



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Relação entre o grau de lesão cardíaca e níveis de expressões gênicas, expressões proteicas e ácidos graxos.

Mariana Fehr Nicácio

Projeto apresentado ao Departamento de Estatística da Universidade de Brasília para a obtenção do título de Bacharel em Estatística.

Brasília
2015

Mariana Fehr Nicácio

Relação entre o grau de lesão cardíaca e níveis de expressões gênicas, expressões proteicas e ácidos graxos.

Projeto apresentado ao Departamento de Estatística da Universidade de Brasília para a obtenção do título de Bacharel em Estatística.

Orientadora: Profa. Dra. **Joanlise Marco de Leon Andrade**

Brasília
2015

Dedicatória

Dedico este trabalho em especial aos meus avós, por todo amor que sempre me dedicaram, e aos familiares, por todos os ensinamentos e apoio durante a minha longa caminhada...

Agradecimentos

Consigno meu perpétuo reconhecimento ao estímulo recebido da Universidade de Brasília, através de seus professores, quais tive o prazer de ser aluna, e especialmente à orientadora do presente trabalho, Professora Dra. Joanlise Marco de Leon Andrade e Professora Mestre Maria Teresa Leão Costa, que muito contribuíram para o meu engrandecimento humano e profissional. Agradeço, também, ao professor Dr. George Freitas von Borries por fazer parte da Banca Examinadora desse trabalho.

Devo à minha fé, no Deus que acredito, toda a persistência e força para a realização do desejo de obtenção do título de bacharel em Estatística. Agradeço a minha querida avó senhora Maria Bartira, a minha tia Angélica Lott, a minha irmã Thaís Fehr e minha Mãe Simone Fehr, por fazerem parte dessa jornada e por sempre me lembrarem que de nada vale o sacrifício sem ter com quem compartilhar a vitória.

Agradeço aos amigos que fiz ao longo do curso de Estatística, em especial à Geiziane, Andressa, Luana, Márcia, Agda, Arthur, Victor, Nathália, Guilherme e Matheus, pois, em muito contribuíram para a minha formação e aos meus demais amigos que me acompanharam em diversos momentos dessa caminhada.

Aos colaboradores que estiveram envolvidos no recrutamento de participantes e coleta dos dados utilizados no presente trabalho, professores Dra. Vivian Nogueira Silbiger, Dr. André Ducati Luchessi, Dr. Adriana Augusto Rezende, Dr. Mario Hiroyuki Hirata e Dra. Rosário Domingues Crespo Hirata. Aos médicos Dra. Maria Sanali Moura de Oliveira Paiva e Dr. Itamar Ribeiro de Oliveira. Aos alunos de mestrado Mychelle Kytchia Rodrigues Nunes Duarte e Victor Hugo Augusto Rezende. Aos alunos de iniciação científica da Universidade Federal do Rio Grande do Norte Ayda Maria Quirino Silva dos Santos, Jéssica Nayara Góes de Araújo, Jéssica Cavalcante dos Santos, Katiene Macêdo de Oliveira, Letícia da Costa Araújo, Nathalia Ribeiro Calazans Brandão da Silva e da Universidade Federal de Campina Grande Juliana Marinho de Oliveria.

Resumo

Neste trabalho foram avaliadas as relações de expressões gênicas, expressões proteicas e ácidos graxos com o escore de Friesinger. Com a finalidade de identificar potenciais biomarcadores de aterosclerose. Avaliou-se treze genes, quatro proteínas e dez ácidos graxos, bem com a relação entre o escore de Friesinger e Framingham e os fatores clássicos de risco. O Brasil está entre os 10 países com maiores taxas de mortalidade por doenças cardiovasculares (DCVs) no mundo. Representam coletivamente a principal causa de morte no Brasil. As DCVs constituem uma das principais causas de permanência hospitalar e são responsáveis pela principal alocação de recursos públicos. A maior parte das DCVs é causada pela aterosclerose, doença caracterizada pelo acúmulo de placas de gorduras nas artérias ao longo dos anos. Quando há obstrução das artérias, ocorre isquemia, ou seja, morte celular devido à diminuição do suprimento de oxigênio necessário. O Infarto agudo do miocárdio (IAM) é a doença isquêmica do coração mais comum. Acredita-se que a aterosclerose, e as DCVs de modo geral, sejam causadas por interações entre fatores genéticos, ambientais, demográficos e/ou comportamentais. Fatores de risco amplamente estabelecidos na literatura incluem sexo masculino, idade avançada, hipertensão, diabetes, obesidade, dislipidemia, sedentarismo, tabagismo e histórico familiar. O estudo de marcadores da presença e severidade da aterosclerose é de importância indiscutível. A identificação de biomarcadores menos invasivos e de fácil execução é fundamental para o aprimoramento na conduta clínica, inclusive na detecção de subgrupos de pacientes com aterosclerose precoce. Muitos estudos vêm sendo desenvolvidos a fim de se identificar genes, proteínas e outros componentes moleculares que possam atuar como marcadores de aterosclerose, possibilitando prever o seu desenvolvimento e até mesmo diferenciar situações clínicas antes indistinguíveis. Muitos genes já foram identificados como fatores de risco, geralmente de efeitos moderados. O estudo da expressão gênica e proteica, a partir dos tecidos periféricos sanguíneos, demonstram ser potenciais ferramentas para a detecção de marcadores do processo inflamatório e outros indicadores de lesão. A atuação de ácidos graxos para a manutenção de funções cardíacas e sua desregulação em eventos cardíacos isquêmicos é bastante conhecida na literatura. Portanto, expressões gênicas, expressões proteicas e concentrações de ácidos graxos podem atuar como indicadores de lesão car-

díaca. O escore de Framingham apresentou uma correlação moderada porém significativa com o escore de Friesinger. Destacou-se como sendo uma das variáveis mais significantes apresentadas nos modelos de regressão. Observou-se correlação moderada de Friesinger positiva com os genes ALOX15, MYL4, COX7B, BCL2A1, IL18R1 e MMP9 e com o ácido Gamalinoleico, sendo significativa para o modelo de regressão múltipla proposto apenas o gene ALOX15. Destaca-se o ácido Gamalinoleico em tal estudo pois foi o ácido que apresentou menor p-valor na correlação. Os demais ácidos não apresentaram associação. As variáveis sexo, IMC e triglicérides não se mostraram significantes. Destaca-se no modelo linear generalizado as variáveis gene BCL2A1, COX7B, ácido Gamalinoleico, Framingham, LDL e Idade como sendo significativas para explicar a variabilidade do escore de Friesinger. Um fator importante que justificaria o ajuste não tão adequado são as 78 observações deletadas por falta de informação. Vários estudos evidenciam a importância de se conhecer novos marcadores para DCV precoce e marcadores de aterosclerose menos invasivos. Novos estudos são necessários para elucidar o papel das expressões gênicas, proteicas e ácidos graxos no desenvolvimento da aterosclerose.

Palavras-chave: Doença Cardiovascular, Aterosclerose, Friesinger Framingham, Modelo Linear Múltiplo, Modelo Linear Generalizado, Poisson, Expressões Gênicas, Expressões Proteicas, Ácidos Graxos.

