

LESLIE ECKER FERREIRA

**ANÁLISE DE VARIANTES GENÉTICAS E INTERAÇÕES GÊNICAS QUANTO AO
RISCO DO AVC ISQUÊMICO**

**JOINVILLE
2018**

LESLIE ECKER FERREIRA

**ANÁLISE DE VARIANTES GENÉTICAS E INTERAÇÕES GÊNICAS QUANTO AO
RISCO DO AVC ISQUÊMICO**

Tese de doutorado apresentada como requisito parcial para obtenção de título de Doutor no Programa de Pós-Graduação em Saúde e Meio Ambiente, na Universidade da Região de Joinville. Orientador: Prof. Dr. Paulo Henrique Condeixa de França.
Co-orientador: Dr. Rodrigo Secolin.

**JOINVILLE
2018**

Catálogo na publicação pela Biblioteca Universitária da Univille

F383a Ferreira, Leslie Ecker
Análise de variantes genéticas e interações gênicas quanto ao risco do AVC isquêmico/ Leslie Ecker Ferreira; orientador Dr. Paulo Henrique Condeixa de França; coorientador Dr. Rodrigo Secolin. – Joinville: UNIVILLE, 2018.

93 f. : il. ; 30 cm

Tese (Doutorado em Saúde e Meio Ambiente – Universidade da Região de Joinville)

1. Acidente vascular cerebral – Pacientes – Joinville (SC). 2. Acidente vascular cerebral – Fatores de risco. 3 Fatores hereditários e ambientais. I. França, Paulo Henrique Condeixa de (orient.). II. Secolin, Rodrigo (coorient.). III. Título.

CDD 616.81

Elaborada por Rafaela Ghacham Desiderato – CRB-14/1437

Termo de Aprovação

“Análise de Variantes Genéticas e Interações Gênicas quanto ao Risco do AVC Isquêmico”

por

Leslie Ecker Ferreira

Tese julgada para a obtenção do título de Doutor em Saúde e Meio Ambiente, área de concentração Saúde e Meio Ambiente e aprovada em sua forma final pelo Programa de Pós-Graduação em Saúde e Meio Ambiente.



Prof. Dr. Paulo Henrique Condeixa de França
Orientador (UNIVILLE)

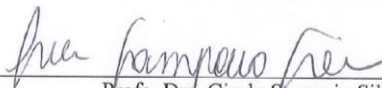


Prof. Dr. Paulo Henrique Condeixa de França
Coordenador do Programa de Pós-Graduação em Saúde e Meio Ambiente

Banca Examinadora:



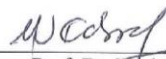
Prof. Dr. Paulo Henrique Condeixa de França
Orientador (UNIVILLE)



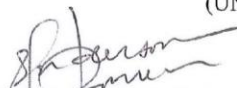
Profa. Dra. Gisele Sampaio Silva
(UNIFESP)



Profa. Dra. Cláudia Vianna Maurer Morelli
(UNICAMP)



Prof. Dr. Norberto Luiz Cabral
(UNIVILLE)



Prof. Dr. Anderson Roman Gonçalves
(UNIVILLE)

Joinville, 20 de março de 2018

AGRADECIMENTOS

A realização desta tese de doutorado contou com importantes apoios e incentivos sem os quais os desafios seriam ainda maiores. Meus profundos agradecimentos às pessoas que participaram, colaboraram, torceram e de alguma forma contribuíram para esse trabalho.

Ao Professor Doutor Paulo Henrique Condeixa de França, pela orientação, apoio, disponibilidade e incentivo ao longo desse trabalho.

Ao Professor Doutor Rodrigo Secolin, pela co-orientação, todo suporte dado desde o início do projeto, disponibilidade e grande colaboração.

À Professora Íscia Lopes-Cendes, pelo grande incentivo, disponibilidade, acesso ao seu Laboratório e principalmente ao apoio financeiro para a realização das genotipagens desse estudo.

Ao Professor Doutor Norberto Cabral, ao incentivo e a disponibilidade dos dados e amostras do Biobanco Univille.

À Professora Dra. Stèphanie Debette, pelo grande incentivo e a imensa disponibilidade para ensinar a fazer ciência.

À equipe do Laboratório de Biologia Molecular da Univille, ao longo desses anos passaram vários alunos e colegas que de alguma forma colaboraram com esse trabalho, meu muito obrigado.

À equipe JOINVASC, em especial a Vivian, que auxiliou nas coletas de dados e pela inestimável ajuda.

À equipe do Laboratório de Genética Molecular da Faculdade de Ciências Médicas da Unicamp, em especial a Amanda Donatti, que me deu o suporte necessário para a realização dos ensaios de PCR em tempo Real.

Ao VINTAGE Team, Vascular and neurological diseases: integrative and genetic epidemiology, do Institut National de la Santé et de la Recherche Médicale e da Université de Bordeaux, por todo suporte, sugestões e acolhimento.

Aos meus colegas de doutorado pelo coleguismo e apoio incondicional.

À toda minha família, meus pais, minhas irmãs e em especial o Manu, por toda a torcida, apoio, e principalmente por estar sempre presente nos momentos mais desafiadores desse trabalho.

**" - Acabou! Palavra tola! Acabou por quê?
Acabou e depois nada, a indiferença plena!
- De que serve o eterno criar,
se a criação nunca acabar? Acabou!"
(Fausto, Goethe)**

RESUMO

O Acidente Vascular Cerebral (AVC) é a doença neurológica mais comum na atualidade e situa-se entre as principais causas de óbito no mundo. O incremento do conhecimento sobre os mecanismos genéticos envolvidos no risco ao AVC, aliados à variabilidade genética populacional e aos fatores clínicos e ambientais associados à doença, são fundamentais à compreensão dos multifatores correlacionados à sua predisposição. Entretanto, resultados inconsistentes derivados de diferentes populações apontaram para a ocorrência de heterogeneidade genética para o AVC. Além disso, variantes genéticas encontradas associadas ao AVC parecem apresentar um efeito alélico fraco. Portanto, pode-se supor que um efeito de interação entre variantes genéticas poderia contribuir para a susceptibilidade ao AVC, em vez das variantes genéticas de modo independente. Neste sentido, objetivou-se investigar a interação entre variantes genéticas associadas à ocorrência do AVC Isquêmico (AVCI) subtipos aterotrombótico e cardioembólico. Foi realizado um estudo caso-controle com base em uma coorte pré-estabelecida em doenças cerebrovasculares, abrangendo 435 pacientes (195 casos aterotrombóticos e 240 cardioembólicos) e 535 indivíduos não acometidos, residentes em Joinville, Brasil. Oito variantes genéticas em genes candidatos (*HDAC9*, *CDKN2B-AS1*, *PITX2*, *ZFH3* e *PDE4D*) foram investigadas via Reação em Cadeia da Polimerase em Tempo Real com emprego de sondas TaqMan® alelo específicas. A análise de interação entre as variantes genéticas foi realizada utilizando-se o método *Multifactor Dimensionality Reduction* (MDR). A fim de avaliar se o conhecimento dos perfis genotípicos melhora a acurácia da predição ao AVCI, foram comparados cenários (modelos de predição) usando-se algoritmo baseado em *random forests*. Para tanto, também foram considerados os fatores de riscos controláveis (tabagismo, hipertensão, hipercolesterolemia e diabetes) e não controláveis (idade e sexo), além de dados clínicos (subtipos de AVCI classificados segundo os critérios do *Trial of Org 10172 in Acute Stroke Treatment* - TOAST e do *Causative Classification System for Ischemic Stroke* - CCS). Foi demonstrado que a variante genética rs2383207*A, próxima ao gene *CDKN2B-AS1*, está associada ao AVCI subtipo aterotrombótico [OR = 2,35 (95% CI 1,79 - 3,08); $p=4.66 \times 10^{-10}$]. Adicionalmente, observou-se um aumento no risco para AVCI aterotrombótico entre os portadores dos genótipos AG e AA num modelo de codominância [OR = 2,14 (95% CI 1,33 - 3,45); OR = 5,26 (95% CI 2,98 - 9,28); respectivamente]. Embora não tenha sido observada associação das variantes genéticas estudadas com AVCI subtipo cardioembólico, testes de interação gênica revelaram sinergismo entre as variantes rs152312 e rs966221, e entre as variantes rs966221 e rs2910829. Uma discreta interação foi observada entre as variantes rs152312 e rs2910829, rs152312 e rs6843082, e rs2910829 e rs6843082, indicando um possível efeito redundante ou antagônico. Adicionalmente, as variáveis mais importantes nos modelos de predição de AVCI foram hipercolesterolemia e hipertensão, atingindo-se acurácias de 86,47% e 90,47% para os subtipos aterotrombótico e cardioembólico, respectivamente, sem aumento significativo na acurácia quando incluídos os resultados genéticos. Estudos genéticos relacionados ao AVC são ainda escassos, especialmente na população brasileira, e a pesquisa realizada contribuiu para a ampliação do entendimento dos fatores envolvidos na predisposição à doença.

Palavras-chave: Acidente Vascular Cerebral Isquêmico; Predisposição Genética, Fatores de Risco, Estudo Caso-Controle.

ABSTRACT

Stroke is the commonest neurologic disease today and is among the leading causes of death worldwide. Increasing knowledge about the genetic mechanisms involved in stroke risk, together with the population genetic variability and the clinical and environmental factors associated with the disease, are fundamental to the understanding of multifactors correlated with its predisposition. However, inconsistent results from different populations pointed to the occurrence of genetic heterogeneity for stroke. In addition, genetic variants found associated with stroke appear to have a weak allelic effect. Therefore, we can hypothesize that an interaction effect among genetic variants could contribute to stroke susceptibility rather than to genetic variants independently. In this sense, the objective was to investigate the interaction between genetic variants associated with the occurrence of ischemic stroke (IS) of atherothrombotic (LAAS-IS) and cardioembolic (CE-IS) subtypes. A case-control study was carried out based on a pre-established cohort of cerebrovascular diseases, covering 435 patients (195 LAAS-IS and 240 CE-IS) and 535 non-affected individuals living in Joinville, Brazil. Eight genetic variants in candidate genes (*HDAC9*, *CDKN2B-AS1*, *PITX2*, *ZFH3*, and *PDE4D*) were investigated via Real-time Polymerase Chain Reaction using allele specific TaqMan® probes. The interaction analysis between the genetic variants was performed using the Multifactor Dimensionality Reduction (MDR) method. In order to evaluate whether the knowledge of the genotypic profiles improves the predictive accuracy of IS, scenarios (prediction models) were compared using algorithms based on random forests. For this, we also considered controllable risk factors (smoking, hypertension, hypercholesterolemia, and diabetes) and non-controllable risks (age and gender), as well as clinical data (IS subtypes were classified according to the criteria of the Trial of Org 10172 in Acute Stroke Treatment - TOAST and the Causative Classification System for Ischemic Stroke (CCS)). The genetic variant rs2383207*A, close to the *CDKN2B-AS1* gene, was shown to be associated with LAAS-IS [OR = 2.35 (95% CI 1.79 - 3.08); $p=4.66 \times 10^{-10}$]. In addition, an increase on LAAS-IS risk was observed among AG and AA genotypes carriers in a codominance model [OR = 2.14 (95% CI = 1.33 - 3.45); OR = 5.26 (95% CI = 2.98 - 9.28); respectively]. Although there was no association of the genetic variants studied with CE-IS, gene interaction tests revealed synergism between variants rs152312 and rs966221, and between rs966221 and rs2910829. A discrete interaction was observed between rs152312 and rs2910829, rs152312 and rs6843082, and rs2910829/rs6843082 indicating a possible antagonistic or redundant effect. Additionally, the most important variables of the prediction models for IS were hypercholesterolemia and hypertension, reaching accuracies of 86.47% and 90.47% for LAAS-IS and CE-IS, respectively, without significant increase in accuracy when genetic results are included. Genetic studies related to stroke are still scarce, especially in the Brazilian population, and the research has contributed to broadening the understanding of the factors involved in predisposition to the disease.

Key-Words: Stroke, Genetic Predisposition, Risk Factors, Case-Control Study.

LISTA DE SIGLAS E ABREVIATURAS

AmpC - *Cyclic Adenosine Monophosphate C* - Monofosfato Cíclico de Adenosina

ANP - *Atrial Natriuretic Peptide* - Peptídeo Natriurético Atrial

ATBF1 – *AT Motif Binding Fator 1* - Fator ligante do motivo AT

APC – *Activated Protein C* - Proteína C Ativada

AUC - *Area Under the Curve* – Área Sobre a Curva

AVC - Acidente Vascular Cerebral

AVCI - AVC Isquêmico

CDKI – *Cyclin-Dependent Kinase Inhibitor* - Inibidor proteico de Quinase Dependente de Ciclina

CDKN2B-AS1- *CDKN2B - Antisense RNA 1* – RNA 1 antisenso do Inibidor proteico de Quinase Dependente de Ciclina 2B

CEO – Acidente Vascular Cerebral Isquêmico subtipo Cardioembólico

FA - Fibrilação atrial

HDAC9 - *Histone Deacetylase 9* – Histona Deacetilase 9

IL6 – *Interleukin-6* - Interleucina 6

LAAS - *Large Artery Atherosclerosis Stroke* - AVC subtipo Aterosclerose de Grandes Vasos

LOOCV - *Leave-One-Out Cross-Validation* - Validação Cruzada usando método *Leave - One- Out*

MDA - *Mean Decrease Accuracy* - Acurácia Média de Decaimento

MDR – *Multifactor Dimensionality Reduction* - Redução de Dimensionalidade Multifatorial

mRNA - Ácido Ribonucleico Mensageiro

NIH - *National Institutes of Health* - Instituto de Saúde Nacional dos Estados Unidos da América

NPPA - *Natriuretic Peptide A* - Gene codificante para Peptídeo Natriurético A

PAI-1 - *Plasminogen Activator Inhibitor-1* - Inibidor de ativação do plasminogênio-1

PDE4D - *Phosphodiesterase 4D* - Fosfodiesterase 4D

PITX2 - *Paired Like Homeodomain 2*

PRC1 - *Protein Regulator of Cytokinesis -1* - Proteína Reguladora de Citocinese 1

PRC2 - *Protein Regulator of Cytokinesis - 2* - Proteína Reguladora de Citocinese 2

qPCR – Real Time - Polymerase Chain Reaction - Reação em Cadeia da Polimerase em Tempo Real

RNA – Ribonucleic Acid - Ácido Ribonucleico

ROC - Receiver Operating Characteristic

TAF – 1- Inibidor do Fator Tecidual 1

TGF- β – Transforming Growth Factor Beta - Fator de crescimento transformante Beta

TOAST - Trial of Org 10172 in Acute Stroke Treatment

TWIST1 - Proteína relacionada ao gene Twist 1

ZFH3 - Zinc Finger Homeobox 3 - Homeobox de Dedo de Zinco

WTCCC2 - Welcome Trust Case Control Consortium 2

LISTA DE FIGURAS

Figura 1 - Tipos de AVC.....	19
Figura 2 - Cascata de Coagulação Sanguínea.....	21
Figura 3 - Variantes Genéticas	27
Figura 4 - Distribuição das frequências alélicas das variantes genéticas do estudo em diversas populações.....	36

LISTA DE TABELAS

Tabela 1 - Fatores genéticos associados ao AVC.....	22
Tabela 2- Variantes Genéticas analisadas no estudo de acordo com o subtipo de AVCI.....	38

SUMÁRIO

1. INTRODUÇÃO	14
2. OBJETIVOS	16
2.1 Objetivo geral	16
2.2 Objetivos específicos	16
3. REVISÃO DE LITERATURA	17
3.1 Breve histórico	17
3.2 Conceito e classificação do Acidente Vascular Cerebral	18
3.3 Epidemiologia	22
3.4 Fatores de risco associados ao AVC	22
3.5 Fatores de riscos genéticos	23
3.6 Variantes Genéticas	27
3.7 Genes relacionados ao AVC Isquêmico	27
3.7.1 <i>HDAC9</i> - Histone Deacetylase 9	28
3.7.2 <i>CDKN2B-AS1</i> - Cyclin-Dependent Kinase Inhibitor 2B	29
3.7.3 <i>PITX2</i> - Paired Like Homeodomain 2	29
3.7.4 <i>ZFH3</i> - Zinc Finger Homeobox 3	30
3.7.5 <i>PDE4D</i> - <i>Phosphodiesterase 4D</i>	30
3.8 Variantes genéticas selecionadas para o estudo	32
3.8.1 Variante rs2107595	32
3.8.2 Variante rs2383207	33
3.8.3 Variante rs6843082	34
3.8.4 Variante rs879324	34
3.8.5 Variantes rs1396476, rs2910829, rs966221, rs152312	35
4. METODOLOGIA	37
4.1 Desenho do estudo e amostras	37
4.2 Critérios de inclusão e exclusão	37
4.3 Fatores clínicos e ambientais	38
4.4 Fatores genéticos	38
4.5 Análise de dados	41
4.6 Análise de interação gênica	41
4.7 Análise de predição	41
	12

5. INTERDISCIPLINARIEDADE	43
6. RESULTADOS E DISCUSSÃO	44
7. CONCLUSÃO	77
8. REFERÊNCIAS	78
ANEXO A – PARECER CONSUBSTANCIADO DO COMITÊ DE ÉTICA	89

1. INTRODUÇÃO

Acidente Vascular Cerebral (AVC) corresponde a segunda maior causa de óbitos no mundo (AHA,2018). Os fatores de risco reconhecidamente envolvidos com o desenvolvimento da doença englobam fatores ambientais, como hábitos alimentares, poluição, tabagismo, alcoolismo, sedentarismo e obesidade, e fatores clínicos, como hipertensão, fibrilação atrial, aterosclerose, problemas circulatórios, diabetes e altas taxas de colesterol. Adicionalmente, avanços recentes na área de biologia molecular identificaram algumas variantes genéticas como fatores de risco envolvidos no desenvolvimento da doença (LIU et al., 2017; MALIK et al., 2016; NINDS, 2016; TRAYLOR et al., 2016). Entretanto, ainda é incipiente o conhecimento acerca das interações gênicas correlacionadas à predisposição ao AVC (YIX et al., 2017).

A pesquisa genética tem potencial para modificar o atendimento clínico dos pacientes com AVC nos próximos anos por meio de contribuições para a medicina personalizada. Isto abrange não só a possibilidade de estimar o risco individual de doença (primária ou recorrente), mas também a capacidade de adaptar todos os procedimentos envolvidos no atendimento ao paciente, incluindo prevenção, diagnóstico, prognóstico, terapêutica e estratégias de reabilitação (FALCONE et al., 2014).

No ano de 2005 foi constituído um banco de dados de registro de casos de todos os eventos de AVC ocorridos na cidade de Joinville (JOINVASC), permitindo dessa forma avaliar as taxas de incidência, letalidade e mortalidade do AVC na cidade (CABRAL, et al., 2016). Em 12 de junho de 2013, o JOINVASC passou a ser amparado através da Lei Municipal Nº 7448. Já no ano de 2014, a Universidade da Região de Joinville – UNIVILLE implementou o primeiro Biobanco de amostras de DNA e informações associadas de pacientes acometidos por AVC do Brasil (Registro CONEP B-018; Parecer CONEP BIOBANCO 005/2014, de 14 de abril de 2014). O Biobanco UNIVILLE foi composto, até outubro de 2016, de 4635 amostras, distribuídas entre casos (1834) e controles (2801), associadas a dados clínicos e epidemiológicos dos pacientes que sofreram AVC e indivíduos não acometidos na cidade de Joinville/SC (FERREIRA *et al.*, 2017).

Dessa forma, o trabalho aqui apresentado pretende contribuir para o entendimento dos fatores de risco ao AVC Isquêmico (AVCI) através da identificação

de variantes genéticas que podem estar associadas entre si e a fatores clínicos e ambientais na população da coorte JOINVASC.

2. OBJETIVOS

2.1 Objetivo geral

Investigar variantes em genes candidatos associados ao Acidente Vascular Cerebral Isquêmico (AVCI) quanto a possíveis interações entre variantes genéticas associadas à predisposição.

2.2 Objetivos específicos

- Investigar a correlação de fatores ambientais e clínicos com a ocorrência de AVCI.
- Verificar a distribuição das frequências genotípicas das variantes em genes candidatos entre controles e pacientes segundo os subtipos de AVCI associados.
- Analisar a interação entre variantes genéticas de genes candidatos em pacientes acometidos conforme os subtipos de AVCI.
- Avaliar o valor preditivo das variantes genéticas investigadas para a ocorrência de AVCI por meio de análises de predição gênica computacional.

3. REVISÃO DE LITERATURA

3.1 Breve histórico

Descrito pela primeira vez por Hipócrates (460 A.C.), como a deprivação repentina da vida, o termo apoplexia foi usado ao longo de muitos anos para descrever o AVC. A terminologia remete a alguém sem sentido – tanto em relação às faculdades sensoriais como também à razão – caracterizada por olhos trêmulos, grandes, vagos e escuros (RODOLPHO, 2015). Hemorragia foi a explicação dada à apoplexia por Robinson em 1732, uma teoria bem aceita na época e que foi fortalecida uma vez que a circulação sanguínea já havia sido descrita por Harvey (1628) estabelecendo uma relação entre hemorragia e perda de pressão sanguínea. Apenas no início do século XIX, o termo apoplexia estava bem conceituado e concretizado que a hemorragia era a causa da alteração do balanço humoral e alteração na pressão sanguínea. Impulsionado pela descoberta da irrigação sanguínea cerebral emergiram questionamentos acerca da conduta médica a ser tomada nessas circunstâncias. Começou-se, então, a tomar medidas de pressão arterial de forma mais frequente e, em consequência, observou-se o rápido aumento de pressão nas hemorragias intracerebrais. Em 1892, Osler recomendou a venosecção após apoplexia para diminuição da pressão arterial, procedimento que mais tarde viria a ser extinto (apud POUND et al., 1997).

Com a popularização da dissecação anatômica e a classificação das doenças conforme relatos de sintomas, doenças começaram a ser explicadas por meio da anormalidade encontrada dentro do corpo *post-mortem*. Dessa forma, a apoplexia teria dois mecanismos distintos: a obstrução sanguínea e a hemorragia, ambas causadas pela degeneração da parede das artérias. No final do século XIX, o termo doença cerebrovascular passou a ser usado e apoplexia caiu em desuso (LAWRENCE et al., 1994).

Em meados da década de 50, novas técnicas e terapias foram desenvolvidas para explorar os mecanismos da patogênese das doenças cerebrovasculares, como as primeiras angiografias, cirurgias cerebrovasculares e o uso de anticoagulantes (DAVIDSON et al., 1952).

Os números absolutos de pessoas com AVC na Inglaterra aumentaram consideravelmente após a II Guerra Mundial, de aproximadamente 20.000 mortes

por ano (1930) para 40.000 (1950) e 80.000 (1980) casos. A explicação para este aumento não é clara, mas provavelmente pode estar associada ao aumento demográfico e a uma maior exposição aos fatores de risco do AVC (CHARLTON et al., 1996).

Finalmente, em 1962, o livro intitulado *Modern Views on 'Stroke' Illness*, produzido pela *Chest and Heart Association*, recomendou o diagnóstico e tratamento da doença por equipes multidisciplinares formadas por enfermeiros, fisioterapeutas domiciliares, fonoaudiólogos, terapeutas ocupacionais e outros profissionais de saúde (CHEST AND HEART ASSOCIATION, 1962). O termo Doença Cerebrovascular foi alterado para Acidente Cerebrovascular e permanece até os dias atuais (POUND et al.,1997).

