardiovasculares, principalmente mortalidade, em indivíduos não-diabéticos com doença coronariana multiarterial. Esta associação não foi observada em indivíduos diabéticos;
  - 2.1 A combinação da informação genética dos polimorfismos estudados foi capaz de melhorar significativamente algoritmos diagnósticos para DM2 baseados apenas em variáveis clínico-laboratoriais conhecidas

Estas conclusões suportam a necessidade de outros estudos, que serão importantes para a elucidação dos mecanismos implicados na associação entre doença coronariana e as variantes genéticas de risco para DM2.

## **7 Referências**

---

Agalliu I, Suuriniemi M, *et al.* Evaluation of a variant in the transcription factor 7-like 2 (TCF7L2) gene and prostate cancer risk in a population-based study. *Prostate*, v.68, n.7, May 15, p.740-7. 2008.

Altshuler D, Hirschhorn JN, *et al.* The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet*, v.26, n.1, Sep, p.76-80. 2000.

Bianchi C, Miccoli R, *et al.* Primary prevention of cardiovascular disease in people with dysglycemia. *Diabetes Care*, v.31 Suppl 2, Feb, p.S208-14. 2008.

Bielinski SJ, Pankow JS, *et al.* TCF7L2 single nucleotide polymorphisms, cardiovascular disease and all-cause mortality: the Atherosclerosis Risk in Communities (ARIC) study. *Diabetologia*, v.51, n.6, Jun, p.968-70. 2008.

Bourassa MG, Yeh W, *et al.* Long-term outcome of patients with incomplete vs complete revascularization after multivessel PTCA. A report from the NHLBI PTCA Registry. *Eur Heart J*, v.19, n.1, Jan, p.103-11. 1998.

Bracken MB. Genomic epidemiology of complex disease: the need for an electronic evidence-based approach to research synthesis. *Am J Epidemiol*, v.162, n.4, Aug 15, p.297-301. 2005.

Broadbent HM, Peden JF, *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*, v.17, n.6, Mar 15, p.806-14. 2008.

Brownlee M. Biochemistry and molecular cell biology of diabetic complications. *Nature*, v.414, n.6865, Dec 13, p.813-20. 2001.

Burcelin R, Knauf C, *et al.* Pancreatic alpha-cell dysfunction in diabetes. *Diabetes Metab*, v.34 Suppl 2, Feb, p.S49-55. 2008.

Burwinkel B, Shanmugam KS, *et al.* Transcription factor 7-like 2 (TCF7L2) variant is associated with familial breast cancer risk: a case-control study. *BMC Cancer*, v.6, p.268. 2006.

Campeau L. Letter: Grading of angina pectoris. *Circulation*, v.54, n.3, Sep, p.522-3. 1976.

Cauchi S, El Achhab Y, *et al.* TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med*, v.85, n.7, Jul, p.777-82. 2007.

Cauchi S, Meyre D, *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*, v.3, n.5, p.e2031. 2008.

Cavalot F, Petrelli A, *et al.* Postprandial blood glucose is a stronger predictor of cardiovascular events than fasting blood glucose in type 2 diabetes mellitus, particularly in women: lessons from the San Luigi Gonzaga Diabetes Study. *J Clin Endocrinol Metab*, v.91, n.3, Mar, p.813-9. 2006.

Cdc. Centers for Disease Control and Prevention. National diabetes fact sheet: general information and national estimates on diabetes in the United States, 2007. Atlanta, GA: Department of Health and Human Services, Centers for Disease Control and Prevention 2008.

Ceriello A. Postprandial hyperglycemia and diabetes complications: is it time to treat? *Diabetes*, v.54, n.1, Jan, p.1-7. 2005.

Chaitman BR, Rosen AD, *et al.* Myocardial infarction and cardiac mortality in the Bypass Angioplasty Revascularization Investigation (BARI) randomized trial. *Circulation*, v.96, n.7, Oct 7, p.2162-70. 1997.

Chimienti F, Devergnas S, *et al.* Identification and cloning of a beta-cell-specific zinc transporter, ZnT-8, localized into insulin secretory granules. *Diabetes*, v.53, n.9, Sep, p.2330-7. 2004.

Ching YP, Pang AS, *et al.* Identification of a neuronal Cdk5 activator-binding protein as Cdk5 inhibitor. *J Biol Chem*, v.277, n.18, May 3, p.15237-40. 2002.

Coutinho M, Gerstein HC, *et al.* The relationship between glucose and incident cardiovascular events. A metaregression analysis of published data from 20 studies of 95,783 individuals followed for 12.4 years. *Diabetes Care*, v.22, n.2, Feb, p.233-40. 1999.

Debniak T, Scott RJ, *et al.* CDKN2A common variants and their association with melanoma risk: a population-based study. *Cancer Res*, v.65, n.3, Feb 1, p.835-9. 2005.

Deeb SS, Fajas L, *et al.* A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. *Nat Genet*, v.20, n.3, Nov, p.284-7. 1998.

DeFronzo RA, Ferrannini E. Insulin resistance. A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease. *Diabetes Care*, v.14, n.3, Mar, p.173-94. 1991.

Diagnosis and classification of diabetes mellitus. *Diabetes Care*, v.31, n.Suppl 1, p.S55-60. 2008.

Dodge HT, Sandler H, *et al.* Usefulness and limitations of radiographic methods for determining left ventricular volume. *Am J Cardiol*, v.18, n.1, Jul, p.10-24. 1966.

Dupuis J, Langenberg C, *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*, v.42, n.2, Feb, p.105-16.

Florez JC. Clinical review: the genetics of type 2 diabetes: a realistic appraisal in 2008. *J Clin Endocrinol Metab*, v.93, n.12, Dec, p.4633-42. 2008.

Foley AC, Mercola M. Heart induction by Wnt antagonists depends on the homeodomain transcription factor Hex. *Genes Dev*, v.19, n.3, Feb 1, p.387-96. 2005.

Folsom AR, Pankow JS, *et al.* Variation in TCF7L2 and increased risk of colon cancer: the Atherosclerosis Risk in Communities (ARIC) Study. *Diabetes Care*, v.31, n.5, May, p.905-9. 2008.

Frayling TM, Timpson NJ, *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*, v.316, n.5826, May 11, p.889-94. 2007.

Fredriksson R, Hagglund M, *et al.* The obesity gene, FTO, is of ancient origin, upregulated during food deprivation and expressed in neurons of feeding-related nuclei of the brain. *Endocrinology*, Jan 24. 2008.

Furukawa Y, Shimada T, *et al.* Polymorphisms in the IDE-KIF11-HHEX gene locus are reproducibly associated with type 2 diabetes in a Japanese population. *J Clin Endocrinol Metab*, v.93, n.1, Jan, p.310-4. 2008.

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, v.447, n.7145, Jun 7, p.661-78. 2007.

Gerken T, Girard CA, *et al.* The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science*, v.318, n.5855, Nov 30, p.1469-72. 2007.

Glucose tolerance and mortality: comparison of WHO and American Diabetes Association diagnostic criteria. The DECODE study group. European Diabetes Epidemiology Group. Diabetes Epidemiology: Collaborative analysis Of Diagnostic criteria in Europe. *Lancet*, v.354, n.9179, Aug 21, p.617-21. 1999.

Goyal A, Mahaffey KW, *et al.* Prognostic significance of the change in glucose level in the first 24 h after acute myocardial infarction: results from the CARDINAL study. *Eur Heart J*, v.27, n.11, Jun, p.1289-97. 2006.

Grant SF, Thorleifsson G, *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet*, v.38, n.3, Mar, p.320-3. 2006.

Habener JF, Kemp DM, *et al.* Minireview: transcriptional regulation in pancreatic development. *Endocrinology*, v.146, n.3, Mar, p.1025-34. 2005.

Haffner SM, Lehto S, *et al.* 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*, v.339, n.4, Jul 23, p.229-34. 1998.

Helgadottir A, Thorleifsson G, *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*, v.316, n.5830, Jun 8, p.1491-3. 2007.

Hoeflich A, Wu M, *et al.* Overexpression of insulin-like growth factor-binding protein-2 in transgenic mice reduces postnatal body weight gain. *Endocrinology*, v.140, n.12, Dec, p.5488-96. 1999.

Holt RI, Simpson HL, *et al.* The role of the growth hormone-insulin-like growth factor axis in glucose homeostasis. *Diabet Med*, v.20, n.1, Jan, p.3-15. 2003.

Horikoshi M, Hara K, *et al.* Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabetologia*, v.50, n.12, Dec, p.2461-6. 2007.

Hossain P, Kowar B, *et al.* Obesity and diabetes in the developing world--a growing challenge. *N Engl J Med*, v.356, n.3, Jan 18, p.213-5. 2007.

Hueb WA, Soares PR, *et al.* Five-year follow-up of the medicine, angioplasty, or surgery study (MASS): A prospective, randomized trial of medical therapy, balloon angioplasty, or bypass surgery for single proximal left anterior descending coronary artery stenosis. *Circulation*, v.100, n.19 Suppl, Nov 9, p.II107-13. 1999.

Huopio H, Shyng SL, *et al.* K(ATP) channels and insulin secretion disorders. *Am J Physiol Endocrinol Metab*, v.283, n.2, Aug, p.E207-16. 2002.

Jin T, Liu L. The Wnt signaling pathway effector TCF7L2 and type 2 diabetes mellitus. *Mol Endocrinol*, v.22, n.11, Nov, p.2383-92. 2008.

Katakami N, Takahara M, *et al.* Accumulation of gene polymorphisms related to plaque disruption and thrombosis is associated with cerebral infarction in subjects with type 2 diabetes. *Diabetes Care*, v.33, n.2, Feb, p.390-5.

Kim WY, Sharpless NE. The regulation of INK4/ARF in cancer and aging. *Cell*, v.127, n.2, Oct 20, p.265-75. 2006.

Kirklin JW, Frye RL, *et al.* Some comments on the indications for the coronary artery bypass graft operation. *Int J Cardiol*, v.31, n.1, Apr, p.23-30; discussion 31-2. 1991.

Knowler WC, Bennett PH, *et al.* Diabetes incidence and prevalence in Pima Indians: a 19-fold greater incidence than in Rochester, Minnesota. *Am J Epidemiol*, v.108, n.6, Dec, p.497-505. 1978.

Koster JC, Permutt MA, *et al.* Diabetes and insulin secretion: the ATP-sensitive K<sup>+</sup> channel (K ATP) connection. *Diabetes*, v.54, n.11, Nov, p.3065-72. 2005.

Lango H, Palmer CN, *et al.* Assessing the combined impact of 18 common genetic variants of modest effect sizes on type 2 diabetes risk. *Diabetes*, v.57, n.11, Nov, p.3129-35. 2008.



Lanz JR, Pereira AC, *et al.* Angiotensinogen M235T polymorphism is associated with coronary artery disease severity. *Clin Chim Acta*, v.362, n.1-2, Dec, p.176-81. 2005.

Lawes CM, Parag V, *et al.* Blood glucose and risk of cardiovascular disease in the Asia Pacific region. *Diabetes Care*, v.27, n.12, Dec, p.2836-42. 2004.

Lee YH, Kang ES, *et al.* Association between polymorphisms in SLC30A8, HHEX, CDKN2A/B, IGF2BP2, FTO, WFS1, CDKAL1, KCNQ1 and type 2 diabetes in the Korean population. *J Hum Genet*, v.53, n.11-12, p.991-8. 2008.

Lyssenko V, Jonsson A, *et al.* Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med*, v.359, n.21, Nov 20, p.2220-32. 2008.

Lyssenko V, Lupi R, *et al.* Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *J Clin Invest*, v.117, n.8, Aug, p.2155-63. 2007.

Macdonald PE, De Marinis YZ, *et al.* A K ATP channel-dependent pathway within alpha cells regulates glucagon release from both rodent and human islets of Langerhans. *PLoS Biol*, v.5, n.6, Jun, p.e143. 2007.

Majithia AR, Florez JC. Clinical translation of genetic predictors for type 2 diabetes. *Curr Opin Endocrinol Diabetes Obes*, v.16, n.2, Apr, p.100-6. 2009.

Malerbi DA, Franco LJ. Multicenter study of the prevalence of diabetes mellitus and impaired glucose tolerance in the urban Brazilian population aged 30-69 yr. The Brazilian Cooperative Group on the Study of Diabetes Prevalence. *Diabetes Care*, v.15, n.11, Nov, p.1509-16. 1992.

Marquezine GF, Pereira AC, *et al.* TCF7L2 variant genotypes and type 2 diabetes risk in Brazil: significant association, but not a significant tool for risk stratification in the general population. *BMC Med Genet*, v.9, p.106. 2008.

Mcperson R, Pertsemlidis A, *et al.* A common allele on chromosome 9 associated with coronary heart disease. Science, v.316, n.5830, Jun 8, p.1488-91. 2007.

Meigs JB, Cupples LA, *et al.* Parental transmission of type 2 diabetes: the Framingham Offspring Study. Diabetes, v.49, n.12, Dec, p.2201-7. 2000.

Meigs JB, Shrader P, *et al.* Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med, v.359, n.21, Nov 20, p.2208-19. 2008.

Milicevic Z, Raz I, *et al.* Natural history of cardiovascular disease in patients with diabetes: role of hyperglycemia. Diabetes Care, v.31 Suppl 2, Feb, p.S155-60. 2008.

Miyake K, Yang W, *et al.* Construction of a prediction model for type 2 diabetes mellitus in the Japanese population based on 11 genes with strong evidence of the association. J Hum Genet, v.54, n.4, Apr, p.236-41. 2009.

Mohlke KL, Boehnke M, *et al.* Metabolic and cardiovascular traits: an abundance of recently identified common genetic variants. Hum Mol Genet, v.17, n.R2, Oct 15, p.R102-8. 2008.

Mori H, Ikegami H, *et al.* The Pro12 -->Ala substitution in PPAR-gamma is associated with resistance to development of diabetes in the general population: possible involvement in impairment of insulin secretion in individuals with type 2 diabetes. Diabetes, v.50, n.4, Apr, p.891-4. 2001.

Morrish NJ, Wang SL, *et al.* Mortality and causes of death in the WHO Multinational Study of Vascular Disease in Diabetes. Diabetologia, v.44 Suppl 2, Sep, p.S14-21. 2001.

Myakishev MV, Khripin Y, *et al.* High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers. Genome Res, v.11, n.1, Jan, p.163-9. 2001.

Nathan DM, Cleary PA, *et al.* Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med*, v.353, n.25, Dec 22, p.2643-53. 2005.

Nielsen J, Christiansen J, *et al.* A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development. *Mol Cell Biol*, v.19, n.2, Feb, p.1262-70. 1999.

Njolstad PR, Molven A, *et al.* Diabetes genetics. A seventh sense for the successful sequel of 'come together'. "The genotypes and phenotypes of diabetes". Bergen, Norway. April 22-26, 2009. 2nd meeting of 'EASD Study Group on Genetics of Diabetes' and 44th annual meeting of the 'Scandinavian Society for the Study of Diabetes'. *JOP*, v.10, n.4, p.466-71. 2009.

Omori S, Tanaka Y, *et al.* Association of CDKAL1, IGF2BP2, CDKN2A/B, HHEX, SLC30A8 and KCNJ11 with susceptibility to type 2 diabetes in a Japanese population. *Diabetes*, Dec 27. 2007.

Pascoe L, Tura A, *et al.* Common variants of the novel type 2 diabetes genes CDKAL1 and HHEX/IDE are associated with decreased pancreatic beta-cell function. *Diabetes*, v.56, n.12, Dec, p.3101-4. 2007.

Pasternack JJ. *Human Molecular Genetics*. Hoboken, New Jersey: John Wiley & Sons. 2005. 630 p.

Pearson ER, Donnelly LA, *et al.* Variation in TCF7L2 influences therapeutic response to sulfonylureas: a GoDARTs study. *Diabetes*, v.56, n.8, Aug, p.2178-82. 2007.

Pereira AC. *Estudo do polimorfismo genético C242T no gene da p22phox e a incidência de eventos cardiovasculares na doença arterial coronariana.* (Tese). Departamento de Cardiopneumologia, Faculdade de Medicina da Universidade de São Paulo, São Paulo, 2008. 200 p.

Perry JR, Frayling TM. New gene variants alter type 2 diabetes risk predominantly through reduced beta-cell function. *Curr Opin Clin Nutr Metab Care*, v.11, n.4, Jul, p.371-7. 2008.

Pesaro AE, Nicolau JC, *et al.* Influence of leukocytes and glycemia on the prognosis of patients with acute myocardial infarction. *Arq Bras Cardiol*, v.92, n.2, Feb, p.84-93. 2009.

Pinney SE, Simmons RA. Epigenetic mechanisms in the development of type 2 diabetes. *Trends Endocrinol Metab*, v.21, n.4, Apr, p.223-9.

Pires de Sousa AG, Pereira AC, *et al.* Derivation and external validation of a simple prediction model for the diagnosis of type 2 diabetes mellitus in the Brazilian urban population. *Eur J Epidemiol*, v.24, n.2, p.101-9. 2009.

Poulsen P, Kyvik KO, *et al.* Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia*, v.42, n.2, Feb, p.139-45. 1999.

Prokopenko I, Langenberg C, *et al.* Variants in MTNR1B influence fasting glucose levels. *Nat Genet*, v.41, n.1, Jan, p.77-81. 2009.

Rane SG, Dubus P, *et al.* Loss of Cdk4 expression causes insulin-deficient diabetes and Cdk4 activation results in beta-islet cell hyperplasia. *Nat Genet*, v.22, n.1, May, p.44-52. 1999.

Rijkelijkhuizen JM, Nijpels G, *et al.* High risk of cardiovascular mortality in individuals with impaired fasting glucose is explained by conversion to diabetes: the Hoorn study. *Diabetes Care*, v.30, n.2, Feb, p.332-6. 2007.

Ryan TJ, Bauman WB, *et al.* Guidelines for percutaneous transluminal coronary angioplasty. A report of the American Heart Association/American College of Cardiology Task Force on Assessment of Diagnostic and Therapeutic Cardiovascular Procedures (Committee on Percutaneous Transluminal Coronary Angioplasty). *Circulation*, v.88, n.6, Dec, p.2987-3007. 1993.

Safian RD, Feldman T, *et al.* Coronary angioplasty and Rotablator atherectomy trial (CARAT): immediate and late results of a prospective multicenter randomized trial. *Catheter Cardiovasc Interv*, v.53, n.2, Jun, p.213-20. 2001.

Sakura H, Wat N, *et al.* Sequence variations in the human Kir6.2 gene, a subunit of the beta-cell ATP-sensitive K-channel: no association with NIDDM in white Caucasian subjects or evidence of abnormal function when expressed in vitro. *Diabetologia*, v.39, n.10, Oct, p.1233-6. 1996.

Schafer SA, Tschritter O, *et al.* Impaired glucagon-like peptide-1-induced insulin secretion in carriers of transcription factor 7-like 2 (TCF7L2) gene polymorphisms. *Diabetologia*, v.50, n.12, Dec, p.2443-50. 2007.

Schwanstecher C, Meyer U, *et al.* K(IR)6.2 polymorphism predisposes to type 2 diabetes by inducing overactivity of pancreatic beta-cell ATP-sensitive K(+) channels. *Diabetes*, v.51, n.3, Mar, p.875-9. 2002.

Scott LJ, Mohlke KL, *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science*, v.316, n.5829, Jun 1, p.1341-5. 2007.

Serrano M, Hannon GJ, *et al.* A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature*, v.366, n.6456, Dec 16, p.704-7. 1993.

Shu L, Matveyenko AV, *et al.* Decreased TCF7L2 protein levels in type 2 diabetes mellitus correlate with downregulation of GIP- and GLP-1 receptors and impaired beta-cell function. *Hum Mol Genet*, v.18, n.13, Jul 1, p.2388-99. 2009.

Shu L, Sauter NS, *et al.* Transcription factor 7-like 2 regulates beta-cell survival and function in human pancreatic islets. *Diabetes*, v.57, n.3, Mar, p.645-53. 2008.

Silva JA, Escobar A, *et al.* Unstable angina. A comparison of angioscopic findings between diabetic and nondiabetic patients. *Circulation*, v.92, n.7, Oct 1, p.1731-6. 1995.

Sladek RRG, Rung J, Dina C, Shen L, Sladek R, Rocheleau G, *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*, v.445, n.7130, Feb 22, p.881-5. 2007.

Sones FM, Jr. Cine-coronary arteriography. *Ohio Med*, v.58, Sep, p.1018-9. 1962.

Sousa AG, Marquezine GF, *et al.* TCF7L2 polymorphism rs7903146 is associated with coronary artery disease severity and mortality. *PLoS One*, v.4, n.11, p.e7697. 2009.

Staiger H, Machicao F, *et al.* Pathomechanisms of type 2 diabetes genes. *Endocr Rev*, v.30, n.6, Oct, p.557-85. 2009.

\_\_\_\_\_. Polymorphisms within novel risk loci for type 2 diabetes determine beta-cell function. *PLoS ONE*, v.2, n.9, p.e832. 2007.

Staiger H, Stancakova A, *et al.* A candidate type 2 diabetes polymorphism near the HHEX locus affects acute glucose-stimulated insulin release in European populations: results from the EUGENE2 study. *Diabetes*, v.57, n.2, Feb, p.514-7. 2008.

Stamler J, Vaccaro O, *et al.* Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes Care*, v.16, n.2, Feb, p.434-44. 1993.

Steinthorsdottir V, Thorleifsson G, *et al.* A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet*, v.39, n.6, Jun, p.770-5. 2007.

Stratton IM, Adler AI, *et al.* Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. *BMJ*, v.321, n.7258, Aug 12, p.405-12. 2000.

Tavares V, Hirata RD, *et al.* Association between Pro12Ala polymorphism of the PPAR-gamma2 gene and insulin sensitivity in Brazilian patients with type-2 diabetes mellitus. *Diabetes Obes Metab*, v.7, n.5, Sep, p.605-11. 2005.

Tominaga M, Eguchi H, *et al.* Impaired glucose tolerance is a risk factor for cardiovascular disease, but not impaired fasting glucose. The Funagata Diabetes Study. *Diabetes Care*, v.22, n.6, Jun, p.920-4. 1999.

Tonjes A, Stumvoll M. The role of the Pro12Ala polymorphism in peroxisome proliferator-activated receptor gamma in diabetes risk. *Curr Opin Clin Nutr Metab Care*, v.10, n.4, Jul, p.410-4. 2007.

Tschritter O, Fritsche A, *et al.* Increased insulin clearance in peroxisome proliferator-activated receptor gamma2 Pro12Ala. *Metabolism*, v.52, n.6, Jun, p.778-83. 2003.

Tschritter O, Stumvoll M, *et al.* The prevalent Glu23Lys polymorphism in the potassium inward rectifier 6.2 (KIR6.2) gene is associated with impaired glucagon suppression in response to hyperglycemia. *Diabetes*, v.51, n.9, Sep, p.2854-60. 2002.

Tsutsui T, Hesabi B, *et al.* Targeted disruption of CDK4 delays cell cycle entry with enhanced p27(Kip1) activity. *Mol Cell Biol*, v.19, n.10, Oct, p.7011-9. 1999.

Tunstall-Pedoe H, Kuulasmaa K, *et al.* Myocardial infarction and coronary deaths in the World Health Organization MONICA Project. Registration procedures, event rates, and case-fatality rates in 38 populations from 21 countries in four continents. *Circulation*, v.90, n.1, Jul, p.583-612. 1994.

Unoki H, Takahashi A, *et al.* SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat Genet*, v.40, n.9, Sep, p.1098-102. 2008.

Van Dam RM, Hoebee B, *et al.* Common variants in the ATP-sensitive K<sup>+</sup> channel genes KCNJ11 (Kir6.2) and ABCC8 (SUR1) in relation to glucose intolerance: population-based studies and meta-analyses. *Diabet Med*, v.22, n.5, May, p.590-8. 2005.

