

**GUILHERME FIGUEIREDO MARQUEZINE**

**Papel do polimorfismo rs7903146 do gene  
TCF7L2 na população brasileira e sua aplicação  
na predição de risco de diabetes tipo 2**

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

Área de concentração: Cardiologia

Orientador: Prof. Dr. José Eduardo Krieger

**São Paulo**

**2009**

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

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

©reprodução autorizada pelo autor

Marquezzine, Guilherme Figueiredo

Papel do polimorfismo rs7903146 do gene TCF7L2 na população brasileira e sua aplicação na predição de risco de diabetes tipo 2 / Guilherme Figueiredo Marquezzine. -- São Paulo, 2009.

Tese(doutorado)--Faculdade de Medicina da Universidade de São Paulo.

Departamento de Cardio-Pneumologia.

Área de concentração: Cardiologia.

Orientador: José Eduardo Krieger.

Descritores: 1.Diabetes mellitus tipo 2/genética 2.Polimorfismo genético  
3.Fatores de transcrição TCF 4.Predisposição genética para a doença  
5.Grupos de populações continentais

USP/FM/SBD-100/09



## DEDICATÓRIA

À minha esposa, Ana Karina, por me apoiar na realização de um sonho, pela compreensão na minha ausência, pela tolerância pelas horas roubadas de convívio, pelo amor e companheirismo.

Aos meus pais, meus heróis, maiores exemplos de vida e conduta familiar, fonte de sabedoria, força e bom senso intermináveis. Seu incentivo tornou possível a realização de todos os sonhos que me trouxeram até aqui hoje.

Ao meu irmão, Henrique, que, com sua irreverência, ensinou a somente dar a devida importância aos percalços da vida, tornando-os invariavelmente menores e mais simples.

À família, pelo incondicional apoio.

Aos amigos, *and you know who you are*, por participarem e se interessarem, por torcerem e apoiarem, por inspirarem. Diariamente ou eventualmente, pessoalmente ou virtualmente, vocês tornaram o caminho menos sinuoso.

## **AGRADECIMENTOS**

Ao final do processo de preparação de uma tese, após o trabalho planejado e executado, dados todos analisados e revisados, conclusões descritas e discutidas, fica evidente o quanto o trabalho não é realmente “nosso”. Há muito de diversas outras pessoas nessas páginas: dedicação, inspiração, torcida, suor. O desafio de tentar nomear a todos os envolvidos é grande e cruel, pois o risco de deixar de fora pessoas importantes existe. A estes, peço desculpas antecipadamente.

Agradeço, portanto, àqueles que ajudaram a tornar este trabalho possível, das mais variadas formas.

Ao Professor José Eduardo Krieger, por abrir as portas do Laboratório para mim, por seu entusiasmo, atenção e por ensinar, nas pequenas atitudes, noções que vão muito além da ciência.

Ao amigo e orientador informal Alexandre da Costa Pereira, reconhecido por sua competência e enorme capacidade de trabalho, por sua impressionante paciência ao guiar-me, desde o início, passo a passo.

Ao Professor Bernardo Léo Wajchenberg, pelas oportunidades, pelo exemplo de disposição e capacidade de trabalho. O convívio com o senhor pelos corredores de hospital e congressos mudou minha trajetória profissional e ajudou a me trazer até aqui.

À Camila Maciel Oliveira, amiga desde os tempos “do outro lado da rua”, na Endocrinologia do HC, que me recomendou e apresentou ao Laboratório, por todas as vezes que saiu do seu caminho para colaborar como fosse possível.

Ao amigo André Gustavo Pires de Sousa, que também “atravessou a rua”, vindo da Endocrinologia para tantas vezes dividir seu tempo comigo e esta tese.

Às amigas Glória Mota, Marcilene Floriano e Noely Ferreira, cuja dedicação e expertise técnica permitiu a realização dessa tese.

À Maria de Lourdes Junqueira, Isabel Cristina Lameirinhas e, especialmente, Silvana Campos, pela ajuda e compreensão.

Aos amigos do Grupo de Genética do LGCM: Noely, Débora, Isolmar, Alessandra, Amanda, Diogo, Marisa, Samantha, Fábio, Luciana, Marcelo, Rodrigo, por seu apoio e alegria, que tornam o dia-a-dia mais leve.

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

Referências: adaptado de *International Committee of Medical Journals Editors* (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

Lista de abreviaturas, símbolos e siglas

Lista de tabelas

Lista de figuras

Resumo

Summary

1 INTRODUÇÃO.....	1
1.1 Prevalência de diabetes no Brasil e no mundo .....	1
1.2 Custos econômicos do diabetes.....	2
1.3 Fisiopatologia do diabetes .....	3
2 OBJETIVOS.....	10
3 MATERIAIS E MÉTODOS.....	11
3.1 População de Vitória – ES.....	11
3.2 Seguimento da amostra de Vitória – ES .....	12
3.3 População de Ouro Preto – MG.....	13
3.4 Desenvolvimento do Modelo de Risco e Validação.....	14
3.5 População de Itapetininga – SP .....	16
3.6 População do estudo MASS II .....	16
3.7 Limitações Metodológicas das amostras.....	17
3.8 Análise estatística.....	17



3.9 Genotipagem do polimorfismo rs7903146 do gene TCF7L2	18
4 RESULTADOS.....	19
4.1 Vitória e MASS.....	19
4.2 Itapetininga.....	20
4.2 Modelos de Predição de Risco.....	21
5 DISCUSSÃO.....	26
6 CONCLUSÕES.....	34
7 ANEXOS.....	35
8 REFERÊNCIAS.....	51

## LISTA DE ABREVIATURAS, SÍMBOLOS E SIGLAS

ADA – American Diabetes Association

*ADAMTS9 – ADAM metallopeptidase with thrombospondin type 1 motif, 9*

ANOVA – *Analysis of variation*

ATP – Adenosina tri-fostato

AUC – *Area under curve*

*CAMK1D – Calcium/calmodulin-dependent protein kinase 1D*

CDC – *Center for Disease Control*

*CDC123 – Cell division cycle 123 homolog*

CDKAL1 – *CDK5 regulatory subunit associated protein 1-like 1*

*CDKN2A/CDKN2B – Cyclin-dependent kinase inhibitor 2A/2B*

DESIR – *Data from the Epidemiological Study on the Insulin Resistance Syndrome*

DM2 – Diabetes tipo 2

DNA – Ácido desoxiribonucleico

EUA – Estados Unidos da América

*FTO – Fat mass and obesity associated (fatso)*

GJ – Glicemia de jejum

GLP-1 – *Glucagon-like peptide 1*

GWAS – *Genome-wide association scan*

HDL-colesterol – Lipoproteínas de alta densidade

HHEX – *Hematopoietically expressed homeobox*

HNF1A – *Hepatocyte nuclear factor 1 homeobox A*

HNF1B - *Hepatocyte nuclear factor 1 homeobox B*

HOMA%S – *Homeostatic model assessment method for insulin sensitivity*

HOMA-B% – *Homeostatic model assessment method for insulin secretion*

IBGE – Instituto Brasileiro de Geografia e Estatística

IC95% – Intervalo de confiança de 95%

IDF – *International Diabetes Federation*

IGF2BP2 – *Insulin-like growth factor binding protein 2*

IMC – Índice de massa corpórea

JAZF1 – *JAZF zinc finger 1*

KCNJ11 – Potassium inwardly-rectifying channel, subfamily J, member 11

kg/m<sup>2</sup> – kilogramas por metro quadrado

KIR6.2 – Inwardly rectifying potassium channel

LDL-colesterol – Lipoproteínas de baixa densidade

*LGR5 – Leucine-rich repeat-containing G protein-coupled receptor 5*

LRP – Leucine-responsive Regulatory Protein

MASS II – ***Medicine, Angioplasty or Surgery Study II***

mg/dL – miligramas por decilitro

MODY - *Maturity-Onset Diabetes of the Young*

MONICA – *Multinational **MON**itoring of trends and determinants in **CA**rdiovascular disease*

OMS – Organização Mundial de Saúde

PCR-RFLP *assay* – *Polymerase chain reaction-restriction fragment length polymorphism*

PPARG – Peroxisome proliferator-activated receptor gamma

Pro12Ala – Substituição da Prolina por Alanina no Códon 12

ROC – *Receiver operating characteristic*

RSA – tipo de enzima de restrição

SLC30A8 – *Solute carrier family 30 (zinc transporter), member 8*

SNP – *Single nucleotide polymorphism*

SPSS – *Software* de análise estatística

TCF-4 – *Transcription factor 4*

TCF7L2 – *Transcription factor 7 like-2*

THADA – *Thyroid adenoma associated*

TSPAN8 – *Tetraspanin 8*

US\$ – Dólares americanos

WFS1 – Gene *Wolfram syndrome 1 (wolframin)*

Wnt – Combinação das siglas Wg (wingless) e INT (gene de integração)

β – Beta

## LISTA DE TABELAS

Tabela 1: Características da população do Estudo MASS-II	35
Tabela 2: Características da população geral de Vitória	36
Tabela 3: Associação dos genótipos com diabetes tipo 2	37
Tabela 4: Regressão logística para risco de diabetes tipo 2	38
Tabela 5: Características da amostra da população adulta de Itapetininga conforme genótipo	39
Tabela 6: Características das crianças da amostra de Itapetininga conforme genótipo	40
Tabela 7: Variáveis do metabolismo glicêmico em crianças conforme genótipo	41
Tabela 8: Características iniciais dos diabéticos e não-diabéticos em Vitória (População de Derivação) e Ouro Preto (População de Validação)	42
Tabela 9: Análise univariada de fatores de risco para diabetes	43
Tabela 10: Variáveis de risco e Análise de Regressão Logística Múltipla	44
Tabela 11: Desempenho de diferentes limiares no escore de risco nos dois modelos preditivos,	45
Tabela 12: Desempenho dos dois modelos preditivos do risco de	47

diabetes tipo 2

## LISTA DE FIGURAS

Figura 1–A: Curvas ROC para a pontuação de risco nos dois modelos preditivos	48
Figura 1–B: Curva ROC para a pontuação de risco na população de validação.	48
Figura 2: Médias das Glicemias em jejum e após seguimento de 5 anos nos tercís do modelo 1 de pontuação.	49
Figura 3: Curvas ROC comparando modelo baseado em variáveis demográficas com o mesmo modelo acrescido do genótipo TT TCF7L2	50

## RESUMO

Os polimorfismos do gene TCF7L2 têm sido fortemente associados com risco de desenvolvimento de diabetes mellitus em populações de diversas origens étnicas. No presente estudo, investigou-se se esta associação se confirma em diferentes populações brasileiras e qual seu efeito sobre o desempenho de um modelo de predição de risco de diabetes mellitus quando a informação genética foi acrescentada às variáveis clínicas e laboratoriais iniciais.

Concluiu-se que, apesar de haver sido confirmado a associação do polimorfismo rs7903146 ao diabetes tipo 2 na população brasileira, a inclusão desta informação no modelo teve desempenho equivalente ao modelo baseado unicamente em variáveis clínicas.



## **SUMMARY**

Recently, polymorphisms of gene TCF7L2 have been strongly associated with Type 2 Diabetes risk in populations of diverse genetic backgrounds. In this study we investigated if this association is present in different Brazilian populations and how the inclusion of genetic information to a diabetes risk prediction model based on clinical and laboratorial variables would affect its performance. We concluded that, even though the TCF7L2 rs7903146 polymorphism is associated to Type 2 Diabetes risk in the Brazilian population, that inclusion of such information to a diabetes risk prediction model based on clinical variables lead to equivalent performance.

## 1) INTRODUÇÃO

### 1.1) Prevalência de diabetes no Brasil e no mundo

O diabetes tipo 2 (DM2) é uma doença crônica que afeta milhões de pessoas em todo o mundo civilizado. A Organização Mundial da Saúde (OMS) estima em 180 milhões o número de diabéticos em todo o mundo, enquanto a International Diabetes Federation (IDF) estimou, em 2007, este número em 246 milhões. Recentemente, temos observado uma clara aceleração nas taxas de incidência e prevalência mundial da doença, relacionada à epidemia de obesidade. Em 1993, por exemplo, havia cerca de 7,8 milhões de diabéticos nos Estados Unidos (EUA), segundo estimativas do Centro de Controle e Prevenção de Doenças (Center for Disease Control and Prevention, CDC, [www.cdc.gov](http://www.cdc.gov)), número equivalente a 3,1% da população nacional. Em 2008, um comunicado à imprensa do mesmo CDC foi publicado revendo essas estimativas para um total de 24 milhões de diabéticos entre diagnosticados ou não, praticamente 8% da população total daquele país(1).