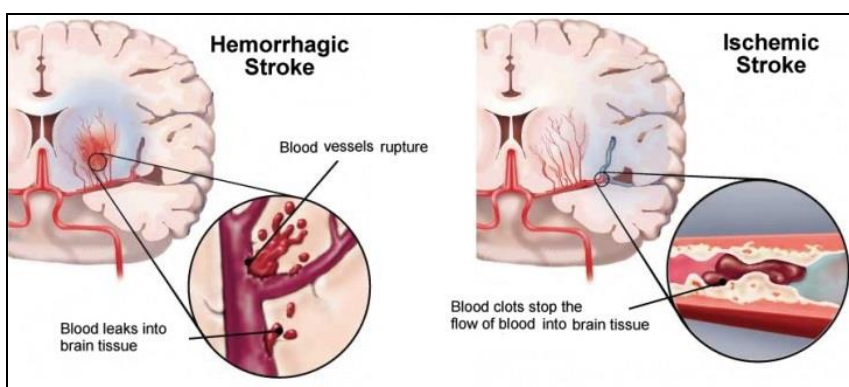
3.2 Conceito e classificação do Acidente Vascular Cerebral

Um complexo conjunto de sintomas de deficiência neurológica focal resultantes de lesões cerebrais provocadas por alterações da irrigação sanguínea define o AVC. Por sua vez, estas alterações se devem a dois mecanismos patológicos que comprometem a função cerebral: oclusão de vasos sanguíneos (originando isquemia e/ou enfarte de tecidos) ou ruptura vascular. Portanto, o AVC pode ser classificado como isquêmico ou hemorrágico (Figura 1), de acordo com mecanismo que o originou (CORREIA, 2011).

O AVCI é causado pela obstrução do fluxo sanguíneo para uma ou mais áreas do cérebro e ocorre em 87% dos casos de AVC. O AVCI pode ser classificado em subtipos por dois métodos, pelo *Trial of Org 10172 in Acute Stroke Treatment* (TOAST) e pelo *Causative Classification of Stroke* (CCS). Conforme os critérios estabelecidos pelo TOAST, o AVCI pode ser subclassificado como aterotrombótico, lacunar, cardioembólico, de outras causas ou criptogênico. O subtipo aterotrombótico envolve estenose ou oclusão significativa (>50%) de uma grande artéria cerebral ou ramo da artéria cortical; o subtipo lacunar corresponde à oclusão de pequenas artérias, ocasionando infarto lacunar; o subtipo cardioembólico é resultante de uma oclusão ocasionada por uma embolia cardíaca; o subtipo criptogênico engloba casos de etiologia não determinada pelos métodos diagnósticos atuais; e o subtipo "outras causas" reúne pacientes que sofreram AVC determinado por causas raras, como vasculopatias não ateroscleróticas,

hipercoagulação ou deficiências hematológicas (ADAMS et al., 1993) (Figura 1). De acordo com o sistema de classificação CCS, os subtipos de AVCI definidos pelo TOAST são subclassificados em três categorias: evidente, provável e possível baseado no diagnóstico por imagem e histórico do paciente (AY, 2010).

Figura 1 - Tipos de AVC (Disponível em: <https://com-neurology-stroke.sites.medinfo.ufl.edu/for-patients/types-of-stroke/>)



O processo isquêmico é caracterizado pela insuficiência de irrigação sanguínea, impedindo o tecido atingido de liberar dióxido de carbono e receber nutrientes e oxigênio. Quando o tecido cerebral é privado de sangue arterial por longos períodos, podem instalar-se lesões graves e irreversíveis no cérebro. A isquemia deve-se ao estreitamento das artérias provocado por aterosclerose ou deposição de gordura; o estreitamento demasiado das artérias, por sua vez, leva a formação de coágulo, que é mediado pelo sistema hemostático (CORREIA, 2011).

O bloqueio de vasos e artérias cerebrais dispara uma cascata de eventos patológicos denominados isquêmico-induzidos, incluindo desequilíbrio da homeostase de íons, excitotoxicidade neuronal e aumento de cálcio intracelular, despolarização peri-infarto, geração de radicais livres, peroxidação lipídica e alteração da síntese proteica que, atualmente, é considerada como uma lesão neuronal irreversível (CORREIA, 2011).

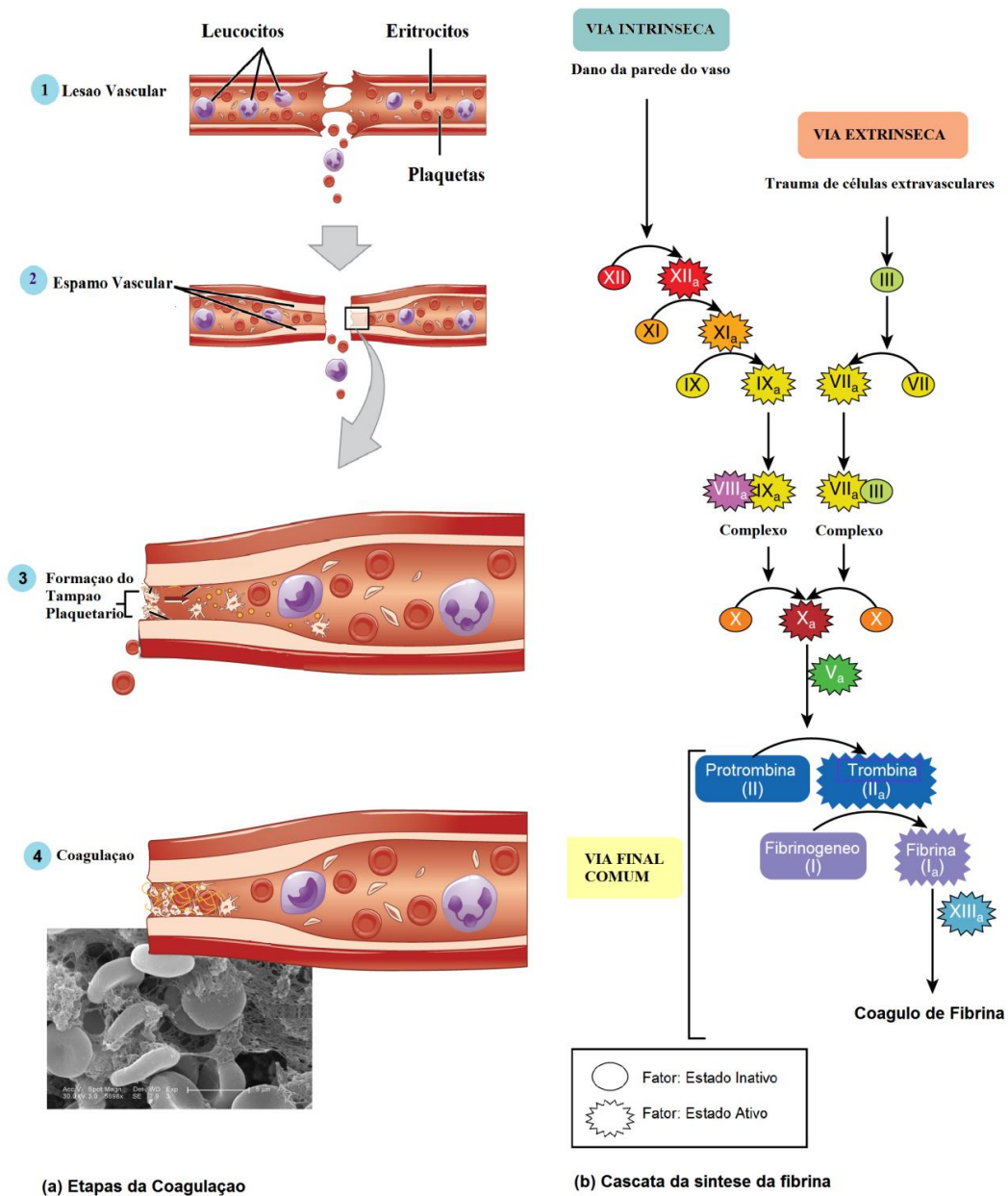
O AVC hemorrágico (AVCH) é reconhecido pela ruptura de vasos sanguíneos cerebrais, e em consequência há extravasamento de sangue e interrupção do suprimento de oxigênio para a área afetada. Os AVCs hemorrágicos são

classificados de acordo com o local que originou a hemorragia. Quando o sangramento dos vasos ocorre internamente à pia-máter do cérebro, é denominada hemorragia intracerebral, enquanto o AVC hemorrágico subaracnoide ocorre entre a pia-máter e a aracnoide. Hemorragias subdurais são definidas acima da região aracnoide, porém dentro da dura-máter, e ainda quando o sangramento ocorre acima da dura-máter, porém internamente à caixa craniana, é caracterizado por hemorragia epidural (CAPLAN, 2006).

O processo hemostático consiste num equilíbrio organizado dos fatores protrombóticos e antitrombóticos na corrente sanguínea (Figura 2). Os mecanismos protrombóticos são responsáveis pela ativação da coagulação e adesão/agregação plaquetária. Este mecanismo é regulado por vários anti-coagulantes naturais (mecanismos antitrombóticos), como a antitrombina, a proteína C ativada (APC), a proteína S, o inibidor do fator tecidual (TAF-1), as células endoteliais e o sistema fibrinolítico. No caso do sistema fibrinolítico e das células endoteliais, a ação anticoagulante é caracterizada pela degradação do coágulo de fibrina (CORREIA, 2011).

Após lesão vascular, o sistema hemostático é ativado dando início à cascata da coagulação. A cascata da coagulação, por sua vez, pode ser iniciada por duas vias, a intrínseca e a extrínseca. A ativação do sistema intrínseco requer os fatores VIII, IX, XI, XIII de coagulação. Em contraste, a via extrínseca é disparada pela ativação do fator VII e do fator tecidual. Ambas as vias convergem num sistema comum envolvendo os fatores X, V e II (protrombina) de coagulação. Como resultado destes processos, forma-se a trombina que leva a conversão do fibrinogênio em fibrina. Simultaneamente, ocorrem a adesão das plaquetas, com exposição das proteínas da matriz extracelular, e a ativação dessas pela trombina. Finalmente, após agregação de ambos os sistemas se forma um trombo de plaquetas e de fibrina. Por outro lado, o processo de fibrinólise é responsável pela degradação da fibrina e remoção do coágulo. Durante esta fase, o plasminogênio é convertido em plasmina, molécula responsável pela degradação da fibrina e proteínas da matriz extracelular. O sistema fibrinolítico é regulado pelo inibidor de ativação do plasminogênio-1 (PAI-1) e 2-antiplasmina (PEZZINI, 2012).

Figura 2 - Cascata de Coagulação Sanguínea – (a) O processo de coagulação é disparado por quatro etapas: (1) Lesão Vascular; (2) Espasmo Vascular; (3) Formação do Tampão Plaquetário e (4) Coagulação - (b) A cascata da coagulação é iniciada por duas vias, a intrínseca e a extrínseca, que requerem fatores de coagulação específicos e ao fim convergem em uma única via (adaptado de OpenStax College,2013).



3.3 Epidemiologia

Anualmente, 17 milhões de pessoas sofrem AVC, sendo 15 milhões do tipo isquêmico. Cinco milhões vão a óbito e outras cinco milhões perdem suas aptidões naturais. São estimados 750.000 novos casos anuais de AVC nos EUA, resultando entre 150.000 a 200.000 mortes (AHA, 2018, 2008). Segundo a última estatística oficial do Ministério da Saúde (DATASUS, 2011), O AVC foi responsável por 8,9% de todos os óbitos ocorridos no país em 2004 e representa as maiores taxas de prevalência, morbidade e mortalidade. Em acréscimo, este cenário não deve se modificar até a próxima década.

Ademais, pacientes com debilitação irreversível atingem uma taxa maior que 15% (MANSO & VICENTE, 2013), envolvendo sequelas como comprometimento motor homolateral, afasia e dificuldades leves de articulação da fala, problemas visuais e auditivos, além da alteração na qualidade de vida dos pacientes com risco de desenvolvimento de depressão (SILVA et al., 2015). No Brasil, os custos de um paciente com AVC derivados de internações hospitalares podem representar até US\$ 1900,00 (CHRISTENSEN, et al 2009).

3.4 Fatores de risco associados ao AVC

Os fatores de risco atribuídos ao AVC podem ser nomeados como Fatores de Riscos Controláveis, Fatores de Riscos Não Controláveis e Fatores de Risco Novos. Ao menos um terço de todas as doenças cardiovasculares e AVC são atribuídas a cinco fatores de riscos considerados controláveis: tabagismo, etilismo, hipertensão, dislipidemia e obesidade. Outros fatores, como inatividade física, diabetes mellitus, baixo *status* econômico, depressão, *stress* psicossocial (*stress* crônico, isolamento social e ansiedade), e hipertrofia do ventrículo esquerdo, também são considerados riscos controláveis, entretanto, com menor impacto. Nos países com menor desenvolvimento social, além dos fatores apontados acima, a subnutrição também está associada ao risco dessas doenças. Idade avançada, histórico familiar, gênero e etnia são considerados fatores de riscos não controláveis. Entre os novos fatores de riscos incluem-se excesso de homocisteína no sangue, processos inflamatórios e coagulação sanguínea anormal (WHO, 2004). Desse modo, o controle dos fatores

de risco pode implicar diretamente na redução da incidência e na recidiva do AVCI (PIRES, GAGLIARDI & GORZONI, 2004).

O AVCI tem como principal fator de risco a hipertensão presente, aproximadamente, em 50% dos casos, configurando um risco relativo de seis vezes para o desenvolvimento de uma doença cerebrovascular. Entretanto, devido a hipertensão ser um fator de risco modificável, caso haja tratamento regular, os riscos do desenvolvimento do evento tornam-se passíveis de redução (PIRES, GAGLIARDI & GORZONI, 2004).

As cardiopatias e a fibrilação atrial (FA) crônica, também representam importantes fatores de risco para o AVC (PIRES, GAGLIARDI & GORZONI, 2004). A FA pode aumentar o risco de ocorrência da doença em cinco vezes; entretanto, uma vez realizada a intervenção com anticoagulantes pode-se obter uma redução de dois terços desse risco (HUGHES & LIP, 2008). Outros fatores, como diabetes mellitus e o acúmulo de gordura, contribuem para o aceleração do processo aterosclerótico devido a diminuição do calibre das artérias, o que poderá promover alterações no fluxo sanguíneo (PIRES, GAGLIARDI & GORZONI, 2004).

O impacto dos fatores de risco, tanto modificáveis, quanto não modificáveis difere entre os subtipos de AVC (OHIRA *et al.*, 2006). Cita-se, como exemplo, o diabetes e o tabagismo que são fatores preditivos importantes para o subtipo lacunar (O'DONNELL *et al.*, 2010).

3.5 Fatores de riscos genéticos

A genética ocupa um importante papel dentre os fatores de riscos ao AVC (Tabela 1), por meio de quatro mecanismos moleculares distintos. Primeiramente, cita-se as doenças monogênicas raras, nas quais um único gene específico é atribuído à doença. Muitas vezes trata-se de uma síndrome familiar na qual o AVC é a primeira manifestação (Ex: arteriopatia cerebral autossômica dominante e leucoencefalopatia - CADASIL). Após, citam-se as doenças atribuídas a um único gene, nas quais pode ocorrer AVC.(Ex: anemia falciforme). Em seguida, as associações genéticas contribuintes para a manifestação de fatores de riscos que podem levar ao AVC, como diabetes, hipertensão e FA. E, finalmente, variantes genéticas podem contribuir de forma modesta para o aumento do risco ao AVC, em

um modelo multifatorial da doença, na qual a associação de variantes genéticas pode explicar a heterogeneidade do AVC (BOEHME et al., 2017; CHAUHAN & DEBETTE, 2016).

Tabela 1. Fatores genéticos associados ao AVC

Doença	Modo de herança	Gene / Proteína	Mecanismo do AVC	Manifestações Clínicas Comuns
Doenças ligadas a genes únicos como causa primária do AVC				
Arteriopatia cerebral autossômica dominante com infartos subcorticais e leucoencefalopatia (CADASIL)	Autossômica dominante	<i>NOTCH3</i> / NOTCH3	Lacunar	AVC Isquêmico, leucoencefalopatia, miogranearia, manifestações psiquiátricas e demência
Arteriopatia cerebral autossômica recessiva com infartos subcorticais e leucoencefalopatia (CARASIL)	Autossômica recessiva	<i>HTRA1</i> / HtrA serinapeptidase-1	Lacunar	AVC Isquêmico, leucoencefalopatia, calvície prematura e espondilose
Angiopatia amiloide familiar	Autossômica dominante	<i>APP</i> / β - precursor de proteína amiloide	Ruptura de pequenos vasos corticais	Hemorragia lobar, “microbleeds”, leucoencefalopatia demência e formações amiloides
Doenças associadas à mutações de colágeno 4 (<i>COL4A1</i>)	Autossômica dominante	<i>Col4A1</i> / α 1-colágeno em cadeia tipo- 4	Ruptura de vasos corticais e subcorticais	Hemorragias superficiais e profundas, aneurisma intracranial, hematúria e doença cística renal
Doenças genéticas que incluem AVC como manifestação clínica				
Ehlers–Danlos tipo 4	Autossômica Dominante	<i>Col3A1</i> / Pró-colágeno tipo III	Dissecção arterial	Ruptura aórtica, renal, esplênica, ilíaca; Aneurismas e pseudoaneurismas; Ruptura visceral e muscular; ruptura uterina durante a gravidez
Doença de Fabry	Ligada ao Cromossomo X	<i>GAL</i> / α -galactosidase	AVCI-LASS; AVCI- CEO	AVC Isquêmico e vasculopatias, angioqueratomas, opacidade da córnea e cataratas, neuropatias e doenças renais
Síndrome de Marfan	Autossômica dominante	<i>FBN1</i> / fibrilina 1	Dissecção arterial; AVCI- CEO	AVC Isquêmico, dissecção arterial, escoliose, <i>pectus excavatum</i> , dilatação aórtica, disfunção valvular / insuficiência cardíaca e <i>ectopia lentis</i>

Encefalopatia mitocondrial com deficiência de ácido láctico e episódios de AVC	Maternal	DNA Mitocondrial (<i>MTTL1</i>) / codificado mitocondrialmente em tRNA leucina 1(UUA/G)	Deficiência energética e AVC metabólico	AVC Isquêmico que não é observado com limites vasculares, baixa estatura, desenvolvimento tardio, pouca visão, miopatia e <i>diabetes mellitus</i>
Anemia Falciforme	Autossômica recessiva	<i>HBB</i> / β -globina (subunidade-hemoglobina)	AVCI-LASS; AVCI- CEO Síndrome Moya-Moya	AVC Isquêmico, crises de dor, crises vasculares e infecções bacterianas
Síndrome da Deficiência Muscular Lisa Multissistêmica.	Autossômica dominante	<i>ACTA2</i> / α -actina	Síndrome Moya-Moya	AVC isquêmico, doença arterial coronária, aneurisma aórtico torácico e Síndrome Moya-Moya
Variantes Genéticas Comuns				
<i>TSPAN2</i>	Variante comum	<i>TSPAN2</i> / tetraspanina-2	Desenvolvimento vascular e aterosclerose	AVCI-LASS
<i>FOXF2</i>	Variante comum	<i>FOXF2</i> / Fator de transcrição <i>forkhead</i>	AVCI- CEO	Anomalias de músculo liso e cobertura de vasos cerebrais. Todos os AVCs, AVC lacunar e doença prematura e extensa de massa branca
<i>ABO</i>	Variantes comuns e raras	<i>ABO</i> / Proteínas de grupo sanguíneo	Trombose	Trombose e AVCI
<i>HDAC9</i>	Variantes comuns e raras	<i>HDAC9</i> / <i>Deacetilases de histonas</i>	Aterosclerose	AVCI-LASS
<i>PITX2</i>	Variantes comuns e raras	<i>PITX2</i> / <i>Pituitary Homeobox Transcript Variant 2</i>	Desenvolvimento do Nó-Sinoatrial e regulação de canais iônicos, modulação da condução cardíaca e fibrilação atrial	AVCI- CEO e FA
<i>ZFHX3</i>	Variantes comuns e raras	<i>ZFHX3</i> / <i>Zinc Finger Homeobox</i>	FA	AVCI- CEO e FA

Fonte: Traduzido de Boehme et al. (2017).

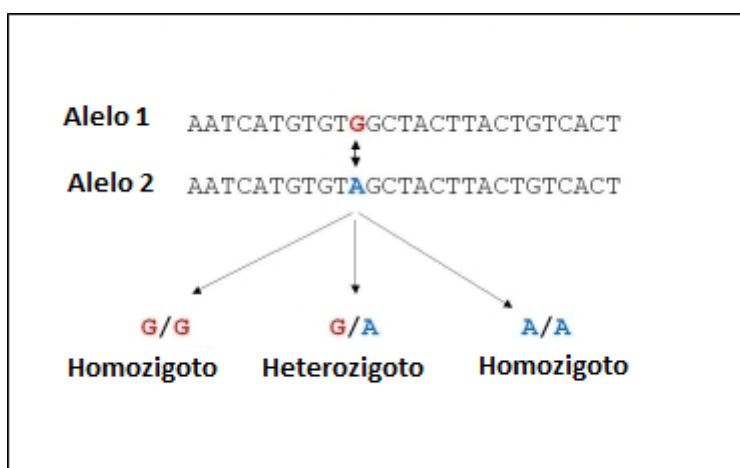
3.6 Variantes Genéticas

Variantes genéticas ocorrem com uma frequência maior que 1% na população. Sua vez, remetem para uma alteração de substituição de um nucleotídeo (Adenina, Timina, Citosina ou Guanina) (Figura 3) e são encontradas a cada 0,3-1 kilobase no genoma humano. Numerosas Variantes Genéticas não apresentam efeitos na função celular, mas alguns, principalmente aqueles que afetam a região funcional ou a regulação de genes, podem predispor indivíduos a doenças ou influenciar o prognóstico e a resposta perante tratamentos (BROWN, 2002).

Variantes Genéticas também são amplamente exploradas como marcadores moleculares em estudos de associação genética de doenças; no entanto, ainda há poucos estudos que apontam para uma correlação direta com o risco ao AVCI. Entretanto, a investigação de associações positivas entre genes e AVCI são fundamentais para validar o quanto esses efeitos são consistentes a riscos atribuídos (CHAUHAN &, DEBETTE, 2016).

Figura 3 – Exemplo de Variante Genética

Fonte: <http://www.traitgenetics.com> (Acesso em 21 de agosto de 2017).



3.7 Genes relacionados ao AVC Isquêmico

Os avanços na área da genética têm proporcionado variados estudos correlacionando genes ao risco do AVCI (TRAYLOR et al., 2012; RANNIKMÄE et al., 2017; MALIK et al., 2017; LINDGREN & MAGUIRE, 2016; RUTTEN-JACOBS et al., 2016; NINDS, 2016). No entanto, ainda são incipientes os estudos demonstrando a interação entre variantes genéticas e a influências destas, isoladamente, em cada subtipo de AVCI.

Por esse motivo, pretendeu-se nesse trabalho avaliar variantes genéticas em subtipos de AVCI (aterotrombótico e cardioembólico) bem caracterizados e as possíveis interações gênicas entre elas. Tais variantes localizam-se próximas ou nos próprios genes que serão descritos a seguir.

3.7.1 HDAC9 - Histone Deacetylase 9

Histonas exercem um papel essencial na deacetilação, regulação transcricional, progressão do ciclo celular e eventos do desenvolvimento. A acetilação de histonas altera a estrutura cromossomal e afeta fatores de transcrição que acessam o DNA. A proteína codificada pelo gene *HDAC9* possui uma sequência proteica homóloga a membros da família de histonas desacetilases. Ademais, mecanismos vasculares e neuronais têm sido propostos para explicar o aumento de risco de AVC associado ao gene *HDAC9* (BELLENGUEZ et al., 2012; HAN et al., 2013).