Van Hoek M, Dehghan A, *et al.* Predicting type 2 diabetes based on polymorphisms from genome-wide association studies: a population-based study. *Diabetes*, v.57, n.11, Nov, p.3122-8. 2008.

Van Vliet-Ostaptchouk JV, Onland-Moret NC, *et al.* HHEX gene polymorphisms are associated with type 2 diabetes in the Dutch Breda cohort. *Eur J Hum Genet*, Jan 30. 2008.

Vigorita VJ, Moore GW, *et al.* Absence of correlation between coronary arterial atherosclerosis and severity or duration of diabetes mellitus of adult onset. *Am J Cardiol*, v.46, n.4, Oct, p.535-42. 1980.

Wajchenberg BL. Beta-cell failure in diabetes and preservation by clinical treatment. *Endocr Rev*, v.28, n.2, Apr, p.187-218. 2007a.

Wajchenberg BL. Postprandial glycemia and cardiovascular disease in diabetes mellitus. *Arq Bras Endocrinol Metabol*, v.51, n.2, Mar, p.212-21. 2007b.

Waller BF, Palumbo PJ, *et al.* Status of the coronary arteries at necropsy in diabetes mellitus with onset after age 30 years. Analysis of 229 diabetic patients with and without clinical evidence of coronary heart disease and comparison to 183 control subjects. *Am J Med*, v.69, n.4, Oct, p.498-506. 1980.

Wang X, Xiao Y, *et al.* A role for the beta-catenin/T-cell factor signaling cascade in vascular remodeling. *Circ Res*, v.90, n.3, Feb 22, p.340-7. 2002.



Weedon MN, McCarthy MI, *et al.* Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. *PLoS Med*, v.3, n.10, Oct, p.e374. 2006.

Wei FY, Nagashima k, *et al.* Cdk5-dependent regulation of glucose-stimulated insulin secretion. *Nat Med*, v.11, n.10, Oct, p.1104-8. 2005.

Wei M, Gaskill SP, *et al.* Effects of diabetes and level of glycemia on all-cause and cardiovascular mortality. The San Antonio Heart Study. *Diabetes Care*, v.21, n.7, Jul, p.1167-72. 1998.

Yang Q, Khoury MJ, *et al.* How many genes underlie the occurrence of common complex diseases in the population? *Int J Epidemiol*, v.34, n.5, Oct, p.1129-37. 2005.

Zeggini E, McCarthy MI. TCF7L2: the biggest story in diabetes genetics since HLA? *Diabetologia*, v.50, n.1, Jan, p.1-4. 2007.

Zeggini E, Scott LJ, *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*, v.40, n.5, May, p.638-45. 2008.

Zeggini E, Weedon MN, *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*, v.316, n.5829, Jun 1, p.1336-41. 2007.

## **Apêndices**

---

# TCF7L2 Polymorphism rs7903146 Is Associated with Coronary Artery Disease Severity and Mortality

André Gustavo P. Sousa<sup>1,2</sup>, Guilherme F. Marquezine<sup>1</sup>, Pedro A. Lemos<sup>3</sup>, Eulogio Martinez<sup>3</sup>, Neuza Lopes<sup>4</sup>, Whady A. Hueb<sup>4</sup>, José E. Krieger<sup>1</sup>, Alexandre C. Pereira<sup>1\*</sup>

**1** Laboratory of Genetics and Molecular Cardiology, Heart Institute (InCor), University of São Paulo Medical School, São Paulo, Brazil, **2** Clinical Medicine Department, Federal University of Rio Grande do Norte, Natal, Brazil, **3** Hemodynamic Laboratory, Heart Institute (InCor), University of São Paulo Medical School, São Paulo, Brazil, **4** Coronary Artery Disease Service, Heart Institute (InCor), University of São Paulo Medical School, São Paulo, Brazil

## Abstract

**Background:** TCF7L2 polymorphisms have been consistently associated with type 2 diabetes mellitus in different populations and type 2 diabetes mellitus is a major risk factor for cardiovascular disease, especially coronary artery disease. This study aimed to evaluate the association between TCF7L2 polymorphism rs7903146 and coronary artery disease in diabetic and non-diabetic subjects.

**Methods and Results:** two populations were studied in order to assess severity of coronary artery disease and cardiovascular events incidence. Eight-hundred and eighty nine subjects who were referred for cardiac catheterization for coronary artery disease diagnosis were cross-sectionally evaluated for coronary lesions (atherosclerotic burden) and 559 subjects from the MASS-II Trial were prospectively followed-up for 5 years and assessed for major cardiovascular events incidence. As expected, rs7903146 T allele was associated with diabetes. Although diabetic patients had a higher prevalence of coronary lesions, no association between TCF7L2 genotype and coronary lesions was found in this subgroup. However, non-diabetic individuals carrying the T allele were associated with a significantly higher frequency of coronary lesions than non-diabetic non-carriers of the risk allele (adjusted OR = 2.32 95%CI 1.27–4.24, p=0.006). Moreover, presence of multi-vessel coronary artery disease was also associated with the CT or TT genotypes in non-diabetics. Similarly, from the prospective sample analysis, non-diabetics carrying the CT/TT genotypes had significantly more composite cardiovascular end-points events than CC carriers (p=0.049), mainly due to an increased incidence of death (p=0.004).

**Conclusions:** rs7903146 T allele is associated with diabetes and, in non-diabetic individuals, with a higher prevalence and severity of coronary artery disease and cardiovascular events. name of registry site (see list below), registration number, trial registration URL in brackets.

**Clinical Trial Registration Information:** Medicine, Angioplasty, or Surgery Study (MASS II): Unique identifier: ISRCTN66068876.

**Citation:** Sousa AGP, Marquezine GF, Lemos PA, Martinez E, Lopes N, et al. (2009) TCF7L2 Polymorphism rs7903146 Is Associated with Coronary Artery Disease Severity and Mortality. PLoS ONE 4(11): e7697. doi:10.1371/journal.pone.0007697

**Editor:** Cuilin Zhang, National Institute of Child Health and Human Development/National Institutes of Health, United States of America

**Received:** May 14, 2009; **Accepted:** September 28, 2009; **Published:** November 17, 2009

**Copyright:** © 2009 Sousa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work has been supported by FAPESP (www.fapesp.br) grant # 2007/54138–2. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: lbmpereira@incor.usp.br

## Introduction

Diabetes mellitus (DM) is a common disease present in about 200 million of individuals worldwide and its prevalence has been increasing [1]. People with DM have a higher risk of cardiovascular (CV) complications, including coronary artery disease (CAD). Actually, diabetes confers an equivalent risk to previous myocardial infarction of future cardiovascular disease (CVD) [2].

Genome-wide association studies have disclosed chromosomal regions containing DM susceptibility genes in chromosome 10q, which were later ascribed to intronic variations in the transcription factor 7-like 2 (TCF7L2) gene and TCF7L2 variants have been consistently associated with type 2 diabetes (T2D) in different

populations [3], including the Brazilian population [4]. In addition, associations between TCF7L2 polymorphisms and other conditions such as prostate, breast and colon cancers [5,6,7] and even cardiovascular events [8] have been studied.

Thus, we aimed to study the association between the rs7903146 polymorphism, the most widely studied TCF7L2 genetic marker, and CVD in individuals with and without diabetes. For this, we have used two independent populations: 1) individuals referred to cardiac catheterization for suspected coronary artery disease; 2) individuals with established multivessel coronary artery disease from the MASS-II Trial (ISRCTN66068876), whom were followed-up for 5 years for CV events incidence (all cause mortality, myocardial infarction, recurrent angina, and new cardiac procedures).

## Methods

### Patient populations

**Individuals referred for cardiac catheterization for the diagnosis of CAD.** A cross-sectional study of severity of coronary artery disease was performed at the Laboratory of Hemodynamic of the Heart Institute (InCor), São Paulo, Brazil. All patients had a suspected clinical diagnosis of angina pectoris and stable angina and no patient enrolled in this study was currently experiencing an acute coronary syndrome. Patients with previous cardiac events and/or previous diagnostic or therapeutic coronary catheterization, acute ischemic events, heart failure class III–IV, hepatic dysfunction, familiar hypercholesterolemia, previous heart or kidney transplantation, and in anti-retroviral treatment were excluded from the study. Of a total of 1,451 patients consecutively seen, 167 were excluded because they have been previously submitted to cardiac bypass surgery, 140 were excluded because they have already received coronary stents, 78 have been previously submitted to angioplasty, 163 due to a diagnostic cardiac catheterism before this evaluation and 7 because of genotyping missing data. The coronary lesions were evaluated by an expert team by analyzing angiogram images, based on standard protocols. Patients answered a clinical questionnaire that covered questions regarding their personal medical history, family history of CAD, sedentary behavior, smoking status, hypertension, obesity, dyslipidemia, diabetes, and current medical treatment. During the interview, anthropometric measures were collected (weight and height), and a blood sample during cardiac catheterization for genomic DNA extraction was obtained. All patients signed an informed consent form and the study has been approved by the Ethics Committee of the Heart Institute of the University of Sao Paulo-Brazil (research protocol # 1081/07).

**MASS II Study population.** The participants of this second population were selected from the prospective, randomized, controlled clinical trial MASS II. The MASS II trial was designed to compare medical treatment, angioplasty/stent placement, and surgical myocardial revascularization in patients with stable multivessel CAD with preserved left ventricular function. Briefly, 2,076 candidates who were indicated for myocardial revascularization were evaluated from May, 1995 to May, 2000. Of these, 611 patients were eligible and met all entry criteria to be randomly assigned to one of the three therapeutic groups. The inclusion criteria were symptomatic multivessel coronary disease, preserved left ventricular function, and the presence of coronary lesions (70% of stenosis) amenable to angioplasty. Then patients were followed-up for five years and were assessed for cardiovascular events such as mortality, myocardial infarction, angioplasty, myocardial revascularization surgery and the combination of them. All patients from the MASS-II study were repeatedly evaluated in a single-center (Heart Institute, Brazil) by the same medical team for the five years of follow-up, according to a standard protocol. All events (except non-hospital deaths) were preferentially attended at the Heart Institute [9,10,11].

All patients signed an informed consent form and the study has been approved by the Ethics Committee of the Heart Institute of the University of Sao Paulo-Brazil (research protocol # 1081/07).

### Coronary artery disease severity score

A score was created in order to assess the severity of the CAD. Twenty coronary segments were analyzed for the construction of the score. Each epicardial vessel or main branch was divided into 3 segments (proximal, medial and distal) except secondary branches

of the right coronary artery that were divided in 2 segments, proximal and distal. The score was defined as the sum of each coronary segment analyzed: no lesion or irregularities — 0 points, less than 50% lesion — 0.3 points, 50–70% lesion — 0.6 points, 70–90% lesion — 0.8 points and 90–100% lesion — 0.95 points. Use and characteristics of this score has been previously described [12].

### TCF7L2 rs7903146 polymorphism genotyping

Genomic DNA was extracted from peripheral blood leukocytes by means of a standard salting-out procedure. PCR primers used were: 5' ACA ATT AGA GAG CTA AGC ACT TTT TAG GTA 3' and 5' GTG AAG TGC CCA AGC TTC TC 3'. The studied polymorphism was detected by polymerase chain reaction-restriction fragment length polymorphism assay (PCR-RFLP). Briefly, a 30-cycle PCR was carried out in a PTC-DNA Engine Tetrad2 using a 10  $\mu$ L reactive solution containing 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 100  $\mu$ M of each dNTP, 0.3 U Easy Taq DNA Polymerase, 5 pmol of each primer and 1  $\mu$ L of genomic DNA template. PCR products were digested with 1 U of *RsaI* restriction enzyme and visualized by 3% agarose gel electrophoresis. The quality control for these assays was assessed by randomly selecting 40 samples that were re-genotyped by an independent technician. Observed concordance between genotyping assays was 100%.

### Statistical analysis

Hardy–Weinberg equilibrium for the distribution of genotypes was estimated by the Chi-square test. Chi-square tests, t-tests, and analysis of variance were used for baseline comparisons. The odds ratios for different association models were calculated with 95% confidence interval (CI) by multiple logistic regression with confounders determined by a backward conditional elimination method for a significance level below 0.05. Long-term survival comparisons were conducted for the entire study group and for the subgroups of diabetic and non-diabetic patients and TCF7L2 genotype. Logistic regression was used to estimate the cross-sectional association of these variables with each of the MASSII trial endpoints (death, myocardial infarction, and recurrent ischemia requiring revascularization), as well as the combined endpoint after the 5-year follow-up period. Survival curves were calculated with the Kaplan-Meier method, and differences between the curves were evaluated with the log-rank statistic. We assessed the relationship between baseline variables and composite end points by using a Cox proportional hazards survival model. A value of  $p < 0.05$  was considered significant for comparisons. SPSS version 13 program for Windows was used for statistical analysis.

The authors had full access to the data and take responsibility for its integrity. All authors have read and agree to the manuscript as written.

## Results

### Cross-sectional analysis - Assessment of CAD Severity

In this first part, we studied a population referred for cardiac catheterization for the diagnosis of CAD at the Laboratory of Hemodynamics, InCor, Brazil. Eight hundred and ninety six consecutive patients submitted for the first time only to coronary angiography to study a clinical suggestive diagnosis of coronary artery disease were selected. This study population had 52.4% of male individuals, with a mean age of 59.5 years. The prevalence of hypertension, diabetes, obesity (BMI  $> 30.0$  kg/m<sup>2</sup>), smoking, heart failure (HF) class I-II, peripheral artery disease (PAD) and

**Table 1.** Demographic and laboratorial characteristics of non-diabetic and diabetic subjects from the cross-sectional population according to TCF7L2 genotype.

	Non-diabetic subjects			Diabetic Subjects		
	CC (n = 284)	CT + TT (n = 324)	P value*	CC (n = 135)	CT+TT (n = 153)	P value*
Age (years)	58.4±11.1	59.4±10.4	0.229	61.0±9.6	60.3±9.4	0.507
Male gender (%)	53.5	57.1	0.376	45.9	46.4	0.935
Smoking (%)	35.6	41.7	0.123	29.6	36.6	0.210
Sedentarism (%)	94.0	90.1	0.214	94.8	93.5	0.303
BMI (kg/m <sup>2</sup> )	27.8±5.3	27.3±4.6	0.087	28.9±5.5	28.6±5.4	0.672
Hypertension (%)	70.1	65.7	0.254	80.7	86.3	0.205
Congestive Heart Failure (%)	10.2	8.6	0.508	6.7	15.7	0.016
Chronic Kidney Disease (%)	1.8	1.5	0.770	5.2	3.9	0.755
Glucose (mg/dL)	104.2±21.1	103.9±19.2	0.288	172.1±69.1	177.7±73.2	0.522
Total cholesterol (mg/dL)	229.1±47.3	230.3±53.3	0.510	231.2±47.4	223.9±47.0	0.248
LDL-c (mg/dL)	145.4±38.5	147.1±43.1	0.698	154.0±42.3	132.4±41.2	0.002
HDL-c (mg/dL)	46.0±13.4	44.7±12.9	0.577	42.9±10.4	44.6±13.0	0.394
Triglycerides (mg/dL)	164.7±104.5	163.3±102.5	0.831	223.4±192.5	210.2±188.1	0.602

(\*) Comparison between genotypes groups in non-diabetic and diabetic subjects, calculated by chi-square test for categorical variables and by t-student test for continuous variables; BMI - body mass index.  
doi:10.1371/journal.pone.0007697.t001

chronic renal failure (CRF) was 73%, 32.1%, 28%, 37.1%, 1.8%, 9.2% and 2.7%, respectively. The mean values of fasting glycemia, total cholesterol, LDL-cholesterol, HDL-cholesterol and triglycerides measurements were 130.2 mg/dl, 229.0 mg/dl, 145.0 mg/dl, 44.8 mg/dl and 181.4 mg/dl, respectively.

In the entire population, the rs7903146 genotype frequencies were: CC = 46.8% (n = 419); CT = 42.6% (n = 382); TT = 10.6% (n = 95). For the CC, CT and TT genotypes, respectively, the prevalence of diabetes was 32.2%, 30.9% and 36.8%. There was no difference in genotype frequencies for age, sex, smoking, BMI, total cholesterol, HDL- and LDL-cholesterol, triglycerides or systolic and diastolic blood pressure, hypertension, APD, HF and CRF.

Although clinically suspected, 267 patients had no identifiable coronary lesions by angiography. As expected, diabetic patients had a higher prevalence of obstructive coronary lesions (~80% of individuals), independent of the genotype they carry (p = 0.93). Although there were no differences in the prevalence of coronary lesions between the three TCF7L2 genotypes in the entire population, non-diabetic patients that carried allele T (non-diabetics CT and TT individuals) had a significantly higher

frequency of coronary lesions as compared to non-diabetic individuals with the CC genotype. Allelic analysis also revealed a significant association between allele T and coronary lesions in this subgroup (OR = 1.37 per allele, 95%CI = 1.05–1.78, p = 0.025). The demographic and laboratorial characteristics of non-carriers of T allele and carriers are shown in Table 1. The individuals CT/TT had an almost 2.5 fold higher chance of having obstructive coronary lesions when compared to group 1 (Table 2). This increased odd of having an increased atherosclerotic burden was still present even after adjustment for age, sex, BMI, smoking, hypertension, glycemia, total cholesterol, LDL-c, HDL-c and triglycerides and use of statins, aspirin and beta-blockers.

Regarding atherosclerosis burden, non-diabetic carriers of the T allele had a significantly higher chance of having any coronary lesion (OR 1.51, p = 0.01); a significantly higher prevalence of three-vessel coronary disease as opposed to having normal coronary arteries (OR 1.76, p = 0.01) (Table 3). In addition, considering the severity score divided in tertiles, having the CT or TT genotypes was also associated with an increased odd of being in the second or third tertile of atherosclerosis burden (Table 4). By linear regression, the presence of one T allele increased the score

**Table 2.** Prevalence and risk estimative of coronary lesions according to TCF7L2 genotype.

	Genotypic groups	Prevalence (n)	OR	Adjusted OR*	P value
Diabetic subjects	CC	80.7% (109)	1.00	-	0.727
	CT+TT	79.1% (121)	0.902 (0.51–1.61)	-	
Non-diabetic subjects	CC	60.6% (172)	1.00	1.00	0.006†
	CT+TT	70.1% (227)	1.52 (1.09–2.13)	2.32 (1.27–4.24)	
All subjects	CC	67.1% (281)	1.00	1.00	0.112†
	CT+TT	73.0% (348)	1.325 (0.99–1.76)	1.47 (0.91–2.36)	

(\*) Adjusted for sex, BMI, smoking, hypertension, glycemia, total cholesterol, LDL-c, HDL-c, triglycerides.

(†) p value for adjusted OR.

doi:10.1371/journal.pone.0007697.t002

**Table 3.** Distribution of coronary lesions in non-diabetic patients with coronary lesions according to genotypes.

	CC % (n)	CT + TT % (n)	OR*	P value	Adjusted OR†	P value‡
Normal	43.7 (124)	34.0 (110)	1.00		1.00	
One-vessel disease	20.8 (59)	27.8 (90)	1.72 (1.13–2.61)	0.01	2.09 (1.08–4.04)	0.03
Two-vessel disease	19.4 (55)	16.0 (52)	1.07 (0.67–1.68)	0.79	1.47 (0.75–2.90)	0.26
Three-vessel disease	16.2 (46)	22.2 (72)	1.76 (1.12–2.77)	0.01	2.06 (1.00–4.25)	0.05

(\*) Odds Ratio calculated by comparison between one-vessel, two-vessel, and three-vessel disease versus normal coronary arteries by multiple logistic regression.

(†) Adjusted for sex, BMI, smoking, hypertension, glycemia, total cholesterol, LDL-c, HDL-c, triglycerides and use of statins, aspirin and beta-blockers; (‡) p value for adjusted OR.

doi:10.1371/journal.pone.0007697.t003

value (analyzed here as a continuous variable) in 1.4 points ( $p = 0.038$ ). After adjustment for potential confounders, including glycemia, some, but not all, unadjusted comparisons were no longer statistically significant (Tables 3 and 4).

#### Prospective Analysis - Incidence of Cardiovascular Events in the MASS-II Trial

In order to evaluate if TCF7L2 genotype was also associated with cardiovascular events in a population of individuals already with extensive coronary artery disease, we studied the patients from the MASS-II Trial. This study population included 559 eligible individuals, of whom 69.2% were male, with a similar mean population age of 59.7 years and mean BMI of 27.1 kg/m<sup>2</sup>. Hypertension was present in 59.6%, diabetes in 30.9% and smoking in 33.9% of studied individuals. The mean value of fasting glycemia, total cholesterol, LDL-cholesterol, HDL-cholesterol and triglycerides measurements were 129.6 mg/dl, 223.4 mg/dl, 147.5 mg/dl, 37.3 mg/dl and 194.0 mg/dl, respectively. Genotype counts and frequency were as follow: CC = 159 (28.4%), CT = 336 (60%), and TT = 65 (11.6%). The characteristics of carriers of T allele and non-carriers patients are shown in Table 5.

Kaplan-Meier curve analysis is shown in Figure 1. The group with rs7903146 genotype CT or TT was associated with a significantly higher incidence of cardiovascular events (Figure 1A). In absolute values, CT/TT group had 47 composite events, compared to 12 events in the CC group. However, it can be observed that the TCF7L2 polymorphism is not associated with the composite end point of cardiovascular events (cardiac death, myocardial infarction, and refractory angina requiring revascularization or new cardiac catheterization) in diabetic subjects (Figure 1B). Decompounding the end points in non-diabetic individuals, we observed that the association between TCF7L2 genotype and cardiovascular events occurred mainly because an increased all-cause mortality incidence in individuals harboring the T allele (number of deaths: CT/TT = 27 vs CC = 4,  $p$  value =

0.004). Although incidence of myocardial infarction, myocardial revascularization, percutaneous coronary intervention (PCI) and stroke were higher in CT/TT group in non-diabetics, we were not able to reach statistical significance ( $p > 0.10$ ). Finally, after adjustment for several established risk factors for cardiovascular disease listed in Table 5, only TCF7L2 genotypes CT or TT ( $p = 0.047$ ), age ( $p = 0.003$ ) and smoking ( $p = 0.013$ ) were significantly and independently associated with cardiovascular events incidence.

#### Discussion

The study data showed a significant association between TCF7L2 rs7903146 genotypes and coronary artery disease in non-diabetic individuals, in both independent populations. Interestingly, we were not able to demonstrate this association in diabetic patients.

It has been well established for many decades that diabetes increases the risk of CVD. About 65% of deaths in individuals with type 2 diabetes are related to heart disease or stroke [13,14]. These data are concordant with the results of this study, in which diabetic individuals had a higher severity of CAD and cardiovascular events incidence than non-diabetics. Considering this information and since there were no associations between rs7903146 polymorphism and established CV risk factors, other than diabetes itself, we hypothesize that the association between TCF7L2 genotype and cardiovascular phenotypes occurs due to glycemic homeostasis changes. Nevertheless, given the scenario of established diabetes, coronary disease risk will be increased by itself, despite TCF7L2 genotype status.

Genetic variants in the gene encoding for TCF7L2 have been associated with type 2 diabetes and impaired cell function, but the mechanisms still remain unknown. The gene TCF7L2 is located on chromosome 10 and encodes the transcription factor4 (TCF-4) and single nucleotide polymorphism (SNP) rs7903146 strongly predicts prevalent and future T2DM in several independent populations [3]. The risk T allele was associated with impaired

**Table 4.** Tertile distribution of severity artery coronary lesions score (SCS) in non-diabetic patients according to TCF7L2 genotypes.