No Brasil, os dados disponíveis são mais antigos. Em 1992 foi publicado o Estudo Multicêntrico sobre a Prevalência de Diabetes no Brasil (2), conhecido como “Censo Brasileiro de Diabetes”, com dados coletados de 1986 a 1988 em nove capitais brasileiras (Porto Alegre, São Paulo, Rio de Janeiro, Brasília, Salvador, Recife, João Pessoa, Fortaleza e Belém). A prevalência de DM2, na época, baseada em medidas de glicemia realizadas

em domicílios sorteados ao acaso, foi de 7,6% em indivíduos de 30 a 69 anos de idade. Quase 10 anos depois, entre 1996 e 1997, novo estudo realizado em Ribeirão Preto, seguindo a mesma metodologia do Censo Brasileiro de Diabetes, estimou a prevalência de diabetes naquele município em 12,1% na mesma faixa etária(3). Isso representaria um aumento de quase 60% na prevalência de diabetes avaliada pelos dois estudos.

O impacto desta escalada é enorme e preocupante, não só no que diz respeito à saúde pública. O diabetes é uma doença crônica que afeta indivíduos, famílias, sistemas de saúde e a economia dos países. Na grande maioria das vezes, especialmente em países mais pobres, é relacionada com menor expectativa de vida. Dados do IDF de 2007 estimam em 3,8 milhões as mortes por diabetes, cerca de 6% da mortalidade global. Ainda segundo o IDF, utilizando-se dados da OMS sobre o número de anos de vida perdidos por pessoa, seriam mais de 25 milhões de anos de vida por ano, além de mais 23 milhões de anos de vida perdidos por invalidez e redução na qualidade de vida causada pelas complicações do diabetes.

### 1.2) Custos econômicos do diabetes

Além do enorme impacto nos níveis pessoal e social, o custo econômico do diabetes e suas complicações é preocupante e vem crescendo exponencialmente. Recentemente, a *American Diabetes Association* (ADA) publicou um relatório referente aos custos envolvidos com o diabetes nos EUA no ano de 2007(4): o custo total estimado ficou em US\$

174 bilhões. Destes, US\$ 58 bilhões, foram estimados como impacto por redução da produtividade nacional e US\$ 116 bilhões (dois terços) foram gastos em excesso de despesas médicas, dos quais US\$ 27 bilhões foram alocados para o controle da glicemia, US\$ 58 bilhões em tratamento das complicações atribuíveis ao diabetes e US\$ 31 bilhões em excesso de gastos médicos em geral. Do total de gastos médicos, pelo menos metade é dedicada ao tratamento intra-hospitalar, enquanto medicamentos e equipamento ficam com 12% e honorários médicos com 9%. Diabéticos, em média, custam anualmente 2,3 vezes mais do que pessoas que não têm diabetes. Na presença de doença renal crônica terminal, o custo é até quatro vezes superior ao de um diabético sem complicações. De maneira interessante, à medida que se analisam países mais pobres, proporcionalmente mais recursos são gastos com medicações orais e medidas que evitam as complicações mortais da hiperglicemia. Na América Latina, por exemplo, o custo de medicações hipoglicemiantes orais é estimado em até mais de 50% do total gasto com a doença(5). No Brasil, não há dados exatos recentes disponíveis.

### 1.3) Fisiopatologia do diabetes

Entende-se por diabetes mellitus a doença cuja principal característica é a hiperglicemia. Entretanto, esta é uma característica comum a diversas doenças muitas vezes bastante diferentes entre si. Apesar de a hiperglicemia ser a mais marcante característica do diabetes, especialmente no tipo 2, pode-se dizer que ela reflete a quebra na homeostase entre duas

propriedades: capacidade relativa de secreção de insulina pela célula  $\beta$  pancreática e grau de resistência à ação da insulina em tecido hepático e muscular; a insulina, peptídeo secretado pela célula  $\beta$  pancreática em resposta à elevação glicêmica no plasma, aumenta a captação de glicose pelos tecidos periféricos e suprime a produção e liberação hepática de glicose. Dessa forma, ocorrem habitualmente picos e vales alternantes nos níveis de insulina e glucagon que ocorrem pra manter a homeostase glicêmica. A homeostase, ou tolerância glicêmica, depende de três ações altamente coordenadas para ser mantida: estímulo à secreção insulínica; supressão da produção hepática de glicose, mediada pela insulina; captação da glicose por tecidos periféricos, também mediada pela presença e ação da insulina.

#### Genética do diabetes tipo 2

Embora, há muito, seja observado clinicamente que o diabetes sofre importante influência de fatores ambientais, especialmente a dieta, também é observado, há décadas, que existe forte influência genética na sua determinação, conforme demonstraram estudos familiares, e em gêmeos, que mostram intensa agregação familiar e alta taxa de concordância entre gêmeos dizigóticos, ao redor de 70%, o dobro do observado entre irmãos(6, 7). Como gêmeos monozigóticos dividem o ambiente, tanto intra-uterino quanto extra-uterino e 100% da carga genética (ao contrário dos gêmeos dizigóticos, que têm somente 50% dos seus genes semelhantes), taxas de concordância em excesso daqueles dizigóticos são utilizadas para distinguir contribuintes genéticos dos não-genéticos. Em indivíduos acima dos 60

anos, a taxa de concordância referente ao diabetes foi de 35-58% em gêmeos monozigóticos, comparada com 17-20% em gêmeos dizigóticos(8, 9). A inclusão de intolerância à glicose na análise aumentou a taxa de concordância em gêmeos monozigóticos para 88%(10). A presença de diabetes do tipo 2 em um parente de primeiro grau aumenta o risco de diabetes em 2 a 4 vezes. Filhos de pai e mãe diabéticos têm 80% de risco de se tornarem diabéticos ao longo da vida(6, 11, 12). Outra evidência da influência da genética na patogenia do diabetes é a grande diferença de incidência e prevalência em populações etnicamente diversas. Os índios Pima, por exemplo, têm prevalência de diabetes ao redor de 50%, marcada por intensa resistência à insulina, maior do que a observada em caucasianos. Negros e hispânicos norte-americanos também se encontram em risco aumentado(11). Na população brasileira, existe o estudo prospectivo, ainda em andamento, do *Japanese-Brazilian Diabetes Study Group*, que começou a acompanhar, em 1993, a primeira (nascidos no Japão) e a segunda gerações (nascidos no Brasil) de japoneses imigrantes e residentes em Bauru, Estado de São Paulo. Este importante estudo confirma que a população mais jovem, exposta ao estilo de vida urbano e a hábitos alimentares ocidentais, tem prevalência aumentada de diabetes, doenças metabólicas(13) e doença cardiovascular(14), principalmente relacionadas a alterações dietéticas que incluem maior consumo de gorduras, especialmente as do subtipo trans(15). Outro estudo nacional recente, de base familiar, estimou a herdabilidade de diversos fatores de risco cardiovasculares em população geral de uma cidade do interior de Minas

Gerais. A herdabilidade da glicemia de jejum, ou seja, o quanto esta característica é determinada por fatores genéticos, após ajustes, foi de 32,3%, com P estatístico altamente significativo ( $< 0,00001$ )(16). Apesar de estes dados corroborarem a influência genética na determinação do diabetes tipo 2, observa-se que, mesmo em populações de alto risco genético para diabetes, há fatores ambientais que exercem enorme impacto e papel definitivo nessa equação. Isso é classicamente observado na diferença de prevalência de diabetes entre índios Pima americanos e mexicanos, sendo a prevalência dos últimos menos de 20% daquela dos índios Pima que vivem nos Estados Unidos (6,9% vs. 38%)(17).

Entretanto, mesmo com a clara influência genética de determinantes do diabetes tipo 2, a busca pelo isolamento desses fatores ainda está em andamento. Isso é típico de doenças complexas, nas quais apesar de um indivíduo poder carregar múltiplos fatores de risco genéticos que o predisponham à doença, pode ser necessária a associação a fatores ambientais para que o processo patogênico seja iniciado(18).

A procura por genes candidatos por muito tempo foi focada em genes codificadores de proteínas possivelmente envolvidas na fisiopatologia do diabetes tipo 2, como desenvolvimento pancreático, síntese, secreção ou ação da insulina(19). Até 2006, somente variantes nos genes *KCNJ11* (*potassium inwardly-rectifying channel, subfamily J, member 11*) e *PPARG* (*peroxisome proliferator-activated receptor gamma*) haviam sido relacionados de forma consistente com o diabetes do tipo 2. O gene *KCNJ11* codifica o *inwardly rectifying potassium channel (KIR6.2)*,

subunidade do canal de potássio ATP-dependente na célula beta pancreática. Mutações ativadoras desse gene levam a uma forma grave e precoce de diabetes chamado diabetes permanente neonatal, por insuficiência de secreção insulínica que tende a responder a tratamento com sulfoniluréias, drogas que agem bloqueando o canal de potássio e permitindo a despolarização celular e secreção insulínica, enquanto que mutações inativadoras do gene são associados com hipoglicemia hiperinsulinêmica também neonatal(20). Já o gene PPARG codifica um membro da família dos receptores nucleares ativados por proliferadores de peroxissomos, que modulam a expressão gênica de diversos outros fatores proliferativos e metabólicos, além de serem alvo das drogas pertencentes à classe das tiazolidinedionas. Carreadores dos alelos de risco do polimorfismo Pro12Ala desse gene apresentam importante resistência à insulina em fígado, músculo e gordura, aumentando o risco de diabetes na população, dessa forma.

#### TCF7L2 e Mecanismos Fisiopatológicos no Diabetes tipo 2

O gene *transcription factor 7 like-2* ou TCF7L2, antigamente referido como TCF-4 (*transcription factor 4*), é membro de uma família de fatores de transcrição na qual se encontram outros genes, não-homólogos, TCF-1 e TCF-2, também chamados, respectivamente, de HNF1A e HNF1B (*hepatocyte nuclear factor 1 homeobox A e B*), responsáveis por quadros de diabetes monogênico (MODY, ou Maturity-Onset Diabetes of the Young). O TCF7L2 é amplamente expresso e intimamente envolvido com a cascata de



sinalização Wnt, cujo termo advém da combinação de Wg (*wingless*) e *INT*, genes que codificam proteínas de integração, e podem ser pronunciados *wint*. O gene *wingless* foi originalmente identificado em estudos de embriogênese em *Drosophila*. A transdução da via Wnt é uma das vias centrais que controlam o crescimento e diferenciação dos organismos(21). Em um dos principais ramos da via Wnt, chamado via canônica, Wnts se ligam a receptores frizzled em conjunto com correceptores da família LRP (Leucine-responsive Regulatory Protein). A ativação resultante da via previne a fosforilação da  $\beta$ -catenina e sua degradação subsequente. A  $\beta$ -catenina estabilizada transloca-se para dentro do núcleo onde ela interage com fatores de transcrição da família TCF para ativar a expressão de genes-alvo do TCF7L2. Em conclusão, acredita-se que a regulação controlada desse mecanismo de sinalização leva a proliferação e diferenciação normais em tecidos-chave para a patogênese do diabetes tipo 2, como o adiposo e, possivelmente, o pâncreas endócrino(22-24).