A proteína codificada por *HDAC9* inibe a miogênese e está envolvida no desenvolvimento do coração, mas não no mecanismo específico que pode levar à predisposição do processo patológico em grandes artérias sistêmicas. (MARTIN et al., 2009).

O gene *HDAC9* está localizado no cromossomo 7p21-p15 (ZHOU, et al, 2001) e muitos estudos associam o *locus* ao risco do AVC. Adicionalmente, pode levar ao aumento de risco por alteração da resposta isquêmica cerebral. Além disso, tem efeitos na sobrevivência neuronal, entretanto, a forma como pode levar à predisposição ao risco do AVC não está clara ainda. Inibidores de *histonas deacetilases* também têm sido sugeridos para o tratamento de AVCI (BELLENGUEZ et al., 2012; ISGC, 2012; ZHOU et al, 2017).

Análises por RT-PCR do *HDAC9*, detectou sua expressão gênica preferencialmente em tecidos do cérebro, coração e pâncreas. Entretanto,

isoformas da proteína codificada apresentam expressão tecidual variável (ZHOU et al., 2001).

3.7.2 *CDKN2B-AS1*- Cyclin-Dependent Kinase Inhibitor 2B

O gene *CDKN2B-AS1* está localizado no conjunto de genes *CDKN2B/CDKN2A* na região cromossomal 9p21, o qual está fortemente relacionado à suscetibilidade genética de doenças cardiovasculares e, também a uma série de tipos de câncer, aneurisma intracranial, diabetes tipo 2, periodontite, doença de Alzheimer, endometriose e glaucoma. O gene *CDKN2B-AS1* interage com os complexos repressivos *PRC1* e *PRC2* e sua função pode ser regulada por um fator epigenético transcricional repressor (ZHU et al., 2015; HUANG et al., 2015).

Destaca-se também que o gene *CDKN2B-AS1* tem uma associação indireta com o AVC, uma vez que esse gene contribui para o desenvolvimento de HAS (BAYOGLU et al., 2016). Esse gene codifica a proteína CDKI, que está envolvida na regulação do ciclo celular, no processo de senescência e apoptose celular, sendo que desregulações nessas funções podem contribuir para aterosclerose (BAYOGLU et al., 2016).

Variante próxima ao gene *CDKN2B-AS1* têm sido extensivamente identificadas como preditores de doenças cardiovasculares, cerebrovasculares e tumores cerebrais (HUANG et al., 2014; KREMER, 2015). Um estudo realizado com indivíduos turcos revelou a associação entre variantes do gene *CDKN2B-AS1* com o AVC, sendo essa explicada pelo fato desse gene estar associado à aterosclerose ou por ser um fator ativo no desenvolvimento cardiovascular (BAYOGLU et al., 2016).

3.7.3 *PITX2*- Paired Like Homeodomain 2

O gene *PITX2* codifica um fator de transcrição *homeobox* bicoidal RIEG/PITX. A proteína codificada atua como um fator de transcrição e regula a expressão do gene codificante para procolagênio lisil-hidroxilase. Esta proteína tem papel na diferenciação terminal de fenótipos celulares somatotróficos e lactotróficos, está envolvida também no desenvolvimento do olho, dentes e

órgãos abdominais, além de funcionar como um regulador da transcrição envolvido na atividade basal do hormônio regulado da prolactina (WNUK et al., 2011).

3.7.4 ZFH3 - Zinc Finger Homeobox 3

Este gene codifica um fator de transcrição com múltiplos *homeodomínios* e motivos do tipo *zinc-finger* e regula a diferenciação miogênica e neuronal. A proteína codificada suprime a expressão do gene da alfa-fetoproteína por ligação a um motivo intensificador, rico em adenina e timina. A proteína também tem demonstrado regulação negativa ao inibidor do ciclo celular dependente de ciclina-1A (também conhecido como p21Cip1). O gene *ZFH3* funciona como um supressor do tumor em vários tipos de câncer e variantes deste gene estão associadas com a fibrilação atrial (MARTIN et al., 2014).

Chou et al. (2013) analisaram 74 variantes genéticas candidatas a predisposição a AVCI e arteriosclerose. Dentre os achados, variantes no gene *ZFH3* demonstraram uma associação significativa à predisposição da aterosclerose, um fator de risco ao AVCI.

A interação desses dois genes, *PITX2* e *ZFH3*, que são marcadores de doenças coronarianas, parece também estar ligada à ocorrência de fibrilação atrial, que pode ter por consequência um AVC do tipo cardioembólico (TRAYLOR et al., 2012).

3.7.5 PDE4D - Phosphodiesterase 4D

A proteína codificada pelo gene *PDE4D* tem atividade 3'-5' fosfodiesterase e degrada o AmpC, que atua como uma molécula de transdução de sinal em vários tipos de células. A família PDE4 é codificada por 4 genes conhecidos como *PDE4A* (1:19p13.2), *PDE4B* (1p31), *PDE4C* (19p13.1) e *PDE4D* (5p12) que através de *splicing* alternativo originam mais de 25 formas distintas de PDE4, tornando essa a família mais extensa dentre as fosfodiesterases.

Esse gene utiliza diferentes promotores para gerar múltiplas variantes de transcritos de *splicing* alternativo que codificam proteínas funcionais

(WENINGER et al., 2014). Através de suas isoformas, PDE4 atua na regulação do diâmetro da artéria, contribuindo para a homeostase de AMPc em células endoteliais e do músculo lisos. Por outro lado, a inibição de PDE4D causa a dilatação de artérias cerebrais. Adicionalmente, *PDE4* é expresso no músculo liso vascular (em humanos, bovinos e aorta de ratos) e em células do endotélio. A inibição de *PDE4D* tem sido relatada também para fins de reduzir a interação entre leucócitos e células do endotélio com objetivo de diminuir a expressão de moléculas de adesão e, conseqüentemente, a resposta a estímulos pró-inflamatórios. Em AVCI, a inibição de PDE4 promove a proteção da barreira hemato-encefálica e reduz a inflamação e a trombose (MUNSHI et al., 2012).

Estudos prévios (HOUSLAY & ADAMS, 2003; ICHIKI, 2006; LORENOWICZ et al., 2007) indicam que as variações genéticas do *PDE4D* podem influenciar na modulação da proteína *PDE4D* e na degradação do AmpC, resultando em modificações das propriedades neuronais e possivelmente provocar danos permanentes nos neurônios após um AVCI (MUNSHI et al., 2012).

De forma complementar, em um estudo realizado pelo consórcio europeu DeCODE (GRETARSDOTTIR et al., 2003), mutações no gene *PDE4D* demonstraram ser promissoras para a identificação da susceptibilidade ao AVCI, especialmente para os subtipos cardioembólico e aterotrombótico. Adicionalmente, no mesmo estudo foi evidenciado que haplótipos apresentam diferentes níveis de risco aos subtipos de AVCI. Dessa forma a isoforma PDE4D7 e a PDE4D9 apresentam baixa expressão de acordo com do haplótipo do indivíduo. Além disso PDE4D está envolvido na patogênese do evento possivelmente através da aterosclerose.

Rosand et al. (2006), observou que após a publicação de Gretarsdottir et al. (2003), nove novos estudos foram publicados, na tentativa de replicar os achados descritos, entretanto apenas cinco deles obtiveram resultados replicados. Entretanto, um total de 11 variantes genéticas no gene PDE4D, foram investigados em diferentes subtipos de AVCI, no entanto nenhum demonstrou associação significativa com AVCI. Dessa forma Rosand et al. (2006), sugerem que a associação proposta por Gretarsdottir et al. (2003), deve ser vista com cautela.

Variantes gênicas localizadas no gene *PDE4D* também foram investigadas na população de Andhra Pradesh, no sul da Índia, corroborando a hipótese da participação do *PDE4D* na patogenia do AVCI (MUNSHI et al., 2012). Controversamente, as mutações associadas a AVCI descritas no estudo, não foram reproduzidas nas populações coreanas e russas (BEVAN, et al., 2005).

3.8 Variantes genéticas selecionadas para o estudo

Na última década, estudos de GWAS tem permitido a identificação de inúmeras variantes genéticas com associações ($p \leq 5 \times 10^{-8}$) para doenças complexas, incluindo o AVC. Para associação ao AVC, mais de 100 *loci* gênicos já foram reportados (MACARTHUR et al., 2017). Entretanto para a validação dessas variantes, estudos de replicação em diferentes populações devem ser executados.

Nesse sentido, para o presente estudo foram selecionadas 8 variantes genéticas relatadas em um estudo de meta-análise (TRAYLOR et al., 2012), selecionado no ano de 2014, como sendo o mais atual na época da construção do projeto dessa tese de doutorado. A distribuição das frequências genotípicas dessas variantes genéticas em algumas populações no mundo é mostrada na figura 4. Entretanto, até o nosso conhecimento, essas variantes não foram ainda estudadas na população brasileira em relação ao AVCI.

3.8.1 Variante rs2107595

Recentemente, um estudo de GWAs realizado pelo *Welcome Trust Case Control Consortium 2* (WTCCC2) revelou uma nova associação genética com AVCI do tipo aterotrombótico na população europeia (TRAYLOR, 2012). A variante rs2107595 está localizada no cromossomo 7p21, na porção 3' do gene *HDAC9* e a 100 kb do gene *TWIST1*. Azghandi et al. (2015) investigaram a expressão gênica desses genes em células periféricas mononucleadas humanas, e os resultados indicaram que a rs2107595 é somente regulada pelo *mRNA* de *HDAC9* e, que o efeito desse *locus* no risco ao AVC pode ser mediado através do aumento da expressão do *HDAC9*.

A variante genética rs2107595 também foi pesquisada na população asiática em um estudo com 1632 controles e 816 pacientes que tiveram AVC e os resultados apontaram que a variante não está associada ao AVCI, embora esteja associada aos níveis de colesterol total e triglicerídeos dos pacientes, o que explicaria a aterogênese (SU et al., 2015). De acordo com o banco de dados 1000 *Genome Browser*, a frequência alélica do menor alelo (MAF) para a variante rs2107595 da população chinesa é 0,311, enquanto que na população europeia é 0,200. Essa diferença de frequências demonstra que há uma influência desproporcional da variante ao risco do AVC entre populações.

3.8.2 Variante rs2383207

Uma análise realizada com múltiplas coortes identificou variantes genéticas localizadas no cromossomo 9p21 associadas a doenças coronárias e infarto do miocárdio. No entanto, até o momento não foram identificadas regiões codificantes neste *locus*, apenas um longo transcrito não codificante no *locus* do *CDKN2B*, cujos níveis de expressão estão relacionados com aterosclerose. Por sua vez, variantes genéticas no *locus* 9p21 também têm sido reportadas como fatores de risco para aterosclerose e ambas as condições clínicas, doenças coronárias e AVCI.

Matarin et al. (2008) demonstraram que o haplótipo do cromossomo 9p21 representa um risco aumentado em 1,4 a 1,7 vezes para AVCI entre norte-americanos. Em adição, as variantes rs1800796 do gene *IL6* e rs2383207 próximo ao gene *CDKN2A/CDKN2B*, localizados no cromossomo 9p21, foram associados ao risco de AVCI em homens índios do oeste da África (GONG et al., 2012).

Helgadottir et al. (2008) encontraram as mesmas variantes correlacionadas ao risco de AVCI, com associação discreta (OR: 1.15; *p-value* = 0.015) na população da Islândia e, uma vez excluídos os casos que tiveram concomitância com doenças cardiovasculares, a associação foi ainda mais atenuada. Estudos similares com outras populações (Bélgica, EUA e China) falharam em demonstrar associações das variantes rs1800796 e rs2383207 com AVCI (HECKMAN, 2013; LEMMENS, et al., 2009; SH-Y et al., 2014).

Inúmeros estudos investigarem a relação do *locus* 9p21 e o risco ao AVCI. No entanto, os resultados obtidos ainda são controversos devido ao tamanho das amostras, estratificação populacional, controles inadequados e/ou heterogeneidade de fenótipos, o que revela a necessidade de uma maior investigação do papel de variantes genéticas, como a rs2383207, no AVCI.

3.8.3 Variante rs6843082

Em 2012, um estudo de GWAs com amostras de ancestralidade europeia demonstrou que a variante genética rs6843082, localizada próxima ao gene *PITX2*, apresentava forte associação com AVCI subtipo cardioembólico ($p = 2.8 \times 10^{-16}$) (TRAYLOR et al., 2012). O gene *PITX2* possui um importante papel no desenvolvimento do coração, (MOMMERSTEEG et al., 2007), o que poderia auxiliar no entendimento dessa associação.

Entretanto, até a presente data, poucos estudos relataram a replicação desses achados. Um estudo realizado com a população chinesa demonstrou que a variante rs6843082 estava associado ao aumento de pressão arterial sistólica e níveis elevados de colesterol total. No entanto, não foi possível demonstrar a associação com AVCI (SU et al., 2015).

A variante *rs6843082* também tem sido correlacionada como preditor de desfecho clínico de ablação de cateter no tratamento de Fibrilação Atrial, configurando um importante marcador molecular para essa condição clínica (MOHANTY et al., 2016). O mecanismo molecular dessa associação é explicado pelo fato de que *PITX2* regula o aumento da expressão do gene *NPPA*, que codifica o peptídeo natriurético atrial (ANP) expresso no miocárdio. Em consequência, o ANP inibe a proliferação e fibrose de miofibroblastos induzidos por TGF- β . É possível que, modulando a expressão de *PITX2*, a variante rs6843082 proteja o tecido cardíaco da remodelação estrutural e da fibrose miocárdica através da via *PITX2* / *NPPA* / TGF- β (LUBITZ et al., 2014).

3.8.4 Variante rs879324

Assim como a variante *rs6843082*, a variante *rs879324* foi atribuída como fator de risco genético ao AVCI cardioembólico através do estudo de

Traylor et al. (2012) ($p=2.0 \times 10^{-8}$) e esses resultados ainda não foram replicados em outras populações.

A variante está situada no cromossomo 16q.22 na porção intrônica do gene *ZFHX3*, que codifica o fator de transcrição ATBF1, e foi correlacionada com a FA. No entanto, o exato papel molecular dessa variante na FA ainda é desconhecido (XUE et al., 2016).

3.8.5 Variantes rs1396476, rs2910829, rs966221, rs152312

Ao contrário das outras variantes citadas anteriormente, variantes do gene *PDE4D* são bem conhecidas como fatores de predisposição e indicadores de prognóstico ao AVCI.

PDE4D está localizado no cromossomo 5q12 que abrange uma região de 1,6 Mb e contém 24 exons, 8 *splicing* alternativos e centenas de variantes. Seu produto gênico pertence a uma superfamília de fosfodiesterases (família PDE4) e desempenha um papel importante na degradação de AmpC. A proliferação e migração de células de músculos lisos vasculares e macrófagos é responsável pela aterosclerose e AmpC está envolvido neste processo (BEVAN et al., 2005).

Em 2003, Gretarsdottir et al. foram os primeiros autores a identificar a associação de SNP83 (rs966221) no gene *PDE4D* com AVC na população islandesa. Desde então, muitos estudos encontraram essa associação em várias populações. Woo et al. (2006) descobriram que variantes do gene *PDE4D* estavam associados com AVCI e, em particular, com AVC cardioembólico em brancos e negros. Milton et al. (2011) realizaram um estudo de coorte na Austrália e também detectaram uma associação positiva de *PDE4D* e AVC cardioembólico. Além disso, várias meta-análises sugeriram que *PDE4D* estava relacionado ao AVCI, especialmente em populações asiáticas e chinesas (MATSUSHITA, et al., 2009; LIANG, et al., 2015; XU, et al., 2010; WANG, et al., 2017).

Figura 4 - Distribuição das frequências alélicas das variantes genéticas do estudo em diversas populações (AFR: Africanos; AMR: Americanos; EAS: Asiáticos do Leste; EUR: Europeus; SAS: Asiáticos do Sul) Fonte: 1000Genome (www.internationalgenome.org/).



4. METODOLOGIA

4.1 Desenho do estudo e amostras

Trata-se de um estudo do tipo caso-controle em que foram utilizadas amostras e informações associadas pertencentes ao biobanco UNIVILLE, compreendendo amostras de DNA, dados clínicos e epidemiológicos derivados de pacientes acometidos por AVCI e controles saudáveis não consanguíneos, ambos residentes em Joinville/SC.

Os subtipos de AVCI foram classificados de acordo com os critérios TOAST (ADAMS, et al., 1993). As amostras controles foram coletadas conforme descrito por Cabral et al (2009). Os controles foram definidos pela ausência de histórico de AVC do paciente e familiares, não consanguinidade entre casos e outros controles, pareados preferencialmente por idade e gênero. Todos os controles foram coletados durante o período de visita do paciente.

O estudo foi avaliado e aprovado pelo Comitê de Ética em Pesquisa da UNIVILLE sob o parecer 1.396.172 (ANEXO A).

4.2 Critérios de inclusão e exclusão

Foram incluídos pacientes com diagnóstico de primeiro evento de AVCI e idade igual ou superior a dezoito anos de acordo com os critérios do *National Institute of Health Stroke Scale* (NIHSS) (SCHLEGEL et al., 2003), independentemente das comorbidades ou. Os controles foram definidos como indivíduos não acometidos por qualquer tipo de AVC, sem histórico familiar e não consanguinidade com pacientes do estudo. Todos deveriam possuir amostra de DNA e informações associadas armazenadas no Biobanco UNIVILLE. Foram incluídos pacientes e controles que consentiram, por meio do termo de consentimento livre e esclarecido referente ao biobanco, a utilização de sua amostra de DNA, já coletada e armazenada, em pesquisas subsequentes, sem necessidade de novo contato e obtenção de consentimento específico.

Além dos critérios já descritos, foram excluídas amostras derivadas de casos de AVCH ou por outras causas e amostras com qualidade e quantidade de DNA insuficiente para as análises genéticas necessárias.

4.3 Fatores clínicos e ambientais

Quanto aos fatores de risco controláveis ao AVCI foram considerados: tabagismo (fator ambiental), hipertensão, diabetes e dislipidemia (fator clínico), definidos conforme descrito em CABRAL et al, 2009. Tabagismo foi definido por três categorias (não tabagista, atual ou ex-tabagista), hipertensão arterial sistêmica definida pela história de hipertensão tratada ou uso atual de medicação anti-hipertensiva, diabetes caracterizada pelo histórico de diabetes ou tratamento atual para diabetes mellitus) e, dislipidemia caracterizada pelo tratamento atual para essa condição (CABRAL et al, 2009) Em relação aos riscos não modificáveis foram considerados idade no evento e gênero.

4.4 Fatores genéticos

As análises genéticas foram realizadas em parceria com o Laboratório de Genética Molecular da Faculdade de Ciências Médicas da UNICAMP, sob coordenação da Dra. Íscia Lopes-Cendes.

Primeiramente, alíquotas de DNA de todos os pacientes e controles que foram incluídos no estudo foram avaliadas via espectrofotometria de luz ultravioleta (260nm/280nm) no aparelho para micro volumes EPOCH (Biotech, Winooski, EUA) afim de obter a quantificação dos DNA's que seriam analisados.

As genotipagens das variantes analisadas (Tabela 2) foram realizadas por meio de Reação em Cadeia da Polimerase em Tempo Real (qPCR), empregando-se sondas Taqman® (Thermo Fisher Scientific, Massachusetts, EUA). A presença de duas sondas em cada reação, além dos iniciadores, permite a determinação dos dois possíveis alelos da variante: a sonda correspondente a um alelo é marcada com o *Reporter Dye VIC* e a sonda do outro alelo é marcada com o *Reporter Dye FAM*, que apresentam fluorescências em comprimentos de onda diferentes e são rastreadas durante a

reação. O sinal da fluorescência gerado pela amplificação indica qual(is) alelo(s) está(ão) presente(s) na amostra. Para cada reação utilizaram-se 5,0 μ L 2X TaqMan[®] Universal PCR Master Mix (Life Technologies), 0,17 μ L 20X de sondas específicas TaqMan[®] e 10-20ng de DNA, em um volume reacional total de 7 μ L. Para cada placa, foram utilizados 3 controles negativos de reação (sem amostra de DNA). Todos os procedimentos foram realizados de acordo com as recomendações do fabricante das sondas e utilizando o equipamento Real- Time PCR 7500 (Applied Biosystems, Foster City, EUA). Posteriormente os perfis genotípicos foram obtidos por meio do software (versão 2.06) do próprio equipamento.

Tabela 2- Variantes Genéticas analisadas no estudo de acordo com o subtipo de AVCI.

Subtipo de AVCI	Gene	Variante Genética	Posição cromossômica	Consequência genética	Sonda utilizada
Aterotrombótico	<i>HDAC9</i>	<i>rs2107595</i>	7:19049388	Intergênico	TACTGTGGGACAAAAACATTTTCCC[A/G]CAAAAAAGAATGTACAAGCACACAA
	<i>CDKN2B-AS1</i>	<i>rs2383207</i>	9:22115959	Íntron	TTTTACTCCTGTTCCGATCCCTTC[A/G]GCTAAGCATGATTATTTACTATTTT
Cardioembólico	<i>PITX2</i>	<i>rs6843082</i>	4:111718067	Intergênico	TGGTGATGAGTGGTGCATAACAGCC[A/G]TTTCAGCTCTGAGGCTATCAAATAG
	<i>ZFH3</i>	<i>rs879324</i>	16:73068678	Íntron	TTCTGCCTCTGGGAAGGTTTTCACA[A/G]TCAGGTGAGCACCCCCACCTGTACG
	<i>PDE4D</i>	<i>rs1396476</i>	5:59396642	Íntron	GGCATATATCAGCTCTTAGGAAATA[G/T]GTTACACATACACACTGTGGCATG
		<i>rs2910829</i>	5:59469899	Íntron	AATGGATGCATCTAGTTTGGGAAAT[A/G]TTGTGTAGATGTTATGAATGTCCAG
		<i>rs966221</i>	5:59502520	Íntron	AAGGATCTGCTGCTGGATAAACCAC[A/G]TTTTACTATCTATCCTAAAATGATG
<i>rs152312</i>		5:59787816	Íntron	GACTGCTCATTAAAAGAAAATTGTC[A/G]GGAGGGATAGCATTGGGAGATATAC	

4.5 Análise de dados

Os dados relativos aos fatores clínicos e ambientais e os resultados das análises genéticas foram organizados em planilhas para fins de testes estatísticos. Para a determinação das frequências genotípicas e alélicas, do Equilíbrio de Hardy-Weinberg e da associação das variantess estudadas foram utilizados os programas *Plink* (versão 3, 29 de junho de 2007 -Free Software Foundation, Human Longevity, Inc. Data Science team, San Diego, CA). e *SNPStats* (Disponível em <https://www.snpstats.net/snpstats>).

4.6 Análise de interação gênica

Para cada subtipo de AVCI considerado no estudo foi realizada uma análise de interação gênica entre as variantess correlatas utilizando-se o método *Multifactor Dimensionality Reduction* (MDR) (HAHN et al., 2003), o qual colapsa uma matriz de dados de interação de muitas variantes com um grande número de dimensões em uma matriz unidimensional, permitindo as estimativas de interação por meio de validação cruzada. Os resultados foram avaliados em termos de valores de p , corrigidos por permutação do MDR utilizando-se 100.000 replicações.