	CC % (n)	CT + TT % (n)	OR*	P value	Adjusted OR†	P value‡
1st Tertile ( $\leq 0.9$ )	38.7 (110)	28.4 (92)	1.00		1.00	
2 <sup>nd</sup> Tertile (0.91–2.2)	29.9 (85)	36.1 (117)	1.65 (1.11–2.44)	0.01	1.004 (0.53–1.91)	0.99
3 <sup>rd</sup> Tertile ( $\geq 2.2$ )	31.3 (89)	35.5 (115)	1.55 (1.04–2.28)	0.03	1.31 (0.69–2.47)	0.40

(\*) Odds Ratio calculated by comparison between tertiles by multiple logistic regression (reference group first tertile).

(†) Adjusted for sex, BMI, smoking, hypertension, glycemia, total cholesterol, LDL-c, HDL-c, triglycerides, use of statins, aspirin and beta-blockers; (‡) p value for adjusted OR.

doi:10.1371/journal.pone.0007697.t004

**Table 5.** Demographic and laboratorial characteristics of non-diabetic subjects from the MASS-II population according to TCF7L2 genotype.

	Non-diabetic subjects			Diabetic subjects		
	CC (n = 120)	CT + TT (n = 266)	P value*	CC (n = 38)	CT + TT (n = 135)	P value*
Age (years)	59.6±9.2	59.1±9.4	0.836	62.7±6.4	59.8±9.1	0.071
Male gender (%)	68.3	71.4	0.537	60.5	65.2	0.280
Smoking (%)	34.2	36.5	0.663	31.6	25.9	0.479
BMI (kg/m <sup>2</sup> )	27.6±3.9	26.4±4.0	0.966	27.9±4.4	28.0±4.6	0.880
Hypertension (%)	56.7	53.8	0.595	73.7	66.7	0.673
Previous MI (%)	39.2	50.0	0.048†	26.3	45.9	0.030
Glucose (mg/dL)	104.0±17.4	107.2±21.3	0.551	161.3±63.6	183.0±79.7	
Total cholesterol (mg/dL)	220.5±44.4	225.9±48.3	0.409	221.0±49.6	219.2±50.6	0.848
LDL cholesterol (mg/dL)	147.3±38.2	149.9±44.7	0.349	144.7±41.9	141.2±46.8	0.688
HDL cholesterol (mg/dL)	37.5±9.4	37.0±10.7	0.783	37.2±11.9	38.1±10.6	0.673
Triglycerides (mg/dL)	186.2±109.9	185.0±94.6	0.191	213.3±156.2	205.4±141.5	0.765

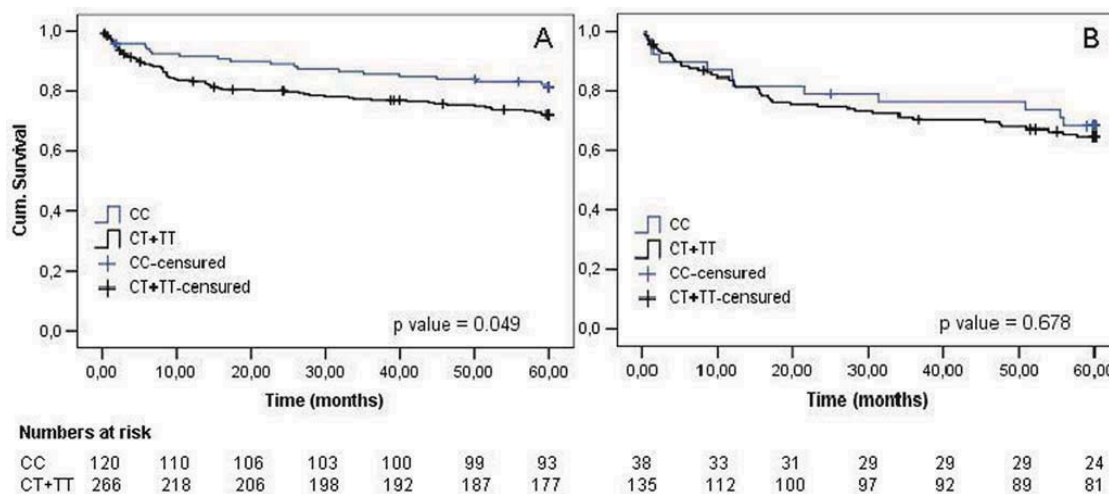
(\*) calculated by qui-square test for categorical variables and by t-student test for continuous †variables; (†) p value <0.05; BMI - body mass index; MI - myocardial infarction.

doi:10.1371/journal.pone.0007697.t005

insulin secretion and incretin effects [15], trending to postprandial hyperglycemia. Data from MONICA/KORA surveys revealed the T allele at rs7903146 as inversely associated with log-transformed HOMA-%Beta (as a measure of basal insulin secretion), but no association was found with insulin resistance and metabolic syndrome. Therefore, we hypothesize that even non-diabetic carriers of CT/TT genotypes have increased postprandial glycemia and consequently, higher CAD tendency, severity and cardiovascular events incidence.

Actually, a meta-analysis of published data from 20 studies with almost 100,000 subjects followed-up for 12.4 years confirmed a

significant association between 2-hour glucose levels after an oral glucose overload and incident cardiovascular events [16]. In this same study, there was a suggestive trend (p value equal to 0.056) between fasting glucose and cardiovascular events. These data support the hypothesis that even non-diabetic degrees of fasting and mainly postprandial hyperglycemia are associated with CVD. Therefore, if T allele is associated with postprandial hyperglycemia, we hypothesized T allele would be associated with cardiovascular events. Plasma glucose is a continuous risk factor for CVD, in both diabetic and non-diabetic people, the risk extending below impaired fasting and impaired glucose tolerance



**Figure 1.** Kaplan-Meier Curve of TCF7L2 genotypes and composite cardiovascular end points in non-diabetic (A) and diabetic patients (B) after 5 years of follow-up. In non-diabetic patients, individuals CT and TT had a significant higher incidence of composite cardiovascular endpoints (death, non-fatal myocardial infarction, refractory angina requiring revascularization or new cardiac catheterization – PCI). Although diabetic individuals had a higher incidence of cardiovascular events than non-diabetics, a presence of T allele was not associated with cardiovascular endpoints.

doi:10.1371/journal.pone.0007697.g001



cutoffs. Indeed, some authors consider that dysglycemia should be added to the list of established continuous cardiovascular risk factors [17]. Most cardiovascular risk factors are affected directly by an acute increase of glycemia: increase in LDL oxidation and endothelial dysfunction (vasoconstriction and decreased vasodilating response), increased production of collagen from the mesangial cell, activation of blood coagulation, increase in blood pressure, increase in the circulating levels of intracellular adhesion molecule 1 (ICAM-1), increase in inflammation (production of plasma interleukin-6, interleukin-18 and tumor necrosis factor- $\alpha$  - TNF- $\alpha$ ) and increase in oxidative stress [17,18]. Unfortunately, there was no information about glycated haemoglobin or postprandial glycemia in our study and we were not able to evaluate this correlation. Regarding fasting glycemia, there is no difference between genotypes in non-diabetic individuals, but we were not able to evaluate the effects of the studied genotype in longitudinal glucose-related variables. In addition, after adjustment for other established cardiovascular risk factors, including baseline fasting glycemia, genotypes CT and TT continued to be related to CAD severity and CV events, mainly death.

Moreover, some studies have observed that transcriptional activation of TCF-4 is related to the nuclear factor- $\kappa$ B (NF- $\kappa$ B) signaling pathway, which regulates inflammatory signaling pathways [19] and this could be hypothesized as an additional pleiotropic mechanism operant in the *milieu* of the vascular wall.

It should be remembered that the diabetic subgroups from both populations have a smaller number of individuals and one can not exclude that this small number of individuals associated with the higher prevalence of coronary artery disease and incidence of cardiovascular events in this particular subgroup of individuals has led to reduced statistical power to detect the association we were able to describe on non-diabetic individuals. However, even a relatively small sample size of non-diabetic subjects was able to provide an adequate statistical power analysis, in part due to the relatively large effect size. Despite this, it is not possible to exclude completely the possibility of false positive results. Considering it is a genetic study of association, a limitation is the possibility of other non evaluated variable that, by acting as a confounder, could play a significant role in the findings. One should remember, however, that multiple potential confounder adjustments were made in the presented analysis. Finally, results in the same direction in two different and independent populations may actually be indicative of a true association.

Previous myocardial infarction was significantly more prevalent in genotypes CT and TT in non-diabetic individuals from the MASS-II population. However, this factor was not independently associated with CV events, like TCF7L2 genotype, smoking and aging. Therefore, the higher prevalence of a previous myocardial

infarction at baseline does not completely explain such association. Actually, these data contribute to the idea that non-diabetic patients carriers of the rs7903146 allele T presents a higher CAD severity.

Bielinski *et al* [8] have recently published a study in which TCF7L2 SNPs rs7903146, rs12255372, rs7901695, rs11196205 and rs7895340 were assessed with a similar objective, using more than 13,000 subjects from the ARIC study population. In their report, they have not demonstrated a significant association between any TCF7L2 SNP with incident coronary disease, ischemic stroke, CVD, prevalent peripheral artery disease (PAD) or all-cause mortality in the full cohort nor when stratified by race or diabetic status. Although divergent results, there are large methodological differences between the two studies: in the ARIC study, all patients with prevalent coronary disease were excluded from analysis and, consequently, the patients were free of prevalent CVD at baseline. On the other hand, our study populations contain patients with suspected CAD (cross-sectional study population) or with established multi-vessel CAD plus stable angina (MASS-II cohort) at baseline and these differences in the *a priori* risk of coronary artery disease in these populations could explain the apparent conflicting results. Here as well, it should be acknowledged that the interpretations of the present study are derived from two different and independent cohorts of patients, straightening the conclusions of the present work.

In conclusion, these data demonstrate an association between coronary artery disease severity and rs7903146 polymorphism mainly in non-diabetic individuals with increased cardiovascular risk. In people with diabetes, we were not able to show this association. Additionally, it was found that non-diabetic allele T carriers also had higher cardiovascular event incidence, mainly due to death. Due to the limitations already discussed, more studies are necessary to confirm this association and to evaluate the intrinsic mechanisms underlining this process. If this association is confirmed, other studies will also be useful to establish the best strategy to use this information for the better risk stratification of individuals.

## Acknowledgments

We would like to thank the work of two anonymous reviewers for insightful suggestions during the review process.

## Author Contributions

Conceived and designed the experiments: ACP. Performed the experiments: AGS. Analyzed the data: AGS GFM. Contributed reagents/materials/analysis tools: PAL EM NL WAH JEK. Wrote the paper: AGS.

## References

1. IDF (2004) International Diabetes Federation - Diabetes Atlas.
2. Booth GL, Kapral MK, Fung K, Tu JV (2006) Relation between age and cardiovascular disease in men and women with diabetes compared with non-diabetic people: a population-based retrospective cohort study. *Lancet* 368: 29–36.
3. Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, et al. (2007) TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med* 85: 777–782.
4. Marquzine GF, Pereira AC, Sousa AG, Mill JG, Hueb WA, et al. (2008) TCF7L2 variant genotypes and type 2 diabetes risk in Brazil: significant association, but not a significant tool for risk stratification in the general population. *BMC Med Genet* 9: 106.
5. Agalliu I, Saurinemi M, Prokunina-Olsson L, Johannesson B, Collins FS, et al. (2008) Evaluation of a variant in the transcription factor 7-like 2 (TCF7L2) gene and prostate cancer risk in a population-based study. *Prostate* 68: 740–747.
6. Burwinkel B, Shanmugam KS, Hemminki K, Meindl A, Schmutzler RK, et al. (2006) Transcription factor 7-like 2 (TCF7L2) variant is associated with familial breast cancer risk: a case-control study. *BMC Cancer* 6: 268.
7. Folsom AR, Pankow JS, Peacock JM, Bielinski SJ, Heiss G, et al. (2008) Variation in TCF7L2 and increased risk of colon cancer: the Atherosclerosis Risk in Communities (ARIC) Study. *Diabetes Care* 31: 905–909.
8. Bielinski SJ, Pankow JS, Folsom AR, North KE, Boerwinkle E (2008) TCF7L2 single nucleotide polymorphisms, cardiovascular disease and all-cause mortality: the Atherosclerosis Risk in Communities (ARIC) study. *Diabetologia* 51: 968–970.
9. Hueb W, Soares PR, Gersh BJ, Cesar LA, Luz PL, et al. (2004) The medicine, angioplasty, or surgery study (MASS-II): a randomized, controlled clinical trial of three therapeutic strategies for multivessel coronary artery disease: one-year results. *J Am Coll Cardiol* 43: 1743–1751.
10. Pereira AC, Lopes NH, Soares PR, Krieger JE, de Oliveira SA, et al. (2006) Clinical judgment and treatment options in stable multivessel coronary artery disease: results from the one-year follow-up of the MASS II (Medicine, Angioplasty, or Surgery Study II). *J Am Coll Cardiol* 48: 948–953.
11. Schettert IT, Pereira AC, Lopes NH, Hueb WA, Krieger JE (2006) Association between platelet P2Y12 haplotype and risk of cardiovascular events in chronic coronary disease. *Thromb Res* 118: 679–683.



## TCF7L2 and CAD

12. Lanz JR, Pereira AC, Lemos PA, Martinez E, Krieger JE (2005) Angiotensinogen M235T polymorphism is associated with coronary artery disease severity. *Clin Chim Acta* 362: 176–181.
13. Milicevic Z, Raz I, Beattie SD, Campaigne BN, Sarwat S, et al. (2008) Natural history of cardiovascular disease in patients with diabetes: role of hyperglycemia. *Diabetes Care* 31 Suppl 2: S155–160.
14. Morrish NJ, Wang SL, Stevens LK, Fuller JH, Keen H (2001) Mortality and causes of death in the WHO Multinational Study of Vascular Disease in Diabetes. *Diabetologia* 44 Suppl 2: S14–21.
15. Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, et al. (2007) Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *J Clin Invest* 117: 2155–2163.
16. Coutinho M, Gerstein HC, Wang Y, Yusuf S (1999) The relationship between glucose and incident cardiovascular events. A metaregression analysis of published data from 20 studies of 95,783 individuals followed for 12.4 years. *Diabetes Care* 22: 233–240.
17. Wajchenberg BL (2007) Postprandial glycemia and cardiovascular disease in diabetes mellitus. *Arq Bras Endocrinol Metabol* 51: 212–221.
18. Ceriello A (2005) Postprandial hyperglycemia and diabetes complications: is it time to treat? *Diabetes* 54: 1–7.
19. Wang X, Xiao Y, Mou Y, Zhao Y, Blankestijn WM, et al. (2002) A role for the beta-catenin/T-cell factor signaling cascade in vascular remodeling. *Circ Res* 90: 340–347.

1 **Title:** Genetic Variants of Diabetes Risk and Incident Cardiovascular Events in Chronic  
2 Coronary Artery Disease

3

4 **Short title:** Genetic Variants of Diabetes and Cardiovascular Events

5

6 **Authors:** André Gustavo P. Sousa<sup>1,2</sup> M.D., Neuza H. Lopes<sup>3</sup> M.D., Ph.D, Whady A. Hueb<sup>3</sup>  
7 M.D., Ph.D, José Eduardo Krieger<sup>1</sup> M.D., Ph.D, Alexandre C. Pereira<sup>1</sup> M.D., Ph.D

8

9 **Affiliations:** 1) Laboratory of Genetics and Molecular Cardiology, Heart Institute,  
10 University of São Paulo Medical School, São Paulo, Brazil; 2) Clinical Medicine  
11 Department, Federal University of Rio Grande do Norte, Natal, Brazil; 3) Coronary Artery  
12 Disease Service, Heart Institute, University of São Paulo Medical School, São Paulo, Brazil

13

14 **Corresponding author:**

15 André Gustavo P. Sousa, MD

16 Clinical Medicine Department

17 Federal University of Rio Grande do Norte, Natal, Brazil

18 **Address:** Av Nilo Peçanha, 620, Petrópolis, Natal – RN, Brazil, 59012-300

19 **Telephone/Fax:** 55 – 84 – 3342-9706

20 **E-mail:** agpsousa@ig.com.br

21

22

23 **Clinical Trial Registration Information:** Medicine, Angioplasty, or Surgery Study

24 (MASS II): URL: <http://www.controlledtrials.com>. Unique identifier: ISRCTN66068876

25

---

26 **Abstract**

27 Objective: To determine whether information from genetic risk variants for diabetes is  
28 associated with cardiovascular events incidence.

29 Methods: We genotyped single-nucleotide polymorphisms at 10 loci associated with type-2  
30 diabetes in 425 subjects from the MASS-II Study, a randomized study in patients with  
31 multi-vessel coronary artery disease. Genotype scores using information from genetic  
32 markers were created. The discriminatory ability of genotype scores was assessed for  
33 diabetes by calculating the area under the ROC curve (AUC). Performance of genetic  
34 models relative to major cardiovascular events incidence was analyzed through Kaplan-  
35 Meier curve comparison and Cox Hazard Models.

36 Results: Genetic information was also able to predict 5-year incidence of major  
37 cardiovascular events and overall-mortality in non-diabetic individuals. Non-diabetic  
38 individuals with high genetic risk had a similar incidence of events then diabetic  
39 individuals (cumulative hazard of 29.3, 33.0 and 35.4 for genotype scores *versus* 35.1% of  
40 diabetic subjects). All derived scores were associated with diabetes. The accuracy measured  
41 by the AUC ranged from 0.59 to 0.62. The addition of combined genetic information to  
42 clinical information significantly improved AUC for the diagnosis of diabetes (AUC =  
43 0.63, 0.62 and 0.65 versus 0.58).

44 Conclusions: Combined information of genetic variants for diabetes risk is associated to  
45 major cardiovascular events incidence, including overall mortality, in non-diabetic  
46 individuals with coronary artery disease.

47

48 **Keywords:** diabetes, genetics, prediction score, coronary disease; SNP

49

50  
51  
52  
53  
54

## Introduction

55           Type-2 diabetes mellitus (T2DM) is a common disease present in about 200  
56 million individuals around the world and its prevalence has been increasing [1]. Genome-  
57 wide association studies (GWASs) have identified novel susceptibility genes for T2DM in  
58 different populations [2,3,4,5,6,7]. Meta-analysis totaling over 32,000 individuals  
59 confirmed the association of polymorphisms in nine different genes with the risk of  
60 developing T2DM [4]: *PPARG*, *SLC30A8*, *HHEX*, *TCF7L2*, *KCNJ11*, *IGFBP-2*, *CDKAL1*,  
61 *CDKN2A/B* and *FTO*. Most polymorphisms cited have, individually, only small effects on  
62 T2DM risk, with odds ratios around 1.3. Considering that the genetic susceptibility for  
63 complex diseases such as T2DM is caused by multiple genetic variants, predictive tools  
64 based on information from a single genetic polymorphism will be of limited value [8].  
65 Nevertheless, studies have found that the predictive value for T2DM diagnosis can be  
66 improved by combining multiple common low-risk variants [8,9,10,11,12,13,14]. The  
67 importance of predicting the risk of an individual to present diabetes is related to the  
68 possibility of early diagnosis and the potential avoidance of T2DM-related complications,  
69 particularly cardiovascular disease. Despite this fact, fewer studies have concentrated  
70 efforts in determining whether knowledge of T2DM-associated SNPs in individuals with  
71 established cardiovascular disease can be of any good. Recently, a study had shown the  
72 association of polymorphism rs7903146 of *TCF7L2* gene and cardiovascular events  
73 incidence [15], but it did not assess whether the combination of multiple risk variants  
74 beyond *TCF7L2* is useful in predicting cardiovascular events.

75           In this study, we tested whether the combination of multiple genetic risk variants  
76 known to predict T2DM can be used to predict cardiovascular events in individuals with

76 chronic coronary artery disease.

---

77 **Methods**

78

79 Study population

80

81

82 The MASS II trial was a prospective, randomized and controlled clinical trial that  
83 was designed to compare medical treatment, percutaneous coronary intervention (PCI), and  
84 coronary artery bypass graft (CABG) surgery in patients with stable multi-vessel CAD and  
85 preserved left ventricular function. Briefly, individuals who were indicated for myocardial  
86 revascularization were evaluated from May, 1995 to May, 2000 and were randomly  
87 assigned to one of the three therapeutic groups. The inclusion criteria were symptomatic  
88 multi-vessel coronary disease, preserved left ventricular function, and the presence of  
89 coronary lesions (70% of stenosis) amenable both to angioplasty and surgery. Enrolled  
90 patients were followed up for five years and were assessed for cardiovascular events  
91 (overall-mortality, myocardial infarction, PCI, CABG) or the combination of them. The  
92 diagnosis of diabetes mellitus was based on previous treatment for diabetes or ADA criteria  
93 [16]. Individuals that did not meet criteria for diabetes according ADA were included in the  
94 non-diabetic group at baseline. Other studies that used this same population for similar  
95 analysis have been previously published [17,18,19,20].

95

96 We genotyped single-nucleotide polymorphisms (SNP) at 10 loci associated with  
97 type-2 diabetes in 425 subjects from the MASS-II Study and a 5-years follow-up was  
98 performed. Other 186 individuals who were initially randomized were excluded because  
99 incomplete genotypic information from all studied genetic markers. A comparison of  
100 baseline characteristics between included and excluded individuals was performed and  
101 there were no significant differences between them (data not shown). All subjects gave  
102 written informed consent, and the Ethics Committee of the Heart Institute of the University

102 of São Paulo-Brazil approved the study. All procedures were performed in accordance with  
103 the Helsinki Declaration of 1975 (revised 1983).

104

#### 105 Selection of SNPs, risk alleles, and genotyping

106 A large-scale meta-analysis performed for T2DM in populations of European  
107 ancestry [4] reported consistent associations between T2DM and ten SNPs: *KCNJ11*  
108 rs5219, *PPARG* rs1801282, *TCF7L2* rs7903146, *SLC30A8* rs13266634, *HHEX* rs1111875,  
109 *CDKAL1* rs7754840, *IGFBP2* rs4402960, *FTO* rs8050136, *CDKN2A/B* rs10811661 and  
110 rs9300039. We selected these genetic markers for the present study.

111 Genomic DNA was extracted from peripheral blood leukocytes by means of a  
112 standard salting-out procedure. Genotyping assays were run as submicroliter PCR-based  
113 assays on array tape, which is a continuous plastic tape that is used in conjunction with a  
114 flexible configuration of dispensing, pipetting, sealing and detection modules manufactured  
115 by Douglas Global Array. Thermocycling was performed in sealed master tape in aqueous  
116 immersion in the thermocycler. A modified allele-specific PCR assay as described by  
117 Myakishev [21] was used. We obtained an average genotyping success rate of more than  
118 95% and an average genotyping accuracy of 100% by regenotyping 32 samples. All SNPs  
119 were in Hardy–Weinberg equilibrium ( $P > 0.001$ ).

120

#### 121 Genotype-Score Construction

122 In order to evaluate if the utilization of information from all ten SNPs is  
123 associated with T2DM in our sample, we created three different genetic scores. First, we  
124 considered the number of genes for which an individual was homozygous for the risk allele  
125 (Score 1). Score 2 was created as the total number of risk alleles. Finally, to create the third

126 score, a multiple logistic regression model was fitted using baseline-T2DM status as the  
127 dependent variable and each SNP in its best fitting univariate model (additive, dominant or  
128 recessive) was included as independent variable. The regression coefficient values of each  
129 SNP were multiplied by 10 and rounded to the nearest integer (See E-Table 1 with these  
130 values in Supplemental Material). Score 3 was calculated as the sum of these values. A  
131 clinical score that incorporated identified risk factors including sex, age, body mass index  
132 (BMI), arterial hypertension and smoking status was separately built similarly to Score 3  
133 and the incremental value of combining genetic information to the clinical risk factor score  
134 was also evaluated (E-Table 2 in Supplemental Material).