Em humanos, foi demonstrado que pessoas homozigotas para as variantes que predispõem ao diabetes apresentam quadro de intolerância à glicose administrada por via oral, mas não intravenosa, o qual sugere a presença de um defeito no eixo enteroinsular, dependente de incretinas, entre as quais uma das principais é o *glucagon-like peptide 1* (GLP-1). O eixo enteroinsular é a rede de comunicação neural e endócrina entre intestino delgado e pâncreas que potencializa a liberação de insulina em resposta à alimentação. Há relatos de que a heterodimerização do TCF7L2 com a  $\beta$ -catenina poderia modular a produção de proglucagon intestinal, um

dos precursores do GLP-1(25). Também há dados que demonstram outras características metabólicas importantes em portadores do alelo T de risco para o polimorfismo rs7903146, como reduzida capacidade de secreção insulínica pela célula  $\beta$  pancreática, reduzido efeito incretina e aumento na taxa de gluconeogênese hepática, especialmente nos homozigotos para o alelo de risco. A expressão do TCF7L2 em ilhotas humanas foi 5 vezes maior em diabéticos, particularmente nos homozigotos para o alelo T, o que foi associado à redução da secreção de insulina estimulada por glicose nessas células, reforçando o envolvimento do gene na patogênese do diabetes(26). Uma metanálise mostrou o efeito dos polimorfismos do alelo T do polimorfismo TCF7L2 rs7903146 sobre o risco de diabetes tipo 2 em diversas populações mundiais, obtendo um *odds ratio* médio de 1,46 (1,42 – 1,51) ( $P = 5,4 \times 10^{-140}$ )(27)

## 2) OBJETIVOS

Os objetivos desse trabalho consistem de a) confirmar a associação entre as variantes do gene TCF7L2 e diabetes tipo 2 em populações brasileiras, que têm como característica serem geneticamente mais heterogêneas que as de origem norte-européias, devido à intensa e recente miscigenação local; b) testar utilidade de modelo de predição de risco de diabetes com base em variáveis clínicas somente e compará-lo com um que inclua variáveis laboratoriais e ainda com um terceiro que disponha de informações genéticas, para que se possa quantificar o impacto da inclusão de outras variáveis sobre o poder preditivo final.

### 3) MATERIAIS E MÉTODOS

Para a avaliação da relação da população brasileira com os polimorfismos do gene TCF7L2, utilizam-se dados previamente coletados e armazenados em laboratório, provenientes de estudos anteriores. Esses estudos abordaram amostras de quatro diferentes populações, que serão descritas abaixo.

#### 3.1) População de Vitória – ES

Um estudo transversal sobre fatores de risco para doenças cardiovasculares foi realizado na população urbana de Vitória, ES, usando-se as diretrizes do Projeto WHO-MONICA (Multinational **MON**itoring of trends and determinants in **C**ardiovascular disease)(28). O estudo pesquisou e analisou dados socioeconômicos em uma amostra probabilística de residentes adultos do município, entre 25 e 64 anos. Segundo o censo de 1996 realizado pelo IBGE (Instituto Brasileiro de Geografia e Estatística), a população de Vitória, naquele ano, era de 264.874 habitantes. A amostragem foi feita em quatro estágios: por bairro, setor censitário, sorteio de quarteirões para escolha de casas e sorteio de indivíduos em cada casa. A pesquisa foi realizada com um indivíduo de cada casa sorteada, e o sorteio foi feito randomicamente. Foram selecionadas 2.268 casas na cidade de Vitória e todas elas foram visitadas. Ao indivíduo selecionado em cada uma dessas casas foram explicados os

objetivos da pesquisa e ele foi convidado a participar, caso consentisse. Todos os pacientes desse e de outros estudos citados abaixo assinaram o termo de consentimento livre e esclarecido, assim como os projetos foram previamente aprovados pelos respectivos comitês de ética de cada centro organizador. Os indivíduos selecionados foram convidados a comparecer à Clínica de Investigação Cardiovascular no Hospital Universitário para serem submetidos a testes clínicos e laboratoriais. Da amostra total de mais de 2.200 pessoas, 1.507 compareceram e foram entrevistados e examinados, com a coleta de dados como peso, altura, circunferência abdominal, pressão arterial, presença de hábitos como tabagismo, consumo de álcool e quantidade de atividade física realizada rotineiramente, além da presença de comorbidades como diabetes e hipertensão. O sedentarismo foi definido como a realização de menos do que 1 hora de exercícios ao dia, pelo menos 3 vezes na semana. Glicemia, colesterol total, frações de lipoproteínas e triglicérides foram dosados por técnicas-padrão em amostras de sangue, após jejum de 12 horas, coletadas durante a visita à clínica(29).

### 3.2) Seguimento da amostra de Vitória – ES

Uma reavaliação dos indivíduos da amostra de Vitória foi realizada 5 anos após a inicial. Todos os participantes da primeira parte do estudo foram convidados a passar por nova avaliação, destes 655 concordaram e foram reavaliados segundo a mesma metodologia e mesma avaliação clínica e laboratorial de 5 anos antes.

### 3.3) População de Ouro Preto – MG

Ouro Preto é uma pequena cidade do interior do sudeste brasileiro, com 37.603 habitantes. Em um estudo transversal, 930 domicílios foram randomicamente selecionados dos 33 distritos da cidade, por amostragem simples, levando-se, porém, em conta a proporcionalidade da densidade populacional de cada distrito. Em cada casa sorteada, o indivíduo maior de 15 anos de idade, cujo aniversário estivesse mais próximo da data da entrevista, foi selecionado para inclusão no estudo. Os procedimentos relacionados à entrevista, exame físico e coleta de exames laboratoriais, seguiram as mesmas diretrizes e padrões do Projeto WHO-MONICA utilizado também no estudo de Vitória. Estudos prévios já foram publicados utilizando-se esta mesma base de dados(30).

Todos os participantes sem diabetes conhecido foram incluídos. O grau de tolerância à glicose foi classificado segundo o critério modificado da ADA de 2008(31). Definiu-se disglícemia como glicemia em jejum maior que 100 mg/dL. Indivíduos com GJ  $\geq$  126 mg/dL, um diagnóstico provisório de diabetes segundo a ADA, foram considerados diabéticos nesse estudo. Como em outros estudos, indivíduos com 34 anos de idade ou menos foram excluídos da análise devido à possibilidade desse grupo incluir outras síndromes hiperglicêmicas não-diabetes tipo 2. Os indivíduos foram classificados em 5 classes socioeconômicas (aqui agrupados em 4 classes), com base em dados do nível de educação do chefe da família e do tipo e número de equipamentos domésticos na casa(32). A classificação de subgrupos étnicos foi baseada em pigmentação do abdômen, cor e forma

dos cabelos, tipo e conformação do nariz e lábios e história familiar, determinada e de acordo com a opinião de dois médicos. Deve ser notado que tal estratificação de subgrupos étnicos foi baseada em critérios morfológicos, e não é uma estratificação étnica verdadeira, que envolve um grupo de características genéticas, lingüísticas e culturais. Em particular, já foi demonstrado que, na população brasileira, os critérios morfológicos não são altamente concordantes com a procedência. Esta classificação, entretanto, foi validada e previamente utilizada em outros estudos(33-35).

#### 3.4) Desenvolvimento do Modelo de Risco e Validação

Fatores de risco e variáveis laboratoriais clinicamente reconhecidos como associados ao diabetes tipo 2 foram selecionados para o desenvolvimento de um *escore* de risco. Esses fatores de risco selecionados incluíram idade, sexo, diagnóstico de hipertensão e nível de atividade física. Dados laboratoriais incluídos foram colesterol total, LDL-colesterol (lipoproteínas de baixa densidade), HDL-colesterol (lipoproteínas de alta densidade), triglicérides e ácido úrico. Análise univariada por regressão logística foi realizada com diabetes tipo 2 ( $GJ \geq 126$  mg/dL) como variável dependente. Na regressão logística múltipla, o critério para inclusão de variáveis foi  $P < 0,1$ . Essas variáveis foram incluídas utilizando-se do método por eliminação do tipo *stepwise backward*, com diabetes tipo 2 como variável dependente. Coeficientes de regressão logística foram utilizados para estimar a razão de chances para cada variável independente no modelo. As variáveis contínuas foram categorizadas em bandas. Para criar

o modelo final, variáveis com um valor P menor que 0,05 foram consideradas, e, para cada variável significativa na análise de regressão logística múltipla, um escore foi calculado pela multiplicação do coeficiente de regressão por 10, arredondando-se pelo integral mais próximo. Finalmente, um escore de soma foi calculado para cada participante pela adição do escore de cada variável significativa no modelo de risco. Sensibilidade, especificidade, valores preditivos positivo e negativo e acurácia foram comparados para diferentes *cutoffs* do escore. Além disso, a proporção entre sujeitos que precisariam testes subseqüentes (aqueles indivíduos que tiveram valores de escore maiores que o valor de *cutoff*) foi comparada entre os diversos grupos.

Para avaliar o desempenho do modelo, uma curva ROC (*receiver operating characteristic*) foi montada para analisar probabilidades obtidas com a regressão logística, sendo a área sob a curva utilizada para mensurar o poder de discriminação de indivíduos de alto e baixo risco. Curvas ROC foram desenhadas para o escore de risco de diabetes, sendo a sensibilidade desenhada no eixo y e a taxa de falso-positivo (1 - especificidade) desenhada no eixo x. Quanto mais discriminatório o teste, mais inclinada a parte da curva voltada pra cima e maior a área sobre a curva. O ponto ótimo foi determinado tentando-se conseguir a máxima sensibilidade e especificidade. Para avaliar a performance do escore de risco na população de Ouro Preto (validação externa), os métodos utilizados foram similares ao do teste exploratório.



### 3.5) População de Itapetininga – SP

A amostra de Itapetininga consiste de 487 indivíduos, entre 2 e 18 anos de idade, selecionados aleatoriamente em 89 escolas públicas na cidade de Itapetininga, SP. Também foram incluídos os pais desses jovens, os quais com estes compuseram 487 trios familiares, aumentando o número total da amostra para 1.461 pessoas. Os pais assinaram termo de consentimento informado referente tanto a eles próprios quanto aos filhos. Fatores de risco cardiovasculares foram avaliados utilizando-se um questionário-padrão direcionado tanto aos pais quanto aos filhos, além de um detalhado exame físico. Amostras de sangue foram coletadas após jejum de 12 horas e as medidas avaliadas segundo procedimentos-padrão laboratoriais. Para a estimativa da sensibilidade à insulina em jejum (HOMA%S) foi utilizado o “HOMA Calculator”, versão 2.2, desenvolvido pela Oxford University.

### 3.6) População do estudo MASS II

O estudo MASS II (Medicine, Angioplasty or Surgery Study) envolveu 611 pacientes com doença arterial coronária documentada e fração de ejeção esquerda normal(36). Utilizaram-se somente informações das características iniciais desta população. A presença de diabetes não foi necessária para o alistamento nesse protocolo. Foi possível conseguir dados de genótipo para o polimorfismo rs7903146 em 560 pacientes, dos quais 190 eram diabéticos segundo critérios da ADA ou tratamento prévio.

### 3.7) Limitações metodológicas das amostras

Apesar das múltiplas amostras e considerável número de participantes em cada uma delas, a metodologia utilizada na realização deste estudo apresenta algumas limitações que merecem consideração. Para o diagnóstico de diabetes foi utilizado uma única dosagem de glicemia plasmática em jejum, enquanto que o recomendado é que seja confirmadas eventuais alterações com outra amostra de sangue. Também não foram realizados testes orais de tolerância à glicose. Informações sobre a história familiar de diabetes ou outras comorbidades não foram incluídas no questionário.

Apesar dessas limitações terem sido reconhecidas antes do início do trabalho, elas não foram julgadas limitantes à realização e interpretação dos dados resultantes do trabalho, possivelmente por envolverem somente uma pequena fração do número total de participantes.

### 3.8) Análise estatística

A versão 13 do programa SPSS foi utilizado para a análise estatística. O ajuste para distribuição normal foi avaliado utilizando-se o teste de Kolmogorov-Smirnov. Para avaliar as diferenças entre diversas características, o teste *t* de Student ou ANOVA foi utilizado para variáveis contínuas e o  $\chi^2$  foi usado para variáveis categóricas. Valores de P menores que 0,05 em testes de dois caminhos foram considerados significantes. Dados foram mostrados como média  $\pm$  desvio-padrão para variáveis

contínuas (escala) e como porcentagens para variáveis ordinais ou categóricas.