4.7 Análise de predição

A fim de avaliar se os perfis genotípicos relativos as variantes genéticas avaliados melhoram a acurácia de predição de AVCI foram comparados dois cenários: o primeiro levando em conta somente as variáveis clínicas e

ambientais e o segundo cenário incluindo as variantes. Cada cenário foi estimado com um modelo de predição usando um algoritmo baseado em random forests, o qual constrói uma “floresta” contendo 5.000 árvores de classificação. A validação dos dados é dada por amostragem usando *Leave-One-Out Cross-Validation* (LOOCV) e os resultados avaliados em termos de valores de área sobre a curva (*Area Under the Curve - AUC*) a partir de uma curva ROC (*Receiver Operating Characteristic*) para cada cenário. Além disso, as variáveis foram avaliadas em termos de importância para o modelo utilizando o grau de decaimento da acurácia média (*Mean Decrease Accuracy - MDA*). Neste caso, os valores de MDA indicam o quanto a acurácia do modelo diminui caso a variável x seja retirada do cálculo, e quanto mais importante uma variável x for para o modelo de predição, maior será este valor (HASTIE et al., 2009). Foram incluídas nos modelos propostos apenas as variáveis clínicas e ambientais que apresentaram diferença significativa entre os casos e controles do estudo.

5. INTERDISCIPLINARIEDADE

Este estudo interconecta as áreas de saúde e meio ambiente, uma vez que investigou fatores ambientais, aos quais indivíduos acometidos e não acometidos pelo evento estão expostos, e os relacionou com fatores clínicos.

Ressalta-se que fatores ambientais constituem riscos modificáveis e, dessa forma, seus impactos podem ser prevenidos, uma vez conhecido o seu exato papel na doença. Ademais, a elucidação da contribuição genética no desenvolvimento do AVCI pode também contribuir com a elucidação de mecanismos epigenéticos correlacionados a fatores ambientais que, por sua vez, podem alterar a expressão da influência de variantes genéticas.

6. RESULTADOS E DISCUSSÃO

Conforme as normas do Programa de Pós-Graduação em Saúde e Meio Ambiente da UNIVILLE, este capítulo será apresentado na forma de dois artigos científicos, listados abaixo:

1 - Ferreira LE, França PHC, Nagel V, Venancio V, Safanelli J, Reis FID, Furtado L, Martins RK, Weiss G, Oda E, Lopes-Cendes I, Pontes-Neto O, Cabral NL. Joinville stroke biobank: study protocol and first year's results. *Arq Neuropsiquiatr*. 2017 Dec;75(12):881-889. doi: 10.1590/0004-282X20170157. PubMed PMID: 29236892.

2- Ferreira LE, Secolin R, Lopes-Cendes, Cabral NL, França PHC. Association and Interaction of genetic variants with stroke among Brazilians patients. Será submetido à revista *Human Mutation* (ISSN: 1098-1004).

Joinville stroke biobank: study protocol and first year's results

Biobanco de acidente cerebrovascular de Joinville: protocolo de estudos e resultados dos primeiros anos

Leslie Ecker Ferreira¹, Paulo Henrique Condeixa de França¹, Vivian Nagel¹,
Vanessa Venancio¹, Juliana Safanelli¹, Felipe Ibiapina dos Reis¹, Luis Furtado², Rui
Kleber Martins³, Gustavo Weiss⁴, Elder Oda⁵, Iscia Lopes-Cendes^{6,7}, Octavio
Pontes-Neto³, Norberto Luiz Cabral¹

ABSTRACT

Aiming to contribute to studies that use detailed clinical and genomic information of biobanks, we present the initial results of the first Latin American Stroke Biobank. Methods: Blood samples were collected from patients included in the Joinville Stroke Registry and four Brazilian cities. Demographic socio-economic data, cardiovascular risk factors, Causative Classification System for Ischemic Stroke, Trial of Org 10172 in Acute Stroke Treatment and National Institutes of Health scores, functional stroke status (modified Rankin) and brain images were recorded. Additionally, controls from both geographic regions were recruited. High-molecular-weight genomic DNA was obtained from all participants. Results: A total of 2,688 patients and 3,282 controls were included. Among the patients, 76% had ischemic stroke, 12% transient ischemic attacks, 9% hemorrhagic stroke and 3% subarachnoid hemorrhage. Patients with undetermined ischemic stroke were most common according the Trial of Org 10172 in Acute Stroke Treatment (40%) and Causative Classification System for Ischemic Stroke (47%) criteria. A quarter of the patients were under 55 years of age at the first-ever episode. Conclusions: We established the Joinville Stroke Biobank and discuss its potential for contributing to the understanding of the risk factors leading to stroke.

Keywords: stroke; genetics, biobank.

RESUMO

Com o objetivo de contribuir para estudos que utilizam informações clínicas e genômicas de biobancos, apresentamos os resultados iniciais do primeiro Biobanco Latinoamericano em Acidente Vascular Cerebral (AVC). Métodos: Foram coletadas amostras de sangue de pacientes recrutados pelo Registro de AVC de Joinville e posteriormente de quatro cidades brasileiras. Foram registrados dados socioeconômicos demográficos, fatores de risco cardiovasculares, *Causative Classification System* (CCS), *Trial of Org 10172 in Acute Stroke Treatment*, *National Institutes of Health*, estado funcional (Rankin modificado) e imagens cerebrais. Adicionalmente, foram recrutados controles das regiões geográficas correspondentes. Obteve-se DNA genômico de todos participantes. Resultados: Foram incluídos 2688 pacientes e 3282 controles. Entre os pacientes, 76% tiveram AVC isquêmico, 12% ataques isquêmicos transitórios, 9% AVC hemorrágico e 3% hemorragia subaracnóidea. Os casos indeterminados foram os mais frequentes e classificados de acordo com TOAST (40%) e CCS (47%). Um quarto dos pacientes tinham menos de 55 anos no primeiro evento. Conclusões: Estabelecemos o *Joinville Stroke Biobank*, e discutimos aqui seu potencial na compreensão dos fatores de risco do AVC.

Palavras-chave: acidente vascular cerebral; genética; biobancos.

According to the Global Burden of Disease Study, in 2013, there were almost 25.7 million stroke survivors (71% with ischemic stroke (IS)), 6.5 million deaths from stroke (51% died from IS), 113 million disability-adjusted life years due to stroke (58% due to IS) and 10.3 million new strokes (67% IS) worldwide¹. The INTERSTROKE study defined the

¹Universidade da Região de Joinville, Joinville SC, Brasil;

Universidade Federal do Ceará, Sobral CE, Brasil;

³Universidade de São Paulo, Ribeirão Preto SP, Brasil;

⁴Universidade Federal do Mato Grosso do Sul, Campo Grande MS, Brasil;

⁵Universidade Federal do Rio Grande do Sul, Porto Alegre RS, Brasil;

⁶Universidade de Campinas, Campinas SP, Brasil;

⁷Instituto Brasileiro de Neurociência e Neurotecnologia, Campinas SP, Brasil.

Correspondence: Leslie Ecker Ferreira; Rua Paulo Malschitzki, 10; 89219-710 Joinville SC; E-mail: leslie.ferreira@univille.br **Conflict of interest:** There is no conflict of interest to declare. **Support:** CNPq - (Grant Number: 402396-2013/08) and Coordination for the Improvement of Higher Education Personnel (CAPES) (PROCAD 183768). Received 20 May 2017; Received in final form 15 July 2017;

population-attributable risks for ischemic and hemorrhagic strokes in 22 countries, concluding that 10 classical cardiovascular risk factors are associated with 90% of the risk of stroke². However, twins and familial aggregation studies suggest that the risk of stroke has a substantial genetic component³. Although genetics, and more recently genomics, play an increasingly large role in the practice of medicine, the daily clinical care of patients suffering from stroke has not significantly been affected by the advances in these fields. Therefore, one can assume that there is still much knowledge about stroke genetics and genomics needing to be translated into clinical practice.

Currently, there are tremendous efforts to understand the genetic basis of both rare and common cardiovascular and stroke disorders through strategies such as genome-wide association studies and next-generation sequencing studies⁴. In this scenario, it has been pointed out that well-established stroke biobanks with close collaboration between clinicians and geneticists are essential for supplying the biological and clinical information required for these large-scale studies^{4,5,6}. To address these challenges specifically, recent cooperative efforts have been launched such as the Stroke Genetics Network and Bio-Repository of DNA in Stroke^{7,8}.

As defined by Brazilian regulatory guidelines, a biobank represents an organized collection of human biological material and associated data that are prospectively collected and stored for research purposes, in accordance with predefined technical, ethical and operational standards, under institutional responsibility and management. We present the study protocol and current status of the Joinville Stroke Biobank (JSB), the first Latin American DNA biobank of stroke.

METHODS

Participants and samples

The current biobank originated from two sequential initiatives to study epidemiological aspects of stroke in recent years in Brazil⁹. The phenotypic data were extracted from the Joinville Stroke Registry, which is a population-based registry that has been ongoing since 2005. In 2013, the registry became supported by a municipal law¹⁰. The beginning of case and control data ascertainment, blood extraction and DNA storage took place in 2010. These samples and data, which are still being collected and processed in Joinville, constitute phase I of the JSB, which has no deadline for completion.

Phase II was launched in 2015, when the JSB started to receive blood samples and phenotypic data from four other Brazilian cities. This constitutes a task force of the Brazilian Consortium of Stroke Research, which is sponsored by the Brazilian Ministry of Health and the National Council for Scientific and Technological Development, to define stroke

incidence trends, case-fatality proportions and the prevalence of classic cardiovascular risk factors according to the WHO Steps criteria¹¹ in different geopolitical scenarios than Joinville, which has a higher Human Development Index than other cities. Data and samples will be collected until 2017.

Central site and contributing sites

In JSB phase I, inpatients from five hospitals and out-patient data (mild strokes) from the town of Joinville (515,288 inhabitants, 2010 census) were included, whereas JSB phase II covered the following cities: Sobral, north-east region (three hospitals; 147,135 inhabitants), Campo Grande, central region (eight hospitals; 774,202 inhabitants), Sertãozinho, south-east region (two hospitals; 101,784 inhabitants) and Canoas, south region (three hospitals; 323,827 inhabitants), whose geographic locations are shown in the Figure. The central site, as defined by the International Stroke Genetics Consortium is located in the facilities at the University of the Region of Joinville.

Inclusion and exclusion criteria

Since the beginning of both phases, data and blood samples have been collected from each case and related controls. Patients included transient ischemic attacks, subarachnoid hemorrhages, ischemic strokes and hemorrhagic strokes. We included first-ever or recurrent patients, regardless of age and sex, residing in one of the five cities. All patients had at least one cranial tomography. Stroke confirmation and diagnosis were performed as previously reported⁹. In brief, the phenotype of each stroke patient

Figure. Localization of cities, according to Brazilian mesoregions, corresponding to central and contributing sites of the Joinville stroke biobank.



includes demographic information, social class (according to the Brazilian Criteria of Economic Classification based on the National Household Sample Survey)¹², years of education, type of work (manual or not), family history of stroke, cardiovascular risk factors, biochemical tests and functional status (modified Rankin scale)¹³. Furthermore, IS subtypes were classified according to the modified Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria and Causative Classification System for Ischemic Stroke (CCS) criteria^{14,15}. Patients with IS of undetermined cause due to negative investigation or cryptogenic stroke were classified according to the Embolic Strokes of Undetermined Source international work group¹⁶. A neuroradiologist classified all hemorrhagic strokes as lobar and non-lobar subtypes.

Controls were matched by age and gender to patients, and individuals with a previous history of stroke or blood ties to patients were excluded. In cases of a positive family history of stroke, blood samples of all related patients and non-stroke relatives were searched for and nominated as nuclear cases.

Phenotyping methods

Table 1 shows the diagnostic examination and the categories of data retrieved by questionnaires. After 2012, the CCS criteria for phenotyping IS patients were also employed. The neurologists of all cities in the contributing sites were trained by stroke neurologists from the Joinville Stroke Registry using the platform available on the CCS website¹⁵. After the TOAST diagnosis, the clinical history, physical examination findings and the results of diagnostic work-up were reviewed for determining the CCS classification. Ischemic stroke patients with incomplete diagnostic examination were classified as a possible IS subtype or received an IS diagnosis as undetermined with incomplete evaluation. All Joinville IS diagnoses were blinded for adjudication by two authors (FIR and NLC).

Data and blood sample collection

Patient selection and blood collection were performed by neurologists and nurses, respectively. Research nurses invited the controls to enroll from among the accompanying persons willing to participate who fulfilled the inclusion criteria. After clarifying the aims of the JSB and the corresponding roles of all parties involved, blood sampling (phase I: 12 mL by venous puncture employing common EDTA-containing vacuum tubes; phase II: finger puncture followed by drying and storage on an FTA Elute Card[†], Whatman, Kent, United Kingdom) of patients and controls was performed by research nurses. Baseline data included self-declared skin color, collected in strict accordance with the mandatory procedures established by the Brazilian Institute of Geography and Statistics for purposes of population census. Demographics, socio-economic data, cardiovascular risk factors, National Institutes of Health Stroke Scale (clinical stroke severity) and

functional outcome can be extracted from the Joinville Stroke Registry, which also includes causes and dates of death. The follow-up of patients is being carried out by phone calls by previously trained research nurses in one, three, six, nine and 12 months in the first year and once a year for the following four years after stroke diagnosis. Patients are asked about regular visits to a general physician, blood pressure control, results of glycated hemoglobin and cholesterol levels, and smoking habits.

An identification code, which comprises numbers that denote the date of recruitment, contributing site and entry order, was assigned to each patient enrolled. Controls were identified using numbers that keep their relationship with each matched patient. These identification codes link clinical data and demographics with deposited DNA samples and only the biobank managers or authorized personnel have access to the key linking samples and phenotypic data.

Processing of samples, DNA extraction and quality control

Blood samples from phase I were maintained at 4°C until DNA extraction. Whole blood samples were centrifuged for 10 minutes at 2,500 x g for separation of the buffy coat, which contains most of the white blood cells. Thereafter, genomic DNA was obtained by the classic “phenol-chloroform” procedure for total DNA extraction. The DNA was resuspended in 200 µL of buffer TE (10 mM Tris-HCl; 1 mM EDTA; pH 8.0). For phase II, we adopted the procedure for large-amount DNA extraction, starting with three paper punches of 3 mm circles, as recommended by the manufacturer of the FTA Elute Card[†]. The purity and yield of DNA extracted were evaluated by spectrophotometry at 260 and 280 nm (Epoch, Biotek Instruments, Winooski, USA). Long-term storage was considered for A_{260}/A_{280} ratios in the range 1.8–2.0. Finally, DNA was stored in aliquots at -80°C.

Ethical considerations

The JSB development plan and the informed consent forms were both approved by the Institutional Review Board of the University of the Region of Joinville and by the Brazilian National Commission for Ethics in Research (protocol 25000.142907/2013-07), which holds the final approval prerogative for biobanks in Brazil. The adopted standards and procedures strictly follow the ethical rules that came into force in 2011 and are described in Resolution 441/11 of the National Health Council and in Ordinance 2.201 of the Brazil Ministry of Health, which correspond to the national guidelines for biorepositories and biobanks of human biological material for research purposes. The JSB also follows the international ethical regulations. All volunteers (patients and controls) are enrolled after written informed consent has been given freely by the participants or

Table 1. Demographic and diagnostic examination data available in the Joinville Stroke Biobank.

Category	Description	Cases	Controls
Sociodemographic	Date of birth/age	Y	Y
	Gender	Y	Y
	Skin color	Y	Y
	City of residence	Y	Y
	Address	Y	Y
	Social status	Y	N
	Education level	Y	N
Clinical and biochemical data	First-ever event?	Y	N
	Date of enrollment	Y	N
	Hour of enrollment	Y	N
	Creatinine	Y	N
	Triglyceride	Y	N
	Uric acid	Y	N
	Blood glucose	Y	N
	Cholesterol (hdl/ldl)	Y	N
	National Institutes of Health Stroke Scale (admission)	Y	N
	Barthel Index	Y	N
	Rankin Index	Y	N
	Oxfordshire Community Stroke Project classification (Banford)	Y	N
	TOAST classification	Y	N
CCS classification	Y	N	
Premorbid risk factors	Transient ischemic attack	Y	Y
	Hypertension	Y	Y
	Diabetes mellitus	Y	Y
	Dyslipidemia	Y	Y
	Acute myocardial infarction	Y	Y
	Congestive heart failure	Y	N
	Atrial fibrillation	Y	N
	Smoking	Y	Y
	Alcohol consumption **	Y	N
	Drug consumption	Y	N
Physical activity	Y	N	
Diagnostic examination	Brain CT	Y	N
	Brain MRI	Y	N
	Digital angiography	Y	N
	Extracranial vessels evaluation		
	Ultrasound Doppler	Y	N
	CT - angiography	Y	N
	MR - angiography	Y	N
	Intracranial vessels evaluation		
	Transcranial Doppler	Y	N
	CT - angiography	Y	N
	MR - angiography	Y	N
	Embolic sources		
	ECG	Y	N
Transthoracic echocardiography	Y	N	
Transesophageal echocardiography	Y	N	
24 h Holter	Y	N	
Outcome measures	Rankin scale (30 days, 6 months, once a year up to 5 years)	Y	N
	N Death (cause, date) by death certificate	Y	N

TOAST: Trial of Org 10172 in Acute Stroke Treatment criteria¹⁶; CCS: Causative Classification System criteria¹⁷. CT: computed tomography; MRI: magnetic resonance imaging. Y: yes (available); N: no (not available). *Self-declared skin color; **current moderate or heavy drinker was defined

as 50 g/day, which is equivalent to 500 mL or two glasses of wine, 1000 mL of beer or 5 units of spirits, or being intoxicated at least once a week. (Y):Yes; (N): No.

their legal representatives. The freedom to refuse participation or withdraw at any time is guaranteed, and the confidentiality of personal data is ensured in all circumstances. Free access to information and associated implications is guaranteed to participants, thereby respecting the expression of individual will.

RESULTS

Since 2010, a total of 5,970 DNA samples have been registered belonging to patients (2,688) and controls (3,282). The mean ages of patients and controls were 64.6 ± 19.3 and 56.9 ± 14.9 years, respectively. Men were more represented among the patients (64%; 1,731/2,688), while women were more prevalent among controls (68%; 2,226/3,282).

Table 2 shows the distribution of patients and controls per city, as well as their self-declared skin color. White

individuals predominated in the south (88%) and south-east (55%) regions and brown in the north-east (53%) and central (41%) regions.

Table 3 shows the distribution of stroke diagnoses of patients in the different cities. Most of the patients (76%; 2,031/2,688) were IS, followed by 12% (319/2,688) transient ischemic attacks, 9% (243/2,688) hemorrhagic stroke and 3% (95/2,688) subarachnoid hemorrhage. A quarter of all stroke patients were aged ≤ 55 years.

The distribution of IS cases by TOAST criteria (Table 4) was: 20% (426/2,031) cardioembolic; 22% (440/2,031) small-arteryocclusion; 16% (317/2,031) large-artery atherosclerosis; 40% undetermined (807/2,031) and 2% (41/2,031) other determined cause. After 2012, patients were also classified by CCS criteria (56%; 1,503/2,688). As expected, among IS patients, the most common IS subtype was undetermined, not only by TOAST criteria (40%; 807/2,031), but also by CCS criteria (47%; 718/1,503). Out of those 807 undetermined IS patients

Table 2. Baseline characteristics of stroke patients and controls in the Joinville Stroke Biobank database.

City	Campo Grande	Canoas	Joinville	Sertãozinho	Sobral	All
Demography						
Cases (n)	507	187	1619	175	200	2688
Mean age (\pm SD) (years)	65.2 (14.9)	64.5 (14.3)	64.3 (14.1)	64.9 (15.1)	67.1 (16.1)	64.6 (19.3)
Male [n(%)]	290 (57.2)	100 (53.5)	1185 (73.2)	56 (32)	100 (50)	1731 (64.4)
Skin color [% (n/N)]						
White	46.9 (238/507)	49.7 (93/187)	88.2 (1428/1619)	55.4 (97/175)	29.5 (59/200)	71.2 (1915 /2688)
Black	7.7 (39/507)	16.0 (30/187)	5.4 (87/1619)	13.1 (23/175)	16.0 (32/200)	7.8 (211/2688)
Brown	41.4 (210/507)	20.9 (39/187)	6.0 (97/1619)	29.1 (51/175)	53.5 (107/200)	18.8 (504/2688)
Native American	2.0 (10/507)	2.7 (5/187)	0.2 (4/1619)	1.1 (2/175)	0.5 (1/200)	0.8 (22/2688)
Yellow	1.4 (7/507)	0.5 (1/187)	0.1 (2/1619)	1.1 (2/175)	0.5 (1/200)	0.5 (13/2688)
Not-declared	0.6 (3/507)	10.2 (19/187)	0.1 (1/1619)	0.0	0.0	0.9 (23/2688)
Controls (n)						
Controls (n)	180	24	2586	276	216	3282
Mean age (\pm SD) years)	50.2 (14.6)	64.56 (13.1)	68.3 (14.7)	50.2 (13,9)	66.3 (15,5)	56.9 (14.9)
Male [n(%)]	61 (33.8)	14 (58.3)	865 (33.5)	65 (23.5)	51 (23.6)	1056 (32.2)
Skin color [% (n/N)]						
White	53.9 (97/180)	62.5 (15/24)	90.4 (2338/2586)	46.7 (129/276)	32.9 (71/216)	80.7 (2650/3282)
Black	0.6 (1/180)	25.0 (6/24)	2.7 (71/2586)	6.9 (19/276)	6.5 (14/216)	3.4 (111/3282)
Brown	40.0 (72/180)	8.3 (2/24)	6.4 (166/2586)	44.2 (122/276)	47.7 (103/216)	14.2 (465/3282)
Native American	0.6 (1/180)	4.2 (1/24)	0.0 (1/2586)	0.4 (1/276)	0.9 (2/216)	0.2 (6/3282)
Yellow	0.6 (1/180)	0.0	0.1 (3/2586)	1.4 (4/276)	0.0	0.2 (8/3282)
Not-declared	4.4 (8/180)	0.0	0.3 (7/2586)	0.4 (1/276)	12.0 (26/216)	1.3 (42/3282)

Table 3. Major stroke diagnosis of patients in the Joinville Stroke Biobank database per city [% (n/N)].

Variable	City					
	Campo Grande	Canoas	Joinville	Sertãozinho	Sobral	All
IS	16.5 (335/2031)	7.5 (152/2031)	60.9 (1238/2031)	5.6 (112/2031)	9.5 (194/2031)	75.5 (2031/2688)
TIA	20.6 (66/319)	1.6 (5/319)	73.1 (233/319)	4.7 (15/319)	0	11.9 (319/2688)
HS lobar	-	-	11.7 (28/243)	0	-	-
HS non-lobar	-	-	7.8 (19/243)	0	-	-
HS non-classified	34.6 (84/243)	11.1 (27/243)	20.1 (49/243)	13.9 (34/243)	0.8 (2/243)	9.1 (243/2688)
SAH	23.1 (22/95)	3.5 (3/95)	54.6 (52/95)	14.6 (14/95)	4.2 (4/95)	3.5 (95/2688)
Young adult (≤ 55 years old)	4.3 (118/2688)	0.4 (13/2668)	15.7 (424/2688)	1.9 (52/2688)	1.9 (53/2688)	24.5 (660/2688)
Nuclear cases	0	0	0.4 (7/1619)	0	0	0.4 (7/2688)
All cases	18.9 (507/2688)	6.9 (187/2688)	60.2 (1619/2688)	6.5 (175/2688)	7.5 (200/2688)	2688

IS: ischemic stroke; TIA: transient ischemic attack; HS: hemorrhagic stroke; SAH: subarachnoid hemorrhage.

classified by TOAST criteria, 55% (446/807) were patients with an incomplete evaluation, 42% (340/807) were cryptogenic and 3% (21/807) were undetermined with two or more possible causes. Out of those 718 undetermined IS subtypes classified by CCS criteria, 25% (370/1,503) were IS patients with incomplete evaluation, 17% (262/1,503) were unknown other cryptogenic, 8% (61/1,503) were unknown cryptogenic embolism and 1% (25/1,503) were unclassified.