135

#### 136 Statistical Analysis

137 Chi-square tests, t-tests, and analysis of variance were used for baseline  
138 comparisons. For multiple comparisons, a more restrictive significance threshold for p  
139 value was used by a Bonferroni's correction for multiple comparisons. In order to evaluate  
140 the models performance, a receiver operating characteristic (ROC) curve was built and the  
141 area under ROC curve (AUC) was used to measure the discriminative power. Areas under  
142 the ROC curves (AUC) for the clinical score and the genotypic scores added to the clinical  
143 score were compared by using the ROCKIT software ([http://www-radiology.uchicago.edu/  
144 kr1/KRL\\_ROC/software\\_index6.htm](http://www-radiology.uchicago.edu/kr1/KRL_ROC/software_index6.htm)).

145 Long-term survival comparisons were conducted for the entire study group and  
146 for the subgroups of diabetic and non-diabetic patients at baseline. Cox regression was used  
147 to estimate the prospective association of these variables with each of the endpoints (death,  
148 myocardial infarction, and recurrent ischemia requiring revascularization), as well as the  
149 combined endpoint after 5-year follow-up. Survival curves were calculated with the



150 Kaplan-Meier method, and differences between the curves were evaluated with the log-rank  
151 statistic. Statistical analyses were performed with the use of SPSS software, version 13.0.

152

## 153 **Results**

154

155           In the MASS-II population, a total of 425 individuals were genotyped for all of  
156 the ten studied SNPs, from whom 134 had diabetes (31.5%). Clinical characteristics of  
157 participants at baseline are provided in Table 1 and characteristics of the 10 risk loci are  
158 shown in Supplemental Material (E-Table 3). Diabetic and non-diabetic subjects differed in  
159 relation to hypertension prevalence. Therefore, all further tests were corrected for this  
160 variable as a possible confounder. Only *TCF7L2* gene was singly associated with T2DM in  
161 preliminary analysis, however, after correction for multiple comparisons and considering a  
162 p value more restrictive ( $p < 0.005$ ), *TCF7L2* polymorphism did not reach statistical  
163 significance.

164           In order to evaluate if genotype scores were associated with cardiovascular events  
165 in the MASS-II population, we also performed a division of individuals in three groups  
166 based on score tertiles. Although diabetic individuals had a higher incidence of  
167 cardiovascular events than non-diabetic subjects (35.1% vs 24.4,  $p = 0.022$ ), it was  
168 observed that the genotype score groups were not associated with the composite end point  
169 of cardiovascular events (cardiac death, myocardial infarction, and refractory angina  
170 requiring revascularization or new percutaneous coronary intervention – PCI) in individuals  
171 with already established diabetes. However, the groups with higher values in all three  
172 genotype scores were associated with a significantly higher incidence of cardiovascular  
173 events in non-diabetic individuals (Kaplan-Meier curve analysis shown in Figure 1).

174 Interestingly, there were no significant differences between the estimates of the probability  
175 of survival from diabetic individuals (regardless of their genetic score) and non-diabetic  
176 individuals with high genetic risk scores ( $p$  value  $> 0.4$ ). After 5-years of follow-up,  
177 diabetic individuals presented a cumulative incidence of the combined end-point of 35.1%  
178 against non-diabetic individuals that presented 29.3%, 33.0% and 35.4% for higher tertiles  
179 of Score 1, 2 and 3, respectively (Table 2). These data suggest that non-diabetic individuals  
180 whom have high genetic risk score values have similar outcomes as individuals with  
181 already established diabetes.

182           Moreover, resembling the already established behavior of diabetic individuals  
183 regarding treatment options for chronic coronary artery disease [22], non-diabetic  
184 individuals with high genetic risk score values showed similar treatment-specific patterns  
185 of long-term response. Non-diabetic individuals with high genetic risk scores, as well as  
186 diabetic individuals, submitted to percutaneous coronary interventions presented  
187 significantly higher event rates than non-diabetic individuals with low risk scores,  
188 exemplified in Figure 2 by Score 3. Information regarding diabetes or genetic risk-score in  
189 non-diabetic individuals was not able to long-term stratify individuals submitted to CABG  
190 surgery or medical treatment (See also E-Figure 1 in Supplemental Material for details of  
191 other Scores).

192           After multivariate adjustment of a Cox proportional hazards model for several  
193 established risk factors for cardiovascular disease, all derived scores were significantly and  
194 independently associated with cardiovascular events ( $p$  value  $< 0.05$ ). Interestingly, even  
195 when fasting glycemia values were included as covariate in the Cox proportional models,  
196 both Scores 2 and 3 remained significantly associated to composite events incidence (Table  
197 3). In addition, Scores 2 and 3 were also independently associated to overall mortality. We

198 also evaluated 250 patients who completed the 5-years follow up and from whom complete  
199 biochemical data is available. Of these, 38 subjects became diabetic at the end of the  
200 follow-up period. In the remaining 212 patients (still non-diabetic after 5-year follow-up),  
201 both Scores 2 and 3 were also associated with composite cardiovascular end points ( $p =$   
202 0.035 and 0.036, respectively) suggesting that the observed association was not entirely due  
203 to diabetes emergence in individuals with high score values.

204 As it had already been shown that polymorphisms in *TCF7L2* gene were  
205 associated with cardiovascular events [15], genotypic information from *TCF7L2* was  
206 excluded from the genotype scores to assess whether the positive association of combined  
207 information happened only because of *TCF7L2* influence. However, even when *TCF7L2*  
208 data were excluded, genotype Score 2 and 3 remained significantly associated with the  
209 composite cardiovascular end-point in non-diabetic individuals (See E-Figure 2 in  
210 Supplemental Material). In addition, genotype Score 2 and Score 2 without *TCF7L2*  
211 information were also significantly associated with all-cause mortality ( $p$  value = 0.009 and  
212 0.036, respectively).

213 In order to assess T2DM prevalence, a multivariate logistic regression was  
214 performed including clinical T2DM risk factors (age, gender, BMI, arterial hypertension  
215 and smoking) and combined genetic information (three scores) and only the three genotype  
216 scores and arterial hypertension were independently associated with T2DM (See E-Table 4  
217 in Supplemental Material). An increase in the risk of type 2 diabetes by a factor of 1.31 per  
218 one point of Score 1 ( $P = 0.002$ ), 1.18 per one point of Score 2 ( $P = 0.003$ ) and 1.11 per one  
219 point of Score 3 ( $P = <0.0001$ ) was observed. Based on tertiles of genotype score values we  
220 defined low genetic risk groups (lowest values of scores) and high genetic risk groups  
221 (highest values of scores). As expected, the percentage of individuals with diabetes was

222 significantly higher in the high genetic risk groups than in the low risk groups, regardless of  
223 the score used (See E-Table 5 in Supplemental Material for detailed information).

224           The performance of the predictive risk scores can be observed by the ROC  
225 curves in E-Figure 3A (Supplemental Material). All three scores achieved a significant  
226 performance with a discriminatory ability of 0.60 according to the area under ROC curve  
227 (AUC). The comparisons of the three AUCs revealed that there are no significant  
228 differences between them (*P* values for comparisons: Score 1 vs Score 2:  $p = 0.825$ ; Score 1  
229 vs Score 3:  $p = 0.408$ ; Score 2 vs Score 3:  $p = 0.264$ ). Even after the exclusion of *TCF7L2*  
230 data from the analysis, all three scores remained significantly associated with T2DM (AUC  
231 = 0.58,  $p$  value = 0.006; AUC = 0.58,  $p$  value = 0.008 and AUC = 0.59,  $p$  value = 0.003 for  
232 Scores 1, 2 and 3 without *TCF7L2* data, respectively) (See E-Figure 4 in Supplemental  
233 Material).

234           If only clinical variables (combination of gender, age, obesity, smoking and  
235 hypertension) were used to assess T2DM diagnosis in the MASS-II population, the AUC of  
236 this model would be low (AUC = 0.58,  $p = 0.011$ ). The inclusion of other variables such as  
237 triglycerides and HDL levels in this clinical model did not significantly improve its  
238 performance (AUC = 0.57,  $p$  value > 0.05). Therefore, these laboratory variables were not  
239 included. However, the addition of genotype scores significantly improved the clinical  
240 model's performance (E-Figure 3B, Supplemental Material) according to the AUC:  
241 addition of Score 1 = 0.63,  $p = 0.024$  (for the comparison with the clinical risk factor  
242 score); addition of Score 2 = 0.62,  $p = 0.029$ ; addition of Score 3 = 0.65,  $p = 0.038$ .

243

244

245

246 **Discussion**

247

248           The results of our study provide additional information regarding the importance  
249 of genetic risk variants for T2DM. Our results extend previous by showing evidence that  
250 the utilization of combined genetic information for T2DM risk may still provide valuable  
251 information for cardiovascular events prediction. To our knowledge, this is the first study to  
252 evaluate the association between combined common genetic variants for type 2 DM in the  
253 scenario of major cardiovascular events incidence in chronic coronary artery disease.

254           Our findings confirm previous studies in this more specific population. Derived  
255 scores had a performance, assessed by the area under the ROC curve (AUC), ranging  
256 between 0.59 and 0.62 without using information from clinical variables. These values are  
257 comparable to the results from most other studies in different populations [8,9,10,12,13,14].  
258 In addition, inclusion of genetic information was able to improve AUC for diagnosis of  
259 diabetes compared to using only clinical variables. These results are in the same direction  
260 as other studies, in which the addition of information from several polymorphisms  
261 increases the ability to predict T2DM over that of demographic variables alone [10,12,13].  
262 The clinical model's AUC value in this study was comparable to the study by Van Hoek *et*  
263 *al.* [13]. Clinical models from other studies [10,11,12] have achieved better performance  
264 (AUC ranging from 0.72-0.9) possibly because they included other variables such as values  
265 of triglycerides, HDL, fasting glucose and family history of diabetes. Even in those cases,  
266 the inclusion of genetic information has provided a significant but slight increase in AUC.  
267 In our study, the inclusion of laboratory variables did not improve the performance of the  
268 clinical model. Unfortunately, family history data were not available. It is possible that the  
269 inclusion of family history data could significantly increase the performance of our clinical

270 model, potentially reducing the impact of genetic information for diabetes prevalence  
271 discrimination. It is not so clear whether this would influence the described ability to  
272 predict cardiovascular events in this clinical scenario.

273           Different from other studies where T2DM prediction was the main outcome and  
274 the study samples were mainly composed by individuals from the general population, we  
275 have explored the potential use of information associated with inherited risk for diabetes, a  
276 major cardiovascular risk factor, in the somehow more constrained scenario of multi-vessel  
277 chronic coronary artery disease. In this respect, the MASS-II Trial is an exquisite tool by  
278 allowing for the test of this hypothesis in a sample of individuals with established chronic  
279 coronary artery disease, prospectively followed for 5-years. Our data suggest that scores  
280 created to evaluate the T2DM diagnosis are also able to predict cardiovascular events in  
281 non-diabetic individuals. Moreover, the proposed genetic risk scores could also stratify  
282 non-diabetic individuals in risk strata thus allowing one to identify non-diabetic individuals  
283 with similar risk than diabetic individuals. This stratification was above and beyond  
284 established cardiovascular risk factors. Interestingly, we were not able to find similar  
285 results in diabetic individuals, suggesting that once diabetes is established, cardiovascular  
286 risk is high, regardless of the underlying causes of this derangement. Even after adjustment  
287 for established risk factors for cardiovascular disease, Scores 2 and 3 were independently  
288 associated with cardiovascular events (HR for increase of 1 point: Score 2 = 1.16, p =  
289 0.022; Score 3 = 1.08, p = 0.015). Similar results were found regarding all-cause mortality  
290 after decompounding the end points in non-diabetic individuals. Our data suggest that a  
291 subgroup of individuals potentially harboring an increased risk of becoming diabetic could  
292 be targeted for more precocious and rigorous preventive efforts, similar to what is  
293 advocated for diabetic individuals with coronary artery disease.

294 Previously, Haffner et al. [23] have demonstrated that incidence rates of  
295 myocardial infarction in non-diabetic individuals with prior myocardial infarction were  
296 comparable to diabetic subjects without previous myocardial infarction. Interestingly, in  
297 our work, individuals who had the highest values of genotype scores had rates of  
298 cardiovascular events comparable to those who had diabetes. In addition, data from the  
299 BARI study [24] have found that diabetic patients with multi-vessel coronary artery disease  
300 who underwent coronary revascularization therapy with angioplasty had higher event rates  
301 than those that underwent CABG or medical treatment. Similarly, in our data, non-diabetic  
302 patients with the highest scores also had this pattern according to the treatment option they  
303 were randomized to.

304 We did not observe any association between score tertiles and cardiovascular risk  
305 factors at baseline in the studied sample. It is thus reasonable to speculate that the observed  
306 association between combined T2DM risk alleles and cardiovascular events occurs due to  
307 elevations in glycemia, even in a normal, fasting or postprandial, range. Indeed, a recent  
308 meta-analysis confirmed the significant association between 2-hour glucose levels after an  
309 oral glucose overload and incident cardiovascular events [22]. Coutinho et al. carried out a  
310 meta-analysis of 20 studies for a total of more than 90,000 non-diabetics individuals and  
311 have found an exponential relationship between the risk of cardiovascular events and both  
312 fasting and postload plasma glucose levels. Such relationship was present even at below  
313 diagnostic glycemia levels for diabetes, impaired fasting glycemia or impaired glucose  
314 tolerance [22]. In addition, systematic evaluation of databases such as DECODE have  
315 shown that this relationship between glycemia and cardiovascular risk begins within normal  
316 glycemia range, with a linear relationship and shows no indication of a threshold [25].  
317 These data support the hypothesis that even non-diabetic levels of fasting and mainly

318 postprandial hyperglycemia are associated with cardiovascular events incidence and could  
319 explain the reason by non-diabetic individuals had similar cardiovascular events than  
320 diabetic subjects. In addition, considering that the association of glucose with CVD risk is  
321 well established, it would be quite possible that fasting glucose may capture most of the  
322 risk conferred by the genetic variants in non-diabetic individuals. Indeed, after including  
323 this variable in the models all three genotype scores had a reduction in the magnitude of  
324 effects. Nevertheless, both Scores 2 and 3 remained associated with composite  
325 cardiovascular end point and all-cause mortality. In addition, even after exclusion of  
326 subjects that became diabetics at end of follow-up period, the association between  
327 cardiovascular events and Scores 2 and 3 persisted.

328           Some potential limitations of our study should be acknowledged. Sample size was  
329 relatively small. However, even a relatively small sample size of non-diabetic subjects was  
330 able to provide a clear understanding of the relative importance of combined genetic  
331 information as a predictor of cardiovascular events due to the relatively large effect size of  
332 the derived genotype scores. On the other hand, there were only a limited number of  
333 diabetic individuals and one cannot exclude that this has led to reduced statistical power to  
334 detect the association that we were able to describe in non-diabetic individuals. The MASS-  
335 II clinical trial was designed for specific purposes and this work was an observational study  
336 that used data from MASS-II. Therefore, it is not possible to exclude that residual  
337 unmeasured characteristics or confounding factors may influence the results. In addition,  
338 considering the specific characteristics of the studied population, the internal and external  
339 validation of results may be limited. To date, there are approximately 25 risk variants  
340 associated with T2DM [6,26] and only 10 polymorphisms were used in the present study. It  
341 is possible, therefore, that if other variables were included in risk prediction models, the



342 results would be even more significant. Another limitation is that we did not assess  
343 postprandial glicemia at baseline in our study. Thus, it is possible that similar information  
344 could be obtained through an oral glucose tolerance test at baseline. As a result, this study  
345 is not able to confirm that the findings are a result of hyperglycemia, or postprandial  
346 hyperglycemia, even in normal levels. Nevertheless, it is interesting to note that genetic  
347 tests may potentially provide more time for the institution of secondary prevention and do  
348 not require to be repeated periodically, characteristics that may allow for the development  
349 of more cost-effective management programs for this specific population.

350         In conclusion, combined information on genetic variants for diabetes risk was  
351 associated with major cardiovascular events incidence in non-diabetic individuals with  
352 coronary artery disease. In addition, our data suggest that non-diabetic individuals with  
353 higher values of genotype scores had an incidence of cardiovascular events similar to  
354 individuals with established diabetes. Other studies are necessary to confirm our results in  
355 similar and different populations.

356

357 **Acknowledgments:** The authors had full access to all of the data in the study and take  
358 responsibility for the integrity of the data and the accuracy of the data analysis. WAH, NL  
359 and JEK had contributions to conception and design of study and acquisition of data; AGPS  
360 and ACP extensively participated in analysis and interpretation of data. All authors had  
361 participation in drafting the article or revising it critically and final manuscript version was  
362 approved by all of them.

363

364 **Competing interests:** The authors have declared that no competing interests exist.

365

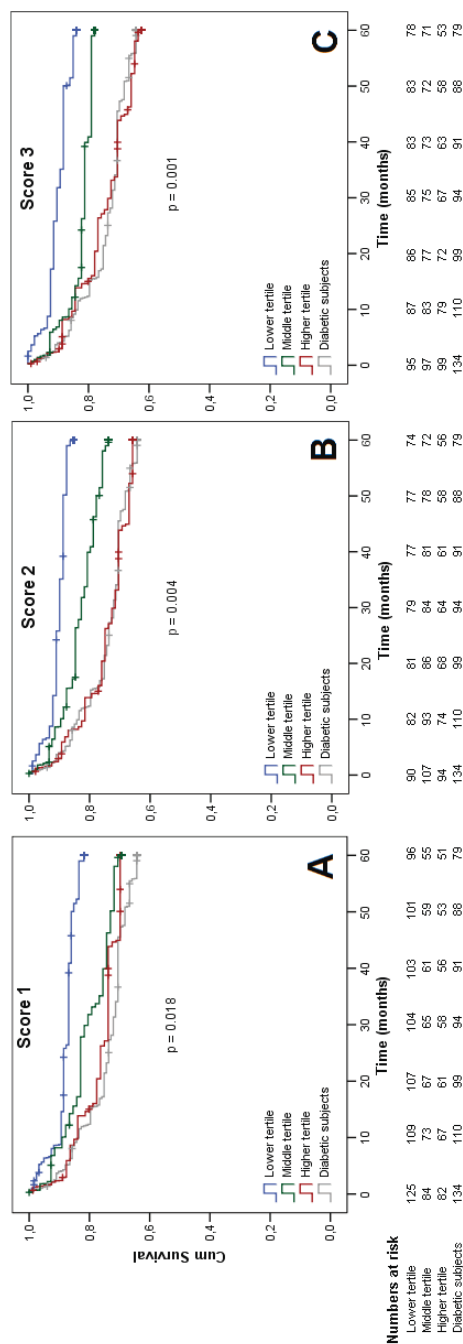
366

367 **References**

- 368 1. IDF (2004) International Diabetes Federation - Diabetes Atlas.  
369 2. (2007) Genome-wide association study of 14,000 cases of seven common diseases and  
370 3,000 shared controls. *Nature* 447: 661-678.  
371 3. Saxena R, Voight BF, Lyssenko V, Burt NP, de Bakker PI, et al. (2007) Genome-wide  
372 association analysis identifies loci for type 2 diabetes and triglyceride levels.  
373 *Science* 316: 1331-1336.  
374 4. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, et al. (2007) A genome-wide  
375 association study of type 2 diabetes in Finns detects multiple susceptibility variants.  
376 *Science* 316: 1341-1345.  
377 5. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. (2007) A genome-wide  
378 association study identifies novel risk loci for type 2 diabetes. *Nature* 445: 881-885.  
379 6. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, et al. (2008) Meta-analysis of  
380 genome-wide association data and large-scale replication identifies additional  
381 susceptibility loci for type 2 diabetes. *Nat Genet* 40: 638-645.  
382 7. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, et al. (2007)  
383 Replication of genome-wide association signals in UK samples reveals risk loci for  
384 type 2 diabetes. *Science* 316: 1336-1341.  
385 8. Lango H, Palmer CN, Morris AD, Zeggini E, Hattersley AT, et al. (2008) Assessing the  
386 combined impact of 18 common genetic variants of modest effect sizes on type 2  
387 diabetes risk. *Diabetes* 57: 3129-3135.  
388 9. Cauchi S, Meyre D, Durand E, Proenca C, Marre M, et al. (2008) Post genome-wide  
389 association studies of novel genes associated with type 2 diabetes show gene-gene  
390 interaction and high predictive value. *PLoS One* 3: e2031.  
391 10. Lyssenko V, Jonsson A, Almgren P, Pulizzi N, Isomaa B, et al. (2008) Clinical risk  
392 factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med* 359:  
393 2220-2232.  
394 11. Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, et al. (2008) Genotype score  
395 in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med*  
396 359: 2208-2219.  
397 12. Miyake K, Yang W, Hara K, Yasuda K, Horikawa Y, et al. (2009) Construction of a  
398 prediction model for type 2 diabetes mellitus in the Japanese population based on 11  
399 genes with strong evidence of the association. *J Hum Genet* 54: 236-241.  
400 13. van Hoek M, Dehghan A, Witteman JC, van Duijn CM, Uitterlinden AG, et al. (2008)  
401 Predicting type 2 diabetes based on polymorphisms from genome-wide association  
402 studies: a population-based study. *Diabetes* 57: 3122-3128.  
403 14. Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, et al. (2006) Combining  
404 information from common type 2 diabetes risk polymorphisms improves disease  
405 prediction. *PLoS Med* 3: e374.  
406 15. Sousa AG, Marquezine GF, Lemos PA, Martinez E, Lopes N, et al. (2009) TCF7L2  
407 polymorphism rs7903146 is associated with coronary artery disease severity and  
408 mortality. *PLoS One* 4: e7697.  
409 16. (2008) Diagnosis and classification of diabetes mellitus. *Diabetes Care* 31: S55-60.  
410 17. Hueb W, Soares PR, Gersh BJ, Cesar LA, Luz PL, et al. (2004) The medicine,  
411 angioplasty, or surgery study (MASS-II): a randomized, controlled clinical trial of

- 412 three therapeutic strategies for multivessel coronary artery disease: one-year results.  
413 J Am Coll Cardiol 43: 1743-1751.
- 414 18. Pereira AC, Lopes NH, Soares PR, Krieger JE, de Oliveira SA, et al. (2006) Clinical  
415 judgment and treatment options in stable multivessel coronary artery disease: results  
416 from the one-year follow-up of the MASS II (Medicine, Angioplasty, or Surgery  
417 Study II). J Am Coll Cardiol 48: 948-953.
- 418 19. Schettert IT, Pereira AC, Lopes NH, Hueb WA, Krieger JE (2006) Association between  
419 platelet P2Y12 haplotype and risk of cardiovascular events in chronic coronary  
420 disease. Thromb Res 118: 679-683.
- 421 20. Hueb W, Lopes NH, Gersh BJ, Soares P, Machado LA, et al. (2007) Five-year follow-  
422 up of the Medicine, Angioplasty, or Surgery Study (MASS II): a randomized  
423 controlled clinical trial of 3 therapeutic strategies for multivessel coronary artery  
424 disease. Circulation 115: 1082-1089.
- 425 21. Myakishev MV, Khripin Y, Hu S, Hamer DH (2001) High-throughput SNP genotyping  
426 by allele-specific PCR with universal energy-transfer-labeled primers. Genome Res  
427 11: 163-169.
- 428 22. Coutinho M, Gerstein HC, Wang Y, Yusuf S (1999) The relationship between glucose  
429 and incident cardiovascular events. A metaregression analysis of published data  
430 from 20 studies of 95,783 individuals followed for 12.4 years. Diabetes Care 22:  
431 233-240.
- 432 23. Haffner SM, Lehto S, Ronnema T, Pyorala K, Laakso M (1998) Mortality from  
433 coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects  
434 with and without prior myocardial infarction. N Engl J Med 339: 229-234.
- 435 24. (1996) Comparison of coronary bypass surgery with angioplasty in patients with  
436 multivessel disease. The Bypass Angioplasty Revascularization Investigation  
437 (BARI) Investigators. N Engl J Med 335: 217-225.
- 438 25. (2003) Is the current definition for diabetes relevant to mortality risk from all causes  
439 and cardiovascular and noncardiovascular diseases? Diabetes Care 26: 688-696.
- 440 26. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, et al. New genetic loci  
441 implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk.  
442 Nat Genet 42: 105-116.
- 443  
444  
445  
  