### 3.9) Genotipagem do polimorfismo rs7903146 do gene TCF7L2

Foi extraído DNA genômico de leucócitos periféricos pelo procedimento de *salting-out*. Os *primers* utilizados foram : 5' ACA ATT AGA GAG CTA AGC ACT TTT TAG GTA 3' and 5' GTG AAG TGC CCA AGC TTC TC 3'. Resumidamente, o polimorfismo estudado foi detectado em ensaio *polymerase chain reaction-restriction fragment length polymorphism* (PCR-RFLP assay). Uma reação em cadeia de polimerase de 30 ciclos foi realizada no termociclador PTC-DNA Engine Tetrad<sub>2</sub> usando-se 10 µL de solução reagente que continha Tris-HCl 10 mM, KCl 50 mM, MgCl<sub>2</sub> 2,5mM, 100 µM de cada dNTP, 0,3 U Easy Taq Polimerase de DNA, 5 pM de cada *primer* e 1 µL de DNA genômico das amostras. Os produtos da PCR foram digeridos com 1 U da enzima de restrição RSA e visualizados por eletroforese em gel de agarose a 3%. O controle de qualidade para esses ensaios foi realizado por amostragem randômica, selecionando-se 40 amostras que foram submetidas à genotipagem por um técnico independente. A concordância observada entre ensaios de genotipagem foi 100%.

## 4) RESULTADOS

A seguir relatar-se-ão os resultados observados nas diferentes etapas deste trabalho, separados por amostras populacionais.

### 4.1) Vitória e MASS

Na amostra da população do estudo MASS-II, um total de 560 indivíduos foram genotipados para a variante do gene TCF7L2 rs7903146 e a proporção genotípica para esta variante estava em equilíbrio de Hardy-Weinberg. Houve predominância do sexo masculino (69,1%) e uma alta prevalência de DM2 de 31,0%. O IMC médio foi 27,1%. De acordo com esta classificação de peso, 66,2% desta população estava com sobrepeso ou era de obesos (Tabela 1). Homozigotos para o alelo T de risco, genótipo TT, revelaram aumento do risco de ter diabetes tipo 2 ( $p = 0,01$ ). Esses dados estão sumarizados na Tabela 3. Após ajuste para outras variáveis potencialmente associadas ao risco de diabetes tipo 2, a presença do genótipo TT permaneceu ainda significativamente associado à doença, aumentando em 2,0 vezes o risco ( $p = 0,01$ ) (Tabela 4).

Em relação à população geral de Vitória, havia 1.440 amostras de DNA disponíveis que passaram pelo processo de genotipagem para a variante rs7903146. As características desta população foram mostradas na Tabela 2. A média de idade foi 44,8 anos (variação entre 23-65). A prevalência de diabetes foi de 7,8%. Obesidade estava presente em 19,3%

das pessoas, mas 55,3% estava na faixa do sobrepeso ( $\text{IMC} \geq 25 \text{ kg/m}^2$ ). O IMC médio foi  $26,2 \pm 4,9 \text{ kg/m}^2$ . Indivíduos não-diabéticos tinham um IMC menor (25,9 contra  $30,6 \text{ kg/m}^2$  nos diabéticos). A prevalência geral do genótipo TT foi de 10,1%, enquanto no grupo de não-diabéticos foi de 9,9% e discretamente mais alta entre diabéticos (12,1%), diferença que não foi estatisticamente significativa ( $p = 0,46$ ). De interesse, o número total de indivíduos diabéticos portadores do genótipo TT, apesar do grande número inicial de pessoas da amostra, foi de somente 14 pessoas. Portanto, o poder estatístico de detecção desta associação nesta amostra da população geral foi somente de 24%.

Embora diversos estudos tenham observado associação do rs7903146 com alguns componentes da síndrome metabólica, não se observou qualquer associação significativa entre o alelo de risco T e tais características (Tabelas 1 e 2).

#### 4.2) Itapetininga

As características antropométricas e laboratoriais dos adultos da amostra de Itapetininga, de acordo com seu genótipo, estão mostradas na Tabela 5. Os adultos portadores do genótipo TT apresentaram significativo aumento de risco de DM2, mesmo após ajuste para idade ( $p < 0,001$ ), o que confirma relatos prévios da literatura. Entretanto, diferenças entre classes de genótipos em outras características como peso, distribuição de gordura e variáveis metabólicas não foram estatisticamente significantes.

Em crianças, os genótipos da variante rs7903146 não mostraram qualquer efeito estatisticamente significativo em variáveis laboratoriais como glicemia, níveis de colesterol ou triglicérides ou níveis de proteína C reativa, como demonstrado na Tabela 6. Também não foram notadas diferenças em pressão arterial, comprimento ao nascer, peso ao nascer ou altura na infância em modelo não-ajustado. Quando feitos ajustes para peso ao nascer e para sexo, o genótipo TT foi significativamente associado com maiores comprimentos ao nascer ( $p = 0,04$ ), diferença esta que não sofreu influência do genótipo da mãe. Não houve crianças diabéticas na amostra; somente uma apresentou glicemia em jejum maior que 110 mg/dL. Finalmente, nós compararam-se variáveis envolvidas no metabolismo glicêmico em dois grupos de genótipos em crianças desta amostra da população geral. Como demonstrado na Tabela 7, não houve qualquer diferença significativa em níveis de glicemia basal ou níveis de insulina de jejum relacionadas às classes de genótipo. Calcularam-se os índices HOMA para secreção de insulina pela célula beta e de resistência à insulina (HOMA-B% para função de célula beta e HOMA-IR para resistência à insulina(37)). Enquanto nenhum dos dois tenha alcançado significância estatística, houve tendência de as crianças portadoras do genótipo homocigoto para o alelo de risco T do polimorfismo rs7903146 apresentarem valores inferiores de HOMA-IR ( $p = 0,09$ ), o que sugere menor resistência à insulina.

#### 4.3) Modelos de Predição de Risco

Na população urbana de Vitória, da amostra total de 1.507 pessoas sem diabetes conhecido, 345 sujeitos foram excluídos da análise porque tinham menos de 35 anos de idade (inclusos 5 diabéticos). Também, em 8 dos 1.232 sujeitos restantes os dados sobre glicemia estiveram ausentes, restando 1.224 indivíduos para a análise final. Nesses indivíduos, a prevalência de diabetes tipo 2 foi de 9,6% (n = 118). Na população de Ouro Preto, de validação, da amostra inicial de 939 indivíduos, 474 foram excluídos da análise (35 pessoas eram sabidamente diabéticas; 348 pessoas tinham menos de 35 anos de idade; 91 não tinham dados sobre a glicemia), restando 465 indivíduos não-diabéticos na amostra final. A tabela 8 mostra as características clínicas e laboratoriais dos indivíduos diabéticos e não-diabéticos na amostra final de Vitória (população de derivação do modelo) e na amostra de Ouro Preto (população de validação externa). Deve ser observado que a população de Ouro Preto é mais velha e tem uma prevalência mais alta de hipertensão, dois fatores de risco reconhecidos para diabetes tipo 2.

Abaixo demonstrar-se-á o processo de desenvolvimento dos escores de risco, suas comparações e avaliação de performance. A Tabela 9 mostra os resultados de uma regressão logística univariada dos potenciais fatores de risco de diabetes. Após eliminação *stepwise* das variáveis não-significantes, dois modelos de predição foram criados: A) Modelo que contém somente variáveis antropométricas e/ou demográficas, chamado Modelo 1; B) Modelo no qual, além dos preditores presentes no Modelo 1, também constavam variáveis laboratoriais (chamado Modelo 2). Os valores

do coeficiente  $\beta$ , razão de chances com intervalo de confiança de 95% (IC95%) e valor p de cada variável independente nos dois modelos podem ser vistos na Tabela 10. No modelo 1, cujo valor máximo era 36, aumento da idade, aumento do IMC e hipertensão foram significativamente associados com diabetes tipo 2. No modelo 2, além dessas variáveis, colesterol total acima de 240 mg/dL e triglicérides acima de 150 mg/dL também se associaram significativamente com diabetes, e seu valor máximo possível era 45. A performance de ambos os escores de risco pode ser observada na Figura 1A, que mostra a curva ROC, e na Tabela 11. Os maiores valores na curva ROC, que obtiveram a mais alta sensibilidade e especificidade, foram 17,5 e 21,5 nos modelos 1 e 2, respectivamente. Não obstante os intervalos de confiança se sobreporem, o Modelo 2 conseguiu um desempenho discretamente superior, calculado pela área sob a curva. Entretanto, como o objetivo principal em se criar um escore de risco é seu potencial uso como uma ferramenta de triagem populacional para o risco de diabetes, o uso do escore mais simples, que emprega somente variáveis demográficas e clínicas, é preferível. O valor de *cutoff* de soma do escore  $\geq 18$  mostrou sensibilidade próxima ao valor pré-especificado (cerca de 75%). A utilização deste *cutoff* deve reduzir a necessidade de maiores investigações laboratoriais para cerca de 37% de toda a população (Tabela 11). Conquanto aumentar o valor de *cutoff* reduza a porcentagem de testes subseqüentes, isso também diminui a sensibilidade. Portanto, como a sensibilidade foi pré-especificada em 75%, o valor de *cutoff*  $\geq 18$  pontos foi escolhido para avaliação. O Modelo 1 alcançou uma boa performance com



capacidade discriminatória de cerca de 77% (AUC) e se correlacionou com dados basais de glicemia de jejum ( $r = 0,303$ ;  $p < 0,0001$ ).

Ainda que o escore de performance tenha sido menor na população de Ouro Preto do que na de Vitória, a área sob a curva no estudo de Ouro Preto foi igual a 0,72, indicando boa performance geral nesta população de validação (Figura 1B). Sensibilidade, especificidade, valores preditivos e o percentual de indivíduos que precisariam testes subsequentes devido ao *cutoff*  $\geq 18$  são mostrados na Tabela 11.

Posteriormente, dos 267 indivíduos com glicemia inicial menor do que 100 mg/dL e que foram acompanhados por 5 anos em Vitória, 28 evoluíram para disglucemia de jejum (glicemia de jejum acima de 100 e abaixo de 126 mg/dL) e somente 13 para diabetes (glicemia de jejum  $\geq 126$  mg/dL). Indivíduos do Modelo 1 com uma pontuação menor que 18 tiveram incidência de disglucemia de cerca de 8,7%. Essa incidência foi 15,3% no grupo com valores de escore maiores que 18 ( $p = 0,12$ ), o que demonstra uma tendência de o Modelo 1 predizer a incidência de disglucemia em 5 anos. Além disso, o maior tercil de escores do Modelo 1 em condições basais, cujo valor de *cutoff* é coincidentemente o mesmo encontrado pela análise da AUC ( $> 18$  pontos), tem média de glicemia de jejum inicial e após 5 anos significativamente mais altas (Figura 2), o que indica, mesmo após um período de 5 anos, que o escore do Modelo 1 continuava sendo significativamente correlacionado com a glicemia de jejum inicial.

Finalmente testou-se o desempenho do Modelo 1, baseado em variáveis clínicas e demográficas, quando acrescido da informação genética

proveniente da genotipagem dos polimorfismos do gene TCF7L2. A inclusão do genótipo TT no modelo não causou praticamente nenhuma melhora na área sob a curva ROC (AUC = 0,776; IC95% 0,73-0,82 em ambos os modelos). As características de desempenho dos dois modelos, com e sem informação genética, são mostrados na Tabela 12. Houve discreta melhora na especificidade quando o genótipo TT do gene TCF7L2 foi incluído, entretanto, a sensibilidade diminuiu. Praticamente não houve diferença entre as curvas ROC dos dois modelos (Figura 3).

## 5) DISCUSSÃO

Desde a descrição da associação de variantes comuns do gene TCF7L2 e diabetes tipo 2(38), achado prontamente replicado em diversas partes do mundo, as atenções têm-se voltado, principalmente, para dois aspectos práticos: o mecanismo através do qual esses polimorfismos genéticos aumentam tão significativamente o risco da doença e o modo como utilizar essa informação para criar melhores mecanismos de predição do risco. O objetivo do presente trabalho, além de confirmar a associação na geneticamente particular população brasileira, se aproxima do segundo. Se o uso de informações genéticas puder aprimorar modelos de predição de doença, auxiliando a identificação dos indivíduos que mais se beneficiariam de estratégias de prevenção e tratamento mais completas – e caras, poder-se-ia esperar a otimização dos recursos disponíveis para a saúde pública, um importante problema em praticamente todo o mundo.