DISCUSSION

Up to October 2016, DNA from 5,970 blood samples distributed among patients (2,688) and controls (3,282) had been extracted and stored. Since it is believed that most of the genetic variants identified in stroke studies will have small, individual effects on disease risk, efforts aimed at the successful discovery of the impact of genes have required large sample sizes (usually involving tens or hundreds of thousands of cases and controls) to achieve sufficient statistical power¹⁰. In addition, as with any case-control study, controls should be representative of cases^{17,18}.

With this in mind, the JSB comprises well-defined phenotypes and samples of high-quality DNA. The priority recruitment of spouses or partners as case-matched controls represents a suitable strategy for minimizing differences due to environmental effects to which the participants (patients and controls) have been exposed. As controls are collected in the same geographic region as patients, it is reasonable to expect a homogeneous distribution of ethnic backgrounds between patients and controls. On the other hand, despite having inflated influences due to shared variants and environmental effects between patients and controls, the JSB may also provide relevant information on familial aggregation analysis. It is also noteworthy to clarify that the JSB has a forecast of unlimited continuity as, among other

factors, the Joinville Stroke Registry is supported by a municipal law. This gives rise to the fact that if there are cases of stroke in individuals previously categorized as a control (since stroke events may happen later in life), their biobank status changes, but maintains the traceability of the change. This allows the possibility of evaluation of the impact on results of any studies already performed with the data and samples of those individuals who no longer belong to the control group. As far as we know, this is the first stroke biobank in Latin America, a region with a massive race mixture.

Under TOAST, IS cases are assigned to one of three main subtypes (small-artery occlusion, large-artery atherosclerosis, cardioembolic); to a rare cause (e.g., carotid or vertebral artery dissection); or remain undetermined (because of more than one potential cause, incomplete investigation or no apparent cause despite complete investigation/cryptogenic stroke). The proportion in the undetermined category varies, but can be substantial (up to 40% or more), and is usually excluded from ischemic subtype-specific analyses^{19,20}. In our cohort, between 22% (TOAST criteria) and 25% (CCS criteria) were undetermined IS due to incomplete investigation. All patients with incomplete investigation were assigned this diagnosis. In fact, the highly diverse sociodemographic scenarios at sites contributing to the JSB, including difficulties related to health-care system access and imaging diagnostic workup, might have influenced significantly. Nevertheless, independent raters may perform blinded readjudications of all JSB IS diagnoses whenever considered necessary.

Globally, large DNA biobanks have been contributing to the understanding of the different genetic architectures and pathophysiological aspects of stroke. In this regard, the UK Biobank (European and British Asian), Brains-SA (India and Sri Lanka), Bio-Repository of DNA in Stroke – Middle East (Qatar)²¹, China Kadoorie Biobank²² and other

Table 4. Ischemic stroke subtypes diagnosis of patients in the Joinville Stroke Biobank database [% (n/N)].

Variable	City					
	Campo Grande	Canoas	Joinville	Sertãozinho	Sobral	All
IS subtype by TOAST						
Large-artery atherosclerosis	6.9 (22/317)	5.1 (16/317)	81.8 (259/317)	2.8(9/317)	3.4 (11/317)	15.8 (317/2031)
Cardioembolic	12.2 (52/426)	3.2 (14/426)	78.3 (333/426)	2.8(12/426)	3.5 (15/426)	20.2 (426/2031)
Small-artery occlusion	17.1 (75/440)	4.3 (19/440)	6.9 (307/440)	0.9(4/440)	7.9 (35/440)	21.7 (440/2031)
Other determined cause	17.1(7/41)	0	82.9 (34/41)	0	0	2.2(41/2031)
Undetermined						
Cryptogenic/ESUS	5.6 (19/340)	16.2 (55/340)	51.8 (176/340)	2.6 (9/340)	23.8 (81/340)	16.8 (340/2031)
Incomplete investigation	35.6(159/446)	10.8 (48/446)	24.6 (110/446)	17.5 (78/446)	11.4 (51/446)	22.1 (446/2031)
Two or more determined causes	4.7 (1/21)	0	90.6 (19/21)	0	4.7 (1/21)	1.2 (21/2031)
IS subtype by CCS						
Large-artery atherosclerosis						
Evident	5.3 (8/150)	10.7 (16/150)	76.7 (115/150)	0	7.3 (11/150)	9.9 (150/1503)
Probable	0	0	92.1(35/38)	7.9(3/38)	0	2.4 (38/1503)
Possible	12(3/25)	0	64 (16/25)	24 (6/25)	0	1.5 (25/1503)
Cardioembolic						
Evident	11.7 (18/154)	7.1 (11/154)	75.3 (116/154)	3.2 (5/154)	2.6 (4/154)	10.2 (154/1503)
Probable	0	0	100 (40/40)	0	0	2.5 (40/1503)
Possible	15.7 (8/51)	0	49.1 (25/51)	13.6(7/51)	21.6 (11/51)	3.3 (51/1503)
Small-artery occlusion						
Evident	6.8(12/175)	10.3 (18/175)	6.5 (113/175)	5.7(1/175)	17.7 (31/175)	11.5 (175/1503)
Probable	15.1 (10/66)	0	8.2(54/66)	3.1 (2/66)	0	4.3 (66/1503)
Possible	2.1(1/48)	0	87.5 (42/48)	2.1(1/48)	8.3(4/48)	3.1 (48/1503)
Other determined cause						
Evident	11.1 (4/36)	0	86.2(31/36)	0	2.7 (1/36)	2.3 (36/1503)
Probable	0	0	100(2/2)	0	0	1.2 (2/1503)
Possible	0	0	0	0	0	0
Undetermined						
Unknown cryptogenic embolism	0	0	80.3(49/61)	0	19.7 (12/61)	4.1 (61/1503)
Unknown other cryptogenic	6.1 (16/262)	20.2(53/262)	44.6 (117/262)	26.7(7/262)	26.3 (69/262)	17.4 (262/1503)
Incomplete evaluation	27.6 (102/370)	12.9 (48/370)	24.3 (90/370)	21.3 (79/370)	13.8 (51/370)	24.6 (370/1503)
Unclassified	4(1/25)	24(6/25)	64 (16/25)	4 (1/25)	4 (1/25)	1.7 (25/1503)
Total IS by TOAST (since 2010)	16.5 (335/2031)	7.5 (152/2031)	60.9 (1238/2031)	5.6 (112/2031)	9.5 (194/2031)	75.5 (2031/2688)
Total IS by CCS (since 2012)	12.2 (183/1503)	10.1 (152/1503)	57.3 (861/1503)	74.5 (112/1503)	12.9 (195/1503)	72.0(1503/2100)

IS: ischemic stroke; TOAST: Trial of Org 10172 in Acute Stroke Treatment criteria¹⁶; ESUS: Embolic Strokes of Undetermined Source; CCS: Causative Classification System criteria¹⁷.

collaborative initiatives, such as the International Stroke Genetics Consortium, have contributed to the identification of several loci associated with IS subtypes.

Genome-wide association study approaches have been adopted for investigating many complex diseases, such as stroke, and have been effective in identifying new genetic variants associated with the risk of disease. However, for stroke studies, genome-wide association studies have so far been applied principally in European, North American and Japanese populations^{3,23}.

In particular, four loci (*PITX2*, *HDAC9*, *ZFH3* and 12q24.2) were convincingly implicated by genome-wide association studies as associated with IS in European-ancestry patients^{24,25}. Recently, a novel locus (rs12122341) was identified at 1p13.2 (near *TSPAN2*), associated with large-artery atherosclerosis stroke, and the locus 12q24 (near *ALDH2*) was associated with small-artery occlusion²⁶. In addition, the gene *TRPV3* was associated with cardioembolic stroke after replication of exome sequencing analysis, demonstrating that two polymorphisms were associated with cardioembolic stroke risk in two cohorts,

with rs151091899 being the most significant²⁷. Therefore, replication studies in other populations are determinants in understanding the significance of such polymorphisms as stroke risk factors. The Brazilian population is one of the most heterogeneous populations in the world, comprising an admixture of Native Americans, Europeans and Africans²⁸. The admixture process occurred through different means in Brazilian geographic regions. For instance, the Native American contribution is more pronounced in northern Brazil, the African contribution is more elevated in the north-east, and the south features a European predominance with few Native American and African influences²⁹. Consequently, the JSB, unlike other biobanks^{8,22,23}, comprises an ideal admixed population with great ethnic variability linked to epidemiological and clinical data, which may allow confirmation of polymorphisms found

in specific populations, contributing to global understanding of genetic mechanisms involved in stroke.

In conclusion, the clinical, environmental and epidemiological issues related to the different types of stroke, combined with the large genetic variability captured by the JSB, may increase the chance of identifying new and relevant factors predisposing individuals to stroke (and recovery) in Brazil. Our aim is to continue patient assessment and blood collection, as well as begin genotyping samples in the near future. Some initial studies are already being conducted in the Laboratory of Molecular Genetics at the University of Campinas with the support of the Brazilian Institute of Neuroscience and Neurotechnology and the use of a genome-wide association genotyping platform (Affymetrix™ 6.0).

References

- Feigin VL, Krishnamurthi RV, Parmar P, Norrving B, Mensah GA, Bennett DA et al. Update on the global burden of ischemic and hemorrhagic stroke in 1990–2013: the GBD 2013 study. *Neuroepidemiology*. 2015;45(3):161-76. <https://doi.org/10.1159/000441085>
- O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P et al. Risk factors for ischemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet*. 2010;376(9735):112-23. [https://doi.org/10.1016/S0140-6736\(10\)60834-3](https://doi.org/10.1016/S0140-6736(10)60834-3)
- Ikram MA, Seshadri S, BisJC, Fornage M, DeStefano AL, Aulchenko YS et al. Genomewide association studies of stroke. *N Engl J Med*. 2009;360(17):218-28. <https://doi.org/10.1056/NEJMoa0900094>
- Musunuru K, Hickey KT, Al-Khatib SM, Delles C, Fornage M, Fox CS et al. Basic concepts and potential applications of genetics and genomics for cardiovascular and stroke clinicians: a scientific statement from the American Heart Association. *Circ Cardiovasc Genet*. 2015;8(1):216-42. <https://doi.org/10.1161/HCG.000000000000020>
- Dichgans M. Genetics of ischaemic stroke. *Lancet Neurol*. 2007;6(2):149-61. [https://doi.org/10.1016/S1474-4422\(07\)70028-5](https://doi.org/10.1016/S1474-4422(07)70028-5)
- Hewitt J, Walters M, Padmanabhan S, Dawson J. Cohort profile of the UK Biobank: diagnosis and characteristics of cerebrovascular disease. *BMJ Open*. 2016;6(3):e009161. <https://doi.org/10.1136/bmjopen-2015-009161>
- Meschia JF, Arnett DK, Ay H, Brown RD Jr, Benavente OR, Cole JW et al. Stroke Genetics Network (SIGN) study: design and rationale for a genome-wide association study of ischemic stroke subtypes. *Stroke*. 2013;44(10):2694-702. <https://doi.org/10.1161/STROKEAHA.113.001857>
- Yadav S, Schanz R, Maheshwari A, Khan MS, Slark J, Silva R et al. Bio-repository of DNA in stroke (BRAINS): a study protocol. *BMC Med Genet*. 2011;12:34. <https://doi.org/10.1186/1471-2350-12-34>
- Cabral NL, Cougo-Pinto PT, Magalhaes PS, Longo AL, Moro CH, Amaral CH et al. Trends of stroke incidence from 1995 to 2013 in Joinville, Brazil. *Neuroepidemiology*. 2016;46(4):273-81. <https://doi.org/10.1159/000445060>
- Prefeitura Municipal de Joinville. Lei Nº 7.448, de 12 de junho de 2013. Institui o banco de dados Registro de Acidente Vascolar Cerebral — AVC de Joinville e dá outras providências. *Jornal Município*. 28 jun 2013.
- World Health Organization — WHO. WHO Steps Stroke Manual (V2.0): the WHO STEPwise approach to stroke surveillance. Geneva: World Health Organization; 2005.
- Paiva GFS, Silva DBN, Feijó CA. Consumption and socioeconomic classification in Brazil: a study based on the Brazilian family expenditure survey. In: IARIW-IBGE Special Conference on "Income, Wealth and Well-being in Latin America"; 2013 Sep 11-14; Rio de Janeiro, Brazil.
- Sulter G, Steen C, De Keyser J. Use of the Barthel index and modified Rankin scale in acute stroke trials. *Stroke*. 1999;30(8):1538-41. <https://doi.org/10.1161/01.STR.30.8.1538>
- Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL et al. Classification of subtype of acute ischaemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. 1993;24(1):35-41. <https://doi.org/10.1161/01.STR.24.1.35>
- Ay H, Benner T, Arsava EM, Furie KL, Singhal AB, Jensen MB et al. A computerized algorithm for etiologic classification of ischemic stroke: the Causative Classification of Stroke System. *Stroke*. 2007;38(11):2979-84. <https://doi.org/10.1161/STROKEAHA.107.490896>
- Hart RG, Diener HC, Coutts SB et al. Embolic strokes of undetermined source: the case for a new clinical construct. *Lancet Neurol*. 2014;13(4):429-38. [https://doi.org/10.1016/S1474-4422\(13\)70310-7](https://doi.org/10.1016/S1474-4422(13)70310-7)
- Jayasinha SR, Mishra A, Van Daal A, Kwan E. Genetics and cardiovascular disease: design and development of a DNA biobank. *Exp Clin Cardiol*. 2009;14(3):33-7.
- McCarty CA, Chisholm RL, Chute CG, Kullo IJ, Jarvik GP, Larson EB et al. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics*. 2011;4(1):13. <https://doi.org/10.1186/1755-8794-4-13>
- Wolfe C, Rudd A, Howard R et al. Incidence and case fatality rates of stroke subtypes in a multiethnic population: the South London Stroke Register. *J Neurol Neurosurg Psychiatry*. 2002;72:211-6.
- Feigin VL, Lawes CM, Bennett DA, Barker-Collo SL, Parag V. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. *Lancet Neurol*. 2009;8(4):355-69. [https://doi.org/10.1016/S1474-4422\(09\)70025-0](https://doi.org/10.1016/S1474-4422(09)70025-0)

Cotlarciuc I, Khan MS, Maheshwari A, Yadav S, Khan FY, Al-Hail H et al. Bio-repository of DNA in stroke: a study protocol of three ancestral populations. *JRSM Cardiovasc Dis.* 2012;1(4):cvd.2012.12019. <https://doi.org/10.1258/cvd.2012.012019>

21. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol.* 2011;40(6):1652-66. <https://doi.org/10.1093/ije/dyr120>
22. Matsushita T, Umeno J, Hirakawa Y, Yonemoto K, Aschikawa K, Amitani H et al. Association study of the polymorphisms on chromosome 12p13 with atherothrombotic stroke in the Japanese population. *J Hum Genet.* 2010;55(7):473-6. <https://doi.org/10.1038/jhg.2010.45>
23. Bellenguez C, Bevan S, Gschwendtner A, Spencer CC, Burgess AI, Pirinen M et al. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet.* 2012;44(3):328-33. <https://doi.org/10.1038/ng.1081>
24. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE Collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol.* 2012;11(11):951-62. [https://doi.org/10.1016/S1474-4422\(12\)70234-X](https://doi.org/10.1016/S1474-4422(12)70234-X)
25. NINDS Stroke Genetics Network (SiGN); International Stroke Genetics Consortium (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol.* 2016;15(2):174-84. [https://doi.org/10.1016/S1474-4422\(15\)00338-5](https://doi.org/10.1016/S1474-4422(15)00338-5)
26. Carrera C, Jiménez-Conde J, Derdak S et al. Whole exome sequencing analysis reveals TRPV3 as a risk factor for cardioembolic stroke. *Thromb Haemost.* 2016;116(6):1165-71. <https://doi.org/10.1160/TH16-02-0113>
27. Amador MA, Cavalcante GC, Santos NP et al. Distribution of allelic and genotypic frequencies of IL1A, IL4, NFKB1 and PAR1 variants in Native American, African, European and Brazilian populations. *BMC Res Notes.* 2016;9(1):101. <https://doi.org/10.1186/s13104-016-1906-9>
28. Santos NP, Ribeiro-Rodrigues EM, Ribeiro-dos-Santos AK, Pereira R, Gusmão L, Amorim A et al. Assessing individual interethnic admixture and population substructure using a 48-Insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum Mutat.* 2010;31(2):184-90. <https://doi.org/10.1002/humu.21159>

Association and interaction of genetic variants with occurrence of ischemic stroke among Brazilian patients

Leslie Ecker Ferreira¹, Rodrigo Secolin², Iscia Lopes-Cendes², Norberto Luiz Cabral¹, Paulo Henrique Condeixa de França¹.

¹University of Region of Joinville – UNIVILLE and Joinville Stroke Registry, Joinville, Brazil,

²Department of Medical Genetics, School of Medical Sciences, University of Campinas – UNICAMP and Brazilian Institute of Neuroscience and Neurotechnology (BRAINN), Campinas, Brazil.

ABSTRACT

Ischemic Stroke (IS) is a severe and complex disorder of high morbidity and mortality rates associated with clinical, environmental, and genetic predisposing factors. Previous genetic studies have associated genetic variants to IS. However, inconsistent results from different populations pointed to the presence of genetic heterogeneity for IS. In addition, genetic variants found associated with IS have appear to present weak allelic effect. Therefore, we can hypothesize that an interaction effect among genetic variants could contribute to IS rather than genetic variant independently. In this context, we investigated the association and interaction between genetic variants and large-artery atherosclerosis IS (LAAS-IS) and cardioembolic IS (CE-IS) subtypes. We genotyped eight genetic variants in 435 patients (195 LAAS-IS and 240 CE-IS) and 535 controls from a population of Joinville, Santa Catarina, Brazil. Association and interaction analysis were performed by chi-square test in PLINK software and Multifactor-dimensionality Reduction approach in MDR software, respectively. We found an association between rs2383207*A allele, nearby *CDKN2B-AS1*, and LAAS-IS [OR 2.35 (95% CI = 1.79 - 3.08); $p=4.66 \times 10^{-10}$]. In addition, we observed an increase on LAAS-IS risk among AG and AA genotypes carriers in a codominance model [OR 2.14 (95% CI = 1.33 - 3.45); OR 5.26 (95% CI = 2.98 - 9.28); respectively]. We did not find independent association between candidate genetic variants and CE-IS. However, we found an interaction effect in CE-IS among three SNPs (rs2910829, rs966221 and rs152312), with an accuracy of 0.62 ($p=4.3 \times 10^{-5}$). This result could indicate that an interaction effect among SNPs from different genes can contribute to CE-IS risk. Further prediction analysis confirmed that clinical information, such as hypertension and dyslipidemia, presented high accuracy to predict LAAS-IS (86.47%) and CE-IS (90.47%); however, the inclusion of genetic variant information did not increase the accuracy.

Key-Words: Ischemic Stroke, genetic variants, genetic interactions, case-control study; genetic association study.

1 INTRODUCTION

Stroke is a severe disorder with high morbidity, mortality and disability-adjusted life years rates, which is considered a major healthcare problem and a worrisome economic burden. It represents the primary neurological cause of acquired disability in adults and a leading cause of death, provoking almost six million deaths each year worldwide (Feigin et al., 2015; Benjamin et al., 2018). Ischemic stroke (IS) is the most prevalent type of stroke, accounting for 85 to 90% of all cases (Feigin et al., 2015). In Brazilian population, the adjusted annual incidence rate of IS ranges from 62 to 92 per 100,000 inhabitants, with a mixed pattern of cardiovascular risk factors, due to the high prevalence of cardiovascular diseases, dyslipidemia, hypertension, and diabetes (Cabral et al., 2017; Minelli, Fen & Minelli, 2007). Ischemic stroke is usually categorized by etiology as cardioembolic (CE-IS), large artery atherosclerosis (LAAS-IS), lacunar (LAS-IS), and stroke of other determined (SOD-IS) or undetermined (UN-IS) etiology (Adams, Bendixen & Kappelle, 1993).

Risk factors have been associated with IS, including advanced age, hypertension, diabetes mellitus, smoking, and dyslipidemia. However, around 10% of IS events remain unclassified, which had reinforced to the hypothesis of a multifactorial model disease, including genetic and small-effect environmental risk factors and even possible interactions among them (Chauhan & Debette, 2016; O'Donnell et al., 2016). Indeed, genome-wide association studies have pointed to genetic variants as potential biomarkers for prediction of stroke (Zondervan & Cardon, 2007; Zeggini, Scott & Saxena, 2008). More recently, authors have suggested genetic variants in *CDKN2B*, *HDAC9*, *PITX2*, *ZFH3* and *PDE4D* genes as independent risk factors for LAAS-IS and CE-IS (Bai et al., 2014; Bayoglu et al., 2016; Huang et al., 2014; Smith et al., 2009).

Genetic susceptibility has been suggested to be a critical player in the pathogenesis of IS. However, due the small allele effect found in previous studies (Traylor et al., 2012), we may hypothesize that genetic interaction could affect IS rather than genetic variants in an independent way. Therefore, our objective was to investigate the association and interaction among genetic variants for LAAS-IS and CE-IS in a Brazilian cohort. The first part of this manuscript describes the association analysis between each genetic variant and LAAS-IS and CE-IS. The second part describes an interaction statistical analysis among genetic variants which were not found associated with both phenotypes. Finally, the third part describes an additional

prediction analysis in order to evaluate whether the candidate genetic variants combined with clinical variables can improve the accuracy to predict LAAS-IS and CE-IS occurrence.

2 METHODS

2.1 Study population and risk factors

We obtained DNA samples from patients and non-affected IS individuals who were consecutively recruited between March 2010 and October 2015 for inclusion at Joinville Stroke Biobank (JSB), which is linked to Joinville Stroke Registry (Ferreira et al., 2017), an ongoing cohort since 2005, considered the largest epidemiological study on stroke in South America. Joinville is in Southern Brazil (516,288 inhabitants), with four hospitals (one having a stroke unit) and one state-run institutional care facility. The detailed methods of cohort recruitment and data collection procedures have been described previously (Cabral et al., 2009).

For the purpose of this study, the variables considered as risk factors for IS were smoking, defined by three categories - non-smoker, current or former smoker; systemic arterial hypertension, defined by the clinical history of treatment or current regular use of antihypertensive medication; diabetes mellitus, characterized by corresponding clinical history or current therapy; and dyslipidemia, characterized by the current treatment for this condition (Cabral et al., 2009). The non-modifiable risks considered were gender and age on the event.

2.2 Inclusion and exclusion criteria

Patients presenting LAAS-IS or CE-IS diagnosis, classified according to both modified Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria and Causative Classification System for Ischemic Stroke (CCS) criteria, were considered cases (Hackan et al., 2007). Non-affected individuals (controls) were matched by age and gender to cases. Subjects showing blood ties with any case were excluded. A total of 435 patients, distributed between 195 LAAS-IS and 240 CE-IS cases, and 535 controls were included in the study.