446  
  
447  
  
448  
  
449  
  
450  
451  
452

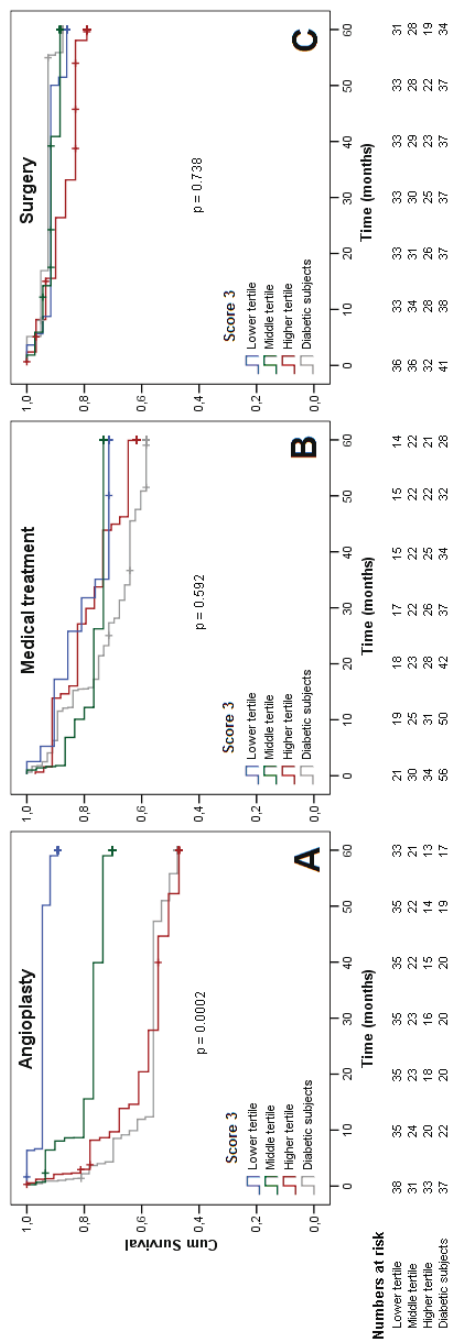
453 Figure 1 – Kaplan-Meier Curves of genotype scores and composite cardiovascular end-point in non-diabetic individuals  
 454 separated by groups according to genotype score tertiles and diabetic subjects after 5 years of follow-up



455

456 Legend: Note diabetic individuals' curves are similar to higher tertiles' curves. P values to pairwise comparisons: A) Lower vs  
 457 middle:  $p = 0.042$ ; lower vs higher:  $p = 0.035$ ; lower vs diabetics:  $p = 0.002$ ; middle vs higher:  $p = 0.909$ ; middle vs diabetics:  $p =$   
 458  $0.381$ ; higher vs diabetics:  $p = 0.482$ . B) Lower vs middle:  $p = 0.052$ ; lower vs higher:  $p = 0.002$ ; lower vs diabetics:  $p = 0.001$ ;  
 459 middle vs higher:  $p = 0.178$ ; middle vs diabetics:  $p = 0.101$ ; higher vs diabetics:  $p = 0.870$ . C) Lower vs middle:  $p = 0.239$ ; lower  
 460 vs higher:  $p = 0.001$ ; lower vs diabetics:  $p = 0.001$ ; middle vs higher:  $p = 0.035$ ; middle vs diabetics:  $p = 0.037$ ; higher vs  
 461 diabetics:  $p = 0.849$ .

462 Figure 2 - Kaplan-Meier curves and composite cardiovascular end-point in non-diabetic subjects separated by groups according  
 463 to Score 3 tertiles and diabetic individuals, in accordance with the kind of treatment received (A - angioplasty, B - medical  
 464 therapy or C - CABG).



465  
 466  
 467 Legend: Note for patients submitted to angioplasty, diabetic subjects and higher tertile non-diabetic individuals had a similar  
 468 end-point incidence. Similar findings can be observed for Scores 1 and 2 (See Supplemental material for detailed information).

469  
 470  
 471  
 472  
 473  
 474

475 Table 1 - Clinical characteristics of diabetic individuals and non-diabetics from MASS-II  
 476 Study

	<b>Diabetes</b> <b>(n=134)</b>	<b>Non-diabetic</b> <b>(n=291)</b>	<b>P value</b>
Age (years)	60.40 ± 8.82	59.31 ± 9.36	0.256
Gender (male %)	63.4%	69.1%	0.250
BMI (kg/m <sup>2</sup> )	27.68 ± 4.44	26.98 ± 4.00	0.109
Total cholesterol (mg/dl)	219.23 ± 50.56	226.31 ± 50.27	0.181
HDL-cholesterol (mg/dl)	38.19 ± 11.27	36.76 ± 9.86	0.200
LDL-cholesterol	143.78 ± 45.67	150.93 ± 44.34	0.141
Triglycerides (mg/dl)	204.78 ± 145.07	190.26 ± 105.24	0.247
Hypertension (%)	67.9%	54.3%	0.008
Smoking (%)	28.4%	35.1%	0.173
Previous MI (%)	42.5%	46.7%	0.419
Three-vessel disease (%)	56.7%	54.6%	0.689

477 Data are shown in mean ± SD or percentage

478

479 Table 2 – Number of composite cardiovascular events and cumulative hazard in diabetic  
 480 subjects and non-diabetic subjects according to genotype score tertiles

	Number of events	Cumulative Hazard (%)	P value
Non-diabetic individuals			
<b>Score 1</b>			0.018*
Lower tertile	22	17.6	
Middle tertile	25	29.1	
Higher tertile	24	29.3	
<b>Score 2</b>			0.004*
Lower tertile	13	14.4	
Middle tertile	27	25.2	
Higher tertile	31	33.0	
<b>Score 3</b>			0.001*
Lower tertile	15	15.8	
Middle tertile	21	21.6	
Higher tertile	35	35.4	
Diabetic individuals	47	35.1	0.022‡

481 (\*) P value for comparison between tertiles and diabetic subjects; (‡) P value for  
 482 comparison between non-diabetic versus diabetic subjects  
 483

484  
 485  
 486  
 487  
 488

Table 3 – Hazard ratio and 95% confidence interval (CI) for composite cardiovascular end-point and mortality according to 1 point increase in each score in the entire population and in diabetic and non-diabetic individuals\*

	Non-diabetic subjects			Diabetic subjects			Entire population		
	HR	95%CI	P value	HR	95%CI	P value	HR	95%CI	P value
<b>Composite Cardiovascular events</b>									
Score 1	1.209	0.99-1.48	0.062	0.996	0.79-1.26	0.974	1.143	0.99-1.32	0.074
Score 2	1.163	1.02-1.32	<b>0.022</b>	1.034	0.87-1.22	0.696	1.113	1.01-1.28	0.033
Score 3	1.079	1.01-1.15	<b>0.015</b>	1.036	0.96-1.12	0.387	1.051	1.01-1.10	0.026
Score 2 no TCF7L2	1.125	0.99-1.29	0.084	1.049	0.88-1.26	0.594	1.104	0.99-1.22	0.061
Score 3 no TCF7L2	1.076	0.99-1.16	0.064	1.007	0.90-1.12	0.896	1.056	0.99-1.12	0.066
<b>Mortality</b>									
	Non-diabetic subjects			Diabetic subjects			Entire population		
	HR	95%CI	P value	HR	95%CI	P value	HR	95%CI	P value
Score 1	1.117	0.85-1.46	0.419	0.861	0.63-1.17	0.349	0.998	0.82-1.22	0.981
Score 2	1.217	1.01-1.46	<b>0.038</b>	0.936	0.75-1.16	0.553	1.081	0.94-1.24	0.261
Score 3	1.110	1.02-1.21	<b>0.017</b>	1.006	0.91-1.12	0.910	1.065	0.99-1.13	0.053
Score 2 no TCF7L2	1.145	0.95-1.38	0.163	0.925	0.74-1.16	0.493	1.034	0.90-1.19	0.638
Score 3 no TCF7L2	1.076	0.97-1.19	0.168	0.996	0.88-1.13	0.953	1.040	0.96-1.12	0.328

(\* ) adjusted by age, gender, arterial hypertension, BMI, total cholesterol, HDL, LDL, triglycerides, previous MI, smoking and fasting glycaemia



### Supplemental Material

This supplementary appendix has been provided by the authors to give readers additional information about their work.

E-Table 1 - Risk variables, multiple logistic regression analysis and punctuation of Score 3 and Score 3 without *TCF7L2*

Gene	SNP	Model	Variants	Score 3		Score 3 no <i>TCF7L2</i>	
				$\beta$ Coefficient	Punctuation ( $\beta \times 10$ )	$\beta$ Coefficient	Punctuation ( $\beta \times 10$ )
KCNJ11	rs5219	Recessive	CT	0.210	0	0.260	0
			TT		2		3
PPARG	rs1801282	Additive	GC	0.431	4	0.458	5
			CC		8		10
TCF7L2	rs7903146	Additive	CT	0.382	4	0.178	0
			TT		8		2
SLC30A8	rs13266634	Recessive	TC	0.193	0	0.312	0
			CC		2		3
HHEX	rs1111875	Recessive	AG	0.332	0	0.203	2
			GG		3		2
CDKAL1	rs7754840	Dominant	GC	0.170	2	0.247	2
			CC		2		4
IGF2BP2	rs4402960	Additive	GT	0.271	3	0.324	0
			TT		6		3
CDKN2A/B	rs10811661	Recessive	CT	0.321	0	0.179	0
			TT		3		2
CDKN2A/B	rs9300039	Recessive	AC	0.148	0	0.118	1
			CC		1		1
FTO	rs8050136	Dominant	CA	0.086	1	0.260	0
			AA		1		3

Legend: beta coefficient estimates the results from a regression analysis performed on variables that have been standardized so that they have variances of 1. The different models the fitting model for each SNP was chosen based on the best model in univariate analyses. The reason to calculate another score after excluding *TCF7L2* is justified by previous study that found positive association between *TCF7L2* risk allele and cardiovascular events. Therefore, it is important to exclude whether the results was totally dependent of *TCF7L2*.

E-Table 2 - Risk variables, multiple logistic regression analysis and points of Clinical Score and Derivative Scores

Variable	Groups	$\beta$ Coefficient	Punctuation ( $\beta \times 10$ )
<b>For Scores 1 and 2</b>			
Gender	Male sex	0.066	1
Smoking	Smoking (+)	0.170	2
Hypertension	Hypertension (+)	0.533	5
Obesity	BMI $\geq 30$ kg/m <sup>2</sup>	0.264	3
Age	55-65yrs	0.347	3
	$\geq 65$ yrs	0.250	3
Score 1	4	0.344	3
	$\geq 5$	0.824	8
Score 2	12-13	0.231	3
	$\geq 14$	0.752	8
<b>For Scores 3 and 3 without TCF7L2</b>			
Gender	Male sex	0.046	0
Smoking	Smoking (+)	0.187	2
Hypertension	Hypertension (+)	0.543	5
Obesity	BMI $\geq 30$ kg/m <sup>2</sup>	0.315	3
Age	55-65yrs	0.447	4
	$\geq 65$ yrs	0.348	4
Score 3	19-22	0.422	4
	$\geq 23$	1.108	11
Score 3 without	17-19	0.696	7
TCF7L2	$> 20$	0.810	8

E-Table 3 – Genetic Factors Predicting Type 2 Diabetes in the MASS-II Population

Gene	Chromo-	SNP	Risk Allele	RA <sub>AFF</sub>	Additive Model		Alternative Model		
					Odds Ratio	P value	Odds Ratio	P value	
<i>KCNJ11</i>	11	rs5219	T	0.33	1.003 (0.738-1.364)	0.984	Recessive	1.398 (0.744-2.627)	0.298
<i>PPARG</i>	3	rs1801282	C	0.93	1.464 (0.837-2.560)	0.182	Recessive	1.444 (0.819-2.548)	0.204
<i>TCF7L2</i>	10	rs7903146	T	0.47	1.511 (1.069-2.135)	0.019	Recessive	2.039 (1.106-3.760)	0.022
<i>SLC30A8</i>	8	rs13266634	C	0.76	1.137 (0.818-1.582)	0.444	Recessive	1.238 (0.816-1.880)	0.315
<i>HHEX</i>	10	rs1111875	G	0.68	1.274 (0.946-1.716)	0.111	Recessive	1.420 (0.939-2.147)	0.096
<i>CDKALI</i>	6	rs7754840	C	0.37	1.198 (0.885-1.622)	0.243	Dominant	1.291 (0.850-1.963)	0.231
<i>IGF2BP2</i>	3	rs4402960	T	0.39	1.237 (0.899-1.702)	0.191	Recessive	1.419 (0.744-2.707)	0.289
<i>CDKN2A/B</i>	9	rs10811661	T	0.85	1.307 (0.880-1.941)	0.184	Recessive	1.406 (0.894-2.212)	0.141
<i>CDKN2A/B</i>	11	rs9300039	C	0.90	1.095 (0.684-1.753)	0.705	Recessive	1.214 (0.718-2.051)	0.470
<i>FTO</i>	16	rs8050136	A	0.40	1.010 (0.743-1.374)	0.948	Dominant	1.167 (0.756-1.801)	0.486

The odds ratios for the risk of type 2 diabetes were calculated with the use of univariate logistic-regression analyses with adjustment for age at participation and sex. The primary genetic models are additive; alternative models are indicated. CI denotes confidence interval, and RAAFF frequency of the risk allele in affected subjects. After Bonferroni's correction for multiple statistical tests (considering a p value more restrictive), no polymorphisms was associated with T2DM.

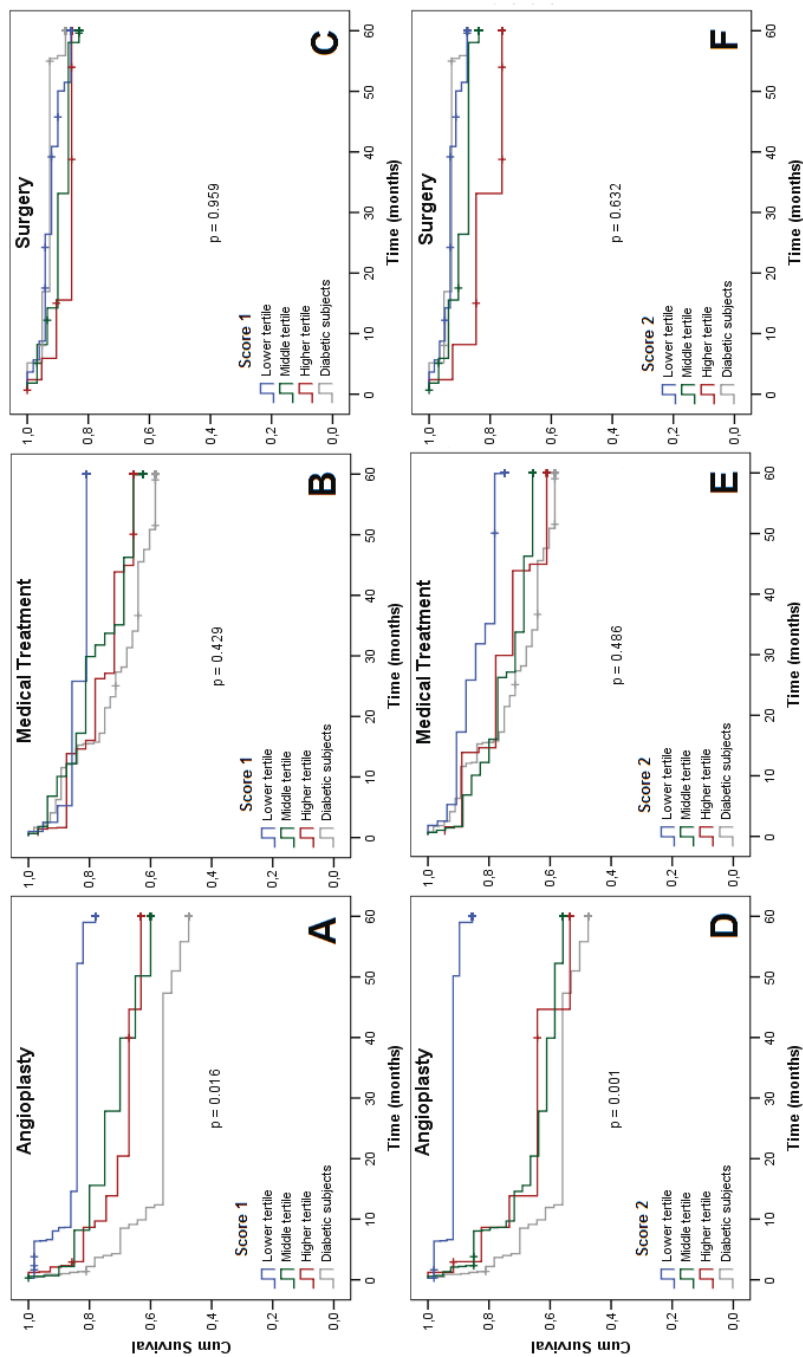
E-Table 4 – Combination of Clinical Factors and Genotype Scores Predicting Type 2 Diabetes in the MASS-II Study calculated by multivariate logistic regression

	<b>Odds Ratio (95%CI)</b>	<b>P Value</b>
<b>Recessive Score (Score 1)</b>		
Male sex	1.048 (0.664-1.655)	0.839
Age, per 1 year	1.012 (0.988-1.037)	0.333
Smoking	1.210 (0.745-1.963)	0.441
Hypertension	<b>1.707 (1.089-2.675)</b>	<b>0.020</b>
Increased body-mass index, per 1 kg/m <sup>2</sup>	1.030 (0.978-1.085)	0.257
Combination of risk alleles Score 1	<b>1.305 (1.105-1.542)</b>	<b>0.002</b>
<b>Additive Score (Score 2)</b>		
Male sex	1.068 (0.677-1.684)	0.778
Age, per 1 year	1.013 (0.988-1.038)	0.311
Smoking	1.169 (0.722-1.892)	0.525
Hypertension	<b>1.674 (1.069-2.622)</b>	<b>0.024</b>
Increased body-mass index, per 1 kg/m <sup>2</sup>	1.032 (0.980-1.087)	0.232
Combination of risk alleles Score 2	<b>1.178 (1.056-1.314)</b>	<b>0.003</b>
<b>Pondered Score (Score 3)</b>		
Male sex	1.048 (0.661-1.660)	0.842
Age, per 1 year	1.014 (0.990-1.040)	0.240
Smoking	1.176 (0.723-1.912)	0.514
Hypertension	<b>1.701 (1.082-2.676)</b>	<b>0.021</b>
Increased body-mass index, per 1 kg/m <sup>2</sup>	1.036 (0.983-1.092)	0.183
Combination of risk alleles Score 3	<b>1.107 (1.052-1.164)</b>	<b>&lt;0.0001</b>

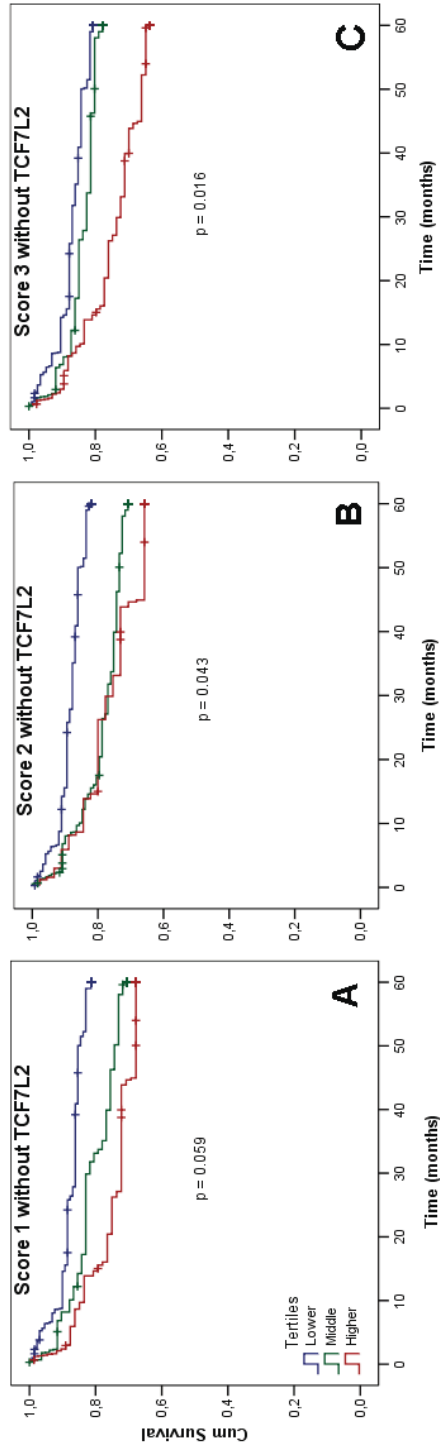
E-Table 5 - Proportion of subjects with T2DM according to genetic risk group (based in tertile of genotypes scores) and risk estimative

	<b>Diabetes (%)</b>	<b>Odds Ratio (95%CI)</b>	<b>P value</b>
<b>Score 1</b>			
Low risk ( $\leq 3$ )	24.2%		
High risk ( $\geq 5$ )	41.4%	2.210 (1.354-3.607)	0.0015
<b>Score 2</b>			
Low risk ( $\leq 11$ )	26.3%		
High risk ( $\geq 14$ )	44.2%	2.214 (1.273-3.852)	0.0049
<b>Score 3</b>			
Low risk ( $\leq 12$ )	21.7%		
High risk ( $\geq 17$ )	42.6%	2.677 (1.569-4.567)	0.0003

E-Figure 1 - Kaplan-Meier Curves of non-diabetic subjects separated by groups according to genotype scores tertiles and diabetic individuals, in accordance with the kind of treatment received (A and D – angioplasty; B and E – medical therapy; C and F – coronary bypass surgery). Score 1 data are shown in boxes A, B and C; Score 2 data in boxes D, E and F.