Sumário

1	Introdução	1
		1
2	Metodologia	5
2.1	Participantes	5
2.1.1	Métodos laboratoriais	7
2.2	Análises Estatísticas	9
2.2.1	Modelo de Regressão Linear Múltipla	9
2.2.2	Análise de Variância	10
2.2.3	Análise dos resíduos	12
2.3	Modelos lineares generalizados	13
2.3.1	Critério de seleção do modelo	13
3	Resultados e Discussão	15
3.1	Análise Descritiva	15
3.1.1	Coefficiente de correlação de Pearson	19
3.2	Regressão Linear múltipla	25
3.2.1	Adequação do modelo	26
3.3	Modelo linear generalizado	28
4	Considerações Finais	31
	Referências Bibliográficas	33
A	Gráficos e Tabela adicionais	39
A.1	Estimativas de Coeficientes de Correlação de Pearson entre expressões gênicas e expressões proteicas	39
A.2	Estimativas de Coeficientes de Correlação de Pearson entre níveis de ácidos graxos	40
A.3	Gráfico de diagnóstico para não linearidade das variáveis dos modelos propostos	41

A.4 Tabela completa das estimativas de Coeficientes de Correlação de Pearson e P-valores	43
---	----

Lista de Figuras

3.1	Gráfico das estimativas dos coeficientes de correlação de Pearson das variáveis de expressões gênicas, expressões proteicas e níveis de ácidos graxos.	19
3.2	Gráfico dos histogramas na diagonal, estimativas dos coeficientes de correlação de Pearson entre os pares acima da diagonal e o diagrama da dispersão bivariada abaixo da diagonal	23
3.3	Gráfico dos histogramas na diagonal, estimativas dos coeficientes de correlação de Pearson entre os pares de expressão gênica e ácidos acima da diagonal e o diagrama da dispersão bivariada abaixo da diagonal	24
3.4	Gráfico dos resíduos para o modelo com os fatores clássicos de risco	26
3.5	Gráfico da distância de Cook	27
3.6	Gráfico dos resíduos para o modelo com ALOX15 e LDL	27
3.7	Gráfico da distância de Cook	28
3.8	Gráfico dos resíduos para o modelo linear generalizado com Distribuição de Poisson	30
3.9	Envelope simulado para os componentes do desvio no modelo de Poisson	30
A.1	Gráfico das correlações	39
A.2	Gráfico das correlações	40
A.3	Gráfico dos resíduos parciais e as variáveis do modelo linear com os fatores clássicos de risco	41
A.4	Gráfico dos resíduos parciais e as variáveis do modelo linear com a variável gene ALOX15	41
A.5	Gráfico dos resíduos parciais e as variáveis do modelo linear generalizado	42

Lista de Tabelas

2.1	Coeficientes da regressão e Razões de chance	7
3.1	Características demográficas e clínicas dos participantes	17
3.2	Médias e desvios padrão de variáveis laboratoriais dos participantes	18
3.3	Tabela das estimativas da correlação de Pearson que possuem p-valor $< 0,1$	20
3.4	Tabela das estimativas da correlação de Pearson que possuem p-valor $< 0,1$	21
3.5	Modelo para Friesinger com alguns fatores de risco conhecidos	25
3.6	Possível Modelo para Friesinger com gene e LDL	25
3.7	Modelo para Friesinger com alguns fatores de risco conhecidos	29
A.1	Tabela das estimativas da correlação de Pearson	43
A.2	Tabela das estimativas da correlação de Pearson	44
A.3	Tabela das estimativas da correlação de Pearson	45
A.4	Tabela das estimativas da correlação de Pearson	46
A.5	Tabela das estimativas da correlação de Pearson	47
A.6	Tabela das estimativas da correlação de Pearson	48
A.7	Tabela das estimativas da correlação de Pearson	49
A.8	Tabela das estimativas da correlação de Pearson	50

Capítulo 1

Introdução

O Brasil está entre os 10 países com maiores taxas de mortalidade por doenças cardiovasculares (DCVs) no mundo. Representam coletivamente a principal causa de morte no Brasil. Segundo a Organização Mundial da Saúde, em 2010 cerca de 30% das mortes no Brasil se deram em decorrência de problemas cardíacos, sendo quase um terço devido especificamente a doenças isquêmicas do coração [24]. As DCVs constituem uma das principais causas de permanência hospitalar prolongada e são responsáveis pela principal alocação de recursos públicos em hospitalizações no Brasil. O custo anual com os casos de DCVs "graves" é significativo. Estima-se que pelo menos R\$ 30,8 bilhões sejam destinados, direta ou indiretamente para o tratamento dessas DCVs para os próximos anos, sendo 36,4% para a área da saúde, 8,4% para o seguro social e reembolso por empregadores e 55,2% como resultado de perda da produtividade[19]. O ônus econômico tem crescido nas últimas décadas, o que leva a um aumento dos gastos no setor de saúde. Entre 1991 e 2000, os custos hospitalares atribuídos às DCVs aumentaram cerca de 176% [18].

A maior parte das DCVs é causada pela aterosclerose, doença caracterizada pelo acúmulo de placas de gorduras nas artérias ao longo dos anos, o que desencadeia um processo inflamatório. Quando há obstrução das artérias, em geral por ruptura e deslocamento de placas ateroscleróticas, ocorre isquemia (interrupção de fluxo sanguíneo), ocasionando um infarto, ou seja, morte celular devido à diminuição do suprimento de oxigênio necessário [24]. O Infarto agudo do miocárdio (IAM) é a doença isquêmica do coração mais comum [23].

Acredita-se que a aterosclerose, e as DCVs de modo geral, sejam causadas por interações entre fatores genéticos, ambientais, demográficos e/ou comportamentais. Fatores de risco amplamente estabelecidos na literatura incluem sexo masculino, idade avançada, hipertensão, diabetes, obesidade, dislipidemia, sedentarismo, tabagismo e histórico familiar. Muitos genes já foram identificados como fatores de risco, geralmente de efeitos moderados [12]. O conhecimento desses fatores associados ao risco é de grande importância para se estabelecer estratégias de prevenção. Consultas médicas regulares são essenciais para o controle de pressão arterial, colesterol, orientação nutricional e avaliação

física.

Por se tratarem de doenças crônicas, geralmente observa-se um período longo entre o desenvolvimento da aterosclerose e o surgimento dos primeiros sintomas de DCVs, o que dificulta o estabelecimento de medidas preventivas e um tratamento precoce da doença. Quando detectadas em sua fase inicial, o tratamento primário e específico é o ideal. Nesse contexto, o estudo de marcadores da presença e severidade da aterosclerose é de importância indiscutível.

Um estudo realizado no Hospital do Coração (HCor) em São Paulo apontou uma preocupante tendência: recentemente, jovens entre 20 e 40 anos estão tendo mais problemas cardiovasculares, como infartos. Segundo o Dr. Ricardo Pavanello, supervisor de cardiologia do HCor e autor do estudo, os casos nesta faixa etária já representam, em média, 12% do total. Há dez anos, esse número não passava de 6%. As razões para esse crescimento, segundo o médico, incluem o aumento no estresse associado ao fumo e a outros fatores de risco, como peso acima do ideal. Tal estudo corrobora a necessidade de identificação de métodos para a detecção precoce de aterosclerose e de previsão do risco de DCVs. Idealmente, o uso de tais ferramentas permitiria uma abordagem individualizada envolvendo um tratamento mais efetivo, eficaz e seguro e com o um menor custo de realização [24].

O escore de Framingham, proposto em 1948, tem sido utilizado por cardiologistas como estimador de risco de eventos de DCVs por ser facilmente calculado com base na ponderação de fatores clássicos de risco, coletados e registrados com mais facilidade. Para se obter o score de Framingham são coletadas informações sobre sexo, idade, pressão arterial, colesterol HDL, colesterol total, tratamento de hipertensão, presença de diabetes e tabagismo. Tal escore vem sendo utilizado também em indivíduos que não apresentam DCVs como uma forma de prevenção primária, em que o risco absoluto estimado é usado para definir um limiar para intervenção [10].