2.3 Genotyping

Firstly, whole blood samples were centrifuged for 10 minutes at 2,500 x g for separation of the buffy coat. Thereafter, genomic DNA was obtained as described elsewhere (Ferreira et al., 2017). DNA was resuspended in 200 µL of buffer TE (10 mM Tris-HCl; 1 mM EDTA; pH 8.0). The purity of DNA and yield of the extraction were evaluated by spectrophotometry (Epoch, BioTek Instruments, Winooski, VT, USA). Samples with A260/A280 ratio inside the range 1.8 – 2.0 were considered for long-term storage at -80°C. DNA was diluted to final concentration of 10 ng/µL immediately before genotyping. Genetic variants previously described in association with stroke were selected to this work: rs2107595 (NC_000007.13:g.19049388G>A) and rs2383207 (NC_000009.11:g.22115959A>G) for LAAS-IS, and rs879324 (NC_000016.10:g.73034779G>A), rs1396476 (NC_000005.10:g.60100815T>C), rs2910829 (NC_000005.10:g.60174072G>A), rs966221 (NC_000005.10:g.60206693A>G), rs6843082 (NC_000004.11:g.111718067G>A), and rs152312 (NC_000005.10:g.60491989G>A) for CE-IS (Traylor et al., 2012). Genotypes were determined by Real Time Polymerase Chain Reaction (qPCR) employing TaqMan® probes (ThermoFisher Scientific, Waltham, MA, USA) specific for all genetic variants of the study. The reactions were performed with 20 ng DNA, 0.17 µL probes 40x, and 2x TaqMan® Master Mix (ThermoFisher Scientific) in a final volume of 7 µL.

2.4 Descriptive statistics and genetic association analysis

The statistical power to detect genetic associations was determined from “A priori: Compute required sample size” statistical power analysis in the G*Power software (Faul et al., 2007) with the following parameters: statistical power = 0.8; logistic regression test; two-tailed; effect size (in terms of OR) = 1.5; and α = 0.05. We found a minimal sample size of $n = 373$ for LAAS-IS group (cases and controls) and $n = 477$ for CE-IS group (cases and controls). Therefore, the study was conducted on 444 individuals for LAAS-IS group (195 cases and 249 controls) and 529 individuals for CE-IS group (240 cases and 285 controls). The same individuals pertained both controls group.

Statistical analysis for the baseline characteristics was performed using SPSS software version 16.0 (SPSS Inc., Chicago, IL, USA). Age of participants were expressed as means and standard deviations (\pm SD) and comparisons between groups were performed using Student's *t*-test. Chi-square (χ^2) test was applied for

evaluating proportions of clinical and environmental variables between cases and controls.

Minor allele frequency (MAF), Hardy-Weinberg equilibrium (HWE), and subsequent genetic association analyzes by χ^2 test were performed by PLINK software version 1.9 (Purcell et al., 2007). Aiming to evaluate the level of allele effect in LAAS-IS and CE-IS, we also estimated the ORs with corresponding 95% confidence intervals (95% CI). Bonferroni adjustment for multiple comparisons were performed on the resulting p -values and $p \leq 0.001$ was considered statistically significant for all analyzes.

2.5 Genetic interaction analysis

Multifactor-dimensionality reduction (MDR) analysis was performed using MDR software version 3.0.2 (Ritchie, Hahn & Moore, 2001), in order to evaluate possible interaction among genetic variants which were not associated with LAAS-IS or CE-IS. In summary, a 10-fold cross-validation was performed on the data set divided into ten equal sizes from which nine subsets served as training sets and one as the test set. The total number of combined factors used was 121. The best factor combination was screened based on the minimum classification error in the training set. The process was repeated ten times. The final candidate model of gene-gene interaction was selected based on maximum cross-validation consistency (CVC) and maximum testing balanced accuracy, which p -values were estimated by a permutation test under 10,000 replicates.

The selected model is plotted in a circle entropy graph based on information theory. Individual and all possible pairing of genetic variants received information percentage scores based on the risk for stroke. A positive percentage score was designated as synergistic interaction (red on the graph), while score of null or below was designated redundant or antagonistic (blue). If paired percentage scores were stronger (for positive values) or weaker (for negative values) than individual score, the former was selected as a better model to describe association with the risk of CE-IS (Ritchie et al., 2001; Moore et al., 2006; Hahn, Ritchie & Moore, 2003).

2.6 Prediction analysis

We performed the random forests method to evaluate whether the combination of genetic variants and clinical information could improve the accuracy

in predicting LAAS-IS and CE-IS. First, we used two different scenarios: one that considers only clinical variables, and a second that encompasses clinical variables and genotypes.

Random forests set m_{try} parameter describes the number of randomly selected variables used to build the classification trees. We applied the Leave-One-Out Cross-Validation (LOOCV) to estimate the best mixture of parameters m_{try} for each of 5,000 trees. In addition, we estimated the importance of each variable for the prediction model by terms of mean decreased accuracy (MDA), which is the difference between the discriminatory accuracy before and after excluding the variable. If the variable is important for the model, the discriminatory accuracy decreases substantially, and the more the accuracy decreases due to exclusion of the variable, more important the variable is considered for the model (Breiman, 2002).

2.7 Ethical considerations

The study was reviewed and approved by the local Institutional Review Board according to Brazilian and international ethical guidelines, under the opinion 1.396.172.

3 RESULTS

3.1 Descriptive analysis

Baseline demographic and clinical characteristics of patients and controls included in the study are summarized in Table 1. As expected, patients presented higher frequency of hypertension (LAAS-IS: 75.3% vs. Controls: 22.8%; CE-IS: 75.8% vs. Controls: 19.3%), dyslipidemia (LAAS-IS: 57.4% vs. Controls: 17.6%; CE-IS: 62.5% vs. Controls: 9.5%), diabetes (LAAS-IS: 62.5 vs. Controls: 10.8%; CE-IS: 62.5% vs. Controls: 9.5%), and smoking habits (LAAS-IS: 34.8% vs. Controls: 15.3%; CE-IS: 13.3% vs. Controls: 12.3%). The mean age of patients and controls did not show significant difference. Men were significantly more represented among LAAS-IS cases (63.1%) when compared with respective controls (44.5%, $p=1.0 \times 10^{-4}$). In contrast, the difference on gender proportion between cases and controls was not significant in CE-IS group ($p=0.66$).

3.2 Association analysis

As shown in Table 2, SNPs rs2107595, rs879324 and rs1396476 were not in HWE; therefore, they were withdrawn of subsequent analysis. We observed an increased risk for LAAS-IS for rs2383207*A allele [OR 2.35 (95% CI = 1.79 - 3.08); $p=4.66 \times 10^{-10}$]. We did not find association between CE-IS and the respective SNPs; therefore, we performed the interaction analysis from CE-IS SNPs.

In order to evaluate the effect of rs2383207 and LAAS-IS clinical variables, we estimated ORs among five different genetic models as shown in Table 3. The increase on LAAS-IS risk was demonstrated among AG and AA genotypes carriers in a codominance model [OR 2.14 (95% CI = 1.33 - 3.45); OR 5.26 (95% CI = 2.98 - 9.28); respectively] and these genotypes showed association with LAAS-IS ($p<0.001$), even after Bonferroni correction. Nevertheless, results were not significant when considering usual risk factors for stroke, as hypertension, dyslipidemia, diabetes, and smoking.

3.3 Interaction analysis

The best interaction model for CE-IS included three SNPs (rs2910829, rs966221 and rs152312) which resulted in a maximum CVC of 10/10 and an accuracy of 0.616. The *p-value* associated with the prediction error and CVC was $p = 4.3 \times 10^{-5}$ after adjustment by permutation test (Table 4).

We observed a synergistic interaction between rs152312 and rs966221, and between rs966221 and rs2910829. Redundant or antagonistic interaction was observed between rs152312 and rs2910829, rs152312 and rs6843082, and rs2910829 and rs6843082 (Figure 1). In addition, absolute percentage values for pairs rs152312/rs6843082 (0.87%) and rs2910829/rs6843082 (0.88%) are higher than rs6843082 alone (0.64%), which could indicate possible redundant or antagonistic interaction effect (Figure 1).

3.4 Prediction analysis

Prediction accuracy only using clinical information was 86.47% for LAAS-IS (sensitivity = 77.86%; specificity = 80.72%; Figure 2A) and 90.47% for CE-IS (sensitivity = 84.81%; specificity = 85.31%; Figure 2B). The inclusion of genotypes of the SNPs in the model had a discrete effect in the prediction accuracy for both LAAS-IS (88.03%; sensitivity = 83.40%; specificity = 83.13%; Figure 2A) and CE-IS (90.84%; sensitivity = 79.50%; specificity = 84.96%; Figure 2B). In fact, the most

important variables in the prediction model were dyslipidemia and hypertension for both LAAS-IS and CE-IS (Figure 3).

4 DISCUSSION

We observed an association of genetic variant rs2383207, located in *CDKN2B-AS1* gene at the chromosome 9p21 locus, with the occurrence of LAA-IS in Joinville population. Previous studies have associated rs2383207 to higher risk for IS in native population of Western of Africa (Akinyemi et al., 2017). In addition, Helgadottir et al. (2008) reported correlation between rs2383207 and IS, but with a weak effect (OR 1.15; $p=0.015$) in Iceland population, and this association was even weaker when cases of cardiovascular diseases and IS concomitance were excluded. Despite similar studies in other populations have failed to show association of rs2383207 with IS (Lin et al., 2010), different genetic variants in the nearby region of *CDKN2B-AS1* also have been found in association with cardiovascular diseases, IS or MI in different populations (Akinyemi et al., 2017; Lin et al., 2011; Matarin et al. 2008), which strengthens the hypothesis that *CDKN2B-AS1* can contribute to IS risk. Therefore, based on our findings and previous studies, we can suggest that different genetic variants in *CDKN2B-AS1* could contribute to IS risk based on specific populations, and further population studies could aid important information to future applications in precision medicine for IS.

CDKN2B-AS1 gene is located at the chromosome 9p21 locus, which has been found in association with the risk of multiple diseases, including coronary heart disease (CHD), myocardial infarction (MI), hypertension, and stroke (Dehghan et al., 2016; Kremer et al., 2015; Matsuoka et al., 2015; Wahlstrand et al., 2009). A recent study has hypothesized that variants on 9p21 locus increase *CDKN2B-AS1* gene expression, which was negatively correlated with *CDKN2B* expression (Johnson et al., 2013). The *CDKN2B-AS1* gene encoded product binds and recruits epigenetic modifiers, such as polycomb repressor complex, to promoter regions of adjacent genes (Hannou et al., 2015; McPherson, 2007). These epigenetic regulations may eventually influence DNA methylation, altering the expression of *CDKN2B* and promoting artery calcification (Zhou et al., 2016).

Our study has two major strengths. First, cases are clinically well-characterized, as both best recognized methods (TOAST and CCS) nowadays for IS subtyping are employed by Joinville Stroke Registry, which is not common in most

similar reports. Second, our study included individuals from the same town in the South of Brazil, encompassing an unique population with substantial stroke rates (Cabral et al., 2017). Previous study showed that this region presented homogeneous genetic background (Kehdy et al., 2015), which reduces the impact of population stratification (Marchini et al., 2004; Gomes et al., 2017). Despite higher incidence of death has been observed among Brazilian patients with CHD due to variants in locus 9p21 (Gioli-Pereira et al., 2012), our study is the first correlating rs2383207 with the occurrence of stroke in Brazil.

However, our study presents results related to rs2383207 for one exclusive subtype of stroke (LAAS-IS). Nevertheless, already reported genome-wide association studies have identified genetic associations in different stroke subtypes, reflecting stroke phenotypic complexity and a range of possibilities of gene-by-environment interactions over different ancestry backgrounds (NINDS et al., 2015).

We also did not find association between CE-IS and selected genetic variants. Since those same variants had already been described in association with CE-IS risk in different populations (NINDS et al., 2015), we can argue that our population could present different genomic ancestry blocks in comparison with those populations, and further genomic population analysis is needed to confirm this hypothesis. In the other hand, we found an interaction effect between rs152312/rs6843082 and rs2910829/rs6843082 in CE-IS sample. In addition, we observed that interactions could result a redundant or antagonistic effect in comparison to rs6843082.

Here we strengthened the hypothesis that LAAS-IS and CE-IS etiology features complex biological mechanisms, including genetic variants and environmental factors. Therefore, instead attempting to explain LAAS-IS and CE-IS etiology based on genetic or environmental factors, we decided to evaluate whether the candidate genetic variant added to the clinical variables could improve the accuracy of predicting LAAS-IS and CE-IS. However, we showed that classic risk factors, such as hypertension and dyslipidemia, have large impact in predicting LAAS-IS and CE-IS, and the inclusion of genetic data retrieved from our study did not change significantly the prediction accuracy. Interesting, the rs2383207 appeared as the third most important variable present in LAAS-IS predicting model (Figure 3). However, due the low increase in prediction accuracy (Figure 2) and the low level of importance compared with hypertension and dyslipidemia (Figure 3), rs2383207 genotyping to aid in LAAS-IS prediction would not be useful in clinical practice. Hence, our results

reinforce the need for strengthening the implementation of prevention policies aimed at modifiable factors, as genetics seems to have less influence in the studied population.

5 CONCLUSION

Both genetic and life-style related factors may increase risk of stroke. Genetic influence driven by genetic variants may indicate predisposition to certain subtypes of stroke. Our findings support that rs2383207 is associated with an increased risk for LAA-IS in a Brazilian population. Therefore, through additional gene-gene and environment-gene interactions analyzes, the knowledge about the genetic effect on the IS risk might be enriched and provide crucial information on the identification of molecular mechanisms related to IS.

6 REFERENCES

- Adams, H. P. Jr., Bendixen, B. H., Kappelle, L. J., Biller, J., Love, B. B., Gordon, D. L., Marsh, E. E. 3rd (1993). Classification of subtype of acute ischaemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*, 24(1), 35-41.
- Akinyemi, R., Arnett, D. K., Tiwari, H. K., Ovbiagele, B., Sarfo, F., Srinivasasainagendra V, SIREN Investigators et al. (2017). Interleukin-6 (IL-6) rs1800796 and cyclin dependent kinase inhibitor (CDKN2A/CDKN2B) rs2383207 are associated with ischemic stroke in indigenous West African Men. *Journal of Neurological Sciences*, 379, 229-235. <https://doi.org/10.1016/j.jns.2017.05.046>.
- Ay, H., Benner, T., Arsava, E. M., Furie, K. L., Singhal, A. B., Jensen, M. B., Sorensen A. G., et al. (2007). A computerized algorithm for etiologic classification of ischemic stroke: the Causative Classification of Stroke System. *Stroke*, 38(11), 2979-2984. <https://doi.org/10.1161/STROKEAHA.107.490896>.
- Bai, Y., Nie, S., Jiang, G., Zhou, Y., Zhou, M., Zhao, Y., Wang, Q. K. et al. (2014). Regulation of CARD8 expression by ANRIL and association of CARD8 single nucleotide polymorphism rs2043211(p.C10X) with ischemic stroke. *Stroke*, 45(2),383-388. <https://doi.org/10.1161/STROKEAHA>.
- Bayoglu, B., Yuksel, H., Cakmak, H. A., Dirican, A., Cengiz, M. (2016). Polymorphisms in the long non-coding RNA CDKN2B-AS1 may contribute to higher systolic blood pressure levels in hypertensive patients. *Clinical Biochemistry*, 49(10-11), 821-827. <https://doi.org/10.1016/j.clinbiochem.2016.02.012>.

Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, et al. (2018). Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association. *Circulation*, 137(12),e67-e492. <https://doi.org/10.1161/CIR.0000000000000558>.

Breiman, L. (2002). Manual on setting up, using, and understanding random forests (Version 3.1). Retrieved from https://www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32.

Cabral, N. L., Freire, A. T., Conforto, A. B., Dos Santos, N., Reis, F. I., Nagel, V., ... Longo, A. L. (2017). Increase of Stroke Incidence in Young Adults in a Middle-Income Country: A 10-Year Population-Based Study. *Stroke*, 48(11), 2925-2930. <https://doi.org/10.1161/STROKEAHA.117.018531>.

Cabral, N. L., Gonçalves, A. R., Longo, A. L., Moro, C. H., Costa, G., Amaral, C. H., et al. (2009). Incidence of stroke subtypes, prognosis and prevalence of risk factors in Joinville, Brazil: a 2 year community based study. *Journal of Neurology, Neurosurgery, and Psychiatry*, 80(7), 755-761. <https://doi.org/10.1136/jnnp.2009.172098>.

Chauhan, G., Debette, S. (2016). Genetic Risk Factors for Ischemic and Hemorrhagic Stroke. *Current Cardiology Reports*, 18(12),124. <https://doi.org/10.1007/s11886-016-0804-z>.

Dehghan, A., Bis, J. C., White, C. C., Smith, A. V., Morrison, A. C., Cupples, L. A., O'Donnell, C. J., et al. (2016). Genome-wide association study for incident myocardial infarction and coronary heart disease in prospective cohort studies: the CHARGE consortium. *PLoS One*, 11, e0144997. <https://doi.org/10.1371/journal.pone.0144997>.

Dichgans, M., Malik, R., König, I. R., Rosand, J., Clarke, R., Gretarsdottir, S., et al., International Stroke Genetics Consortium. (2014). Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke*, 45(1), 24–36. <https://doi.org/10.1161/STROKEAHA.113.002707>.

Faul, F., Erdfelder, E., Lang, A.G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*, 39, 175–191.

Feigin, V. L., Krishnamurthi, R. V., Parmar, P., Norrving, B., Mensah, G. A., Bennett, Stroke Panel Experts Group, et al. (2015). Update on the global burden of ischemic and hemorrhagic stroke in 1990–2013: the GBD 2013 study. *Neuroepidemiology*, 45, 161-176. <https://doi.org/10.1159/000441085>.

Ferreira, L. E., França, P. H. C., Nagel, V., Venancio, V., Safanelli, J., Reis, F. I. D., Cabral, N. L., et al. (2017). Joinville stroke biobank: study protocol and first year's results. *Arquivos de Neuro-psiquiatria*, 75(12), 881-889. <https://doi.org/10.1590/0004-282X20170157>.

Gioli-Pereira, L., Santos, P. C. J., L., Ferreira, N. E., Hueb, W. A., Krieger, J. E., Pereira, A. C. (2012). Higher incidence of death in multi-vessel coronary artery disease patients associated with polymorphisms in chromosome 9p21. *BMC Cardiovascular Disorders*, 12, 61. <https://doi.org/10.1186/1471-2261-12-61>.

Gomes, K. F., Santos, A. S., Semzezem, C., Correia, M. R., Brito, L. A., Ruiz, M. O., Silva, M. E., et al. (2017). The influence of population stratification on genetic markers associated with type 1 diabetes. *Scientific Reports*, 7, 43513. <https://doi.org/10.1038/srep43513>.

Hahn, L. W., Ritchie, M. D., Moore, J. H. (2003). Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. *Bioinformatics*, 19(3), 376–382.

Hannou, S. A., Wouters, K., Paumelle, R., Staels, B. (2015). Functional genomics of the CDKN2A/B locus in cardiovascular and metabolic disease: what have we learned from GWASs? *Trends in Endocrinology and Metabolism*, 26(4), 176–184. <https://doi.org/10.1016/j.tem.2015.01.008>.

Helgadottir, A., Thorleifsson, G., Magnusson, K. P., Grétarsdottir, S., Steinthorsdottir, V., Manolescu, A., Stefansson, K. et al. (2008). The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nature Genetics*, 40(2), 217-224. <https://doi.org/10.1038/ng.72>.

Huang, Y., Ye, H., Hong, Q., Xu, X., Jiang, D., Xu, L., Duan, S. et al. (2014). Association of CDKN2BAS polymorphism rs4977574 with coronary heart disease: a case-control study and a meta-analysis. *International Journal of Molecular Sciences*, 15(10), 17478-17492. <https://doi.org/10.3390/ijms151017478>.

Johnson, A. D., Hwang, S. J., Voorman, A., Morrison, A., Peloso, G. M., Hsu, Y. H., O'Donnell, C. J. et al. (2013). Resequencing and clinical associations of the 9p21.3 region: a comprehensive investigation in the Framingham heart study. *Circulation*, 127(7), 799–810. <https://doi.org/10.1161/CIRCULATIONAHA.112.111559>.

Kehdy, F. S., Gouveia, M. H., Machado, M., Magalhães, W. C., Horimoto, A. R., Horta, B. L., Brazilian EPIGEN Project Consortium, et al. (2015). Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. *Proceedings of the National Academy of Sciences of the United States of America*, 112(28), 8696-8701. <https://doi.org/10.1073/pnas.1504447112>.

Kremer, P. H., Koeleman, B. P., Pawlikowska, L., Weinsheimer, S., Bendjilali, N., Sidney, S., Klijn, C. J., et al., (2015). Evaluation of genetic risk loci for intracranial aneurysms in sporadic arteriovenous malformations of the brain. *Journal of Neurology, Neurosurgery, and Psychiatry*, 86(5), 524-529. <https://doi.org/10.1136/jnnp-2013-307276>.

Lee, S. Y., Chung, Y., Elston, R. C., Kim, Y., Park, T. (2007). Log-linear model-based multifactor dimensionality reduction method to detect gene-gene interactions. *Bioinformatics*, 23(19), 2589-2595. <https://doi.org/10.1093/bioinformatics/btm396>.

- Lin, H. F., Tsai, P. C., Liao, Y. C., Lin, T. H., Tai, C. T., Juo, S. H., Lin, R. T. (2011). Chromosome 9p21 genetic variants are associated with myocardial infarction but not with ischemic stroke in a Taiwanese population. *Journal of Investigative Medicine*, 59(6), 926-30. <https://doi.org/10.2310/JIM.0b013e318214ea49>.
- Lin, H.-F., Tsai, P.-C., Lin, R.-T., Khor, G.-T., Sheu, S.-H., Juo, S.-H. H. (2010). Sex differential genetic effect of chromosome 9p21 on subclinical atherosclerosis. *PLoS One*, 5(11), e15124. <https://doi.org/10.1371/journal.pone.0015124>.
- Matarin, M., Brown, W. M., Singleton, A., Hardy, J. A., Meschia, J. F., ISGS investigators (2008). Whole genome analyses suggest ischemic stroke and heart disease share an association with polymorphisms on chromosome 9p21. *Stroke*, 39(5), 1586-1589. <https://doi.org/10.1161/STROKEAHA.107.502963>.
- Matsuoka, R., Abe, S., Tokoro, F., Arai, M., Noda, T., Watanabe, S., Yamada, Y., et al. (2015). Association of six genetic variants with myocardial infarction. *International Journal of Molecular Medicine*, 35(5), 1451-1459. <https://doi.org/10.3892/ijmm.2015.2115>.
- McPherson, R., Pertsemlidis, A., Kavaslar, N., Stewart, A., Roberts, R., Cox, D., Cohen, J. C., et al. (2007). A common allele on chromosome 9 associated with coronary heart disease. *Science*, 316(5830), 1488–1491. <https://doi.org/10.1126/science.1142447>.
- Minelli, C., Fen, L. F., Minelli, D. P. (2007). Stroke incidence, prognosis, 30-day, and 1-year case fatality rates in Matão, Brazil: a population-based prospective study. *Stroke*, 38(11), 2906-2911. <https://doi.org/10.1161/STROKEAHA.107.484139>.
- Moore, J. H., Gilbert, J. C., Tsai, C. T., Chiang, F. T., Holden, T., Barney, N., White, B. C. (2006). A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *Journal of Theoretical Biology*, 241(2), 252-261. <https://doi.org/10.1016/j.jtbi.2005.11.036>.
- NINDS Stroke Genetics Network (SiGN); International Stroke Genetics Consortium (ISGC) (2015). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol*, 15:175-184. [https://doi:10.1016/S1474-4422\(15\)00338-5](https://doi:10.1016/S1474-4422(15)00338-5).
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Sham, P. C., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3), 559–575. <https://doi.org/10.1086/519795>.
- Ritchie, M. D., Hahn, L. W., Roodi, N., Bailey, L. R., Dupont, W. D., Parl, F. F., Moore, J. H., et al. (2001). Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *American Journal of Human Genetics*, 69(1), 138-147. <https://doi.org/10.1086/321276>.