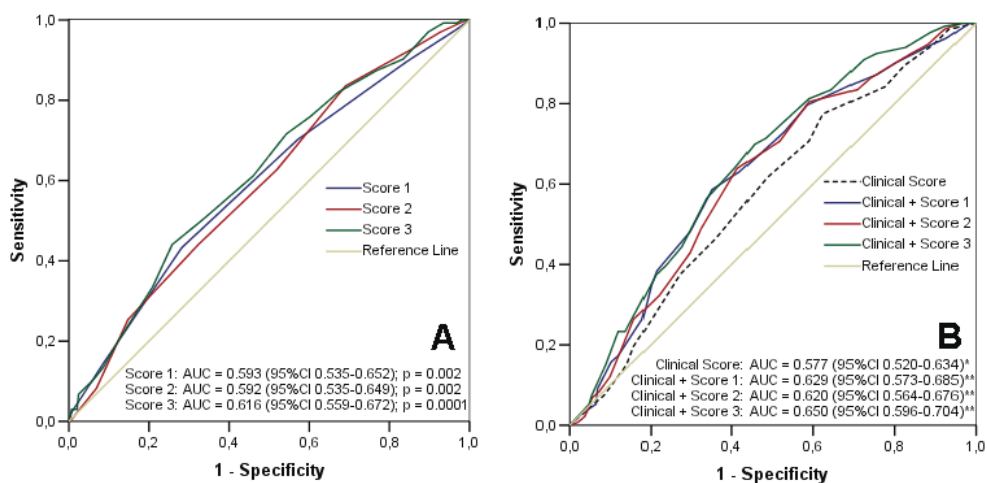


E-Figure 2 – Kaplan-Meier Curves of genotype Scores and genotype Scores without TCF7L2 data and composite cardiovascular end points in non-diabetic individuals separated by groups according to genotype scores tertiles after 5 years of follow-up



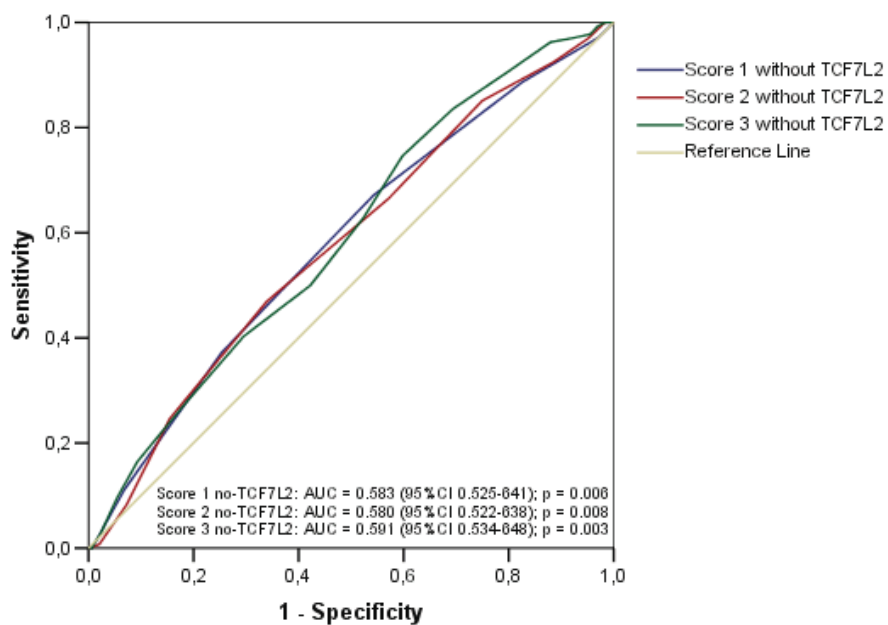


E-Figure 3 – Discriminatory ability for diabetes of genotype scores and clinical risk factors



Legend: A) ROC Curves for T2DM diagnosis for derived scores in the MASS-II population. The p values refer to asymptotic significance; B) Increased area under ROC curves (AUC) after addition of genotypes information into clinical variables score for diabetes diagnosis. Asymptotic significance: (\*) P value = 0.011; (\*\*) p value < 0.0001. P values for comparisons between Clinical score and addition of: Score 1= 0.0243; Score 2 = 0.0292; Score 3 = 0.0380

E-Figure 4 - ROC curves for T2DM diagnosis for three scores without TCF7L2 data in MASS-II population



Eur J Epidemiol (2009) 24:101–109  
DOI 10.1007/s10654-009-9314-2

DIABETES

## Derivation and external validation of a simple prediction model for the diagnosis of type 2 Diabetes Mellitus in the Brazilian urban population

André Gustavo Pires de Sousa · Alexandre Costa Pereira · Guilherme Figueiredo Marquezine · Raimundo Marques do Nascimento-Neto · Silvia N. Freitas · Roney Luiz de C. Nicolato · George Luiz Lins Machado-Coelho · Sérgio L. Rodrigues · José G. Mill · José Eduardo Krieger

Received: 6 August 2008 / Accepted: 31 December 2008 / Published online: 4 February 2009  
© Springer Science+Business Media B.V. 2009

**Abstract** A risk score model was developed based in a population of 1,224 individuals from the general population without known diabetes aging 35 years or more from an urban Brazilian population sample in order to select individuals who should be screened in subsequent testing and improve the efficacy of public health assurance. External validation was performed in a second, independent, population from a different city ascertained through a similar epidemiological protocol. The risk score was developed by multiple logistic regression and model performance and cutoff values were derived from a receiver operating char-

acteristic curve. Model's capacity of predicting fasting blood glucose levels was tested analyzing data from a 5-year follow-up protocol conducted in the general population. Items independently and significantly associated with diabetes were age, BMI and known hypertension. Sensitivity, specificity and proportion of further testing necessary for the best cutoff value were 75.9, 66.9 and 37.2%, respectively. External validation confirmed the model's adequacy (AUC equal to 0.72). Finally, model score was also capable of predicting fasting blood glucose progression in non-diabetic individuals in a 5-year follow-up period. In conclusion, this simple diabetes risk score was able to identify individuals with an increased likelihood of having diabetes and it can be used to stratify subpopulations in which performing of subsequent tests is necessary and probably cost-effective.

**Keywords** Type 2 diabetes · Diabetes prediction model · Risk score · Diabetes Mellitus

### Abbreviations

ADA	American Diabetes Association
AUC	Area under curve
BMI	Body mass index
CVD	Cardiovascular diseases
DBP	Diastolic blood pressure
EPV	Events per variable
FPG	Fasting plasma glucose
HDL-c	High density lipoprotein cholesterol
IFG	Impaired fasting glycemia
IGT	Impaired glucose tolerance
LDL-c	Low density lipoprotein cholesterol
OGTT	Oral glucose tolerance test
PCOS	Polycystic ovarian syndrome
ROC curve	Receiver operating characteristic curve

A. G. Pires de Sousa · A. C. Pereira (✉) · G. F. Marquezine · J. E. Krieger  
Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of São Paulo Medical School, Av Dr. Enéas Carvalho de Aguiar, 44, Instituto do Coração, 10th Floor, São Paulo-SP, 05403-000, Brazil  
e-mail: alexandre.pereira@incor.usp.br

A. G. Pires de Sousa  
Clinical Medicine Department, Federal University of Rio Grande do Norte, Natal, Brazil

R. Marques do Nascimento-Neto  
Arterial Hypertension Institute, Belo Horizonte, Brazil

R. Marques do Nascimento-Neto  
Federal University of Ouro Preto Medical School, Ouro Preto, Brazil

S. N. Freitas · R. L. de C. Nicolato · G. L. L. Machado-Coelho  
Pharmacy Department, Federal University of Ouro Preto, Ouro Preto, Brazil

S. L. Rodrigues · J. G. Mill  
Department of Physiological Sciences, Federal University of Espirito Santo, Vitoria, Brazil

SBP	Systolic blood pressure
T2DM	Type 2 Diabetes Mellitus

## Introduction

Type 2 Diabetes Mellitus (T2DM) is a common disease with worldwide increasing incidence. In the United States, from 1980 to 2005 the number of adults aged 18–79 with newly diagnosed diabetes almost tripled from 493,000 in 1980 to 1.4 million in 2005 [1]. In Brazil, the prevalence of T2DM in subjects aged 30–69 years was 7.6% in 1990 [2] and this value is likely to be increasing mainly because of the increasing prevalence of obesity, change in nutritional patterns and population aging. Some studies have shown that even newly diagnosed diabetic individuals have one or more micro- or macrovascular complications at the time of diagnosis, including cardiovascular disease (CVD) [3]. Therefore, early diagnosis and prompt treatment are essential to avoid complications. However, there are no public health policies aiming at the screening of T2DM in most countries. The American Diabetes Association (ADA) 2008 position statement recommends that screening pre-diabetes and type 2 diabetes in asymptomatic people (using fasting glycemia) should be considered in adults who are overweight or obese and who have at least one additional risk factor for diabetes, such as physical inactivity, first-degree relative with diabetes, members of a high-risk ethnic population (e.g., African American, Latino, Native American, Asian American, and Pacific Islander), women who delivered a baby weighing  $\geq 4,000$  g (9 lb) or with diagnosed gestational Diabetes Mellitus, hypertension, HDL cholesterol level  $< 35$  mg/dl (0.90 mmol/l) and/or a triglyceride level  $\geq 250$  mg/dl (2.82 mmol/l), women with polycystic ovarian syndrome (PCOS), impaired glucose tolerance (IGT) or impaired fasting glycemia (IFG) on previous testing, history of CVD or other clinical conditions associated with insulin resistance (e.g., severe obesity and acanthosis nigricans). In those without these risk factors, testing should begin at age 45. Repeated testing should be carried out at least at 3-year intervals.

Screening of new cases in large populations, however, is a high-cost process. In order to diminish these costs where the screening of almost the entire population is unfeasible, the creation of scores aiming at identifying high-risk groups to have or to develop T2DM could reduce the need of subsequent tests and improve the efficacy of public health assurance. Unfortunately, the utilization of scores from other populations are inadequate mainly because the intensity of risk conferred for diabetes varies between different ethnic groups [4]. Consequently, it is necessary to develop specific risk scores for different populations.

The aim of this study was to develop for the Brazilian population a simple risk score that, using only data that are routinely collected in general medical practice, could be used by health-care workers to select individuals who should be screened for further laboratory confirmation. According to previous studies with the same objectives, this score should be able to identify at least 75% of individuals with diabetes [5].

## Methods

In this study, two different Brazilian urban populations were used. The prediction model was developed based in a database from the city of Vitoria, and it was externally validated using data from another population-based sample, Ouro Preto, Brazil.

### Derivation population—Vitoria, Brazil

A cross-sectional study of risk factors for cardiovascular diseases was performed in the urban population of Vitoria, Brazil, using the WHO-MONICA project guidelines [6]. The study design was based on cross-sectional research methodology and was developed by means of surveying and analyzing socioeconomic and health data in a probabilistic sample of residents from the municipality of Vitoria, ES, Brazil. The sampling plan had the objective of ensuring that the research would be socioeconomically, geographically and demographically representative of the residents of the municipality. The resident population aged 25–64 years was studied. According to the census carried out by IBGE foundation in the year 1996, the resident population of Vitoria was 265,874 inhabitants. The IBGE census is the most trustworthy database of Brazilian population and it is performed once a decade. The sampling was performed in four stages: by district, IBGE census sector, drawing lots to choose homes, and drawing lots to choose an individual from each home. The survey was conducted with just one resident within the age group of the study in each home drawn. The draw was carried out by means of a randomization mechanism. A selection of 2,268 residential homes located in Vitoria was obtained and these were visited. The individual selected at each of these homes was given explanations of the purposes of the research and invited to participate in the study, after obtaining his or her consent. The project received approval from the Ethics Committee of the Biomedical Center of the Federal University of Espírito Santo. The selected individuals were asked to attend on the following day the Cardiovascular Investigation Clinic of the University Hospital to be submitted to clinical and laboratory tests necessary to estimate their cardiovascular risk profile. Of

the total sample, 1,507 individuals attended. Participants were submitted to physical examination emphasizing measurement of height, weight, waist circumference and blood pressure. Major cardiovascular risk factors such as smoking habits, alcohol intake, amount of physical activity, diabetes and hypertension were inquired. Sedentary lifestyle was defined as exercise less than 1 h, three times a week. Blood glucose, total cholesterol, lipoprotein fractions, and triglycerides were assayed by standard techniques in 12 h fasting blood sample collected during the clinic visit [7].

#### Follow-up database

A 5-year follow-up reevaluation was performed in a subgroup of the Vitoria population sample. All participants of the first part of the WHO-MONICA study in Vitoria were invited to be reassessed after 5 years. Of that first sample, 655 participants were re-evaluated through the same methodology and the same clinical and laboratory measures obtained at baseline evaluation.

#### External validation population—Ouro Preto, Brazil

Ouro Preto is a small-sized city located in the metallurgic zone of Minas Gerais State, southeast of Brazil, and it comprises around 37,603 inhabitants. In a cross-sectional study, 930 homes were randomly selected out of 33 city districts on a simple sample basis. Proportionality of population density within each district was systematically considered throughout the selection process. On each sampled home, an individual aging 15 years or more whose birthday was nearby the interview date was selected for study inclusion. Participation in the study was voluntary, and a signed informed consent was obtained from all participants. Blood pressure was measured at home and sociodemographic information as gender, age, education, and economical class were obtained upon personal interview. Clinical conditions such as diabetes, arterial hypertension and dyslipidemia, besides behavior parameters such as smoking or ingestion of alcoholic beverages were also acquired. Subsequently, agreeable volunteers were encouraged to present themselves at the health service of Federal University of Ouro Preto for collection of blood samples and anthropometric measurements. The study was previously approved by the Ethics Review Board of the Federal University of Ouro Preto. Previous studies have been already published using this same database [8].

#### Selection of cases and definitions

All participants without known diabetes were enrolled. Glucose tolerance was classified according to a modified

ADA 2008 criteria [9]. Dysglycemia was defined as having fasting plasma glucose (FPG) levels  $\geq 100$  mg/dl. Subjects with FPG  $\geq 126$  mg/dl (7.0 mmol/l), that is a provisional diagnosis of diabetes according ADA criteria, were classified as T2DM patients in this study. As made in other studies, in order to identify T2DM subjects, the individuals aged 34 years or less were excluded from the analysis due to the possibility of this group having hyperglycemic non-T2DM individuals.

Body mass index (BMI) (weight in kg/height in  $m^2$ ) was calculated and overweight or obesity defined as a BMI  $\geq 25$   $kg/m^2$ . Blood pressure was measured using a standard mercury sphygmomanometer on the left arm in the sitting position, after 5 min rest. Systolic and diastolic blood pressures were calculated from three readings, with a minimal interval of 5 min between readings. In order to evaluate systolic and diastolic blood pressures, the mean of these three readings was considered. The definition of arterial hypertension was the mean systolic blood pressure (SBP) above 140 mmHg or the diastolic blood pressure (DBP) above 90 mmHg, or use of antihypertensive drugs.

The individuals were classified into five socioeconomic classes (here grouped into four classes), on the basis of data on the education level of the head of the family and the type and quantity of domestic equipment in the home (IBGE foundation) [10]. The ethnic morphological subgroup classification was based on phenotype pigmentation of the abdomen, hair color and shape, type and conformation of the nose and lips, and family history, as determined and agreed upon by two examining physicians. It should be noted that our subgroup stratification was based on morphological criteria and not on a true ethnic group stratification. Ethnic group stratification is defined by a group of genetic, linguistic and cultural characteristics. In particular, it has been shown that in the Brazilian population morphological criteria are not highly concordant with ancestry. This classification, however, has been validated and previously used in other studies [11–13].

#### Development of risk score and validation

Risk factors and laboratory analysis known clinically to be associated with diabetes were selected for the development of a risk score. These selected risk factors included age, sex, diagnosis of hypertension and physical activity. Laboratory analysis included total cholesterol, low density lipoprotein cholesterol (LDL-c), high density lipoprotein cholesterol (HDL-c), triglycerides and uric acid. Univariate analyses were performed by logistic regression with T2DM (FPG  $> 126$  mg/dl) as the dependent variable. In the multiple logistic regression, a criterion for including variables with  $P < 0.1$  was used. These variables were included using stepwise backward elimination, with T2DM as the

dependent variable. Logistic regression coefficients were used to estimate odds ratios for each of the independent variables in the model. The continuous variables were categorized into bands. To create the final model, variables with a *P* value less than 0.05 were considered and for each significant variable in the multiple logistic regression analysis, a score was calculated by multiplying the regression coefficient by 10 and rounding to the nearest integer. Finally, a sum score was calculated for each participant by adding the score for each significant variable in the risk model. Sensitivity, specificity, positive and negative predictive values and accuracy were compared for different score cutoffs. In addition, the proportion of subjects that needed subsequent testing (those individuals who have score values above the cutoff value) was compared between these several groups.

In order to evaluate the model performance, a receiver operating characteristic (ROC) curve was built to plot probabilities saved with the logistic regression procedure and the area under ROC curve (AUC) was used to measure the power to discriminate high-risk from low-risk individuals. ROC curves were plotted for the diabetes risk score, the sensitivity was plotted on the *y*-axis and the false positive rate (1—specificity) was plotted on the *x*-axis. The more accurate discriminating the test, the steeper the upward portion of the ROC curve and the higher the area under the curve (AUC). The optimal point was determined aiming to achieve the highest sensitivity and specificity. To evaluate the performance of the risk score in the second population (external validation), the methods used were similar to those for the exploratory set.

#### Statistical analysis

SPSS version 13 program was used for statistical analysis. The goodness of fit for normal distribution was evaluated using the Kolmogorov–Smirnov test. To test for differences in various characteristics, the Student *t* test or ANOVA were used for continuous variables and the  $\chi^2$  test was used for categorical variables. A *P* value less than 0.05 on a two-sided test was considered significant. Data are showed as mean  $\pm$  standard deviation to continuous variables (scale) and as percentage to ordinal or categorical variables.

#### Results

In the Vitoria urban population, of the total sample of 1,507 without known diabetes individuals, 345 subjects were excluded from analysis because they were less than 35-years-old (including five diabetics subjects). Additionally, in 8 of 1,232 subjects the glycemia data were absent,

totalizing 1,224 subjects for final analysis. In these individuals, the prevalence of T2DM was 9.6% ( $n = 118$ ). In the Ouro Preto population (validation), of the initial sample of 939 individuals, 474 individuals were excluded of analysis (35 persons were known diabetics; 348 persons were less than 35-years-old, and 91 individuals had no glycemia registry), totalizing 467 non-diabetics individuals in the final sample. Table 1 shows the clinical and laboratorial characteristics of the T2DM and non-diabetics subjects from Vitoria (derivation population) and from Ouro Preto (external validation population). From these data, it can be observed that the Ouro Preto population is older and has a higher prevalence of hypertension, two previously known risk factors for T2DM.

#### Development of risk scores, comparison of models and performance evaluation

Table 2 shows results from univariate logistic regression analysis of potential risk factors for diabetes. A model containing only demographic and/or anthropometric variables was created. The values of the  $\beta$  coefficient, odds ratio with 95% confidence interval (95%CI) and *P* value of each independent variable in the predictive model can be seen in Table 3. In this model, whose highest possible value is 36, increasing age, increasing BMI and hypertension were significantly associated with T2DM. The performance of the predictive risk score can be observed by the ROC curves in Fig. 1A and in Table 4. The best value in ROC curves that achieved the highest sensitivity and specificity was 17.5. Cutoff value of the sum score  $\geq 18$  showed sensitivity close to the pre-specified value (about 75%). The utilization of this cutoff should reduce the necessity of further laboratory investigation to about 37% of the whole population (Table 4). Although increasing the cutoff value reduces the percentage of subsequent testing, it also decreases the sensitivity. Therefore, because sensitivity  $\geq 75\%$  was pre-specified, the cutoff value of  $\geq 18$  points was chosen for evaluation. The model achieves a good performance with a discriminatory capacity of about 77% (AUC) and it significantly correlates with fasting baseline glycemia ( $r = 0.303$ ;  $P < 0.0001$ ).

#### External validation

Although the score performance in the Ouro Preto population was lower than in Vitoria, the area under the ROC curve in Ouro Preto study was equal to 0.72, indicating a good overall performance in this validation population (Fig. 1B). Sensitivity, specificity, predictive values and the percentage of individuals who needed subsequent testing by a cutoff  $>18$  are shown in Table 4.

**Table 1** Baseline characteristics of diabetics and non-diabetics subjects in Vitoria (derivation population) and Ouro Preto (validation population)

	Derivation population (Vitoria)			Validation population (Ouro Preto)		
	T2DM ( <i>n</i> = 118)	Non-T2DM ( <i>n</i> = 1106)	<i>P</i> value*	T2DM ( <i>n</i> = 29)	Non-T2DM ( <i>n</i> = 438)	<i>P</i> value*
Age (year)	53.4 ± 7.4	48.4 ± 8.3	<0.0001	59.5 ± 11.0	51.9 ± 12.0	0.001
Gender (%)						
Male	39.8	45.8	0.216*	17.2	26.3	0.282
Female	60.2	54.2		82.8	73.7	
Ethnicity (%)			0.54	NA	NA	–
White	33.9	37.7				
Mulatto	55.1	48.2				
Black	6.8	7.3				
Others	4.2	5.9				
Sedentarism (%)	73.3	72.4	0.32	NA	NA	–
Smoking (%)	26.3	22.6	0.098	17.8	21.4	0.97
BMI (kg/m <sup>2</sup> )	30.64 ± 5.89	26.15 ± 4.41	<0.0001	29.50 ± 5.59	26.31 ± 4.72	0.001
Abdominal circumference (cm)	96.91 ± 12.81	86.75 ± 11.48	0.18	97.96 ± 13.3	87.37 ± 12.7	<0.0001
Waist-hip ratio	0.93 ± 0.08	0.88 ± 0.09	0.154	NA	NA	–
SBP (mmHg)	144 ± 23	129 ± 22	0.35	157 ± 25.6	142 ± 27.5	0.006
DBP (mmHg)	90 ± 14	85 ± 14	0.98	91 ± 16.2	89 ± 15.1	0.48
Hypertension (%)	14.6	5.1	<0.0001	89.7	65.5	0.027
Glycemia (mg/dl)	185.34 ± 69.78	99.07 ± 10.93	<0.0001	155.1 ± 41.0	96.45 ± 11.0	<0.0001
Total cholesterol (mg/dl)	243.84 ± 68.52	218.83 ± 44.1	0.01	205.14 ± 43.1	191.21 ± 38.0	0.063
LDL-c (mg/dl)	159.5 ± 40.2	146.4 ± 38.9	0.6	107.68 ± 39.2	102.41 ± 29.2	0.391
HDL-c (mg/dl)	42.94 ± 9.78	45.43 ± 12.06	0.25	63.82 ± 13.9	63.84 ± 14.2	0.994
Triglycerides (mg/dl)	219.50 ± 251.46	138.82 ± 116.46	<0.0001	148.72 ± 56.0	121.0 ± 68.0	0.046
Uric acid (mg/dl)	5.20 ± 1.38	4.82 ± 1.56	0.2	NA	NA	–
Creatinine (mg/dl)	0.97 ± 0.19	0.98 ± 0.2	0.64	NA	NA	–

\* *P* value for comparison between diabetic and non-diabetics subjects in each population

NA not available

### Follow-up analysis

The highest tercile of the selected model scores in baseline conditions, whose cutoff value is coincidentally the same as the best value encountered by ROC curve analysis ( $\geq 18$  points), has a baseline and 5-years follow-up fasting glycemia mean significantly higher (Fig. 2), indicating that even after a 5-year follow-up period the selected model score was still significantly correlated with fasting blood glucose.

### Discussion

In this study, a simple score was developed to evaluate the risk of T2DM in the Brazilian urban population. The utilization of this type of model score of points was able to identify 77.2% of individuals with previously

undiagnosed T2DM. It was also able to identify such individuals without the necessity of laboratory tests in the screening phase.

In the conduction of multivariate analysis, we judged that BMI and abdominal circumference would take the same information (related to obesity) and we chose one of them (in our case, BMI because it has presented higher  $\beta$  coefficient values in univariate regression analysis).