Os primeiros relatos da associação dos polimorfismos do gene TCF7L2 com DM2 vieram de países da Escandinávia e amostras de caucasianos da América do Norte, populações caracterizadas pela alta homogeneidade genética, resultado de múltiplas gerações de cruzamentos entre indivíduos geneticamente semelhantes. Para confirmar os achados na população brasileira altamente diversa, do ponto de vista genético, iniciou-se este estudo examinando-se o genótipo de duas populações brasileiras, com o uso da amostra do Estudo MASS-II e uma amostra da população geral de

Vitória, para o polimorfismo rs7903146, um dos mais associados com DM2 no trabalho original(38). Como demonstrado acima, a associação se confirmou na população do estudo MASS-II, na qual os indivíduos homozigóticos para o alelo de risco T mostraram risco duas vezes maior de desenvolver DM2 (OR = 2,0; p = 0,01). Na população de Vitória não foi possível confirmar estatisticamente esta associação, provavelmente devido à baixa prevalência de diabetes nesta amostra de população geral.

Em outra etapa desse trabalho, buscou-se a relação entre o polimorfismo rs7903146 do gene TCF7L2 em outra população geral, a de Itapetininga, amostra que dispunha de trios formados com alunos da rede escolar da zona rural da cidade e seus respectivos pai e mãe. Em adultos, como se observou em Vitória, a relação entre o polimorfismo e DM2 se confirmou em modelo de herança recessiva, com os homozigotos para o alelo T de risco, apresentando prevalência de DM2 2,48 vezes superior ao observado na população restante (p < 0,001). Em crianças esses dados apontaram para maiores comprimentos, dos homozigotos para o alelo de risco, ao nascimento, um achado inédito – e contrário a dois relatos prévios(39, 40) da literatura, embora os mecanismos envolvidos sejam desconhecidos até o momento. A análise da amostra de crianças também demonstrou tendência à menor resistência à insulina pelo modelo recessivo de herança. Não houve efeito aparente do genótipo sobre valores de HOMA-β%, que avalia a secreção insulínica pela célula β. Na literatura, há um estudo em crianças o qual constatou que o HOMA-β% foi maior em homozigotos para o alelo de risco T quando comparados com homozigotos

para o alelo C(41). Esse resultado vai contra diversos outros relatos em adultos em que se sugere que o mecanismo pelo qual o gene TCF7L2 estaria relacionado com o DM2 seria por um defeito de secreção de insulina, possivelmente modulado por incretinas(25, 42) – e não por aumento da resistência à insulina, como é esperado em casos de diabetes associados ao excesso de peso.

Portanto, no Brasil, a associação entre o polimorfismo rs7903146 do gene TCF7L2 foi confirmada em amostras do Estudo MASS-II e da porção adulta da amostra da população geral de Itapetininga. Não foi estatisticamente significativa a relação entre as variantes do gene e o risco de DM2 na coorte de Vitória.

Com a associação descrita através de um estudo do tipo pesquisa de associações em todo o genoma (*genome-wide association scan*, GWAS) entre o gene TCF7L2 e DM2 comprovada também na população brasileira, pôde-se passar à segunda fase do trabalho, que engloba a tentativa de utilizar a informação genética no desenvolvimento de ferramentas de predição de risco. A informação de que alguns polimorfismos, em especial os do gene TCF7L2, poderiam aumentar o risco de diabetes foi amplamente divulgado nos últimos 2 anos. Alguns estudos, na literatura, demonstraram que a associação da informação de mais variantes genéticas previamente associadas com diabetes poderia aumentar a capacidade de identificação de indivíduos com maior risco. Por exemplo, Weedon *et al.* mostraram dados em que a associação de alelos de risco definidos previamente por uma metanálise, no caso TCF7L2, PPARC e KCNJ11, em indivíduos britânicos,

levou a OR de 5,71 para o risco de diabetes (IC95%, 1,15 a 28,3) quando comparado ao que correm os portadores de alelos selvagens. A área sob a curva ROC, determinada por modelo de regressão logística, foi de 0,58 – resultado considerado fraco no que concerne agregar novas informações preditivas ao modelo(43).

Além de usar a informação genética isoladamente, como feito acima, pode-se associá-la a modelos de avaliação de risco tradicionalmente utilizados. Diversos modelos baseados em variáveis clínicas já foram descritos e validados em diversas situações. As variáveis associadas são definidas por regressão logística e portanto, teoricamente, devem ser avaliadas de forma específica em cada amostra. A obesidade, por exemplo, é reconhecida como fator de risco para o DM2 há décadas, mas quando avaliada por IMC ou circunferência abdominal em indivíduos de etnia asiática é frequentemente subestimada.

Cauchi *et al.*(44) combinaram dados de um modelo de predição baseado em variáveis clínicas com outro combinando o modelo clínico a informações genéticas. Eles analisaram dados provenientes de um grande estudo de associação em todo o genoma (GWAS) na população francesa(45), no qual 15 SNPs, inclusive o TCF7L2, foram identificados como associados ao diabetes tipo 2, embora alguns não tenham sido replicados em outras populações após a descrição inicial. Quando combinada toda a informação disponível em regressão logística ajustada para idade, gênero e IMC para montagem de uma curva ROC, esta apresentou um resultado muito mais expressivo de 0,86, o que demonstra

potencial para uso, na prática clínica. Também foi demonstrado o efeito cumulativo desses polimorfismos sobre o risco. Quanto mais alelos de risco um indivíduo carrega, maior a chance de desenvolvimento de diabetes. Indivíduos com pelo menos 18 alelos de risco (dentre os 30 possíveis) teriam um risco de desenvolvimento de DM2 quase 9 vezes maior que indivíduos-controle. Se confirmado, esse aumento no risco de DM2 devido a fatores genéticos seria maior do que o provocado pela obesidade grave (OR = 7,37), fator de risco mais estabelecido do diabetes(46). Balkau *et al.* estudaram a incidência de diabetes em população adulta avaliada em intervalos de 3 anos durante 9 anos no estudo DESIR (Data from the Epidemiological Study on the Insulin Resistance Syndrome). As variáveis clínicas e laboratoriais preditivas de diabetes nesta população foram circunferência abdominal e hipertensão arterial em ambos os sexos, tabagismo em homens e história familiar de diabetes em mulheres, glicemia em jejum em ambos os sexos e níveis de triglicérides e HDL-colesterol. Das variáveis genéticas estudadas (Glucokinase: *GCK-30* G/A rs1799884, Interleucina 6: *IL6-174* G/C rs1800795, Kir6.2: *KCNJ11* E23K rs5219 e *TCF7L2* rs7903146), nenhuma se relacionou significativamente com diabetes incidente nesta amostra. Se bem que indivíduos carreadores de quatro alelos deletérios tivessem um risco de diabetes incidente (OR) 3,60 vezes maior (95% CI 1.09–11.9) em homens e 3,22 vezes maior (0.62–16.5) em mulheres, diferença estatisticamente significativa em ambos os sexos (respectivamente,  $P < 0,008$  e  $0,03$ ), a inclusão desta informação no modelo clínico-laboratorial de predição modificou muito pouco a curva ROC,

indicando pequena melhora na capacidade preditiva do modelo(47). Ainda outro estudo nessa linha analisou a capacidade de melhorar a predição de diabetes incidente com informação proveniente de testes genéticos utilizados em estudo prospectivo na população geral de Rotterdam(48). Para essa finalidade, foram submetidos ao processo de genotipagem 18 polimorfismos relacionados ao diabetes provenientes de recentes estudos de associação em todo o genoma, na literatura. Quanto à associação com diabetes, 9 dos 18 polimorfismos estudados estavam correlacionados com diabetes, porém a combinação da informação desses com parâmetros clínicos alterou muito pouco a área sob a curva ROC, novamente utilizada para avaliar a capacidade de predição do modelo. Parâmetros genéticos tiveram AUC de 0,6 (IC95% 0,57 – 0,63); 0,66 (0,66 – 0,68) no modelo baseado em parâmetros clínicos; e 0,68 (0,63 – 0,71) no modelo combinado. Portanto, também nessa amostra o poder de predição não foi muito alterado com a adição de variáveis genéticas ao modelo. Recentemente, ainda mais dois estudos com metodologias semelhantes confirmaram esses achados em grandes populações acompanhadas prospectivamente durante 28 e 23 anos(49, 50). O primeiro avaliou 2.377 pacientes acompanhados no Framingham Offspring Study, onde 18 polimorfismos associados a DM2 em estudos recentes(51-53) foram testados e avaliados em modelo de predição baseado no número de alelos de risco presentes, para o que utilizou-se de regressão logística para gerar estatística C o que indica o quanto esse modelo consegue discriminar o risco de diabetes se utilizado sozinho ou em conjunto com modelo baseado em variáveis clínicas. O resultado,



confirmando achados prévios, é de que, apesar do modelo conseguir prever risco de diabetes nessa população, ele melhora muito pouco a capacidade de predição do modelo baseado em variáveis clínicas (Valor da estatística C foi 0,534 sem o escore genético e 0,581 com o escore em modelo não ajustado;  $p = 0,01$  e  $0,900$  em modelo sem escore genético e  $0,901$  em modelo com escore genético quando ajustado para idade, sexo, história familiar, IMC, glicemia de jejum, pressão arterial sistólica, níveis de HDL colesterol e triglicérides;  $p = 0,49$ ). Finalmente, o último trabalho publicado sobre o assunto até o momento utilizou uma grande amostra europeia acompanhada prospectivamente na Suécia (the Malmo Preventive Project,  $n = 16061$ ) e na Finlândia (the Botnia Study,  $n = 2.770$ , totalizando 18831 pessoas seguidas por mais de 23 anos). Foram identificadas as características clínicas associadas ao risco de DM2 nessa população, que foram: história familiar de diabetes, IMC alto, aumento das enzimas hepáticas, tabagismo atual e medidas reduzidas de ação e secreção de insulina. Também foram pesquisadas variantes genéticas de 16 genes associados ao diabetes descritos na literatura (*TCF7L2*, *PPARG*, *FTO*, *KCNJ11*, *NOTCH2*, *WFS1*, *CDKAL1*, *IGF2BP2*, *SLC30A8*, *JAZF1*, *HHEX*, *CDKN2A/CDKN2B*, *CDC123/CAMK1D*, *TSPAN8/LGR5*, *THADA*, *ADAMTS9*). Os 11 primeiros associaram-se com DM2 nesta amostra e desses, 8 associaram-se com medidas que sugerem função prejudicada de células beta pancreática. Entretanto, a adição de variáveis genéticas ao modelo aumentou a área sob a curva ROC de 0,74 para 0,75, apesar do aumento dessa magnitude ter sido estatisticamente significativo ( $p = 1,0 \times$

$10^{-4}$ ).

Portanto, não há quaisquer evidências, até esse momento, de que variáveis genéticas aumentem o valor discriminativo de modelos preditivos de risco de diabetes do tipo 2 a ponto de justificar sua inclusão nos modelos preditivos vigentes. Os melhores modelos de predição de diabetes tipo 2 são baseados em informações clínicas, antropométricas e laboratoriais. Entretanto, as variáveis presentes nesses diferentes modelos são de início relativamente tardio; no presente estudo as variáveis associadas com risco de desenvolvimento de diabetes tipo 2 foram obesidade, hipertensão e idade mais avançada, enquanto outros trabalhos listam também dislipidemias ou mesmo medidas de glicemia alterada. Apesar desses critérios permitirem a estratificação e identificação de indivíduos potencialmente em alto risco de doença, os próprios critérios, em grande parte, já são doenças em si mesmo, tornando a tentativa de identificação precoce e prevenção de doenças de certa forma menos eficiente. Uma ferramenta de predição de risco seria realmente eficaz quando for capaz de identificar indivíduos ainda totalmente saudáveis, porém que apresentem uma combinação genética que os predisponha a apresentar determinadas doenças, porém anos antes que qualquer sintoma se torne aparente. Esta abordagem realmente preditiva permitiria a melhor distribuição de recursos e teria o potencial de aliviar a enorme conta paga atualmente no tratamento de diversas doenças crônicas.