O escore de Friesinger é um marcador de lesão aterosclerótica. Consiste na avaliação e quantificação de placas ateroscleróticas via cateterismo cardíaco comumente medido em pacientes considerados de alto risco. Para a realização de tal procedimento são inseridos cateteres em vasos sanguíneos que são guiados até o coração. É possível observar e pontuar o grau de lesão em artérias coronárias, válvulas cardíacas e câmaras dos pacientes. O escore de Friesinger foi conceituado como um marcador do nível de lesão das placas ateroscleróticas para um diagnóstico mais preciso das condições cardíacas, porém, por necessitar da realização de cateterismo para a sua quantificação é considerado invasivo e, portanto, recomendado apenas para indivíduos com risco de DCVs [12].

Portanto, a identificação de biomarcadores menos invasivos e de fácil execução é fundamental para o aprimoramento na conduta clínica, inclusive na detecção de subgrupos de pacientes com aterosclerose precoce. Assim, muitos estudos vêm sendo desenvolvidos a fim de se identificar genes, proteínas e outros componentes moleculares que possam atuar como marcadores de aterosclerose, possibilitando prever o seu desen-

volvimento e até mesmo diferenciar situações clínicas antes indistinguíveis. O estudo da expressão gênica e proteica, a partir dos tecidos periféricos sanguíneos, demonstram ser potenciais ferramentas para a detecção de marcadores do processo inflamatório e outros indicadores de lesão, que reflitam o dano celular. O sangue periférico fornece informações sobre o estado do paciente em resposta a estímulos de citocinas e outros fatores solúveis liberados pelo endotélio, fígado, rins, coração, e outros tecidos, durante o processo inflamatório que pode preceder o infarto agudo do miocárdio [12].

Cita-se um estudo realizado por Silbiger e colaboradores, envolvendo um total de 84 indivíduos com Síndrome Coronariana Aguda (SCA) nas primeiras 48h após o IAM e 47 indivíduos sem IAM (grupo controle) de ambos os sexos, com idades entre 30 a 65 anos, detectou 599 genes diferencialmente expressos através da avaliação da expressão gênica global utilizando *microarrays* de *RNA_m*. Trinta e três genes foram selecionados e submetidos à validação técnica, dentre os quais, 20 genes foram submetidos à validação biológica em tempo real. Os genes considerados como potenciais marcadores para IAM foram: *Amphiregulin* (AREG), *Triggering Receptor Expressed On Myeloid Cells-Like 4* (TREM4), *BCL2-Related Protein A1* (BCL2A1), *BCL2-Like 1* (BCL2L1), *Carbonic Anhydrase I* (CA1), *Cytochrome C Oxidase Subunit VIIb* (COX7B), *Insulin Receptor Substrate 2* (IRS2), *Potassium Channel, Voltage Gated Subfamily E Regulatory Beta Subunit 1* (KCNE1), *Matrix Metalloproteinase 9* (MMP9), *Enoyl CoA Hydratase Domain Containing 3* (ECHDC3), *Myosin, Light Chain 4* (MYL4), *Interleukin 18 Receptor 1* (IL18R1) e *Arachidonate 15-Lipoxygenase* (ALOX15). Pouco explorados na literatura, esses genes revelam um potencial clínico importante para o desfecho de complicações cardiovasculares [12]. Em outros estudos, realizados em tecidos cardíacos de coelhos induzidos a infartos, observou-se modificações significantes nas expressões de algumas proteínas [13]. A atuação de ácidos graxos para a manutenção de funções cardíacas e sua desregulação em eventos cardíacos isquêmicos é bastante conhecida na literatura [14]. Portanto, expressões gênicas, expressões proteicas e concentrações de ácidos graxos podem atuar como indicadores de lesão cardíaca.

O presente trabalho tem como objetivo principal avaliar as relações entre os escores de Friesinger e Framingham com as expressões gênicas, proteicas e níveis de ácidos graxos com o intuito de se identificar potenciais biomarcadores de aterosclerose. Além disso, foi avaliada a relação entre os dois escores, bem como a associação entre as características bioquímicas e os ácidos graxos de pacientes com DCVs.

Capítulo 2

Metodologia

2.1 Participantes

O delineamento do presente estudo observacional é transversal, que fornece informações sobre características de uma doença e de potenciais fatores de risco em um mesmo período de tempo.

A seleção dos participantes avaliados no presente estudo iniciou-se em 2011, após a aprovação Comitê de Ética em Pesquisa do Hospital Universitário Onofre Lopes (CEP-HUOL) em Natal no Rio Grande do Norte (RN). O atual projeto faz parte de um projeto de maior espectro, intitulado "Transcriptoma em sangue periférico na avaliação de potenciais biomarcadores de doenças cardiovasculares". Como critério de seleção foram incluídos os pacientes de ambos os sexos, com idade entre 25 a 80 anos que seriam submetidos à cinecoronariografia, solicitada previamente por médicos do ambulatório da Unidade Clínica de Aterosclerose (HUOL), para elucidação diagnóstica devido a sintomas sugestivos de doença arterial coronariana. Os demais pacientes foram excluídos por apresentarem um ou mais critérios de exclusão como insuficiência renal crônica e hepática, distúrbio endócrino, com exceção de diabetes, doenças inflamatórias, neoplasia, doença hematológica, hipercolesterolemia familiar conhecida ou presente em membros da família, miocardiopatia, incluindo doença de chagas, valvulopatia, doença congênita e pericardiopatia.[12]

Desta forma, de acordo com os critérios de inclusão e de exclusão, já foram triados 234 pacientes, dos quais, 154 foram selecionados por apresentarem os critérios desejáveis após análises laboratoriais. No presente estudo considerou-se 132 pacientes que possuíam medição do escore de Friesinger.

Para a avaliação dos critérios de inclusão e exclusão os pacientes responderam um questionário e foram submetidos a testes de triagem como: glicemia, colesterol total e frações, ureia, creatinina, ácido úrico, alanina aminotransaminase (ALT) e aspartato aminotransaminase (AST).

O índice de massa corpórea (IMC), calculado dividindo-se o peso em quilogramas pelo quadrado da altura em metros, foi utilizado para determinar as categorias

de peso. Indivíduos que apresentaram o IMC superior a 30 Kg/m^2 foram classificados como obesos; indivíduos com IMC entre 25 Kg/m^2 e 30 Kg/m^2 , com sobrepeso e aqueles com o IMC inferior a 25 Kg/m^2 , com peso normal para homens e mulheres. O peso foi determinado com roupas leves, sem sapatos, utilizando balança mecânica Filizola modelo 31 e a altura mensurada por um antropômetro de 1,92 m.

Um estudo transversal realizado em 2010, caracterizou a associação da Pressão de Pulso (PP) com o escore de Friesinger e o risco em 10 anos de eventos cardiovasculares, dado por Framingham. O estudo mostrou, então, que a pressão de pulso é um marcador bem estabelecido de saúde vascular e um importante fator preditor independente da gravidade da DCV [20]. Portanto, tal variável foi utilizada como fator de risco no presente trabalho.

O paciente que relatasse diagnóstico médico de Diabetes Mellitus em qualquer fase da vida ou que fazia tratamento com hipoglicemiantes orais ou insulina foi considerado como portador da doença.

O escore de Framingham, que estima o risco cardíaco durante um período de 10 anos, com base em informações sobre os fatores clássicos de risco, tais como sexo, idade, pressão arterial sistólica, HDL, colesterol total, tratamento de hipertensão, diabetes e tabagismo, utiliza um sistema baseado em pontos para cada fator de risco. São utilizadas equações diferentes para cada sexo e ambas foram desenvolvidas a partir de um estudo aplicado aos indivíduos que não apresentavam DCVs. A idade, colesterol total, HDL e pressão arterial foram usadas como variáveis contínuas e transformadas para a escala logarítmica, em contraste a variáveis dicotômicas (sim/não), tais como tabagismo, diabetes e tratamento para hipertensão [10].