The procedure used to develop the risk score was equivalent to the methods used in other studies in different populations, although there are differences in the risk variables included [4, 5, 14, 15]. These studies also included subjects aged above 35 or 40 years old in analysis aiming to reduce the probability of including non-T2DM patients. Other studies have included different risk factor variables. The Danish diabetes risk score incorporated age, male sex, known hypertension, low

**Table 2** Univariate analyses of risk factors for diabetes

	Odds ratio	95%CI	$\beta$	P value
Age (year)				
45–54 versus 35–44 year	2.72	1.52–4.85	1.0	0.001
55 or more versus 35–44 year	4.64	2.65–8.13	1.535	<0.0001
Gender				
Female:male	1.276	0.867–1.879	0.24	0.22
Ethnicity (%)				
Mulatto versus white	1.271	0.840–1.924	0.240	0.256
Black versus white	1.030	0.465–2.281	0.029	0.943
Others versus white	0.802	0.305–2.106	–0.221	0.654
Sedentarism (%)	0.891	0.573–1.384	–0.116	0.607
Smoking (%)	1.233	0.799–1.903	0.209	0.345
BMI (kg/m <sup>2</sup> )				
25–29.9 versus <25	1.887	1.066–3.339	0.635	0.029
30 or more versus <25	8.052	4.736–13.689	2.086	<0.0001
Abdominal circumference (cm)				
Men $\geq$ 94 and women $\geq$ 80 versus normal AC	3.378	2.027–5.629	1.217	<0.0001
M 94–101, 9/W 80–87.9 versus normal AC	1.539	0.776–3.053	0.431	0.217
M $\geq$ 102/W $\geq$ 88 versus normal AC	5.206	3.043–8.905	1.650	<0.0001
Hypertension	3.156	2.076–4.799	1.149	<0.0001
CT (mg/dl)				
$\geq$ 240 versus <240	2.259	1.537–3.321	0.815	<0.0001
LDL (mg/dl)				
$\geq$ 130 versus <130	1.861	1.170–2.961	0.621	0.009
HDL (mg/dl)				
<40M/50W versus $\geq$ 40M/50W	1.266	0.844–1.899	0.236	0.254
TG (mg/dl)				
$\geq$ 150 versus <150	2.776	1.886–4.087	1.021	<0.0001
Uric acid (mg/dl)				
$\geq$ 5.5 or versus <5.5	1.161	1.034–1.303	0.149	0.011
Creatinine (mg/dl)	0.768	0.291–2.026	–0.264	0.594

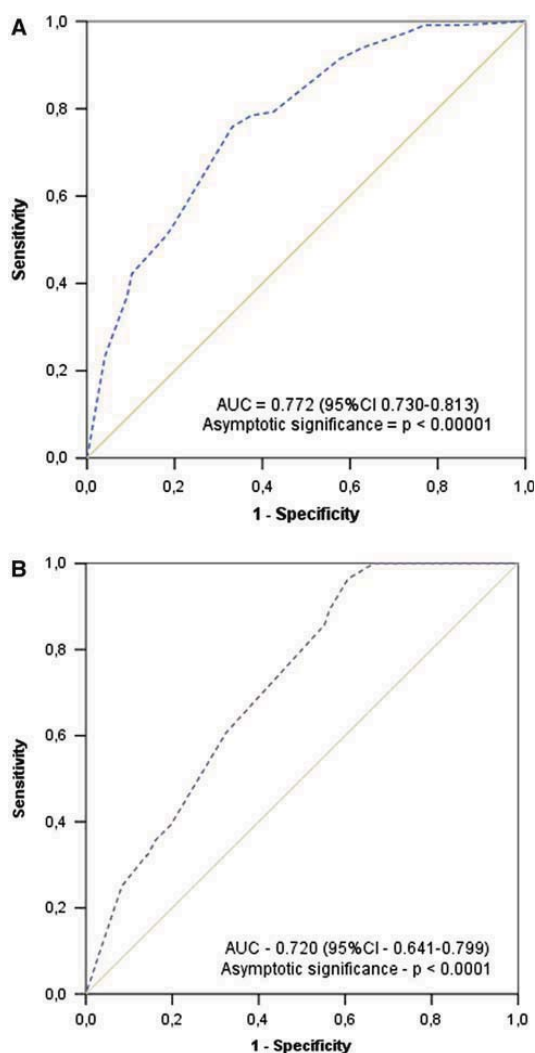
**Table 3** Risk variables and multiple logistic regression analysis

	$\beta$	Odds ratio	95%CI	P value	Risk score ( $\beta \times 10$ )
<i>Model</i>					
Age					
45–54 versus 35–44 year	0.743	2.102	1.150–3.842	0.016	7
55 or more versus 35–44 year	1.227	3.412	1.894–6.145	<0.0001	12
BMI (kg/m <sup>2</sup> )					
25–29.9 versus <25	0.473	1.605	0.899–2.866	0.110	5
30 or more versus <25	1.802	6.060	3.492–10.518	<0.0001	18
Hypertension	0.626	1.870	1.178–2.969	0.008	6

physical activity and family history of diabetes as the main risk contributors [5]. In our study, gender and sedentary lifestyle were not significantly associated with T2DM. Additionally, there are no data on familiar history

for diabetes in the Vitoria database and possibly the inclusion of this variable could further improve the performance of our model. In the Finnish study, a history of high blood glucose was also strongly predictive of T2DM





**Fig. 1** (A) ROC curve for the risk score in the derivative population; (B) ROC curve for the risk score in validation population. *Note:* The asymptotic significance was calculated for each ROC curve in relation to null hypothesis and it does not calculate differences between curves

[15]. This parameter was also not included in the Vitoria database. Other scores have included biochemical measures [16, 17] but the applicability of such scores is related to the availability of this routine biochemical information at the population level. In less developed countries, such as Brazil, the inclusion of these biochemical variables in prediction models is less likely to be cost-effective at the population level.

Unlike other studies that used the oral glucose tolerance test (OGTT) for the diagnosis of T2DM, in this study we

used the value of fasting glycemia above 126 mg/dl (7 mmol/l) because OGTT was not performed in the Vitoria population. Thus, certainly some cases of T2DM remained undiagnosed and this could have potentially influenced the final model estimative. Nevertheless, previous studies such as the 1999 Health Survey for England also used a fasting glucose value  $\geq 126$  mg/dl (7 mmol/l) for diagnosis of diabetes [4, 18].

In our study, the population sample size was comparable to other studies [14]. Steyerberg et al. describe that overoptimistic estimates of predictive performance are a common problem of predictive models, especially in small data sets, in that there are relatively few patients or outcomes in comparison with the number of candidate predictors [19]. Overoptimism of our estimative may have influenced our results. However, the model contains only three variables (age, BMI and hypertension) and more than 100 events (glycemia above 126 mg/dl) and other authors [20, 21] have suggested that the number of events per variable (EPV) has been a criterion for the size of a dataset and this value must be greater than 10. In our study, EPV was almost 40. In addition, one has to contemplate that we have solely used transversal data for model construction (although we were able to test model's performance in a prospective dataset). This strategy has also been used by other successful studies [4, 5].

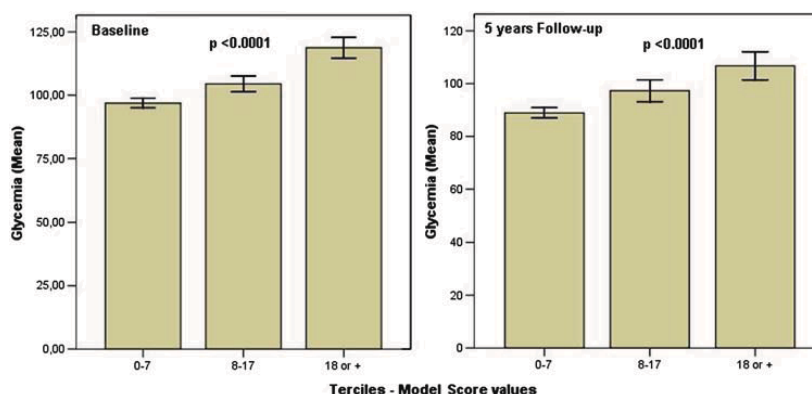
Unfortunately, only a small sample size of the baseline population was available for follow-up analysis. Despite this fact, an association between 5 years glycemia levels and model scores was observed.

The results of the validation study confirmed that the risk score performs well in the prediction of diabetes in a second, independent, population. However, as well as in other studies [22], the performance of models in derivation populations is better than in validation populations. In this scenario, some degree of poorer performance is expected since all models are more accurate in the population in which they were developed. Maybe the use of a different cutoff point for this population could improve results [4]. Interestingly, in the validation dataset more than half of the individuals reached the risk score cutoff value of 18 or more and, consequently, would need subsequent testing. Additionally, in this high risk group negative predictive values are almost 100%. Thus, even in high risk population, this score may be used to exclude patients who do not need further investigation.

In conclusion, this simple score that contains only clinical characteristics and anthropometric variables can be confidently used to select a subpopulation to be screened for T2DM. Because of its low cost and feasibility of execution, this score might optimize health care resources, mainly in developing countries such as Brazil.

**Table 4** Performance of different cutoffs (points above and below the optimal cutoff score) to the risk score in predictive model

Value	Sensitivity	Specificity	PPV	NPV	Accuracy	OR (95%CI)	P value	Needing to additional tests (%)
Derivation data								
<i>Cutoffs</i>								
13	0.7931	0.5743	0.166	0.963	0.60	5.172 (2.77–7.58)	<0.00001	46.1
17	0.7845	0.6251	0.183	0.964	0.64	6.07 (3.28–8.86)	<0.00001	41.5
18	0.7586	0.6685	0.197	0.963	0.68	6.34 (3.53–9.15)	<0.00001	37.2
19	0.5345	0.8024	0.225	0.941	0.78	4.66 (2.83–6.5)	<0.00001	23.0
Validation data								
<i>Cutoffs</i>								
13	0.9643	0.3934	0.095	0.994	0.43	17.50 (2.36–130.8)	0.0001	62.8
17	0.8929	0.436	0.095	0.984	0.46	6.443 (1.92–21.67)	0.0006	58.4
18	0.8571	0.4479	0.093	0.979	0.47	4.87 (1.66–14.27)	0.0015	57.1
19	0.6071	0.6754	0.11	0.963	0.67	3.215 (1.47–7.05)	0.0022	34.2

**Fig. 2** Mean of baseline and 5 years follow-up fasting glycemia of tertiles of model score values. Errors bars are 95%CI. P value was calculated by one way ANOVA

**Acknowledgment** This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo [grant number 07/54138-2]

## References

1. Incidence of diabetes in the population aged 18–79 Years. Cent Dis Control Prev. 2006. <http://www.cdc.gov/diabetes/statistics/incidence>. Accessed Aug 2006.
2. Malerbi DA, Franco LJ. Multicenter study of the prevalence of diabetes mellitus and impaired glucose tolerance in the urban Brazilian population aged 30–69 yr. The Brazilian Cooperative Group on the study of diabetes prevalence. *Diabetes Care*. 1992;15(11):1509–16.
3. Harris MI. Undiagnosed NIDDM: clinical and public health issues. *Diabetes Care*. 1993;16(4):642–52.
4. Ramachandran A, Snehalatha C, Vijay V, Wareham NJ, Colagiuri S. Derivation and validation of diabetes risk score for urban Asian Indians. *Diabetes Res Clin Pract*. 2005;70(1):63–70.
5. Glumer C, Carstensen B, Sandbaek A, Lauritzen T, Jorgensen T, Borch-Johnsen K. A Danish diabetes risk score for targeted screening: the Inter99 study. *Diabetes Care*. 2004;27(3):727–33.
6. Tunstall-Pedoe H, Kuulasmaa K, Amouyel P, Arveiler D, Rajakangas AM, Pajak A. Myocardial infarction and coronary deaths in the World Health Organization MONICA project. Registration procedures, event rates, and case-fatality rates in 38 populations from 21 countries in four continents. *Circulation*. 1994;90(1):583–612.
7. Bisi Molina Mdel C, Cunha Rde S, Herkenhoff LF, Mill JG. Hypertension and salt intake in an urban population. *Rev Saude Publica*. 2003;37(6):743–50.
8. Candido AP, Ferreira S, Lima AA, de Carvalho Nicolato RL, de Freitas SN, Brandao P, et al. Lipoprotein(a) as a risk factor associated with ischemic heart disease: Ouro Preto study. *Atherosclerosis*. 2007;191(2):454–9.
9. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2008;31(Suppl 1):S55–60.
10. Pesquisa por amostra de domicílios 2000: microdados. Rio de Janeiro: Fundação Instituto Brasileiro de Geografia e Estatística; 2001.
11. Alves-Silva J, da Silva Santos M, Guimaraes PE, Ferreira AC, Bandelt HJ, Pena SD, et al. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet*. 2000;67(2):444–61.
12. Ferreira RG, Moura MM, Engracia V, Pagotto RC, Alves FP, Camargo LM, et al. Ethnic admixture composition of two western Amazonian populations. *Hum Biol*. 2002;74(4):607–14.

13. Marqueline GF, Oliveira CM, Pereira AC, Krieger JE, Mill JG. Metabolic syndrome determinants in an urban population from Brazil: social class and gender-specific interaction. *Int J Cardiol.* 2007;129:259–65.
14. Griffin SJ, Little PS, Hales CN, Kinmonth AL, Wareham NJ. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. *Diabetes Metab Res Rev.* 2000;16(3):164–71.
15. Lindstrom J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care.* 2003;26(3):725–31.
16. Stern MP, Williams K, Haffner SM. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med.* 2002;136(8):575–81.
17. Tabaei BP, Herman WH. A multivariate logistic regression equation to screen for diabetes: development and validation. *Diabetes Care.* 2002;25(11):1999–2003.
18. Spijkerman AM, Yuyun MF, Griffin SJ, Dekker JM, Nijpels G, Wareham NJ. The performance of a risk score as a screening test for undiagnosed hyperglycemia in ethnic minority groups: data from the 1999 health survey for England. *Diabetes Care.* 2004;27(1):116–22.
19. Steyerberg EW, Eijkemans MJ, Harrell FE Jr, Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Mak.* 2001;21(1):45–56.
20. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA.* 1997;277(6):488–94.
21. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* 1996;49(12):1373–9.
22. Aekplakorn W, Bunnag P, Woodward M, Sritara P, Cheepudomwit S, Yamwong S, et al. A risk score for predicting incident diabetes in the Thai population. *Diabetes Care.* 2006;29(8):1872–7.

Research article

Open Access

**TCF7L2 variant genotypes and type 2 diabetes risk in Brazil: significant association, but not a significant tool for risk stratification in the general population**GF Marquezine<sup>1</sup>, AC Pereira\*<sup>1</sup>, AGP Sousa<sup>1,2</sup>, JG Mill<sup>3</sup>, WA Hueb<sup>1</sup> and JE Krieger<sup>1</sup>Address: <sup>1</sup>Heart Institute, University of São Paulo Medical School, Sao Paulo, Brazil, <sup>2</sup>Federal University of Rio Grande do Norte, Natal, Brazil and <sup>3</sup>Federal University of Espírito Santo, Vitória, Brazil

Email: GF Marquezine - g.marq2@gmail.com; AC Pereira\* - alexandre.pereira@incor.usp.br; AGP Sousa - agpsousa@ig.com.br; JG Mill - jgmill@npd.ufes.br; WA Hueb - mass@incor.usp.br; JE Krieger - krieger@incor.usp.br

\* Corresponding author

Published: 4 December 2008

Received: 26 June 2008

BMC Medical Genetics 2008, 9:106 doi:10.1186/1471-2350-9-106

Accepted: 4 December 2008

This article is available from: <http://www.biomedcentral.com/1471-2350/9/106>

© 2008 Marquezine et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**Abstract**

**Background:** Genetic polymorphisms of the *TCF7L2* gene are strongly associated with large increments in type 2 diabetes risk in different populations worldwide. In this study, we aimed to confirm the effect of the *TCF7L2* polymorphism rs7903146 on diabetes risk in a Brazilian population and to assess the use of this genetic marker in improving diabetes risk prediction in the general population.

**Methods:** We genotyped the single nucleotide polymorphisms (SNP) rs7903146 of the *TCF7L2* gene in 560 patients with known coronary disease enrolled in the MASS II (Medicine, Angioplasty, or Surgery Study) Trial and in 1,449 residents of Vitoria, in Southeast Brazil. The associations of this gene variant to diabetes risk and metabolic characteristics in these two different populations were analyzed. To assess the potential benefit of using this marker for diabetes risk prediction in the general population we analyzed the impact of this genetic variant on a validated diabetes risk prediction tool based on clinical characteristics developed for the Brazilian general population.

**Results:** SNP rs7903146 of the *TCF7L2* gene was significantly associated with type 2 diabetes in the MASS-II population (OR = 1.57 per T allele,  $p = 0.0032$ ), confirming, in the Brazilian population, previous reports of the literature. Addition of this polymorphism to an established clinical risk prediction score did not increase model accuracy (both area under ROC curve equal to 0.776).

**Conclusion:** *TCF7L2* rs7903146 T allele is associated with a 1.57 increased risk for type 2 diabetes in a Brazilian cohort of patients with known coronary heart disease. However, the inclusion of this polymorphism in a risk prediction tool developed for the general population resulted in no improvement of performance. This is the first study, to our knowledge, that has confirmed this recent association in a South American population and adds to the great consistency of this finding in studies around the world. Finally, confirming the biological association of a genetic marker does not guarantee improvement on already established screening tools based solely on demographic variables.

## Background

Type 2 diabetes mellitus (T2D) is a heterogeneous disease characterized by different degrees of insulin resistance and defects in insulin secretion, both of which are thought to result from the interplay of genetic and environmental factors. The identification of causative genes that predispose to T2D could provide clues to better understand the primary pathogenesis and therefore result in better prevention, diagnosis and treatment of this increasingly prevalent and costly condition. Up until recently, results of research on the genetic field have been mostly elusive.

Genome-wide linkage scans have discovered chromosomal regions containing type 2 diabetes (T2D) susceptibility genes in chromosome 10q, which were later ascribed to intronic variations in the transcription factor 7-like 2 (*TCF7L2*) gene, which were strongly associated with a twofold increase in risk for T2D in an Icelandic population [1]. Two additional cohorts showed similar associations, giving a combined odds ratio (OR) of 1.6 (95% confidence interval (CI) 1.4 to 1.7,  $p = 4.7 \times 10^{-18}$  per allele). This finding has readily been replicated in cohorts of European [1-13], Asian [6,14-16] and African [6,17] descent, showing very similar results. Most evidence for *TCF7L2* has come from case-control studies or intervention trials. There are much fewer studies that analyzed the effect of this polymorphism in the general population [10,18], but they suggest that diabetic carriers of the at-risk allele have more severe  $\beta$ -cell dysfunction and more microvascular complications but less metabolic syndrome features and a more protective lipid profile than non-carriers.

In addition, it is necessary to confirm this association in different populations worldwide. At moment, there are no studies in South American populations. Therefore, this study had two objectives: 1) investigate whether *TCF7L2* variants are associated with diabetes in the Brazilian population; 2) evaluate how significant is the impact of this association in predicting the prevalent diabetes risk in the general population from Brazil.

## Methods

To respond the first objective, we used a cohort of patients from the MASS- II study and to address the second one, we used a sample representative from the general urban population of the city of Vitoria, Brazil, both of which are described below.

### **Multi-vessel coronary artery disease patients (MASS II Study)**

Six-hundred and eleven patients that have documented multi-vessel coronary artery disease and normal left ventricular function were included in the MASS II trial [19]. Here we have solely used information from the baseline

characteristics of the studied population. Diabetes was not necessary for enrollment in this protocol. From these individuals we were able to obtain genotype data for *TCF7L2* on 560 patients, of whom 190 had diabetes mellitus, based on American Diabetes Association criteria or previous treatment for diabetes. All patients gave written informed consent for participating in the study. The Ethics Committee of the Heart Institute approved the trial, and all procedures were performed in accordance with the Helsinki Declaration.

### **General population of Vitoria/ES, Brazil**

A cross-sectional study of risk factors for cardiovascular diseases was performed in the urban population of Vitoria, Brazil, using the WHO-MONICA project guidelines. The study design was based on cross-sectional research methodology and was developed by means of surveying and analyzing socioeconomic and health data in a probabilistic sample of residents aged 25 to 64 years from the municipality of Vitoria, ES, Brazil. The population was randomized and the sample was socioeconomically, geographically and demographically representative of the residents of this municipality. A selection of 2,268 residential homes located in Vitoria was made and these were visited. The project received approval from the Ethics Committee of the Biomedical Center of Universidade Federal do Espirito Santo (UFES). The selected individuals were asked to attend the Cardiovascular Investigation Clinic of the University Hospital for tests to be performed on the following day. Of the total sample, 1,577 individuals attended. Participants were submitted to physical examination. Major cardiovascular risk factors such as smoking habits, alcohol intake, sedentarism, diabetes and hypertension were inquired. Blood glucose, total cholesterol, lipoprotein fractions, and triglycerides were assayed by standard techniques in a 12 hour fasting blood sample.

Body mass index (BMI) (weight in Kg/height in meters<sup>2</sup>) was calculated, overweight defined as a BMI  $\geq 25$  Kg/m<sup>2</sup>, and obesity defined as a BMI  $\geq 30$  Kg/m<sup>2</sup>.

In order to screen individuals at higher risk for type 2 diabetes from the general population to be further investigated, we previously built a prediction model of diabetes risk in this same population (Sousa, AGP, personal communication). This simple risk score model was developed based in a population of 1,224 individuals from the general population without known diabetes aged 35 years or more. Also in this analysis, diabetes diagnosis was based on fasting glycemia equal or above 126 mg/dl (7 mmol/l). This model was externally validated in an independent population from a different city ascertained through a similar epidemiological protocol and items independently and significantly associated with diabetes, and chosen for inclusion in the model, were age, BMI and known

hypertension. Selected variables in this model were age, BMI and hypertension.

The association between rs7903146 genotype and diabetes was measured using a genotype trend genetic model for an additive allelic effect captured by a regression model for diabetes in both population.

For model construction and performance measurement, univariate analyses were performed by logistic regression with T2DM (FPG > 126 mg/dl) as the dependent variable. Logistic regression coefficients were used to estimate odds ratios for each of the independent variables in the model. The continuous variables were categorized into strata. To create the final model, a score was calculated by multiplying the regression coefficient by 10 and rounding to the nearest integer. Finally, a sum score was calculated for each participant by adding the score for each significant variable in the risk model. Sensitivity, specificity, positive and negative predictive values and accuracy were compared for different score cutoffs. In addition, the proportion of subjects that needed subsequent testing (those individuals who have score values above the cutoff value) was compared between these several groups. In order to evaluate the model performance, a receiver operating characteristic (ROC) curve was built to plot probabilities saved with the logistic regression procedure and the area under ROC curve (AUC) was used to measure the power to discriminate high-risk from low-risk individuals.

#### Genotyping of the rs7903146 polymorphism

Genomic DNA was extracted from peripheral blood leukocytes by means of a salting-out procedure. PCR primers used were: 5' ACA ATT AGA GAG CTA AGC ACT TTT TAG GTA 3' and 5' GTG AAG TGC CCA AGC TTC TC 3'. Briefly, the studied polymorphism was detected by polymerase chain reaction-restriction fragment length polymorphism assay (PCR-RFLP). A 30-cycle PCR was carried out in a PTC-DNA Engine Tetrad<sub>2</sub> using a 10 µL reactive solution containing 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 2.5 mM MgCl<sub>2</sub>, 100 µM of each dNTP, 0.3U *Easy Taq* DNA Polymerase, 5 pmol of each primer and 1 µL of genomic DNA template. PCR products were digested with 1U of *RsaI* restriction enzyme and visualized by 3% agarose gel electrophoresis. The quality control for these assays was assessed by randomly selecting 40 samples that were re-genotyped by an independent technician. Observed concordance between genotyping assays was 100%.

#### Statistical Analysis

We used SPSS v.13 program for statistical analysis. The goodness of fit for normal distribution was evaluated using the Kolmogorov-Smirnov test. To test for differences in various characteristics, non-parametric tests were used for continuous variables and the  $\chi^2$  test and Fischer exact

test were used for categorical variables. Hardy-Weinberg equilibrium was studied through the use of Haploview software. For power calculation, we used the PS: Power and Sample Size Calculation from the Department of Biostatistics of Vanderbilt University, available for download at <http://biostat.mc.vanderbilt.edu/twiki/pub/Main/PowerSampleSize/pssetup.exe>. All statistical significance tests calculates are performed as a two-sided test and a value less than 0.05 was considered significant.