Smith, J. G., Melander, O., Lökvist, H., Hedblad, B., Engström, G., Nilsson, P., Lindgren, A. (2009). Common genetic variants on chromosome 9p21 confers risk of ischemic stroke: a large-scale genetic association study. *Circulation. Cardiovascular Genetics*, 2(2), 159-164. <https://doi.org/10.1161/CIRCGENETICS.108.835173>.

Traylor, M., Farrall, M., Holliday, E. G., Sudlow, C., Hopewell, J. C., Cheng, Y. C., International Stroke Genetics Consortium (2012). Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *The Lancet. Neurology*, 11(11), 951-962. [https://doi.org/10.1016/S1474-4422\(12\)70234-X](https://doi.org/10.1016/S1474-4422(12)70234-X).

Wahlstrand, B., Orho-Melander, M., Delling, L., Kjeldsen, S., Narkiewicz, K., Almgren, P., Melander, O. (2009). The myocardial infarction associated CDKN2A/CDKN2B locus on chromosome 9p21 is associated with stroke independently of coronary events in patients with hypertension. *Journal of Hypertension*, 27(4), 769-773. <https://doi.org/10.1097/HJH.0b013e328326f7eb>.

World Health Organization: The WHO STEP Wise Approach to Stroke Surveillance. WHO Steps Stroke Manual (Version 2.0). Geneva, World Health Organization, 2014.

Yi, X., Lin, J., Wang, Y., Zhou, J., Zhou, Q. (2017). Interaction among CYP2C8, GPIIIa and P2Y12 variants increase susceptibility to ischemic stroke in Chinese population. *Oncotarget*, 8(41), 70811–70820. <https://doi.org/10.18632/oncotarget.19991>.

Zeggini, E., Scott, L. J., Saxena, R., Voight, B. F., Marchini, J. L., Hu, T., Altshuler, D. (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nature Genetics*, 40(5), 638–645. <https://doi.org/10.1038/ng.120>.

Zondervan, K. T., Cardon, L. R. (2007). Designing candidate gene and genome-wide case–control association studies. *Nature Protocols*, 2(10), 2492–2501. <https://doi.org/10.1038/nprot.2007.366>.

Zhou S, Zhang Y, Wang L, Zhang Z, Cai B, Liu K, Zhang H, Dai M, Sun L, Xu X, Cai H, Liu X, Lu G, Xu G (2016). CDKN2B methylation is associated with carotid artery calcification in ischemic stroke patients. *J Transl Med*, 14(1):333.

TABLE 1 Baseline demographic and clinical characteristics of the studied population

Risk Factors	Variable	LAAS-IS (N=195)	Controls (N=249)	p-value	CE-IS (N=240)	Controls (N=285)	p-value
Mean age (±SD) (years) †		66.9 (11.6)	61.6 (10.7)	0.49	68.7 (13.7)	61.6 (10.7)	2.40x10 ⁻¹
Gender [n (%)]	Male	123 (63.1)	111 (44.5)	1.00x10 ^{-4*}	108 (45)	122 (42.8)	6.58x10 ⁻¹
	Female	72 (36.9)	138 (55.4)		132 (55)	163 (57.2)	
Hypertension [n (%)]	Yes	147 (75.3)	57 (22.8)	6.14x10 ^{-36*}	182 (75.8)	55 (19.3)	4.59x10 ^{-47*}
	No	48 (24.6)	192 (77.1)		54 (22.5)	230 (80.7)	
Diabetes [n (%)]	Yes	122 (62.5)	27 (10.8)	1.01x10 ^{-17*}	150 (62.5)	27 (9.5)	1.22x10 ^{-16*}
	No	73 (37.4)	222 (89.2)		80 (33.3)	258 (90.2)	
Smoking [n (%)]	Yes	68 (34.8)	38 (15.3)	4.70x10 ^{-8*}	32 (13.3)	35 (12.3)	8.56x10 ^{-10*}
	Ex-Smoking No	58 (29.7) 70 (35.8)	54 (21.7) 157 (63.0)		118 (49.2) 90 (64.3)	60 (21.1) 190 (66.7)	
Dyslipidemia [n (%)]	Yes	112 (57.4)	44 (17.6)	1.89x10 ^{-25*}	160 (62.5)	37 (9.5)	1.25x10 ^{-50*}
	No	83 (42.7)	205 (82.3)		80 (33.3)	248 (90.5)	

† Age-at-diagnosis for cases and age-at-collection for controls. * Statistical significance. LAAS-IS: Large Artery Atherosclerosis Ischemic Stroke; CE-IS: Cardioembolic Ischemic Stroke.

TABLE 2 Results of genetic association analyzes

IS Subtype	Gene	Genetic Variant	RA	F_AFF	F_UNAFF	HWE <i>p-value</i>	OR	SE	<i>p-value</i>	
CE-IS	<i>PITX2</i>	rs6843082	G	0.27	0.30	7.09x10 ⁻²	0.85 (0.65 - 1.12)	0.14	2.34x10 ⁻¹	
	<i>ZFHX3</i>	rs879324	A	0.27	0.25	8.81x10 ^{-4*}	-	-	-	
	<i>PDE4D</i>		rs1396476	G	0.17	0.15	6.99x10 ^{-4*}	-	-	-
			rs2910829	G	0.43	0.45	1.12x10 ⁻¹	0.91 (0.71 - 1.15)	0.12	4.03x10 ⁻¹
			rs966221	A	0.45	0.48	2.53x10 ⁻¹	0.66 (0.52 - 0.85)	0.13	9.24x10 ⁻⁴
			rs152312	A	0.32	0.38	1.52x10 ⁻¹	0.66 (0.89 - 0.48)	0.15	5.91x10 ⁻³
LAAS-IS	<i>CDKN2B-AS1</i>	rs2383207	A	0.58	0.37	1.81x10 ⁻¹	2.35 (1.79 - 3.09)	0.14	4.67x10 ^{-7*}	
	<i>HDAC9</i>	rs2107595	A	0.19	0.49	9.06x10 ^{-7*}	-	-	-	

RA: Risk Allele; OA: Other Allele; F_AFF: Frequency of Cases; F_UNAFF: Frequency of controls; HWE: Hardy-Weinberg Equilibrium; OR: Odds ratio; SE: Standard error; * Statistical significance.

TABLE 3 Relationship of rs2383207 with risk factors for LAAS-IS

Cases vs Controls (adjusted by gender)					
Model 1	Genotype	LAAS-IS n (%)	Controls n (%)	Odds Ratio (95% CI)	p-value
Codominant	GG	36 (18.5)	100 (40.2)	1.00	<0.0001*
	GA	92 (47.2)	114 (45.)	2.14 (1.33 - 3.45)	
	AA	67 (34.4)	35 (14.1)	5.26 (2.98 - 9.28)	
Dominant	GG	36 (18.5)	100 (40.2)	1.00	<0.0001*
	GA+AA	159 (81.5)	149 (59.8)	2.86 (1.83 - 4.48)	
Recessive	GG+GA	128 (65.6)	214 (85.9)	1.00	<0.0001*
	AA	67 (34.4)	35 (14.1)	3.27 (2.04 - 5.24)	
Overdominant	GG+AA	103 (52.8)	135 (54.2)	1.00	0.95
	GA	92 (47.2)	114 (45.8)	1.01 (0.69 - 1.49)	
Log-additive	---	---	---	2.29 (1.72 - 3.04)	<0.0001*
Hypertension (adjusted by gender)					
Model 2	Genotype	LAAS-IS n (%)	Controls n (%)	OR (95% CI)	p-value
Codominant	GG	51 (25.1)	84 (35.1)	1.00	0.03
	GA	96 (47.3)	109 (45.6)	1.45 (0.93 - 2.26)	
	AA	56 (27.6)	46 (19.2)	2.01 (1.19 - 3.39)	
Dominant	GG	51 (25.1)	84 (35.1)	1.00	0.022
	GA+AA	152 (74.9)	155 (64.8)	1.62 (1.07 - 2.45)	
Recessive	GG+GA	147 (72.4)	193 (80.8)	1.00	0.039
	AA	56 (27.6)	46 (19.2)	1.60 (1.02 - 2.49)	
Overdominant	GG+AA	107 (52.7)	130 (54.4)	1.00	0.73
	GA	96 (47.3)	109 (45.6)	1.07 (0.73 - 1.56)	
Log-additive	---	---	---	1.42 (1.09 - 1.84)	0.0083

Diabetes (adjusted by gender)					
Model 3	Genotype	LAAS-IS n (%)	Controls n (%)	OR (95% CI)	p-value
Codominant	GG	22 (22.2)	114 (33.1)	1.00	0.098
	GA	52 (52.5)	153 (44.5)	1.78 (1.02 - 3.10)	
	AA	25 (25.2)	77 (22.4)	1.70 (0.89 - 3.23)	
Dominant	GG	22 (22.2)	114 (33.1)	1.00	0.032
	GA+AA	77 (77.8)	230 (66.9)	1.75 (1.03 - 2.96)	
Recessive	GG+GA	74 (74.8)	267 (77.6)	1.00	0.55
	AA	25 (25.2)	77 (22.4)	1.17 (0.70 - 1.97)	
Overdominant	GG+AA	47 (47.5)	191 (55.5)	1.00	0.15
	GA	52 (52.5)	153 (44.5)	1.39 (0.89 - 2.17)	
Log-additive	---	---	---	1.30 (0.96 - 1.77)	0.094
Dyslipidemia (adjusted by gender)					
Model 4	Genotype	LAAS-IS n (%)	Controls n (%)	OR (95 CI)	p-value
Codominant	GG	26 (23.2)	105 (33.1)	1.00	0.12
	GA	58 (51.8)	142 (44.8)	1.67 (0.99 - 2.84)	
	AA	28 (25)	70 (22.1)	1.64 (0.89 - 3.04)	
Dominant	GG	26 (23.2)	105 (33.1)	1.00	0.041
	GA+AA	86 (76.8)	212 (66.9)	1.66 (1.01 - 2.74)	
Recessive	GG+GA	84 (75)	247 (77.9)	1.00	0.52
	AA	28 (25)	70 (22.1)	1.18 (0.71 - 1.96)	
Overdominant	GG+AA	54 (48.2)	175 (55.2)	1.00	0.19
	GA	58 (51.8)	142 (44.8)	1.33 (0.87 - 2.06)	
Log-additive	---	---	---	1.28 (0.95 - 1.73)	0.099
Smoking (adjusted by gender)					
Model 5	Genotype	LAAS-IS n (%)	Controls n (%)	OR (95% CI)	p-value

Codominant	GG	26 (27.1)	110 (31.6)	1.00	0.62
	GA	44 (45.8)	162 (46.5)	1.04 (0.59 - 1.81)	
	AA	26 (27.1)	76 (21.8)	1.34 (0.71 - 2.52)	
Dominant	GG	26 (27.1)	110 (31.6)	1.00	0.64
	GA+AA	70 (72.9)	238 (68.4)	1.13 (0.67 - 1.90)	
Recessive	GG+GA	70 (72.9)	272 (78.2)	1.00	0.33
	AA	26 (27.1)	76 (21.8)	1.31 (0.77 - 2.22)	
Overdominant	GG+AA	52 (54.2)	186 (53.5)	1.00	0.68
	GA	44 (45.8)	162 (46.5)	0.91 (0.57 - 1.45)	
Log-additive	---	---	---	1.15 (0.84 - 1.59)	0.38

LAAS-IS: Large Artery Atherosclerosis Ischemic Stroke; * Statistical significance.

TABLE 4 Comparison of the best models, prediction accuracies, cross-validation consistencies, and *p*-values identified by generalized multifactor dimensionality reduction analysis for CE-IS.

Model	Training balanced accuracy	Testing balanced accuracy	Sign test <i>p</i>-value
rs966221	0.5578	0.5431	5.42×10^{-3}
rs966221, rs152312	0.6099	0.6081	2.99×10^{-4}
rs2910829, rs966221, rs152312	0.6408	0.6156	4.30×10^{-5}
rs6843082, rs2910829, rs966221, rs152312	0.6738	0.6279	2.96×10^{-4}

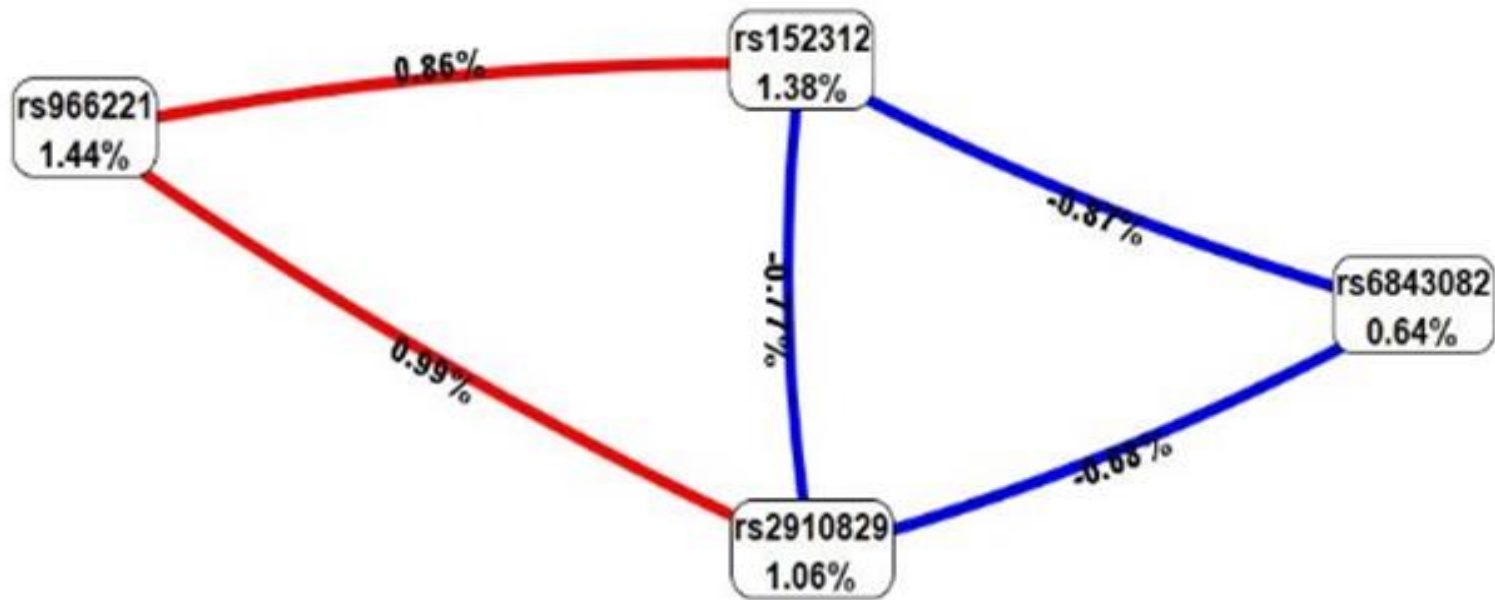


FIGURE 1 - Entropy Graph Circle among genetic variants related to CE-IS (red lines represent the synergistic interactions, while blue lines designate the antagonistic interactions)

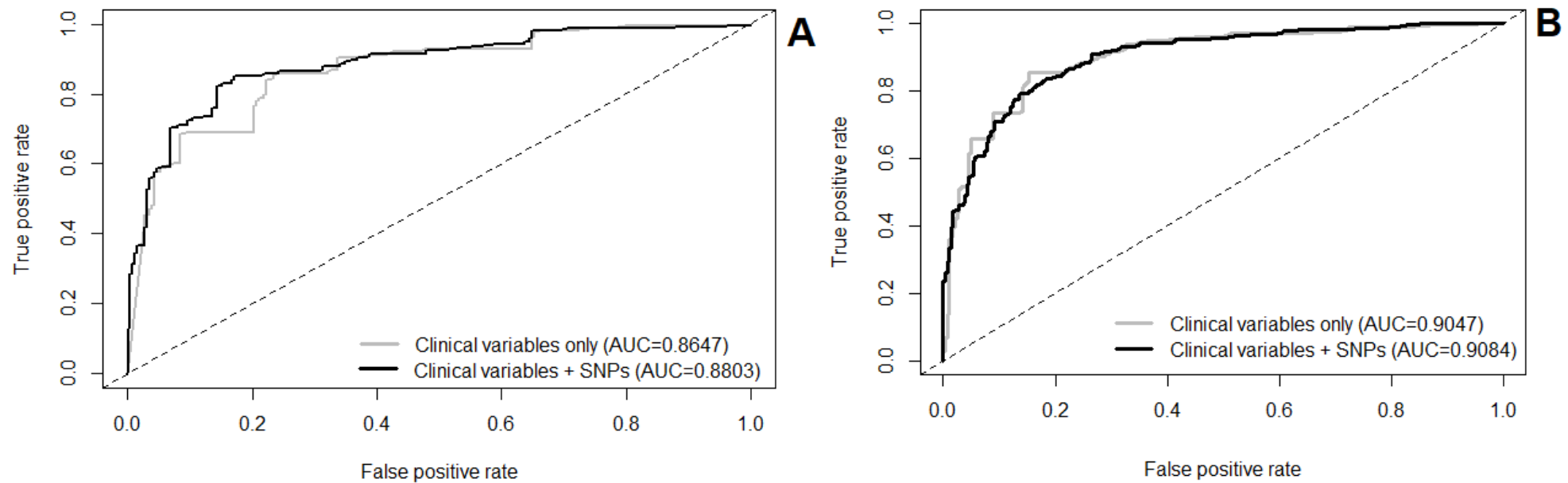


FIGURE 2 Receiver operating characteristic curves showing the accuracy of predicting LAAS-IS (A) and CE-IS (B). True positive rate and false positive rate indicate sensitivity and specificity, respectively. Solid grey lines indicate the scenario considering only clinical variables (hypertension, dyslipidemia, diabetes, and smoking); solid black lines indicate the scenario including clinical variables and genetic variants (rs2383207 for LAAS-IS; rs6843082, rs2910829, rs966221 and rs152312 for CE-IS); Diagonal dashed line indicates a non-informative prediction as control (AUC=0.5)

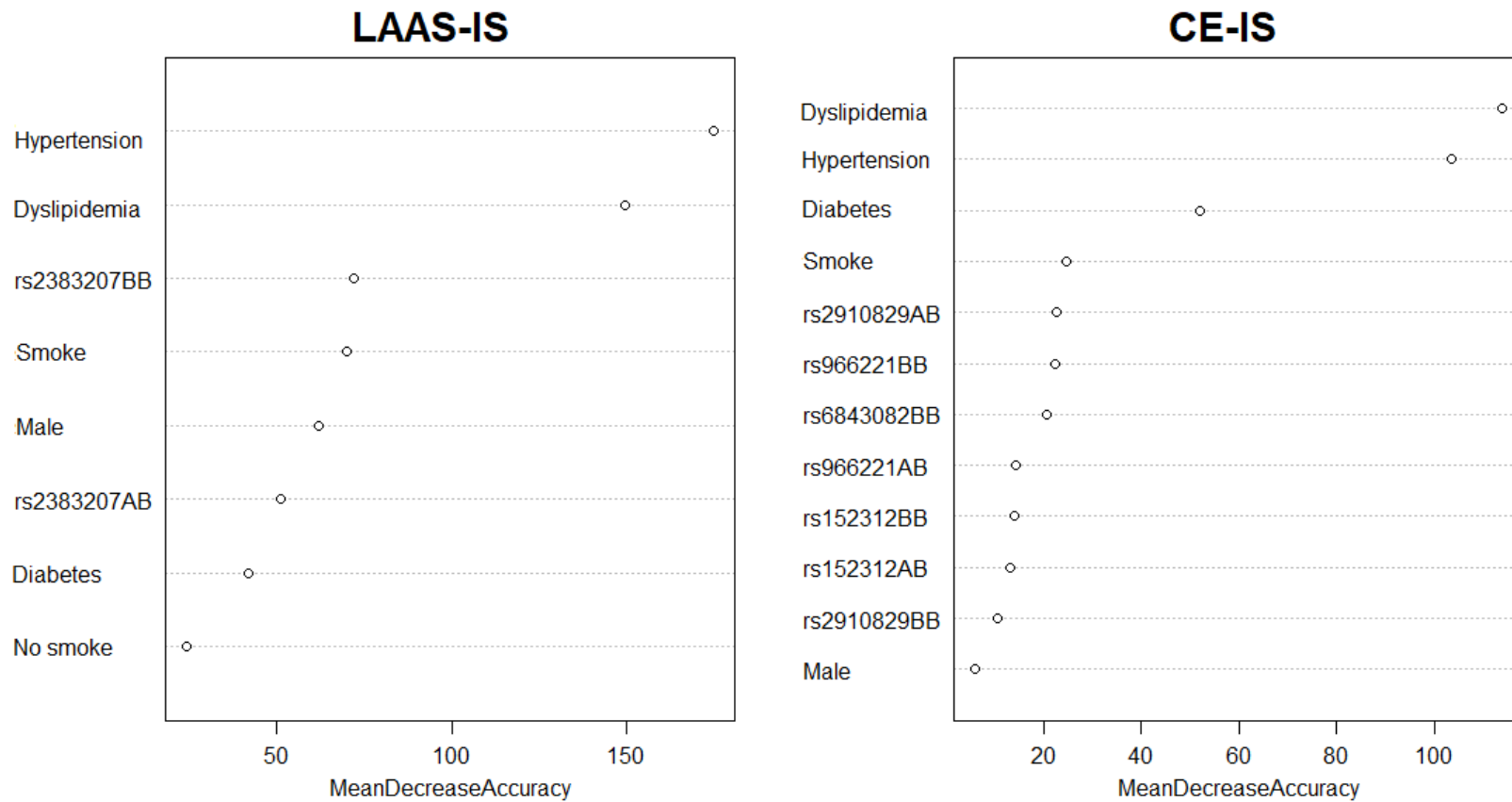


FIGURE 3 Variant importance plot for LAAS-IS and CE-IS. Each point represents the mean decreased accuracy estimate (x-axis), for each clinical variable or genotype selected by the model (y-axis). Letter codes AB and BB, after genetic variants names, indicate heterozygous and the less frequent homozygous genotypes for each variant, respectively.

7. CONCLUSÃO

Visando atingir os resultados propostos nesse estudo apresenta-se abaixo um sumário dos objetivos alcançados. Através da implementação do *Joinville Stroke Biobank* foi possível caracterizar a população incluída no biobanco o que nos auxiliará em futuras investigações genéticas para a compreensão dos fatores de risco para o AVC na população brasileira.

- A variante genética *rs2383207*, próxima ao gene *CDKN2B-AS1* demonstrou-se estar associado a AVCI subtipo aterotrombótico;
- Embora não tenha sido observada associação das variantes genéticas estudadas com AVCI subtipo cardioembólico, testes de interação gênica revelaram uma interação sinérgica entre as variantes *rs152312* e *rs966221*, e entre as variantes *rs966221* e *rs2910829*. Adicionalmente um efeito antagônico ou redundante foi observado entre as variantes *rs152312* e *rs2910829*, *rs152312* e *rs6843082*, e *rs2910829* e *rs6843082*.
- O modelo de predição demonstrou que hipercolestoremia e hipertensão são fatores preditivos para a ocorrência do AVCI subtipo cardioembólico.