As folhas de pontuação para a predição de DCVs usando variáveis categorizadas, foram desenvolvidas a partir dos coeficientes do modelo de regressão de Cox que assume riscos proporcionais. Fornecendo comparação do risco absoluto, em 10 anos, para pessoas com mesma idade e sexo, para diferentes níveis, variando de muito baixo a muito alto, de incidência de DCVs. Tais distinções são arbitrárias, mas fornecem uma base para determinar a necessidade de intervenção clínica [9].

Os coeficientes do modelo de regressão múltipla para a previsão do risco e as razões de chance são apresentados na tabela 2.1. As estatísticas de sobrevivência *Cross-validated* e χ^2 foram utilizados para avaliar o desempenho das equações do modelo. As estatísticas de χ^2 de previsão foram 13,48 em homens e 7,79 para as mulheres. A chance de sobrevivência foi de 0,89 e 0,95 em homens e mulheres, respectivamente [8].

Tabela 2.1: Coeficientes da regressão e Razões de chance

Variáveis	β^*	Razão de Chance	IC 95%
Sexo masculino			
Log idade	3,06	21,35	(14,03;32,48)
Log colesterol total	1,12	3,08	(2,05;4,62)
Log HDL	-0,93	0,39	(0,30;0,52)
Log PS	1,93	6,91	(3,91;12,20)
Log PS com trat.	2,00	7,38	(4,22;12,92)
Tabagismo	0,65	1,92	(1,65;2,24)
Diabetes	0,57	1,78	(1,43;2,20)
Sexo feminino			
Log idade	2,33	10,27	(5,65;18,64)
Log colesterol total	1,21	3,35	(2,00;5,62)
Log HDL	-0,71	0,49	(0,35;0,69)
Log PS	2,76	15,82	(7,86;31,87)
Log PS com trat.	2,82	16,82	(8,46;33,46)
Tabagismo	0,53	1,70	(1,40;2,06)
Diabetes	0,69	2,00	(1,49;2,67)

Notas.: β^* coeficientes estimados do modelo de regressão, HDL—Lipoproteínas de alta densidade e PS—Pressão Sistólica.

O escore de Friesinger foi utilizado como critério de avaliação da extensão das lesões ateroscleróticas nas três principais artérias coronárias e em seus ramos. Tal escore pode ter valores entre 0 a 15 pontos. Sendo os pacientes com escores iguais a zero (0) classificados como sem lesão; entre 1 e 5 pontos, com lesão baixa; de 6 a 10 pontos, com lesão intermediária; e de 11 a 15 pontos, com lesão grave. [12]

Segundo pesquisa publicada no *Jornal Internacional de Clínica Médica (IJCM)*, as correlações significativamente positivas entre o escore de Friesinger (FR) foram colesterol total (p-valor = 0,002), triglicérides (p-valor < 0,001), IMC (p-valor < 0,001), hipertensão (p-valor < 0,001), diabetes (p-valor < 0,001) e tabagismo (p-valor < 0,001) e negativa com colesterol HDL (p-valor < 0,001). Correlações de FR com as variáveis idade e sexo não foram significantes. Assim, o estudo aponta uma ligação significativa entre a extensão da doença arterial coronariana e níveis de determinadas variáveis de lipídios e demais características clínicas [11].

2.1.1 Métodos laboratoriais

A realização de exames para os parâmetros laboratoriais e medições de expressões gênicas, expressões proteicas e níveis de ácidos graxos, foram coletadas amostras de sangue periférico de todos os pacientes do estudo, após jejum de no mínimo 8 horas, utilizando um sistema de coleta a vácuo. Para a realização das dosagens bioquímicas de glicose, colesterol total, HDL, triglicérides, ácido úrico e uréia foram utilizados conjuntos de reagentes da empresa Labtest (Lagoa Santa, MG, BR) através do método de ponto final. Os níveis de AST, ALT e Creatinina foram quantificados através de conjuntos de reagentes também da empresa Labtest pelo método cinético. Todas as dosagens foram realizadas em analisadores bioquímicos semiautomáticos BIO-200 (Bioplus, Lagoa Santa,

MG, BR) e BIO-100 (Bioclin, Quibasa Química Básica Ltda, MG, BR). A quantificação da expressão dos genes TREML4 e ECHDC3 foi realizada pela transcrição reversa do RNAm seguida por PCR (marcador de resposta inflamatória), valores elevados de PCR no plasma foram associados com o risco de ocorrência de episódios agudos de ruptura da placa aterosclerótica e IAM, em tempo real. A partir de 4ml de sangue total, 1 μ g de RNA foi utilizado para a obtenção do cDNA através do conjunto de reagentes utilizando *High Capacity RNA-to-cDNA Master Mix (Life TechnologiesTM, CA, EUA)*, constituído de iniciadores aleatórios (*random primers*), dNTPs, a enzima transcriptase reversa e tampão. A reação de transcriptase reversa foi realizada com o termociclador *MyCyclerThermalCycler (Bio-Rad, CA, EUA)* de acordo com as seguintes etapas: 25° por 5 minutos, 42° por 30 minutos e 85° por 5 minutos. O cDNA obtido foi armazenado a -20°C até a etapa de amplificação.

O sistema *TaqManTM (LifeTechnologiesTM, CA, EUA)* foi utilizado para a amplificação do cDNA realizada por PCR em tempo real. A análise da expressão gênica foi realizada pelo método de quantificação relativa, utilizando o gene *ACTB* (β -actina) como referência endógena, após análise em relação aos outros dois genes controle (*GAPDH* e *18S*) através do programa *Genorm Software e NormFinder*. Os iniciadores e as sondas marcadas com fluoróforo foram fornecidos (em solução 20 vezes concentrada) pelo serviço "*Assay by design*" através dos códigos HS01080584_G1(*TREML4*) e HS00226727_M1(*ECHDC3*) [15].

As reações de PCR em tempo real foram realizadas no termociclador *7500 Fast Real-Time PCR System (Life TechnologiesTM, Foster City, EUA)*. As reações de amplificação foram constituídas de iniciadores a 90nM e sonda marcada com *FAMTM* (fluoróforo 6-carboxifluoresceína) a 250nM. Os demais reagentes foram fornecidos pela *Life TechnologiesTM* em solução 2 vezes concentrada, denominada *TaqMan® Universal Master Mix*, contendo dNTPs, com dUTP, *AmpErase® UNG* (Uracil N- glicosilase), enzima *AmpliTaqGold® DNA Polymerase*, e tampão de reação *Gold PCR Buffer*. O volume final por reação foi 20 μ L e os ensaios foram realizados em duplicata [12].

A programação de PCR em tempo real escolhida é constituída por: (1) um ciclo de 2 minutos a 50°C; (2) um ciclo de 10 minutos a 95°C; 40 ciclos de 15 segundos a 95°C e 1 minuto a 60°C, onde os sinais de fluorescência emitidos pelos fluoróforos das sondas *TaqMan®* foram detectados pelo equipamento nesta etapa. Os dados foram analisados utilizando o programa *7500 Software v2 0.5 (Life TechnologiesTM, CA, EUA)*, que gera curvas semi-logarítmicas dos sinais de amplificação.

A quantificação relativa da expressão gênica foi realizada utilizando o método comparativo com o ciclo *threshold* (C_T), sendo a fórmula utilizada de $2^{-\Delta C_T}$, em que ΔC_T é igual à diferença entre o C_T do gene alvo e o C_T do gene controle [16][12].