## 6) CONCLUSÕES

Conclui-se que o polimorfismo rs7903146 do gene TCF7L2 está associado ao diabetes mellitus tipo 2 em populações brasileiras adultas. A introdução de tal informação a modelos de predição de risco de desenvolvimento de diabetes tipo 2 teve desempenho semelhante aos dos modelos previamente vigentes baseados em variáveis clínicas.

## 7) ANEXOS

Tabela 1: Características da população do Estudo MASS-II

	Total	rs7903146		p
		CC/CT	TT	
Número de pacientes (%)	611 (100)	495 (88,4)	65 (11,6)	
Gênero masculino	423 (69,1)	339 (88,5)	155 (88,1)	,88
Idade (anos)	59,8 ± 9,1	59,7 ± 9,12	58,8 ± 9,6	,45
IMC médio (kg/m <sup>2</sup> )	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
GJ média (mg/dL)	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
Colesterol total (mg/dL)	223,2 ± 47,7	222,9 ± 47,6	221,7 ± 52,5	,85
HDL-C médio (mg/dL)	37,4 ± 10,4	37,5 ± 10,6	36,5 ± 9,9	,49
Triglicérides médio (mg/dL)	195,1 ± 121,0	193,7 ± 118,0	179,9 ± 97,0	,37
Componentes da SM:				
Obesidade (IMC)	129 (21,2)	112 (22,7)	10 (15,6)	,19
Hipertrigliceridemia	344 (57,1)	279 (57,1)	34 (53,1)	,55
HDL-C baixo	447 (78,1)	371 (79,3)	46 (75,4)	,48
Hipertensão arterial	364 (59,6)	292 (59,1)	36 (55,4)	,64
GJ ≥ 110 mg/dL	278 (46,6)	223 (45,9)	35 (54,7)	,18

IMC = Índice de massa corpórea; GJ = Glicemia de jejum; HDL-C = Colesterol de alta densidade de lipoproteínas; Os componentes da Síndrome Metabólica (SM) foram categorizados conforme critérios da National Cholesterol Education Program (NCEP), 2003.

Tabela 2: Características da população geral de Vitória

	Total	rs7903146		p
		CC + CT	TT	
Número pacientes (%)	1577 (100)	1303 (89,9)	146 (10,0)	
Gênero masculino	718 (45,6)	600 (46,0)	64 (43,8)	,61
Idade (anos)	44,8 ± 10,9	44,7 ± 10,9	44,3 ± 10,7	,65
IMC médio (kg/m <sup>2</sup> )	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
GJ média (mg/dL)	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
Colesterol Total(mg/dL)	214,4 ± 47,8	214,0 ± 48,1	215,4 ± 44,0	,74
HDL-C médio (mg/dL)	45,4 ± 12,3	45,2 ± 12,1	44,0 ± 10,4	,27
Triglicérides médio (mg/dL)	137,6 ± 127,9	137,2 ± 131,0	133,0 ± 82,1	,70
PAS (mmHg)	128 ± 2	128 ± 2	129 ± 2	,52
PAD (mmHg)	84 ± 1	84 ± 14	84 ± 14	,56
Componentes SM				
Obesidade visceral	255 (16,2)	216 (16,6)	21 (14,4)	,50
Hipertrigliceridemia	484 (30,7)	391 (30,1)	51 (35,0)	,23
HDL-c baixo	848 (53,8)	705 (54,1)	79 (54,1)	,99
Hipertensão arterial	727 (46,1)	596 (45,7)	75 (51,4)	,19
GJ 110 MG/dL	334 (21,4)	281 (21,7)	29 (20,0)	,63
Síndrome metabólica (SM)	397 (25,4)	326 (25,2)	40 (27,6)	,53

IMC = Índice de massa corpórea; GJ = Glicemia de jejum; HDL-C = Colesterol de alta densidade de lipoproteínas; PAS = Pressão arterial sistólica; PAD = Pressão arterial diastólica  
 Presença de obesidade visceral foi classificada conforme medidas de circunferência abdominal. Os componentes da Síndrome Metabólica (SM) foram categorizados conforme critérios da National Cholesterol Education Program (NCEP), 2003.

Tabela 3: Associação dos genótipos com diabetes tipo 2

rs7903146		DM n (%)	Não-DM n (%)	OR / p
MS2	CC + TC	144 (83,2)	350 (90,7)	OR = 2,0
	TT	29 (16,8)	36 (9,3)	P = 0,01*
VIT	CC + TC	102 (87,9)	1193 (90,1)	OR = 1,5
	TT	14 (12,1)	131 (9,9)	P = 0,46

Modelos ajustados para idade e sexo.

Tabela 4: Regressão logística para risco de diabetes tipo 2

	MASS II		Vitória	
	OR	P	OR	p
Genótipo TT	2,0	,01	1,5	,237
Idade	1,0	,09	1,1	< 0,0001
Gênero feminino	1,2	,33	1,0	,863
Obesidade	1,6	,02	4,8	< 0,0001

Tabela 5: Características da amostra da população adulta de Itapetininga conforme genótipo

	CC/CT	TT	p
n	780 (87,7%)	109 (12,3%)	
Idade	39,67 ± 7,98	40,47 ± 8,70	0,36
Peso	69,62 ± 14,55	69,64 ± 14,28	0,99
Cintura	87,70 ± 13,22	87,44 ± 11,93	0,88
Quadril	97,28 ± 9,69	97,06 ± 9,84	0,86
ICQ	0,89 ± 0,10	0,88 ± 0,15	0,78
IMC	26,32 ± 4,75	27,00 ± 5,43	0,23
Glicose	90,70 ± 27,15	94,92 ± 38,80	0,17
Disglicemia	93 (13,4%)	14 (14,4%)	0,77
DM2	<b>20 (2,9%)</b>	<b>7 (7,2%)</b>	<b>0,02</b>
CT	197 ± 42,10	197 ± 43,38	0,91
LDL-c	123 ± 36,74	125 ± 36,78	0,74
HDL-c	44,41 ± 12,36	44,14 ± 11,21	0,84
Triglicérides	146 ± 104,20	143 ± 90,82	0,78
PA sistólica	120 ± 19,45	121 ± 17,85	0,72
PA diastólica	77 ± 13,06	78 ± 10,36	0,70

ICQ = Índice cintura-quadril; IMC = Índice de massa corpórea em kg/m<sup>2</sup>; DM2 = Diabetes do tipo 2; CT = Colesterol total; LDL-c = colesterol de lipoproteínas de baixa densidade; HDL-c = colesterol de lipoproteínas de alta densidade, todos em mg/dL; PA = Pressão arterial em mmHg.



Tabela 6: Características das crianças da amostra de Itapetininga conforme genótipo

	CC/CT	TT	p
n = 491	437 (87,4%)	54 (10,8%)	
Idade	10,28 ± 3,83	9,16 ± 3,44	0,04
Gênero (masc)	175 (40,0%)	22 (40,7%)	0,92
Comp, ao nascer	47,99 ± 3,55	48,69 ± 3,28	0,17
Peso ao nascer	2963 ± 592	3031 ± 502	0,42
Altura	1,34 ± 0,20	1,30 ± 0,17	0,13
DP Altura	-0,29 ± 1,1	-0,04 ± 0,93	0,12
Peso	36,48 ± 18,10	32,50 ± 14,90	0,12
DP Peso	-0,12 ± 1,22	0,00 ± 1,24	0,48
Cintura	64,98 ± 13,72	62,55 ± 12,27	0,21
Quadril	73,89 ± 15,43	70,80 ± 13,63	0,16
ICQ	0,90 ± 0,07	0,90 ± 0,06	0,78
IMC	18,74 ± 4,63	18,12 ± 4,08	0,35
DP IMC	0,19 ± 1,15	0,23 ± 1,21	0,82
Glicose	85,87 ± 8,15	84,50 ± 7,35	0,24
Disglicemia	18 (94,63%)	1 (5,26%)	0,41
CT	171,60 ± 27,67	176,19 ± 29,77	0,25
LDL-c	105,58 ± 0,36	108,59 ± 0,26	0,36
HDL-c	47,98 ± 10,05	48,98 ± 8,95	0,48
Triglicérides	90,66 ± 45,05	91,41 ± 35,65	0,90
PA sistólica	95,00 ± 13,40	92,87 ± 9,88	0,25
PA diastólica	63,26 ± 9,46	62,31 ± 8,16	0,48
Fibrinogênio	225 ± 98,50	230 ± 94,10	0,75
Proteína C reativa	2,22 ± 7,42	4,38 ± 12,86	0,17

Variáveis contínuas demonstradas como média ± desvio padrão (DP) e as categóricas em número e porcentagem; ICQ = Índice cintura-quadril; IMC = Índice de massa corpórea; CT = Colesterol total; LDL-c = colesterol de lipoproteínas de baixa densidade; HDL-c = colesterol de lipoproteínas de alta densidade; PA = Pressão arterial

Tabela 7: Variáveis do metabolismo glicêmico em crianças conforme genótipo

	CC/CT	TT	p
Glicose	85,81 (85,07 – 86,56)	84,98 (82,86 – 87,09)	0,24
Insulina	2,50 (2,13 – 2,87)	3,13 (2,03 – 4,24)	0,29
HOMA-B%	222,98 (223,19 – 212,20)	205,92 (172,78 – 239,05)	0,33
HOMA-IR	2,91 (2,74 – 3,10)	2,42 (1,87 – 2,97)	0,09

Variáveis apresentadas como Média  $\pm$  Desvio padrão; p = p estatístico.

Tabela 8 - Características iniciais dos diabéticos e não-diabéticos em Vitória (População de Derivação) e Ouro Preto (População de Validação)

	População de Derivação (Vitória)			População de Validação (Ouro Preto)		
	DM2 (n = 118)	Não-DM2 (n = 1106)	p *	DM2 (n = 29)	Não-DM2 (n = 438)	p *
Idade (anos)	53,4 ± 7,4	48,4 ± 8,3	<0,0001	59,5 ± 11,0	51,9 ± 12,0	0,001
Gênero (%)						
Masculino	39,8	45,8	0,216*	17,2	26,3	0,282
Feminino	60,2	54,2		82,8	73,7	
Raça (%)			0,54	ND	ND	-
Branços	33,9	37,7				
Mulatos	55,1	48,2				
Negros	6,8	7,3				
Outros	4,2	5,9				
Sedentarismo (%)	73,3	72,4	0,32	ND	ND	-
Tabagismo (%)	26,3	22,6	0,098	17,8	21,4	0,97
IMC (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
Circunferência Abdominal (cm)	96,91 ± 12,81	86,75 ± 11,48	0,18	97,96 ± 13,3	87,37 ± 12,7	<0,0001
ICQ	0,93 ± 0,08	0,88 ± 0,09	0,154	ND	ND	-
PAS (mmHg)	144 ± 23	129 ± 22	0,35	157 ± 25,6	142 ± 27,5	0,006
PAD (mmHg)	90 ± 14	85 ± 14	0,98	91 ± 16,2	89 ± 15,1	0,48
Hipertensão (%)	14,6	5,1	<0,0001	89,7	65,5	0,027
Glicemia (mg/dl)	185,34 ± 69,78	99,07 ± 10,93	<0,0001	155,1 ± 41,0	96,45 ± 11,0	<0,0001
Colesterol total (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
Triglicérides (mg/dl)	219,50 ± 251,46	138,82 ± 116,46	<0,0001	148,72 ± 56,0	121,0 ± 68,0	0,046
Ácido Úrico (mg/dl)	5,20 ± 1,38	4,82 ± 1,56	0,2	ND	ND	-
Creatinina (mg/dl)	0,97 ± 0,19	0,98 ± 0,2	0,64	ND	ND	-

(\*) valor de p para comparação entre sujeitos diabéticos e não-diabéticos em cada população