#### Results

In the MASS-II population, a total of 560 individuals were genotyped for TCF7L2 gene variant rs7903146 and the genotypic distribution for this variant was in Hardy-Weinberg equilibrium. There was male sex preponderance (69.1%) and a high T2D prevalence of 31.0%. Mean BMI was 27.1 kg/m<sup>2</sup>. According to the used weight classification, 66.2% of this population was overweight or obese (Table 1). The relative frequency of genotypes CC, CT and TT in MASS II population was respectively 28,4%, 60,0% and 11,6% (number of individuals equal to 159, 336 and 65, respectively). Allele and genotype frequencies were not in Hardy-Weinberg equilibrium ( $p = 0,0004$ ) (probably because of the high a priori chance of diabetic individuals in this sample with multi-vessel coronary artery disease). The rate of diabetes increased with an increasing dose of allele T of rs7903146 (OR = 1.57, 95%CI 1.16–2.11,  $p = 0.0032$ ). These data are summarized in Table 3. After adjusting for other covariates potentially associated with T2D risk, presence of allele T was still significantly associated with a 1.61 (95%CI 1.18–2.19) increased risk of presenting T2D ( $p = 0.0025$ ) (Table 4).

In regards to the general population of Vitoria, 1,440 samples of DNA were available and genotyped for TCF7L2 rs7903146 variant. The characteristics of this population are shown in Table 2. Their mean age was 44.8 years (range 23–65). The prevalence of T2D was only 7.9%. Obesity was present in 19.3% of patients, but 55.3% were on the overweight category (BMI  $\geq 25$  kg/m<sup>2</sup>). Mean BMI was  $26.2 \pm 4.9$  kg/m<sup>2</sup>. Non-diabetics individuals had a lower mean BMI (25.9 versus 30.6 kg/m<sup>2</sup> in diabetics). The prevalence of genotype CC was 43,4% ( $n = 614$ ), genotype CT equal to 46,6% ( $n = 660$ ) and overall prevalence of TCF7L2 genotype TT was 10,0% ( $n = 142$ ). Allele and genotype frequencies were in Hardy-Weinberg equilibrium ( $p = 0,07$ ) in this sample from the general population. A slightly higher genotype prevalence was observed in the diabetic sub-population of individuals (12.1%) as compared to non-diabetic individuals (9.9%), but this difference was not statistically significant ( $p = 0.46$ ). Of particular interest, the total number of diabetics that carried the TT genotype, despite the relatively large size of the initial sample, came down to only 14 people. Therefore, the statistical power of detecting this association in the

**Table 1: Characteristics of the MASS-II population**

	Total	rs7903146		p value
		CC/CT	TT	
<b>Number of patients (%)</b>	611 (100)	495 (88.4)	65 (11.6)	
Male	423 (69.1)	339 (88.5)	155 (88.1)	.88
Age (years)	59.8 ± 9.1	59.7 ± 9.12	58.8 ± 9.6	.45
<b>Mean BMI (kg/m<sup>2</sup>)</b>	27.1 ± 4.2	27.1 ± 4.2	27.1 ± 4.2	.97
< 25	201 (33.0)	160 (32.5)	19 (29.7)	.64
25 – 29.9	276 (45.4)	219 (44.4)	34 (53.1)	.18
≥ 30	129 (21.2)	112 (22.8)	10 (15.9)	.21
<b>Mean FPG (mg/dL)</b>	129.5 ± 58.2	127.21 ± 53.7	139.2 ± 75.6	.22
< 110	311 (52.2)	256 (52.7)	29 (45.3)	.26
110 – 125	107 (18.0)	90 (18.6)	12 (18.7)	.96
≥ 126	178 (30.0)	140 (28.8)	23 (36.0)	.24
<b>Total Cholesterol (mg/dL)</b>	223.2 ± 47.7	222.9 ± 47.6	221.7 ± 52.5	.85
<b>Mean HDL (mg/dL)</b>	37.4 ± 10.4	37.5 ± 10.6	36.5 ± 9.9	.49
<b>Mean Triglycerides (mg/dL)</b>	195.1 ± 121.0	193.7 ± 118.0	179.9 ± 97.0	.37
	MS Components:			
<b>Obesity (BMI)</b>	129 (21.2)	112 (22.7)	10 (15.6)	.19
<b>High Triglycerides</b>	344 (57.1)	279 (57.1)	34 (53.1)	.55
<b>Low HDL-c</b>	447 (78.1)	371 (79.3)	46 (75.4)	.48
<b>Hypertension</b>	364 (59.6)	292 (59.1)	36 (55.4)	.64
<b>FPG ≥ 110 MG/dL</b>	278 (46.6)	223 (45.9)	35 (54.7)	.18

**Table 2: Characteristics of the Vitoria general population**

	General	rs7903146		p
		CC + CT	TT	
Number of patients (%)	1577 (100)	1303 (89.9)	146 (10.0)	
Male	718 (45.6)	600 (46.0)	64 (43.8)	.61
Age (years)	44.8 ± 10.9	44.7 ± 10.9	44.3 ± 10.7	.65
<b>Mean BMI (kg/m<sup>2</sup>)</b>	26.3 ± 4.9	26.3 ± 5.0	26.2 ± 4.8	.82
< 25	692 (44.2)	573 (44.3)	65 (44.8)	.91
25 – 29.9	513 (35.7)	458 (35.5)	55 (38.0)	.54
≥ 30	285 (19.4)	260 (20.1)	25 (17.2)	.41
<b>Mean FPG (mg/dL)</b>	105.01 ± 32	105.0 ± 32.3	105.7 ± 29.3	.82
< 110	1130 (78.7)	1014 (78.3)	116 (80.0)	.64
110 – 125	241 (15.4)	200 (15.4)	22 (15.2)	.93
≥ 126	123 (7.8)	102 (7.9)	14 (9.7)	.45
<b>Total Cholesterol (mg/dL)</b>	214.4 ± 47.8	214.0 ± 48.1	215.4 ± 44.0	.74
<b>Mean HDL (mg/dL)</b>	45.4 ± 12.3	45.2 ± 12.1	44.0 ± 10.4	.27
<b>Mean Triglycerides (mg/dL)</b>	137.6 ± 127.9	137.2 ± 131.0	133.0 ± 82.1	.70
SBP (mmHg)	128 ± 2	128 ± 2	129 ± 2	.52
DBP (mmHg)	84 ± 1	84 ± 14	84 ± 14	.56
	MS Components			
Visceral obesity *	255 (16.2)	216 (16.6)	21 (14.4)	.50
High triglycerides	484 (30.7)	391 (30.1)	51 (35.0)	.23
Low HDL-c	848 (53.8)	705 (54.1)	79 (54.1)	.99
Hypertension	727 (46.1)	596 (45.7)	75 (51.4)	.19
FPG ≥ 110 MG/dL	334 (21.4)	281 (21.7)	29 (20.0)	.63
Metabolic syndrome	397 (25.4)	326 (25.2)	40 (27.6)	.53

\*Abdominal circumference above cutoff points according to ATPIII criteria



**Table 3: Genotype association with type 2 diabetes**

rs7903146		DM n (%)	Non-DM n (%)	OR (95%CI)	Allele test (P value‡)
<b>MS2</b>	CC	38 (22.0)	120 (31.1)	1.57 (1.16 – 2.11)	0.0032 (0.0034*)
	CT	106 (61.3)	230 (59.6)		
	TT	29 (16.8)	36 (9.3)		
<b>VIT</b>	CC	45 (40.2)	564 (43.6)	1.126 (0.84–1.51)	0.426
	CT	54 (48.2)	603 (46.6)		
	TT	13 (11.6)	128 (9.9)		

(\*) Adjusted for age and sex.

‡) additive genetic model

studied sample from the general population was only 24%.

Our previous risk model (Sousa AGP, unpublished results) based solely on clinical criteria (age, presence of hypertension and BMI) for predicting diabetes was able to reach a good performance to identify individuals with an increased likelihood of having diabetes. According to this tool, individuals classified as having a higher probability of becoming diabetic (higher score values) should be assessed with further laboratorial tests for diabetes. Interestingly, inclusion of the TT genotype in our model resulted in no improvement on the area under the ROC curve (AUC = 0.776; 95%CI 0.73–0.82 in both models). The performance characteristics of two models (with only demography variables and with *TCF7L2* variables associated to demography variables) are shown in Table 5. There was a slight improvement in specificity when the *TCF7L2* genotype TT was included, however the sensitivity diminished. There was almost no difference among the ROC curves of two models (Figure 1).

Although several studies have observed that the rs7903146 SNP is associated with some of the metabolic syndrome features, we did not observe any significant association between the T risk allele and these characteristics (Table 1 and 2).

## Discussion

Recently Grant et al. [1] have described an association of common genetic variants in *TCF7L2* gene and type 2 dia-

betes risk, a finding which has been promptly replicated in many populations worldwide and made this polymorphism one of the most important discoveries in the pathophysiology of T2D in decades [20-22]. Initially, we were able to confirm the association between polymorphism rs7903146 and T2D in a cohort of patients from the MASS II protocol, whose diabetes prevalence was 31.1%. To our knowledge, this is the first study that confirmed this association in a South American population. The odds ratio of association was similar to other previous studies in different populations [1,2,6,15].

In the second stage, we have genotyped a large sample of individuals from the general population of Brazil, which has gone through intense admixing in its recent history and is characterized by a genetic background that is remarkably heterogeneous when compared to cohorts of European or Asian ancestry. We genotyped a total of 1449 people looking for associations with diabetes and metabolic syndrome components, which also have been associated with this polymorphism in some, but not all, studied populations [8,10]. Even though genotype frequencies for the rs7903146 variants in our population were somewhat similar to those reported previously in different studies, we failed to confirm the association that has been showing great consistency in case-control [23] and general [11,24,25] populations all over the world. We believe this finding is mostly due to the lack of statistical power, because of the low diabetes prevalence and total number of diabetic patients in the tested population than

**Table 4: Logistic regression for type 2 diabetes risk**

	MASS II		Vitoria	
	OR (95%CI)	P value	OR (95%CI)	P value
T Allele	1.57 (1.18–2.19)	0.0025	1.15 (0.84–1.58)	0.391
Age	1.02 (0.99–1.04)	0.097	1.08 (1.06–1.10)	< 0.0001
Female Sex	0.27 (0.02–3.11)	0.29	1.04 (0.69–1.59)	0.837
Obesity*	1.67 (1.08–2.56)	0.021	4.91 (3.25–7.43)	< 0.0001

(\*) BMI equal or greater to 30 kg/m<sup>2</sup>



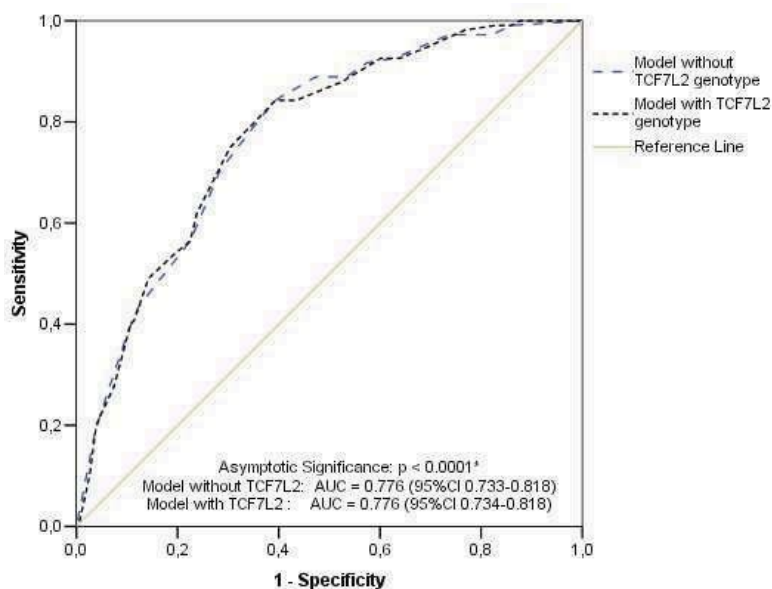
**Table 5: Performance of two predictive models in assessment T2D risk**

Cutoffs Values	Sensitivity	Specificity	Model 1 – Only Demography variables					P value	Needing to Additional Tests(%)
			PPV	NPV	Accuracy	OR (95%CI)			
16.5	0.8426	0.6071	0.19	0.972	0.63	8.27 (3.86–12.68)	< 0.00001	43.7	
18	0.7037	0.7162	0.213	0.957	0.71	5.99 (3.38–8.60)	< 0.00001	32.51	
Model 2 – Demography and TCF7L2 TT Genotype variables									
Cutoffs Values	Sensitivity	Specificity	PPV	NPV	Accuracy	OR (95%CI)	P value	Needing to Additional Tests(%)	
16.5	0.8426	0.5939	0.185	0.972	0.62	7.83 (3.66–12.0)	< 0.00001	44.89	
18	0.7222	0.7081	0.213	0.959	0.71	6.30 (3.51–9.1)	< 0.00001	33.42	

to subtleties of our population such as ethnic variety and genetic heterogeneity.

Most importantly, however, the addition of a known *TCF7L2* genotype in a validated model for predicting prevalent diabetes was not able to improve the performance of the risk score: area under ROC curve in both models equal to 0.776, ie the inclusion of *TCF7L2* genotype is not better than the utilization of only clinical characteristics to predict diabetes risk in our population. A very similar result has been recently published by Balkau *et al* [26] in an incident diabetes study in the French population. The clinical and laboratorial prediction model improved

only modestly after inclusion of *TCF7L2* and interleucin 6 (IL-6) SNPs information. Another example of a practical consequence of applying the predictive model is the value of Needing to Additional Tests (Table 5), which means the proportion from population which would be tested for T2D according to the model. It was observed only a little difference in percentage of individuals that will be further tested according to both models. In addition, as has been demonstrated, the impact of most polymorphisms on determining T2D risk is only modest. In order to improve the accuracy of clinical predictors to assess diabetes risk in a general population, it might be necessary to add several genetic variants in the model [27,28]. Recent work from



**Figure 1**  
**ROC curves for the risk score in the two predictive models. (\*) p value for both models**

van Hoek *et al* [28] combined information from 18 polymorphisms which have been related to T2D and tested them in a general population of elderly persons. The prediction power is modestly enhanced when a model with genetic information is compared to a clinical characteristics model.

Finally, we also looked for a link between the rs7903146 variant of the *TCF7L2* gene and features of the metabolic syndrome such as weight, body mass index, hypertriglyceridemia, hypertension, low HDL-c, as reported earlier [8,18], but could not observe this even when data were adjusted for age, race and abdominal circumference.

### Conclusion

This finding adds to the importance of this polymorphism in the pathogenesis of the disease and to the great consistency in which data are replicated around the world. The use of this particular genetic marker in stratifying individuals regarding the risk of T2D appears to be dependent on the *a priori* risk of the disease in the tested population. One should exercise caution in the indiscriminate use of genetic association results once the confirmation of the biological association of a genetic marker with a common chronic disease does not guarantee meaningful improvement on already established screening tools based solely on demographic variables.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

WH, JGM, and ACP performed the field work. GFM, AGPS, and ACP performed the statistical analyses. GFM, AGPS, and ACP drafted the manuscript under the supervision of WH, JGM and JEK. ACP supervised the study. All authors read and approved the final manuscript.

### References

- Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadóttir A, *et al.*: **Variant of transcription factor 7-like 2 (*TCF7L2*) gene confers risk of type 2 diabetes.** *Nat Genet* 2006, **38**:320-323.
- Cauchi S, Meyre D, Dina C, Choquet H, Sarnon C, Gallina S, Balkau B, Charpentier G, Pattou F, Stetsyuk V, *et al.*: **Transcription factor *TCF7L2* genetic study in the French population: expression in human beta-cells and adipose tissue and strong association with type 2 diabetes.** *Diabetes* 2006, **55**:2903-2908.
- Damcott CM, Pollin TI, Reinhart LJ, Ott SH, Shen H, Silver KD, Mitchell BD, Shuldiner AR: **Polymorphisms in the transcription factor 7-like 2 (*TCF7L2*) gene are associated with type 2 diabetes in the Amish: replication and evidence for a role in both insulin secretion and insulin resistance.** *Diabetes* 2006, **55**:2654-2659.
- Florez JC, Jablonski KA, Bayley N, Pollin TI, de Bakker PI, Shuldiner AR, Knowler WC, Nathan DM, Altshuler D: ***TCF7L2* polymorphisms and progression to diabetes in the Diabetes Prevention Program.** *N Engl J Med* 2006, **355**:241-250.
- Groves CJ, Zeggini E, Minton J, Frayling TM, Weedon MN, Rayner NW, Hitman GA, Walker M, Wiltshire S, Hattersley AT, McCarthy ML: **Association analysis of 6,736 U.K. subjects provides replication and confirms *TCF7L2* as a type 2 diabetes susceptibility gene with a substantial effect on individual risk.** *Diabetes* 2006, **55**:2640-2644.
- Humphries SE, Gable D, Cooper JA, Ireland H, Stephens JW, Hurel SJ, Li KW, Palmen J, Miller MA, Cappuccio FP, *et al.*: **Common variants in the *TCF7L2* gene and predisposition to type 2 diabetes in UK European Whites, Indian Asians and Afro-Caribbean men and women.** *J Mol Med* 2006, **84**:1-10.
- Lehman DM, Hunt KJ, Leach RJ, Hamlington J, Arya R, Abboud HE, Duggirala R, Blangero J, Goring HH, Stern MP: **Haplotypes of Transcription Factor 7-Like 2 (*TCF7L2*) Gene and Its Upstream Region Are Associated With Type 2 Diabetes and Age of Onset in Mexican Americans.** *Diabetes* 2007, **56**:389-393.
- Marzi C, Huth C, Kolz M, Grallert H, Meisinger C, Wichmann HE, Rathmann W, Herder C, Illig T: **Variants of the transcription factor 7-like 2 gene (*TCF7L2*) are strongly associated with type 2 diabetes but not with the metabolic syndrome in the MONICA/KORA surveys.** *Horm Metab Res* 2007, **39**:46-52.
- Mayans S, Lackovic K, Lindgren P, Ruikka K, Agren A, Eliasson M, Holmberg D: ***TCF7L2* polymorphisms are associated with type 2 diabetes in northern Sweden.** *Eur J Hum Genet* 2007, **15**:342-346.
- Melzer D, Murray A, Hurst AJ, Weedon MN, Bandinelli S, Corsi AM, Ferrucci L, Paolisso G, Guralnik JM, Frayling TM: **Effects of the diabetes linked *TCF7L2* polymorphism in a representative older population.** *BMC Med* 2006, **4**:34.
- Saxena R, Gianniny L, Burtt NP, Lyssenko V, Giuducci C, Sjogren M, Florez JC, Almgren P, Isomaa B, Orho-Melander M, *et al.*: **Common single nucleotide polymorphisms in *TCF7L2* are reproducibly associated with type 2 diabetes and reduce the insulin response to glucose in nondiabetic individuals.** *Diabetes* 2006, **55**:2890-2895.
- van Vliet-Ostapchouk JV, Shiri-Sverdlov R, Zernakova A, Strengman E, van Haeften TW, Hofker MH, Wijmenga C: **Association of variants of transcription factor 7-like 2 (*TCF7L2*) with susceptibility to type 2 diabetes in the Dutch Breda cohort.** *Diabetologia* 2007, **50**:59-62.
- Zhang C, Qi L, Hunter DJ, Meigs JB, Manson JE, van Dam RM, Hu FB: **Variant of transcription factor 7-like 2 (*TCF7L2*) gene and the risk of type 2 diabetes in large cohorts of U.S. women and men.** *Diabetes* 2006, **55**:2645-2648.
- Chandak GR, Janipalli CS, Bhaskar S, Kulkarni SR, Mohankrishna P, Hattersley AT, Frayling TM, Yajnik CS: **Common variants in the *TCF7L2* gene are strongly associated with type 2 diabetes mellitus in the Indian population.** *Diabetologia* 2007, **50**:63-67.
- Hayashi T, Iwamoto Y, Kaku K, Hirose H, Maeda S: **Replication study for the association of *TCF7L2* with susceptibility to type 2 diabetes in a Japanese population.** *Diabetologia* 2007, **50**(5):980-984.
- Horikoshi M, Hara K, Ito C, Nagai R, Froguel P, Kadowaki T: **A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population.** *Diabetologia* 2007, **50**:747-751.
- Helgason A, Palsson S, Thorleifsson G, Grant SF, Emilsson V, Gunnarsdottir S, Adeyemo A, Chen Y, Chen G, Reynisdottir I, *et al.*: **Refining the impact of *TCF7L2* gene variants on type 2 diabetes and adaptive evolution.** *Nat Genet* 2007, **39**:218-225.
- Cauchi S, Meyre D, Choquet H, Dina C, Born C, Marre M, Balkau B, Froguel P: ***TCF7L2* variation predicts hyperglycemia incidence in a French general population: the data from an epidemiological study on the Insulin Resistance Syndrome (DESIR) study.** *Diabetes* 2006, **55**:3189-3192.
- Hueb VV, Paulo R, Soares P, Gersh B, César L, Protásio L, Luz P, Puig L, Martinez E, Oliveira S, Ramires J: **The medicine, angioplasty, or surgery study (MASS-II): a randomized, controlled clinical trial of three therapeutic strategies for multivessel coronary artery disease: One-year results.** *Journal of the American College of Cardiology* 2004, **43**:1743-1751.
- Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, Almgren P, Sjogren M, Ling C, Eriksson KF, Lethagen UL, *et al.*: **Mechanisms by which common variants in the *TCF7L2* gene increase risk of type 2 diabetes.** *J Clin Invest* 2007, **117**:2155-2163.
- Zeggini E, McCarthy ML: ***TCF7L2*: the biggest story in diabetes genetics since HLA?** *Diabetologia* 2007, **50**:1-4.
- Schafer SA, Tschritter O, Machicao F, Thamer C, Stefan N, Galloway B, Holst JJ, Dekker JM, T'Hart LM, Nijpels G, *et al.*: **Impaired glucagon-like peptide-1-induced insulin secretion in carriers of**

BMC Medical Genetics 2008, 9:106

<http://www.biomedcentral.com/1471-2350/9/106>

- transcription factor 7-like 2 (TCF7L2) gene polymorphisms.** *Diabetologia* 2007, **50**(12):2443-2450.
23. Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, Weitgasser R, Nejjari C, Patsch W, Chikri M, Meyre D, Froguel P: **TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis.** *J Mol Med* 2007.
  24. De Silva NM, Steele A, Shields B, Knight B, Parnell K, Weedon MN, Hattersley AT, Frayling TM: **The transcription factor 7-like 2 (TCF7L2) gene is associated with Type 2 diabetes in UK community-based cases, but the risk allele frequency is reduced compared with UK cases selected for genetic studies.** *Diabet Med* 2007, **24**:1067-1072.
  25. Vaxillaire M, Veslot J, Dina C, Proenca C, Cauchi S, Charpentier G, Tichet J, Fumeron F, Marre M, Meyre D, et al.: **Impact of common type 2 diabetes risk polymorphisms in the DESIR prospective study.** *Diabetes* 2008, **57**:244-254.
  26. Balkau B, Lange C, Fezeu L, Jean T, de Lauzon-Guillain B, Czernichow S, Fumeron F, Froguel P, Vaxillaire M, Cauchi S, et al.: **Predicting diabetes – clinical, biological and genetic approaches: the D.E.S.I.R. Study.** *Diabetes Care* 2008.
  27. Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, Zeggini E, Rayner NW, Shields B, Owen KR, Hattersley AT, Frayling TM: **Combining information from common type 2 diabetes risk polymorphisms improves disease prediction.** *PLoS Med* 2006, **3**:e374.
  28. van Hoek M, Dehgan A, Witteman JC, van Duijn CM, Uitterlinden AG, Oostra BA, Hofman A, Sijbrands EJ, Janssens AC: **Predicting type 2 diabetes based on polymorphisms from genome wide association studies: a population-based study.** *Diabetes* 2008.

#### Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2350/9/106/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