Para a genotipagem, os polimorfismos foram analisados por discriminação alélica pelo sistema *TaqMan®* utilizando a reação em cadeia da polimerase em tempo real no equipamento *ABI 7500 Fast real time PCR System* e software *7500 SDS* (AppliedBi-

osystems). As condições de amplificações adotadas para todos os 2 SNPs (rs2803495 e rs2803496) se iniciaram a 10 minutos com 95°C, seguida por 40 ciclos de desnaturação por 15 segundos a 92°C e extensão de 90 segundos a 60°C. Os iniciadores e sondas marcadas com fluoróforos FAM e VIC, fornecidos pelo serviço "Assay by design" através dos códigos C_27302616_10 e C_27302614_10 (*Life Technologies*TM, CA, EUA) [12].

2.2 Análises Estatísticas

Medidas descritivas foram obtidas para caracterizar o grupo estudado quanto a variáveis demográficas, clínicas e laboratoriais. Os Coeficientes de correlação de Pearson foram obtidos para todas as combinações da variáveis resposta (score de Friesinger), score de Framingham, expressões gênicas, proteicas e níveis de ácidos graxos. Além disso, avaliou-se a relação entre variável resposta e covariáveis por meio de análises de regressão linear múltipla e modelo linear generalizado com distribuição Poisson.

Para a realização de todas as análises estatísticas foi utilizado o software estatístico R (Rx64 bit - versão: 3.1.3).

2.2.1 Modelo de Regressão Linear Múltipla

A análise de regressão linear múltipla é uma técnica estatística utilizada para modelar a relação entre uma variável resposta quantitativa e variáveis explicativas quantitativas ou qualitativas, chamadas de preditoras ou independentes. O modelo de regressão linear múltipla mais simples descreve a variável resposta aleatória Y em função de p variáveis preditoras (X_1, X_2, \dots, X_p) com n observações, $p+1$ parâmetros lineares. Tal modelo pode ser descrito por:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon, \quad (1)$$

em que $\beta_0, \beta_1, \dots, \beta_p$ são os parâmetros a serem estimados no modelo e ε representa o resíduo, variável aleatória suposta independente e com distribuição de probabilidade $N(0, \sigma^2)$. β_0 representa o intercepto da reta estimada e β_k para $k = 1, 2, \dots, p$ fornece a mudança esperada no valor de Y para o aumento de uma unidade em X_k , mantendo-se os valores das outras variáveis preditoras constantes.

Sob a suposição de que $E(\varepsilon) = 0$, Y também tem distribuição de probabilidade normal e seu valor esperado é dado por:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p.$$

Em termos matriciais o modelo (1) pode ser escrito como:

$$\mathbf{Y}_{n \times 1} = \mathbf{X}_{n \times (p+1)} \boldsymbol{\beta}_{(p+1) \times 1} + \boldsymbol{\varepsilon}_{n \times 1},$$

em que \mathbf{Y} representa o vetor da variável resposta; $\boldsymbol{\beta}$, o vetor de parâmetros; \mathbf{X} , a matriz de variáveis explicativas e $\boldsymbol{\varepsilon}$, o vetor de erros aleatórios.

Supõe-se que $\boldsymbol{\varepsilon}$ seja independente com distribuição de probabilidade $N(0, \sigma^2)$, portanto, sua matriz de variância e covariâncias é dada por:

$$\boldsymbol{\sigma}_{n \times n}^2(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}.$$

Assim, $\boldsymbol{\sigma}_{n \times n}^2(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$ e, por consequência, o vetor aleatório \mathbf{Y} tem $\mathbf{E}(\mathbf{Y})_{n \times 1} = \mathbf{X}\boldsymbol{\beta}$ e a matriz de variância e covariâncias $\boldsymbol{\sigma}_{n \times n}^2(\mathbf{Y}) = \sigma^2 \mathbf{I}$, em que \mathbf{I} é a matriz identidade [4].

Os parâmetros do modelo são estimados pelo método de mínimos quadrados. Tal método estima os parâmetros β 's de modo que minimizem a SQR , ou seja, que representem o menor erro entre o valor estimado pelo modelo de regressão linear múltipla e o verdadeiro valor do parâmetro na população.

A soma de quadrados dos resíduos é dada por:

$$SQR = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \cdots - \beta_p X_{ip})^2.$$

Em termos matriciais, $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ é o vetor das diferenças entre cada valor de \mathbf{y} com o valor correspondente dado pela função em $\boldsymbol{\beta}$ e a solução para se estimar o parâmetro é $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, desde que a inversa de $(\mathbf{X}'\mathbf{X})$ exista.

Sendo assim, o modelo original é redefinido e adota-se, de modo que os valores estimados $\hat{\mathbf{Y}}_{n \times 1} = \mathbf{X}\mathbf{b} \iff \hat{\mathbf{Y}} = \mathbf{H}\mathbf{Y}$, em que $\mathbf{H}_{n \times n} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

De forma similar, o vetor $\boldsymbol{\varepsilon} = (\mathbf{I} - \mathbf{H})\mathbf{Y}$ e matriz residual são redefinidos como

$$\begin{aligned} \boldsymbol{\sigma}_{n \times n}^2(\boldsymbol{\varepsilon}) &= \sigma^2(\mathbf{I} - \mathbf{H}) \text{ e} \\ \mathbf{s}_{n \times n}^2(\boldsymbol{\varepsilon}) &= MQR(\mathbf{I} - \mathbf{H}), \end{aligned}$$

em que MQR é o erro quadrático médio, obtido pela razão entre SQR e seus graus de liberdade.

2.2.2 Análise de Variância

A análise de variância envolve a decomposição da soma de quadrados total (SQT) de \mathbf{Y} em resíduos (SQR) e valores esperados do modelo (SQE) [4].

$$SQT = \mathbf{Y}'\mathbf{Y} - \left(\frac{1}{n}\right) \mathbf{Y}'\mathbf{J}\mathbf{Y} = \mathbf{Y}' \left[\mathbf{I} - \left(\frac{1}{n}\right) \mathbf{J} \right] \mathbf{Y},$$

$$SQR = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{X}\mathbf{b})'(\mathbf{Y} - \mathbf{X}\mathbf{b}) = \mathbf{Y}'(\mathbf{I} - \mathbf{H})\mathbf{Y} \text{ e}$$

$$SQE = \mathbf{b}'\mathbf{X}'\mathbf{Y}' - \left(\frac{1}{n}\right) \mathbf{Y}'\mathbf{J}\mathbf{Y} = \mathbf{Y}' \left[\mathbf{H} - \left(\frac{1}{n}\right) \mathbf{J} \right] \mathbf{Y}.$$

A matriz \mathbf{J} é uma matriz de 1s.

A soma de quadrados total dos resíduos e os valores esperados possuem $(n-1)$, $(n-p-1)$ e (p) graus de liberdade associados a elas. Dividindo às SQ's pelos seus respectivos graus de liberdade e obtemos os quadrados médios.

$$MQE = \frac{SQE}{p} \text{ e}$$

$$MQR = \frac{SQR}{n-p-1}.$$

A expectativa é que $MQR = \sigma^2$ do modelo de regressão linear, sendo $\sigma^2 > 0$.

Testes de hipótese

A hipótese $H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$ é testada utilizando-se a estatística F_0 . Seguindo uma distribuição F , podemos comparar F_0 com o F crítico, denotado por F^c , da tabela F com quantil $(1 - \alpha)$ e graus de liberdade p e $(n - p - 1)$.

$$F_0 = \frac{MQE}{MQR}.$$

Como regra de decisão, se $F \leq F_{(1-\alpha; p; n-p-1)}^c$ H_0 não é rejeitada ou se $p - \text{valor} \leq \alpha$, rejeita-se H_0 ao nível $(100 \times \alpha)\%$ de significância. Note que quanto maior a SQE, maior a chance de rejeitar H_0 , pois menor será a SQR em relação a SQT.