ND – não disponível

Tabela 9 - Análise univariada de fatores de risco para diabetes

	Odds Ratio	95%CI	$\beta$	p
Idade (anos)				
45-54 vs 35-44	2,72	1,52-4,85	1,0	0,001
$\geq 55$ vs 35-44	4,64	2,65-8,13	1,535	<0,0001
Gênero				
Feminino : Masculino	1,276	0,867-1,879	0,24	0,22
Raça (%)				
Mulato vs Brancos	1,271	0,840-1,924	0,240	0,256
Negros vs Brancos	1,030	0,465-2,281	0,029	0,943
Outros vs Brancos	0,802	0,305-2,106	-0,221	0,654
Sedentarismo (%)	0,891	0,573-1,384	-0,116	0,607
Tabagismo (%)	1,233	0,799-1,903	0,209	0,345
IMC (kg/m <sup>2</sup> )				
25-29,9 vs < 25	1,887	1,066-3,339	0,635	0,029
$\geq 30$ vs <25	8,052	4,736-13,689	2,086	<0,0001
Circunf. Abdominal (cm)				
H $\geq$ 94 e M $\geq$ 80 vs NL	3,378	2,027-5,629	1,217	<0,0001
H 94-101,9/M 80-87,9 vs NL	1,539	0,776-3,053	0,431	0,217
H $\geq$ 102/W $\geq$ 88 vs NL	5,206	3,043-8,905	1,650	<0,0001
Hipertensão arterial	3,156	2,076-4,799	1,149	<0,0001
CT (mg/dl)				
$\geq$ 240 vs <240	2,259	1,537-3,321	0,815	<0,0001
LDL-c (mg/dl)				
$\geq$ 130 vs <130	1,861	1,170-2,961	0,621	0,009
HDL-c (mg/dl)				
Baixo vs $\geq$ 40H/50M	1,266	0,844-1,899	0,236	0,254
Triglicérides (mg/dl)				
$\geq$ 150 vs <150	2,776	1,886-4,087	1,021	<0,0001
Ácido Úrico (mg/dl)				
$\geq 5,5$ or vs <5,5	1,161	1,034-1,303	0,149	0,011
Creatinina (mg/dl)	0,768	0,291-2,026	-0,264	0,594

A variável Circunferência abdominal teve diferentes valores categóricos para homens (H) e mulheres (M); 95%IC = Intervalo de confiança;  $\beta$  = coeficiente angular; p = p estatístico; IMC = Índice de massa corpórea; NL = normal; CT = Colesterol total; LDL-c = colesterol de lipoproteínas de baixa densidade; HDL-c = colesterol de lipoproteínas de alta densidade

Tabela 10 – Variáveis de risco e Análise de Regressão Logística Múltipla

	$\beta$	Odds Ratio	95%CI	p	Escore risco ( $\beta \times 10$ )
<b>Modelo 1</b>					
Idade					
45-54 vs 35-44	0,743	2,102	1,150 3,842	0,016	7
$\geq 55$ vs 35-44	1,227	3,412	1,894 6,145	<0,0001	12
IMC (kg/m <sup>2</sup> )					
25-29,9 vs < 25	0,473	1,605	0,899-2,866	0,110	5
$\geq 30$ vs <25	1,802	6,060	3,492-10,518	<0,0001	18
Hipertensão	0,626	1,870	1,178-2,969	0,008	6
<b>Modelo 2</b>					
Idade (anos)					
45-54 vs 35-44	0,805	2,237	1,178-4,246	0,014	8
$\geq 55$ vs 35-44	1,238	3,447	1,844-6,444	<0,0001	12
IMC (Kg/m <sup>2</sup> )					
25,1-29,9 vs < 25	0,408	1,504	0,801-2,821	0,204	4
$\geq 30$ vs <25	1,730	5,639	3,091-10,290	<0,0001	17
Hipertensão	0,512	1,669	1,032-2,697	0,037	5
Colesterol total (mg/dl)					
>240 vs <240	0,517	1,677	1,078-2,607	0,022	5
Triglicérides (mg/dl)					
>150 vs <150	0,587	1,799	1,154-2,803	0,009	6

P = p estatístico; IMC = Índice de massa corpórea

Tabela 11 – Desempenho dos limiares no escore de risco nos dois modelos preditivos

Dados de Derivação								
	S	E	VPP	VPN	A	OR (95%IC)	Valor P	NTA (%)
<b>Modelo 1</b>								
<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
<b>Modelo 2</b>								
<i>Cutoffs</i>								
20	0,7739	0,6475	0,189	0,964	0,66	6,29 (3,43-9,15)	<0,00001	39,3
21	0,7565	0,6586	0,191	0,962	0,67	5,99 (3,33-8,66)	<0,00001	38,1
22	0,7391	0,691	0,203	0,961	0,7	6,34 (3,58-9,1)	<0,00001	35,0
24	0,6174	0,7752	0,226	0,95	0,76	5,56 (3,33-7,8)	<0,00001	26,2
Dados de Validação								
	S	E	VPP	VPN	A	OR (95%IC)	Valor P	NTA (%)
<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

S = Sensibilidade; E = Especificidade; VPP = Valor Preditivo Positivo; VPN = Valor Preditivo Negativo; A = Acurácia; OR = Odds Ratio ou Razão de Chances; NTA = Necessidade de Testes Adicionais.

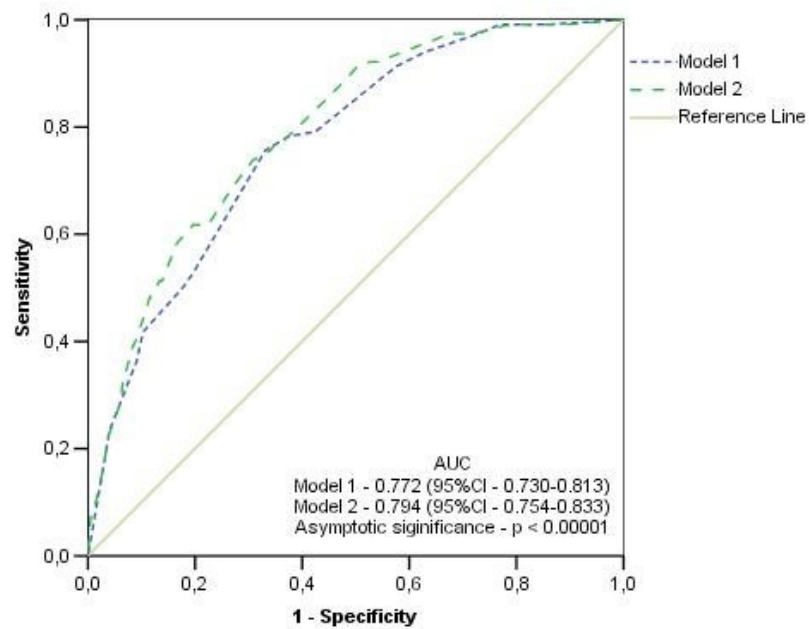
Tabela 12 – Desempenho dos dois modelos preditivos do risco de diabetes tipo 2

Modelo 1 – Somente variáveis demográficas								
<i>Cutoff</i>	S	E	VPP	VPN	A	OR (IC95%)	P	NTA (%)
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
Modelo 2 – Variáveis demográficas e genótipo <i>TCF7L2</i> TT								
<i>Cutoff</i>								
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

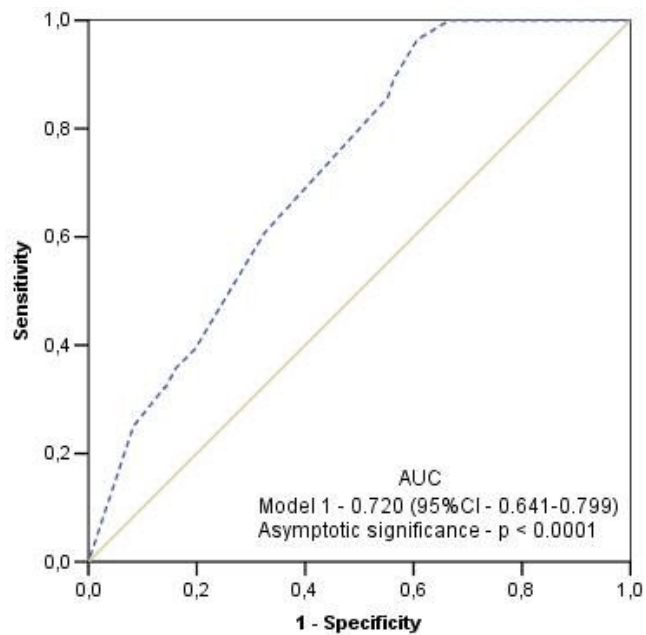
S = Sensibilidade; E = Especificidade; VPP = Valor Preditivo Positivo; VPN = Valor Preditivo Negativo; A = Acurácia; OR = Odds Ratio ou Razão de Chances; NTA = Necessidade de Testes Adicionais.

Figura 1 – A) Curvas ROC para a pontuação de risco nos dois modelos preditivos; B) Curva ROC para a pontuação de risco na população de validação.

A)



B)



Nota: A significância assintomática foi calculada para cada curva ROC em relação à hipótese nula e não para a diferença entre as curvas.



Figura 2 – Médias das Glicemias em jejum e após seguimento de 5 anos nos tercis do modelo 1 de pontuação. Barras de erro são IC95%. Valores de p calculados por ANOVA de um caminho.

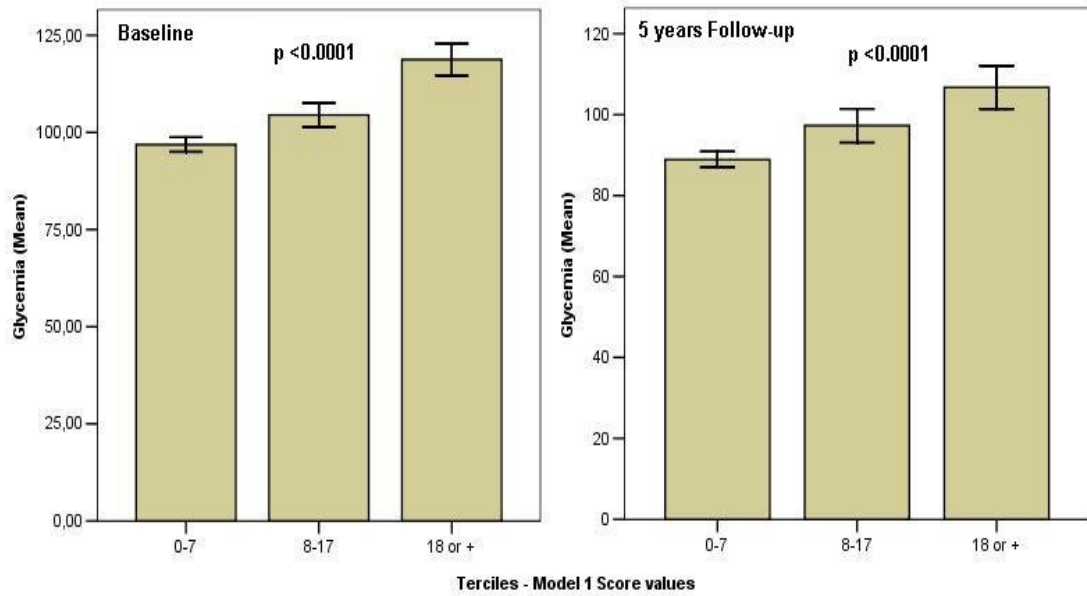
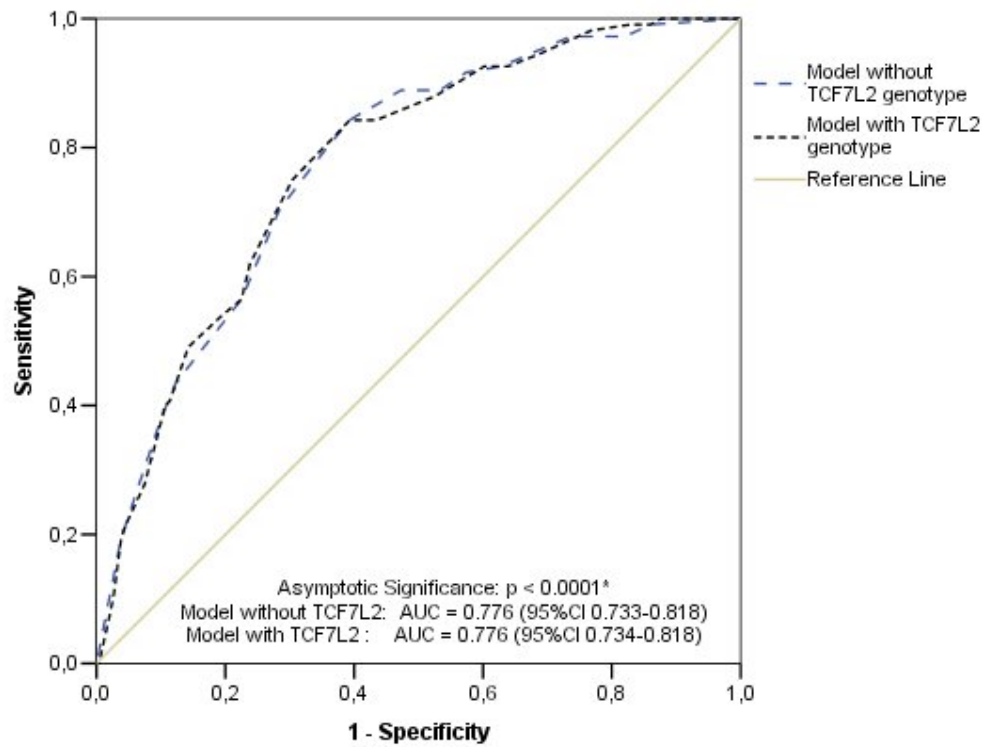