8. REFERÊNCIAS

Ay H. **Advances in the diagnosis of etiologic subtypes of ischemic stroke.** Curr Neurol Neurosci Rep. 2010 Jan;10(1):14-20. doi: 10.1007/s11910-009-0074-x.

ADAMS, H. P. et al. **Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment.** Stroke, v.24, p.35-41, 1993.

AKINYEMI R, et al. **Interleukin-6 (IL-6) rs1800796 and cyclin dependent kinase inhibitor (CDKN2A/CDKN2B) rs2383207 are associated with ischemic stroke in indigenous West African Men.** J Neurol Sci. 2017 Aug 15;379:229-235.

ALBERTS, B. et al. **Biologia molecular da célula.** 4. ed. Porto Alegre: ARTMED, 2004.

ALEGRIA-TORRES, J. A., BACCARELLI, A., BOLLATI V. **Epigenomics.**v. 3, n.3, p. 267-277, 2011.

ANA LUIZA RESENDE GALRÃO **Metilação do gene simportador sódio-iodo (NIS) em tumores de tireóide.** Tese apresentada à Faculdade de Medicina da Universidade de São Paulo para obtenção do título de Doutor em Ciências Programa de Endocrinologia. (2011).

BAYOGLU, B. et al. **Polymorphisms in the long non coding RNA CDKN2B-AS1 may contribute to higher systolic blood pressures levels in hypertensive patients.** Clinical Biochemistry, n. 49, p. 821-827, 2016.

BELLENGUEZ C. et al. **Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke.** Nature genetics.v.44,p. 328–333, 2012.

BEVAN S. et al. **Phosphodiesterase 4D gene, ischemic stroke, and asymptomatic carotid atherosclerosis.** Stroke, v.36 p.949–953, 2005.

BIRD A.P. **Gene number, noise reduction and biological complexity.** Trends Genet. v.11 p.94, 1995.

BOEHME A.K., ESENWA C., ELKIND M.S. **Stroke Risk Factors, Genetics, and Prevention**. Circ Res. v. 3 n.120(3) p.472-495, 2017.

BOTELHO, T. S. et al. **Epidemiology of stroke in Brazil**. Temas em Saúde, v. 16, p. 361- 377, 2016.

Brown T.A. **Genomes**. 2nd edition. Oxford: Wiley-Liss; 2002. Chapter 1, The Human Genome.

CABRAL, N.L. et al. **Trends in stroke incidence, mortality and case fatality rates in Joinville, Brazil: 1995 2006**. Journal of Neurology, Neurosurgery and Psychiatry, v. 8, p. 749-754, 2009.

CANCELA D.M.G. **O acidente vascular cerebral – classificação, principais consequências e reabilitação**. [dissertação] Universidade Lusíada do Porto; 2008.

Caplan, L.R. **Stroke**. 1 ed, ed. A.J. Sumner. Vol. 1. 2006, New York: Demos Medical Publishing. 255

CHARLTON J, et al. **Cardiovascular diseases**. In: **Health of Adult Britain**. London: HMSO/Office for National Statistics, 1996.

CHAUHAN G. DEBETTE S. **Genetic Risk Factors for Ischemic and Hemorrhagic**. Stroke. Curr Cardiol Rep. 2016 Dec;18(12):124.

CHENG X., ROBERTS R.J. **AdoMet-dependent methylation, DNA methyltransferases and base flipping**. Nucleic Acids Res. v.29, p. 3784-3795, 2001.

CHEST AND HEART ASSOCIATION. **Modern views on stroke illness**. London: CHA, 1962.

CHOU S-Y. et al. **Genetic Susceptibility for Ischemic Infarction and Arteriolosclerosis based on Neuropathologic**. Evaluations.Cerebrovascular diseases. v.36 n.3 p.10, 2013.

CORREIA, A.L.F. **Factores geneticos de risco para acidente.** [Dissertacao de mestrado] Departamento de Biologia da Universidade de Aveiro, 2011.

DAVIDSON L.S.P. **The Principles and Practice of Medicine: a textbook for students and doctors,** eds 1-15. Edinburgh: E.and S. Livingstone Ltd, 1952-87.

DICHGANS, M. et al. **Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants.** Stroke v. 45, p. 24–36, 2014.

DONNAN G.A. et al. **Stroke.** Lancet. v.371 p.1612-1623, 2008.

FALCONE, G. et al. **Current concepts and clinical applications of stroke genetics.** The Lancet Neurology, v13, p. 405 - 418, 2014.

FAZZARI M.J., GREALLY J.M., **Epigenomics: beyond CpG islands.** Nature Rev Gen v.5, p.446-455, 2004.

FEINBERG A.P. **The epigenetics of cancer etiology.** Semin. Canc. Biol. (article in press), 2004.

GARRITANO, C. R. et al. **Análise da tendência da mortalidade por acidente vascular cerebral no Brasil no século XXI.** Arq. Bras. Cardiol. [online]. v.98, p. 519-527, 2012.

GOUVEIA, M. E. **Análise do padrão de metilação em genes supressores de tumor na leucemina mieloide cronica.** [Dissertacao de mestrado], Programa de Pos-Graduação em Neuroimunologia, Universidade Federal Fluminense, Niteroi, 2007.

GRETARSDOTTIR S. et al. **The gene encoding phosphodiesterase 4D confers risk of ischemic stroke.** Nat Genet. v.35 p.131–138, 2003.

GWAS Catalog: **The NHGRI-EBI Catalog of published genome-wide association studies;** Disponível em <https://www.ebi.ac.uk/gwas/>. Acesso em 21 de Agosto de 2017.

HAHN L.W., RITCHIE M.D., MOORE J.H. **Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions.** *Bioinformatics*.v.19(3) p.376-82, 2003.

HAN Y. et al. **HDAC9 gene is associated with stroke risk in a Chinese population.** *Exp Biol Med* (Maywood). 2013 Jul;238(7):842-7.

HANDY, D. E., CASTRO, R. & LOSCALZO, J. **Epigenetic modifications: basic mechanisms and role in cardiovascular disease.** *Circulation*, v. 123, p. 2145-56, 2011.

HASTIE T., TIBSHIRANI R., FRIEDMAN J.H. **The elements of statistical learning: data mining, inference, and prediction.** 2nd ed. New York, NY: Springer; 2009.

HELGADOTTIR A. et al. **The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm.** *Nat Genet* 2008;40: 217-224

HOUSLAY M. D., ADAMS D. R. **PDE4 CAMP Phosphodiesterases: Modular enzymes that orchestrate signalling cross-talk, desensitization and compartmentalization.** *Biochem. J.* 370:1–18, 2003.

HUANG Y. et al. **Association of CDKN2BAS polymorphism rs4977574 with coronary heart disease: a case-control study and a meta-analysis.** *Int J Mol Sci*, v. 29 p.17478-92, 2014.

HUGUES, M., LIP, G. Y. H. **Stroke and thromboembolism in atrial fibrillation: A systematic review of stroke risk factors, risk stratification schema and cost effectiveness data.** *Thrombosis and Haemostasis*, v. 99, p. 295-304, 2008.

ICHIKI T. **Role of cAMP response element binding protein in cardiovascular remodeling: good, bad, or both?** *Arterioscler. Thromb. Vasc. Biol.* 26:449–455, 2006.

INTERNATIONAL STROKE GENETICS CONSORTIUM (ISGC) &WELLCOME TRUST CASE CONTROL CONSORTIUM 2 (WTCCC2). **Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke.** Nature genetics. v.44 n.3 p.328-333, 2012.

KOTLEGA, D. et al. **Association between selected gene polymorphisms and statin metabolism, risk of ischemic stroke and cardiovascular disorders.** Postepy Higieny I Medycyny Doswiadczalnej, v. 70, p. 435-447, 2016.

KREMER, P.H., et al. **Evaluation of genetic risk loci for intracranial aneurysms in sporadic arteriovenous malformations of the brain.** J Neurol Neurosurg Psychiatry. v.86, p. 524-529, 2015.

KUNICKI, T. J. **The Influence of Platelet Collagen Receptor Polymorphisms in Hemostasis and Thrombotic Disease.** Arterioscler. Thromb. Vasc. Biol. v.22, p.14 20, 2002.

LAWRENCE C. **Medicine in the making of modern Britain 00-1920.** London: Routledge, 1994.

LIANG W. et al. **Association between Phosphodiesterase 4D (PDE4D) SNP 87 and Ischemic Stroke: A Meta-Analysis.** International Journal of Clinical and Experimental Medicine v.8.2 p. 1715–1725, 2015.

LINDGREN A., MAGUIRE J. **Stroke Recovery Genetics.** Stroke: v.47 n.9 p.2427-34, 2016.

LIPPI G. FRANCHINI M. TARGHER G. **Arterial thrombus formation in cardiovascular disease.** Nature Reviews Cardiology; v.8 p.502 – 12, 2011.

LIU D. et al. **Genetic Variations of Oxidative Stress Related Genes ALOX5, ALOX5AP and MPO Modulate Ischemic Stroke Susceptibility Through Main Effects and Epistatic Interactions in a Chinese Population.** Cell Physiol Biochem. v.43 n.4:1588-1602, 2017.

LORENOWICZ M. J., FERNANDEZ-BORJA M., HORDIJK P. L. **cAMP signaling in leukocyte transendothelial migration.** *Arterioscler. Thromb. Vasc. Biol.* v. 27 p.1014–1022, 2007.

LOTUFO, P.A. **Stroke in Brazil: a neglected disease.** *São Paulo Med J* v.1, p.3-4, 2005.

LUBITZ, S.A. et al. **Novel genetic markers associate with atrial fibrillation risk in Europeans and Japanese.** *J Am Coll Cardiol.* v.63, p.1200–10, 2014.

MACARTHUR J. et al. **The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog).** *Nucleic Acids Research,* v.45 p. D896-D901, 2017.

MALIK R. et al. **Common coding variant in SERPINA1 increases the risk for large artery stroke.** *Proc Natl Acad Sci U S A.,* v.114 n.14 p.3613-3618, 2017.

MALIK R. et al. **Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration.** *Neurology.*v.29,86 n.13 p.1217-26, 2016. Erratum in: *Neurology.* v.20;87 n.12 p:1306, 2016.

MANNHALTER, C. **Biomarkers for arterial and venous thrombotic disorders.** *Hamostaseologie* v. 34, p.115 20, 2014.

MANSO, H. VICENTE, A. M. **Interação entre fatores genéticos e não genéticos na recuperação após um Acidente Vascular Cerebral** *Boletim Epidemiológico do Instituto Nacional de Saúde Doutor Ricardo Jorge* n. 7, p. 11–14, 2013.

MARKUS, H. S. **Stroke genetics: prospects for personalized medicine.** *BMC Medicine,* v.10, n.1,p. 113, 2012.

MARTIN M., KETTMANN R., DEQUIEDT F. **Class IIa histone deacetylases: conducting development and differentiation.** *Int J Dev Biol.* v.53 n.2-3 p.291-301, 2009.

MARTIN R.I. et al. **Chromosome 16q22 variants in a region associated with cardiovascular phenotypes correlate with ZFH3 expression in a transcript-specific manner.** BMC Genet. v.24 p.15:136,2014.

MATARIN M. et al. **Whole genome analyses suggest ischemic stroke and heart disease share an association with polymorphisms on chromosome 9p21.** Stroke.v.39: 1586-1589, 2008.

MATSUSHITA T. et al. **Lack of association between variations of PDE4D and ischemic stroke in the Japanese population.** Stroke.v.40 p.1245–5, 2009.

MINISTERIO DA SAUDE/SE/DATASUS - **Sistema de Informações Hospitalares do SUS SAI. Número total de internações por doença.** Disponível em: http://portal.saude.gov.br/portal/sas/area.cfm?id_area=875; Acessado em Julho, 2014.

MOHANTY S. et al. **Novel association of polymorphic genetic variants with predictors of outcome of catheter ablation in atrial fibrillation: new directions from a prospective study (DECAF).** J Interv Card Electrophysiol, 2015.

MOMMERSTEEG M.T. et al. **PITX2c and Nkx2-5 are required for the formation and identity of the pulmonary myocardium.** Circ Res v.101:902–909, 2007.

MORO, C.H.C. et al. **Trends of the Incidence of Ischemic Stroke Thrombolysis over Seven Years and One-Year Outcome: A Population-Based Study in Joinville, Brazil.** Cerebrovascular Diseases Extra, v. 3, p. 156-166, 2013.

MUNSHI, A. et al. **Association of SNP41, SNP56 and a novel SNP in PDE4D gene with stroke and subtypes.** Gene, p.31-35, 2012.

NINDS Stroke Genetics Network (SiGN) & International Stroke Genetics Consortium (ISGC). **Loci associated with ischaemic stroke and its subtypes**

(SiGN): a genome-wide association study. Lancet Neurol. v.15. p.175-184, 2015.

'DONNELL M.J., et al. **Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study.** Lancet. v.388 n.10046 p.761-75, 2016

OHIRA T., et al. **Risk factors for ischemic stroke subtypes: the Atherosclerosis Risk in Communities study.** Stroke. v.37 n.10 p.2493-8, 2006.

OpenStax College, Anatomy & Physiology, Connexions Web site. Disponível em <http://cnx.org/content/col11496/1.6/>. Acesso em 21 de agosto de 2017.

OU DOT-MELLAKH, T. et al. **Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project.** Br. J. Haematol. v.157 p.230–9, 2012.

PEZZINI, A. **Genetic determinants of juvenile stroke.** Thromb. Res., v.129, p.330 5, 2012.

PIRES, S. L.; GAGLIARDI, R. J.; GORZONI, M. L. **Estudo das frequências dos principais fatores de risco para acidente vascular cerebral isquêmico em idosos.** Arquivos de Neuropsiquiatria, v. 62, n. 3-B, p. 844-851, 2004.

POUND, P., BURY, M., & EBRAHIM, S. **From apoplexy to stroke.** Age and Ageing, v.26 p. 331–337, 1997.

QUENTIEN MH, et al. **Truncation of PITX2 differentially affects its activity on physiological targets.** J Mol Endocrinol. v.46 n. 1 p. 9-19, 2010.

RANNIKMÄE K. et al. **COL4A2 is associated with lacunar ischemic stroke and deep ICH: Meta-analyses among 21,500 cases and 40,600 controls.** Neurology. v. 24,89 n.17 p.1829-1839, 2017.

RODOLPHO, M. **De Physiognomonia Liber: considerações a respeito do ethos e da fisiognomonia em textos da Antiguidade Clássica,** Tese

(Doutorado em Letras Clássicas e Vernáculas) Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo. São Paulo, p. 457. 2015.

RUSH L.J., PLASS C. **Alterations of DNA methylation in hematologic malignancies.** Cancer Letters v.185 p. 1-12, 2002.

RUTTEN-JACOBS L.C. et al. **Association of MTHFR C677T Genotype With Ischemic Stroke Is Confined to Cerebral Small Vessel Disease Subtype.** Stroke. v.47 n.3 p.646-51, 2016.

SCHLEGEL D. et al. Cucchiara BL, Liebeskind DS, et al. **Utility of the NIH Stroke Scale as a predictor of hospital disposition.** Stroke. v.34 n.1 p.134-7.

SILVA, R. C. A., MONTEIRO, G. L., SANTOS, A. G. **O enfermeiro na educação de cuidadores de pacientes com sequelas de acidente vascular cerebral.** Rev. de Atenção à Saúde, v.13, n.45, p. 114-120, 2015.

SINGAL R., GINDER G.D. **DNA methylation.** Blood, v.93, n.2, p.4059-4070, 1999.

SNPS Markers, 2017 – Disponível em <http://www.traitgenetics.com>; Acesso em 21 de ago.de 2017.

SU L, et al. **Association of GWAS-supported loci rs2107595 in HDAC9 gene with ischemic stroke in southern Han Chinese.** Gene, v. 570, n. 2, p. 282–287, 2015

TINTINALLI J.E. et al. **Emergency Medicine: A Comprehensive Study Guide**, 7th Edition: <http://accessmedicine.com>

TRAYLOR M. et al. **Genetic variation at 16q24.2 is associated with small vessel stroke.** Ann Neurol. v.81 n.3 p.383-394, 2017.

TRAYLOR, M. et al. **Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies.** Lancet Neurol. v.11 p.951–962, 2012.

WANG P. et al. **Association between PDE4D rs966221 polymorphism and risk of ischemic stroke: a systematic review and meta-analysis.** *Metab Brain* p.1-9, 2017.

WENINGER S. et al. **Phosphodiesterase 4 interacts with the 5-HT4(b) receptor to regulate cAMP signaling.** *Cell Signal.* v. 26 n.11 p.2573-82, 2014.

WHO. **Risk factors. Atlas Heart.** *Dis. Stroke*, v.15, p.24 25, 2004.

WILLIAMS S.R. et al. **Shared genetic susceptibility of vascular-related biomarkers with ischemic and recurrent stroke.** *Neurology.* v.26;86 n.4 p.351-9, 2016.

WNUK M. et al. **The rs2200733 variant on chromosome 4q25 is a risk factor for ardioembolic stroke related to atrial fibrillation in Polish patients.** *Neurol Neurochir Pol.* v.45 n.2 p.148-52, 2011.

WOO D. et al. **Association of Phosphodiesterase 4D with ischemic stroke - A population-based case-control study.** *Stroke; a journal of cerebral circulation*, v.37 p.371–376, 2006.

WU J.C., SANTI D.V. **On the mechanism and inhibition of DNA cytosine methyltransferases.** *Prog Clin Biol Res* v.198, p.119-129, 1985.

XU X. et al. **Meta-analysis of association between variation in the PDE4D gene and ischemic cerebral infarction risk in Asian populations.** *Neurogenetics.* v.11 n.3 p.327-33, 2010.

XUE Y. et al. **Genetic variant in ZFH3 Geneon 16q22 associated with risk of stroke in Chinese Han population,** *International Journal of Clinical and Experimental Pathology*; v. 9 8; 8650-p8656, 2016.

YI X. et al. **Interaction among CYP2C8, GP1IIa and P2Y12 variants increase susceptibility to ischemic stroke in Chinese population.** *Oncotarget.* v.8 n.41 p.70811-70820,2017.

ZHOU X. et al. **The association between HDAC9 gene polymorphisms and stroke risk in the Chinese population: A meta-analysis.** Scientific Reports. v.7 p.41538, 2017.

ZHU H. et al. **Long non-coding RNA ANRIL is up-regulated in bladder cancer and regulates bladder cancer cell proliferation and apoptosis through the intrinsic pathway.** Biochem Biophys Res Commun. v.467 n.2 p.223-8, 2015.

ANEXO A – Parecer consubstanciado do Comitê de Ética



UNIVERSIDADE DA REGIÃO
DE JOINVILLE UNIVILLE



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: ANÁLISE DE VARIANTES GENÉTICAS E INTERAÇÕES GÊNICAS CORRELACIONADAS AO RISCO DO AVC ISQUÊMICO.

Pesquisador: LESLIE ECKER FERREIRA

Área Temática: Genética Humana:

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP.);

Versão: 2

CAAE: 51144715.9.0000.5366

Instituição Proponente: FUNDACAO EDUCACIONAL DA REGIAO DE JOINVILLE - UNIVILLE

Patrocinador Principal: FUNDACAO EDUCACIONAL DA REGIAO DE JOINVILLE - UNIVILLE

DADOS DO PARECER

Número do Parecer: 1.396.172

Apresentação do Projeto:

Como exposto em parecer consubstanciado nº 1.364.463.

Objetivo da Pesquisa:

Como exposto em parecer consubstanciado nº 1.364.463.

Avaliação dos Riscos e Benefícios:

Como exposto em parecer consubstanciado nº 1.364.463.

Comentários e Considerações sobre a Pesquisa:

Como exposto em parecer consubstanciado nº 1.364.463.

Considerações sobre os Termos de apresentação obrigatória:

A Folha de Rosto apresentada está completa.

Recomendações:

Não se aplica.

Conclusões ou Pendências e Lista de Inadequações:

O projeto "ANÁLISE DE VARIANTES GENÉTICAS E INTERAÇÕES GÊNICAS CORRELACIONADAS AO RISCO DO AVC ISQUÊMICO", sob CAAE LESLIE ECKER FERREIRA teve suas pendências esclarecidas

Endereço: Rua Paulo Malschitzki, nº 10. Bloco B, Sala 17.

Bairro: Zona Industrial

CEP: 89.219-710

UF: SC

Município: JOINVILLE

Telefone: (47)3461-9235

E-mail: comitetica@univille.br



UNIVERSIDADE DA REGIÃO
DE JOINVILLE UNIVILLE



Continuação do Parecer: 1.396.172

pelo (a) pesquisador(a) 51144715.9.0000.5366, de acordo com a Resolução CNS 466/12 e complementares, portanto, encontra-se APROVADO.

Informamos que após leitura deste parecer, é imprescindível a leitura do item "O Parecer do CEP" na página do Comitê no sítio da Univille, pois os procedimentos seguintes, no que se refere ao enquadramento do protocolo, estão disponíveis na página. Segue o link de acesso (<http://community.univille.edu.br/cep/statusparecer/577374>).

Considerações Finais a critério do CEP:

Diante do exposto, o Comitê de Ética em Pesquisa da Universidade da Região de Joinville - Univille, de acordo com as atribuições definidas na Res. CNS 466/12, manifesta-se pela aprovação do projeto de pesquisa proposto.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_622064.pdf	22/12/2015 11:17:20		Aceito
Folha de Rosto	folha_de_rosto_versao02.pdf	22/12/2015 11:15:36	LESLIE ECKER FERREIRA	Aceito
Outros	TTMB.pdf	19/11/2015 19:50:18	LESLIE ECKER FERREIRA	Aceito
Outros	declaracaoanuencia.pdf	19/11/2015 19:48:35	LESLIE ECKER FERREIRA	Aceito
Projeto Detalhado / Brochura Investigador	projetocep.docx	19/11/2015 19:40:22	LESLIE ECKER FERREIRA	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	tcle.pdf	06/11/2015 17:41:06	LESLIE ECKER FERREIRA	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Endereço: Rua Paulo Malschitzki, n° 10. Bloco B, Sala 17.

Bairro: Zona Industrial

CEP: 89.219-710

UF: SC

Município: JOINVILLE

Telefone: (47)3461-9235

E-mail: comitetica@univille.br



UNIVERSIDADE DA REGIÃO
DE JOINVILLE UNIVILLE



Continuação do Parecer: 1.396.172

JOINVILLE, 27 de Janeiro de 2016

Assinado por:
Eleide Abril Gordon Findlay
(Coordenador)

Endereço: Rua Paulo Malschitzki, n° 10. Bloco B, Sala 17.
Bairro: Zona Industrial **CEP:** 89.219-710
UF: SC **Município:** JOINVILLE
Telefone: (47)3461-9235 **E-mail:** comiletica@univille.br

AUTORIZAÇÃO

Nome do autor: Leslie Ecker Ferreira

RG: 3128303

Título da Tese: ANÁLISE DE VARIANTES GENÉTICAS E INTERAÇÕES
GÊNICAS QUANTO AO RISCO DO AVC ISQUÊMICO

Autorizo a Universidade da Região de Joinville – UNIVILLE, através da Biblioteca Universitária, disponibilizar cópias da Tese de minha autoria.

Joinville, 15 de maio de 2018.

Leslie Ecker Ferreira.

Assinatura do aluno