Após a análise do modelo geral, pode-se testar a contribuição de X_k no modelo usando apenas o parâmetro estimado $\hat{\beta}_k$. Sendo a hipótese nula $H_0 : \beta_k = 0$, a estatística de teste é definida como:

$$t = \frac{\hat{\beta}_k - \beta_k}{s_{\hat{\beta}_k}} \sim t_{(n-p-1)}^c, \text{ com } k = 0, 1, \dots, p.$$

Em tal expressão $s_{\hat{\beta}_k}$ representa o erro padrão de b_k calculado pelo modelo.

Caso $t \leq t_{(1-\frac{\alpha}{2}; n-p-1)}^c$, H_0 não é rejeitada. A estatística t segue a distribuição t de Student e o quantil t^c tem $(n-p-1)$ graus de liberdade. Pela suposição de normalidade dos resíduos, aplica-se a definição de estatística studentizada para t padronizando, assim, os resíduos [5].

Sua estimação intervalar com $100 \times (1 - \alpha)\%$ de confiança é:

$$IC(\beta_k) : \left[\hat{\beta}_k \pm t_{(1-\frac{\alpha}{2}; n-p-1)}^c s(\hat{\beta}_k) \right].$$

Coeficiente de determinação

O coeficiente de determinação, representado por R^2 , mede a proporção da variação total de Y associada as variáveis preditoras X_1, \dots, X_p , e é dado por:

$$R^2 = \frac{SQE}{SQT} = 1 - \frac{SQR}{SQT}, \text{ com } 0 \leq R^2 \leq 1.$$

Quando ajusta-se por um número muito grande de variáveis preditoras, a SQE aumenta e, com isso, a SQR diminui em relação ao SQT, gerando um R^2 maior. Sugere-se então corrigir as somas de quadrados pelos seus graus de liberdade, ajustando-se o valor de R^2 para medida conhecida como coeficiente de determinação ajustado (R_a^2) dado por:

$$R_a^2 = 1 - \left(\frac{n-1}{n-p-1} \right) \frac{SQR}{SQT}.$$

Essa medida fornece um percentual de explicação da variação em Y mais consistente e conservador com relação as variáveis preditoras.

2.2.3 Análise dos resíduos

A busca pelo modelo ideal consiste na identificação de variáveis preditoras que melhor expliquem a variável Y . Foram utilizados métodos de seleção de variáveis como "stepwise" e "backward" para se obter um modelo com o número menor das p variáveis preditoras, sendo seus testes de hipóteses significativos e R^2 suficientemente "alto".

Após definido o modelo de regressão, utiliza-se a análise dos resíduos para a avaliação da adequação do modelo aos dados e as suposições de independência, normalidade e homoscedasticidade.

Utilizou-se o gráfico dos resíduos ϵ contra os valores preditos de \hat{Y} . Um perfil aleatório, indica a independência dos erros. O gráfico dos quantis de probabilidade normal é utilizado para se avaliar a normalidade dos resíduos.

Usa-se uma forma alternativa padronizada para analisar os resíduos que leva em consideração suas variâncias. Ao dividir-se ϵ por $Var(\epsilon) = \sigma^2(1 - \mathbf{h})$, obtém-se os resíduos studentizados. Uma forma mais conveniente que possui distribuição t de Student sob normalidade dos dados é expressa por:

$$z_i = \frac{\hat{\epsilon}_i}{\hat{\sigma}_i \sqrt{(1 - h_{ii})}}, \text{ com } i = 1, \dots, n$$

com $(n - k - 1)$ graus de liberdade e h_{ii} o i -ésimo elemento da diagonal da matriz \mathbf{H} simétrica e idempotente. No gráfico de dispersão dos z_i contra \hat{y}_i um perfil aleatório e observações não discrepantes indicam um bom ajuste.

O gráfico das distâncias de Cook também é utilizado para a avaliação das suposições e a adequação do modelo. A distância de Cook mede a influência de cada observação comparando-se as estimativas dos coeficientes de regressão do modelo original

quando uma determinada observação é omitida.

$$D_i = \frac{(\hat{\beta} - \hat{\beta}_i)' \mathbf{X}' \mathbf{X} (\hat{\beta} - \hat{\beta}_i)}{k \hat{\sigma}^2}, \text{ com } i = 1, \dots, n,$$

sendo $\hat{\beta}_i$ o estimador de mínimos quadrados do vetor β , omitindo-se a i -ésima observação e k , o número de colunas de \mathbf{X} .

Valores de D_i maiores que 1 indicam a presença de valores influentes.

Caso a adequação do modelo não satisfaça ao menos um dos critérios estabelecidos, sugere-se reavaliar a utilização do modelo e investigar combinações lineares das próprias variáveis ou mesmo utilizar outros modelos como modelos lineares generalizados.

2.3 Modelos lineares generalizados

Com a finalidade de ampliar a gama de aplicações dos modelos estatísticos lineares, Nelder e Wedderburn (1972) introduziram o conceito do modelo linear generalizado, que acomoda variáveis resposta com distribuições de probabilidade não normais mas que pertencem a família exponencial [4].

Os n valores preditos para a variável Y são respostas independentes que seguem uma distribuição de probabilidade da família exponencial com o valor esperado $E(Y) = \mu$. O preditor linear baseado nas variáveis predictoras X_1, \dots, X_p é denotado por $\mathbf{X}'\beta$. Portanto,

$$\mathbf{X}'_i \beta = \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip},$$

$$\mathbf{X}'_i \beta = g(\mu_i).$$

Modelos lineares generalizados podem ter variâncias σ_i^2 não constantes para as respostas Y_i , mas σ_i^2 deve ser uma função das variáveis predictoras através da resposta média μ_i .

Nos modelos de regressão com distribuição Poisson, pode-se utilizar diferentes funções para $g(\mu_i)$, contanto que a $g(\mu_i)$ seja monótona e diferenciável. Considerando o modelo de regressão linear com erro normal, a função $g(\mu_i) = \ln(\mu_i)$, desde que a relação $\mathbf{X}'_i \beta = \mu_i$ seja verdadeira, e $\mu_i > 0$. Expressa-se o modelo com distribuição Poisson por:

$$Y_i = E(Y_i) + \varepsilon_i \quad i = 1, 2, \dots, n.$$

Os coeficientes do modelo são estimados pelo método de mínimos quadrados.

2.3.1 Critério de seleção do modelo

O Critério de Informação de Akaike - Akaike Information Criterion (AIC) é uma ferramenta utilizada na seleção de modelos. Oferece uma medida relativa da

qualidade do ajuste do modelo estatístico, porém não fornece informações sobre como o modelo se ajusta aos dados.

$$AIC = 2k - 2\ln(L),$$

em que k é o número de parâmetros no modelo e L é o valor da função de máxima verossimilhança do modelo estimado.

Dado um conjunto de modelos é preferido aquele com o menor valor de AIC.

Nelder e Wedderburn propuseram a medida de discrepância "*Deviance*", para avaliar os modelos e medir o ajuste. Calculada a partir dos valores máximos do logaritmo da função de verossimilhança para os modelos saturado e corrente. O modelo saturado atribui toda a variação dos dados ao componente sistemático e o modelo corrente (sob pesquisa) está entre os modelos maximal (maior número de termos) e minimal (menor número de termos) que podem ser considerados para o modelo investigado [3].

Capítulo 3

Resultados e Discussão

3.1 Análise Descritiva

Para as presentes análises foram considerados apenas os indivíduos que possuíam informações sobre o escore de Friesinger, o que corresponde a 130 observações. A Tabela 1 apresenta médias, desvios padrão (DP), frequências e percentuais das variáveis que caracterizam o conjunto de dados.