Figura 3 – Curvas ROC comparando modelo baseado em variáveis demográficas com o mesmo modelo acrescido do genótipo TT TCF7L2



(\*) Valor de p para ambos os modelos

## References

1. Centers for Disease Control and Prevention. National diabetes fact sheet: general information and national estimates on diabetes in the United States, 2007 [Internet]. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2008 [updated Jun 2008; cited Feb 2009]. Available from: <http://www.cdc.gov/Diabetes/pubs/factsheet07.htm/>.
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 Nov;15(11):1509-16.
3. Torquato MT, Montenegro Junior RM, Viana LA, de Souza RA, Lanna CM, Lucas JC, et al. Prevalence of diabetes mellitus and impaired glucose tolerance in the urban population aged 30-69 years in Ribeirao Preto (Sao Paulo), Brazil. *Sao Paulo Med J*. 2003 Nov 6;121(6):224-30.
4. American Diabetes Association. Economic Costs of Diabetes in the U.S. in 2007. *Diabetes Care*. 2008 March 1, 2008;31(3):596-615.
5. International Diabetes Federation. The human, social and economic impact of diabetes. [Internet]. Brussels, Belgium: International Diabetes Federation, 2006 [updated 2006; cited 2009]. Available from: <http://www.idf.org/home/index.cfm?node=41/>.
6. Rewers M, Hamman, R.F. Diabetes in America [Portable Document Format (pdf) on Internet]. 2nd ed. Bethesda, Md: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases (US); 1995. Chapter 9, Risk Factors for Non-Insulin-Dependent Diabetes; [cited Feb 2009]; p. 179-220. Available from: <http://diabetes.niddk.nih.gov/dm/pubs/America/pdf/chapter9.pdf>
7. Rewers M, Hamman RF. Diabetes in America [Portable Document Format (pdf) on Internet]. 2nd ed. Bethesda, Md: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases (US); 1995. Chapter 8, Risk Factors for Insulin-

Dependent Diabetes; [cited Feb 2009]; p. 165-178. Available from:  
<http://diabetes.niddk.nih.gov/dm/pubs/America/pdf/chapter8.pdf>

8. Tattersal RB, Fajans SS. Prevalence of diabetes and glucose intolerance in 199 offspring of thirty-seven conjugal diabetic parents. *Diabetes*. 1975 May;24(5):452-62.

9. Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J, et al. Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia*. 1992 Nov;35(11):1060-7.

10. Newman B, Selby JV, King MC, Slemenda C, Fabsitz R, Friedman GD. Concordance for type 2 (non-insulin-dependent) diabetes mellitus in male twins. *Diabetologia*. 1987 Oct;30(10):763-8.

11. Kenny SJ, Aubert RE, Geiss LS. Diabetes in America [Portable Document Format (pdf) on Internet]. 2nd ed. Bethesda, Md: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases (US); 1995. Chapter 4, Prevalence and Incidence of Non-Insulin-Dependent Diabetes; [Cited Feb 2009]; p. 47–68. Available from:  
<http://diabetes.niddk.nih.gov/dm/pubs/America/pdf/chapter4.pdf>

12. Pierce M, Keen H, Bradley C. Risk of diabetes in offspring of parents with non-insulin-dependent diabetes. *Diabet Med*. 1995 Jan;12(1):6-13.

13. Gimeno SG, Ferreira SR, Franco LJ, Hirai AT, Matsumura L, Moises RS. Prevalence and 7-year incidence of Type II diabetes mellitus in a Japanese-Brazilian population: an alarming public health problem. *Diabetologia*. 2002 Dec;45(12):1635-8.

14. Siqueira AF, Franco LJ, Gimeno SG, Matsumura LK, Abdalla DS, de Barros N, Jr., et al. Macrovascular disease in a Japanese-Brazilian population of high prevalence of metabolic syndrome: associations with classical and non-classical risk factors. *Atherosclerosis*. 2007 Nov;195(1):160-6.

15. Bertolino CN, Castro TG, Sartorelli DS, Ferreira SR, Cardoso MA. [Dietary trans fatty acid intake and serum lipid profile in Japanese-Brazilians in Bauru, Sao Paulo, Brazil]. *Cad Saude Publica*. 2006 Feb;22(2):357-64.

16. de Oliveira C, Pereira A, de Andrade M, Soler J, Krieger J. Heritability of cardiovascular risk factors in a Brazilian population: Baependi Heart Study. *BMC Medical Genetics*. 2008;9(1):32.
17. Schulz LO, Bennett PH, Ravussin E, Kidd JR, Kidd KK, Esparza J, et al. Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the U.S. *Diabetes Care*. 2006 Aug;29(8):1866-71.
18. Taylor A. The genetics of type 2 diabetes: A review. *International Journal of Diabetes and Metabolism*. 2006;14(2):76-81.
19. Koo BK, Cho, Y. M., Park, B. L., Cheong, H. S., Shin, H. D., Jang, H. C., Kim, S. Y., Lee, H. K., Park, K. S. . Polymorphisms of KCNJ11 (Kir6.2 gene) are associated with Type 2 diabetes and hypertension in the Korean population. *Diabetic Medicine*. 2007;24(2):178-86.
20. Sagen JV, Raeder H, Hathout E, Shehadeh N, Gudmundsson K, Baevre H, et al. Permanent Neonatal Diabetes due to Mutations in KCNJ11 Encoding Kir6.2: Patient Characteristics and Initial Response to Sulfonylurea Therapy. *Diabetes*. 2004 October 1, 2004;53(10):2713-8.
21. Chien AJ, Moon RT. WNTs and WNT receptors as therapeutic tools and targets in human disease processes. *Front Biosci*. 2007;12:448-57.
22. Duval A, Rolland S, Tubacher E, Bui H, Thomas G, Hamelin R. The human T-cell transcription factor-4 gene: structure, extensive characterization of alternative splicings, and mutational analysis in colorectal cancer cell lines. *Cancer Res*. 2000 Jul 15;60(14):3872-9.
23. Florez JC. The new type 2 diabetes gene TCF7L2. *Curr Opin Clin Nutr Metab Care*. 2007 Jul;10(4):391-6.
24. Welters HJ, Kulkarni RN. Wnt signaling: relevance to [beta]-cell biology and diabetes. *Trends in Endocrinology & Metabolism*. In Press, Corrected Proof.
25. Yi F, Brubaker PL, Jin T. TCF-4 mediates cell type-specific regulation of proglucagon gene expression by beta-catenin and glycogen synthase kinase-3beta. *J Biol Chem*. 2005 Jan 14;280(2):1457-64.

26. Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, Almgren P, et al. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *J Clin Invest*. 2007 Aug;117(8):2155-63.
27. Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, Weitgasser R, et al. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med*. 2007 Jul;85(7):777-82.
28. 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 Jul;90(1):583-612.
29. Bisi Molina Mdel C, Cunha Rde S, Herkenhoff LF, Mill JG. [Hypertension and salt intake in an urban population]. *Rev Saude Publica*. 2003 Dec;37(6):743-50.
30. 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 Apr;191(2):454-9.
31. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2008 Jan;31 Suppl 1:S55-60.
32. Estatística FIBdGe. Pesquisa por amostra de domicílios 2000: microdados. Rio de Janeiro; 2001.
33. Marquezine 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*. 2008 Sep 26;129(2):259-65.
34. 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 Aug;67(2):444-61.
35. 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 Aug;74(4):607-14.
36. Hueb W, Soares PR, Gersh BJ, Cesar LA, Luz PL, Puig LB, et al. 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*. 2004 May 19;43(10):1743-51.

37. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. 1985 Jul;28(7):412-9.

38. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet*. 2006 Mar;38(3):320-3.

39. Freathy RM, Weedon MN, Bennett A, Hypponen E, Relton CL, Knight B, et al. Type 2 diabetes TCF7L2 risk genotypes alter birth weight: a study of 24,053 individuals. *Am J Hum Genet*. 2007 Jun;80(6):1150-61.

40. Cauchi S, Meyre D, Choquet H, Deghmoun S, Durand E, Gaget S, et al. TCF7L2 rs7903146 variant does not associate with smallness for gestational age in the French population. *BMC Med Genet*. 2007;8:37.

41. Roth CL, Hinney A, Reinehr T, Schreiner F, Nguyen TT, Muller T, et al. TCF7L2 Polymorphism rs7903146 and Predisposition for Type 2 Diabetes Mellitus in Obese Children. *Horm Metab Res*. 2008 Jun 10.

42. Nauck MA, Meier JJ. The enteroinsular axis may mediate the diabetogenic effects of TCF7L2 polymorphisms. *Diabetologia*. 2007 Dec;50(12):2413-6.

43. Weedon MN, McCarthy MI, Hitman G, Walker M, Groves CJ, Zeggini E, et al. Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. *PLoS Med*. 2006 Oct;3(10):e374.

44. Cauchi S, Meyre D, Durand E, Proenca C, Marre M, Hadjadj S, et al. Post genome-wide association studies of novel genes associated with type 2 diabetes show gene-gene interaction and high predictive value. *PLoS ONE*. 2008;3(5):e2031.

45. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. 2007 Feb 22;445(7130):881-5.

46. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA*. 2003 Jan 1;289(1):76-9.
47. Balkau B, Lange C, Fezeu L, Jean T, de Lauzon-Guillain B, Czernichow S, et al. Predicting diabetes - clinical, biological and genetic approaches: the D.E.S.I.R. Study. *Diabetes Care*. 2008 Aug 8.
48. van Hoek M, Dehgan A, Witteman JC, van Duijn CM, Uitterlinden AG, Oostra BA, et al. Predicting type 2 diabetes based on polymorphisms from genome wide association studies: a population-based study. *Diabetes*. 2008 Aug 11.
49. Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, et al. Genotype Score in Addition to Common Risk Factors for Prediction of Type 2 Diabetes. *N Engl J Med*. 2008 Nov 20;359(21):2208-19.
50. Lyssenko V, Jonsson A, Almgren P, Pulizzi N, Isomaa B, Tuomi T, et al. Clinical Risk Factors, DNA Variants, and the Development of Type 2 Diabetes. *N Engl J Med*. 2008 Nov 20;359(21):2220-32.
51. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science*. 2007 Jun 1;316(5829):1341-5.
52. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*. 2008 May;40(5):638-45.
53. Meigs JB, Dupuis J, Herbert AG, Liu C, Wilson PW, Cupples LA. The insulin gene variable number tandem repeat and risk of type 2 diabetes in a population-based sample of families and unrelated men and women. *J Clin Endocrinol Metab*. 2005 Feb;90(2):1137-43.