Observou-se que os participantes possuíam em média 59 anos ($DP = 9,7$) e cerca de 57% eram do sexo masculino. Aproximadamente mais da metade dos indivíduos se declararam como de raça branca, 45 indivíduos de raça parda e de raça negra, 20 indivíduos. A maior parte dos participantes (76,2%) não eram fumantes, não tinham o hábito de consumir bebidas alcoólicas (76,7%) e não praticavam exercícios físicos (63%).

Quarenta e cinco por cento declararam ser de peso normal, seguidos de 34,6% com sobrepeso e 20% obesos. O IMC medido médio foi de 27,5 ($DP = 4,9$), o que caracterizaria sobrepeso. Nas categorias do IMC 40,8% foram considerados obesos; 22,3%, com sobrepeso; 27,7% com peso normal e 9,2% abaixo do normal. Pouco mais de 38% dos participantes apresentavam diabetes.

Oitenta e sete por cento reportaram ser diagnosticados com hipertensão arterial sistêmica (HAS), sendo considerados com HAS os indivíduos com as medidas de pressão maior ou igual a 140/90 mmHg ou que estivessem em uso de medicação anti-hipertensiva. A média da pressão sistólica foi 127,4 mmHg ($DP = 38,8$) E a média da pressão diastólica foi de 100 mmHg ($DP = 32,9$). Por volta de 85% dos participantes usavam medicamento para hipertensão ou já teriam realizado tratamento em algum outro momento quando o diagnosticado de hipertensão foi positivo.

A variável pressão de pulso (PP), considerada um fator predisponente para a severidade de DVCs, foi calculada pela diferença entre pressão arterial sistólica e a pressão arterial diastólica. A média da diferença das duas pressões foi de 26,8 ($DP = 61,7$). Quando $PP < 40$ considera-se uma possível função cardíaca reduzida, $PP = 40$ normal e $PP > 40$ regurgitação das válvulas.

Setenta e cinco por cento do participantes relataram possuir níveis de colesterol

ou triglicéride elevados. Mais de 94% faziam o uso de outros medicamentos para demais doenças diagnosticadas.

O histórico familiar (HF) clínico dos indivíduos também foi avaliado. Foram considerados apenas os relato de DCVs na família nuclear, ou seja, em pais, irmãos ou filhos, com idade abaixo de 55 anos para homens e de 65 anos para as mulheres. Observou-se que 66% dos participantes possuía HF de HAS; 50%, de diabetes; 44,6%, de morte de parente por IAM e 27%, de dislipidemia.

O escore de Friesinger médio foi de 5,5 (DP = 4,4), o que representaria lesão baixa. Em relação a distribuição de frequência das categorias dos níveis de Friesinger, observou-se que 23,1% foram classificados como sem lesão; 27,7%, com grau de lesão baixa; 32,3%, com grau de lesão intermediária e 16,9%, com grau de lesão grave. O escore de Framingham médio foi de 29,8 (DP = 20,6), que corresponderia a uma classificação de risco intermediária mas com uma alta variabilidade nos participantes. Quando o escore de Framingham foi categorizado observou-se que 5,7% dos participantes foram classificados com risco de lesão baixa; 24,1%, com risco de lesão intermediário e 70,1%, com risco alto de desenvolvimento de DCVs. Observou-se, também, 43 participantes sem informação do escore de Framingham.

A Tabela 2 contém médias e desvios padrão da expressão de treze genes, quatro expressões proteicas, níveis de dez ácidos graxos e exames laboratoriais dos participantes do estudo. Entre as expressões gênicas, destacou-se CA1, MYL4 e BCL2L1 com os maiores valores de média e desvios padrão. Os outros genes apresentam valores muito próximos de zero para tais medidas. Entre as expressões proteicas, cita-se vCAM, com valor médio muito superior às demais.

Os ácidos com maiores valores médios e desvios padrão foram os Mirístico, Linoleico, Gamalinoleico, Oleico e Araquidônico. Ressalta-se que para o ácido Mirístico, dois valores foram considerados como discrepantes proveniente de erro de medida e, portanto, foram retirados das análises do presente trabalho. Dos participantes, observou-se que 65 não possuíam as treze expressões gênicas completas, faltando pelo menos uma observação não informada. 46 participantes com valores incompletos para às proteínas e 64 para os ácidos graxos.

Considera-se para os exames laboratoriais os seguintes valores de referência: Glicose normal até 100 mg/dl, elevado de 100 a 139 mg/dl e diabetes quando maior que 140 mg/dl. Colesterol desejável até 200 mg/dl, limiar de 201 até 239 mg/dl e elevado a partir de 240 mg/dl. Triglicérides desejável até 200 mg/dl, elevado maior que 200 mg/dl. Colesterol LDL normal até 130 mg/dl, limiar de 130 a 159 mg/dl, elevado maior que 160 mg/dl.

Tabela 3.1: Características demográficas e clínicas dos participantes

Idade (em anos), média (DP)	59,3	(9,7)
Sexo Masculino, n (%)	n=74	(56,9)
Raça, freq, (%)	n=129	
Branco	64	(49,6)
Pardo	45	(34,9)
Negro	20	(15,5)
Exercício Físico, n (%)	48	(36,9)
Bebida alcoólica, n (%)	30	(23,3)
Cigarro, n (%)	31	(23,8)
Peso (Kg), média (DP)	71,5	(13,3)
Peso, freq, (%)	n=130	
Normal	59	(45,4)
Sobrepeso	45	(34,6)
Obesidade	26	(20,0)
Altura (m), média (DP)	1,61	(0,1)
IMC, média (DP) (Kg/m ²)	27,5	(4,9)
IMC, freq, (%)	n=130	
Abaixo	12	(9,2)
Normal	36	(27,7)
Sobrepeso	29	(22,3)
Obesidade	53	(40,8)
Diabetes, n (%)	50	(38,5)
HAS, n (%)	113	(86,9)
Pressão sistólica (mmHg), média (DP)	127,4	(39,2)
Pressão diastólica (mmHg), média (DP)	100,5	(33,0)
PP, média (DP)	26,8	(61,7)
Dislipidemia, n (%)	98	(75,4)
Hipertensão tratada, n (%)	108	(85,0)
Uso de Medicamentos, n (%)	123	(94,6)
HF de Obesidade, n (%)	18	(13,8)
HF de Diabetes, n (%)	65	(50,0)
HF de HAS, n (%)	86	(66,1)
HF de Dislipidemia, n (%)	35	(26,9)
HF Morte de parente por IAM, n (%)	58	(44,6)
Escore Friesinger, média (DP)	5,5	(4,4)
Escore Friesinger, freq, (%)	n=130	
Sem lesão	30	(23,1)
Baixo	36	(27,7)
Intermediário	42	(32,3)
Grave	22	(16,9)
Escore de Framingham, média (DP)	29,8	(20,6)
Escore Framingham, freq, (%)	n=87	
Baixo	5	(5,7)
Intermediário	21	(24,1)
Alto	61	(70,1)

Notas: HF-Histórico familiar, HAS-Hipertensão arterial sistêmica, IMC-Índice de massa corpórea, IAM-Infarto agudo do miocárdio e PP-(Pressão sistólica - Pressão diastólica).

3.1. ANÁLISE DESCRITIVA

Tabela 3.2: Médias e desvios padrão de variáveis laboratoriais dos participantes

Características	Média	Desvio Padrão
Perfil Laboratorial		
Glicose (mg/dl)	121,6	68,7
Colesterol total (mg/dl)	180,1	48,6
Triglicérides (mg/dl)	165,7	108,7
HDL (mg/dl)	37,5	12,4
LDL (mg/dl)	109,6	43,5
ALT (U/L)	30,4	21,2
AST (U/L)	32,6	18,8
Ureia (mg/dl)	38,8	12,6
Creatinina (mg/dl)	1,0	0,3
Ácido Úrico (mg/dl)	4,8	1,6
Expressões Gênicas ($RNA_m 2e^{-\Delta C_T}$)		
AREG	0,0002	0,0003
TREML4	0,0004	0,0007
BCL2A1	0,0181	0,0096
BCL2L1	0,0363	0,0444
CA1	0,0662	0,0935
COX7B	0,0037	0,0023
IRS2	0,0061	0,0037
KCNE1	0,0001	0,0000
MMP9	0,0172	0,0116
ECHDC3	0,0006	0,0005
MYL4	0,0401	0,0548
IL18R1	0,0003	0,0002
ALOX15	0,0033	0,0054
Expressões proteicas		
E-Selectina	31,4	16,6
vCAM	989,4	272,7
iCAM	147,9	46,1
MMP9	330,7	175,2
Ácidos Graxos		
Mirístico	379,6	147,9
Miristoleico	36,0	18,6
Oleico	150,7	51,1
Linoleico	295,8	121,5
Gamalinoleico	224,8	95,8
Linolênico	34,7	17,5
Araquidônico	112,3	50,9
Eicosapentaenoico	20,0	16,3
Lignocérico	10,3	4,9
Docosaenoico	29,1	16,5

Notas: HDL—Lipoproteínas de alta densidade, LDL—Lipoproteínas de baixa densidade, ALT—Alanina aminotransferase e AST—Aspartato aminotransferase.

3.1.1 Coeficiente de correlação de Pearson

Coeficientes de correlação de Pearson foram estimados entre todos os pares de variáveis incluindo Frisinger, Framingham, expressões gênicas, proteicas e níveis de ácidos graxos. A Figura 3.1 contém informações sobre a matriz de estimativas dos coeficientes de correlação. A cor azul representa associações positivas e a cor laranja, associações negativas. A forma da elipse indica a intensidade da correlação, de maneira que, quanto mais próxima de uma reta for a elipse mais forte é a correlação entre as variáveis.

Para uma melhor visualização dos gráficos, as proteínas E-Selectina Luminex, vCAM Luminex, iCAM Luminex e MMP9 Luminex foram reportadas, respectivamente, como P_1 , P_2 , P_3 e P_4 e o Ácido mirístico, Ácido miristoleico, Ácido oleico, Ácido linoleico, Ácido gamalinoleico, Ácido linolênico, Ácido araquidônico, Ácido eicosapentaenoico, Ácido lignocérico e Ácido docosaenoico como A_1 , A_2 , A_3 , A_4 , A_5 , A_6 , A_7 , A_8 , A_9 e A_{10} , respectivamente.

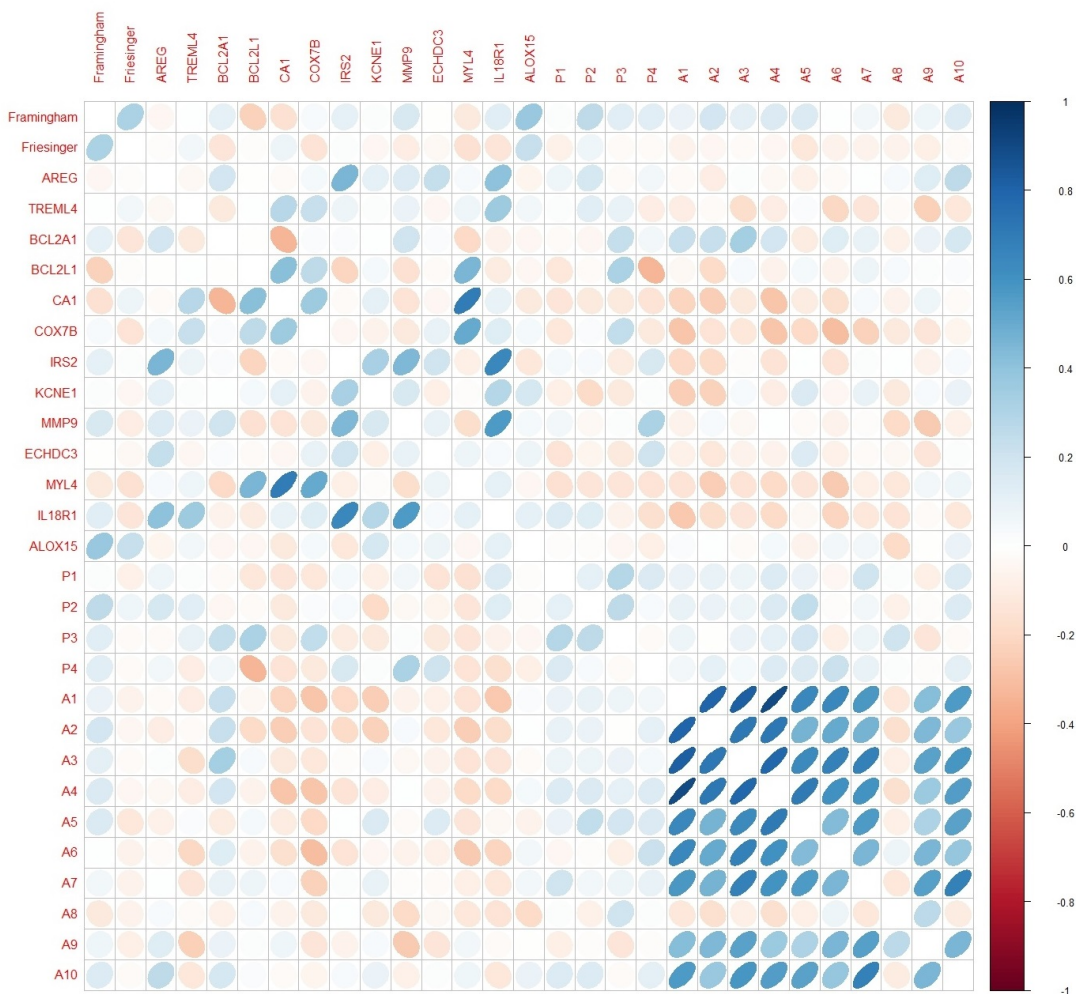


Figura 3.1: Gráfico das estimativas dos coeficientes de correlação de Pearson das variáveis de expressões gênicas, expressões proteicas e níveis de ácidos graxos.

3.1. ANÁLISE DESCRITIVA

As Tabelas 3.3 e 3.4 apresentam estimativas dos coeficientes de correlação de Pearson e p-valores para o teste de hipótese ($H_0 : \rho = 0$). Foram incluídas apenas as estimativas cujos resultados de testes fornecessem p-valores $< 0,1$.

Tabela 3.3: Tabela das estimativas da correlação de Pearson que possuem p-valor $< 0,1$

Variáveis		Correlação	P-valor
Ac. Araquidônico	Ac. Docosaexaenoico	0,68	1×10^{-10}
Ac. Araquidônico	Ac. Lignocérico	0,55	1×10^{-5}
Ac. Eicosapentaenoico	Ac. Lignocérico	0,27	5×10^{-2}
Ac. Gamalinoleico	Ac. Araquidônico	0,57	4×10^{-7}
Ac. Gamalinoleico	Ac. Docosaexaenoico	0,53	3×10^{-6}
Ac. Gamalinoleico	Ac. Lignocérico	0,31	2×10^{-2}
Ac. Gamalinoleico	Ac. Linolênico	0,44	2×10^{-4}
Ac. Lignocérico	Ac. Docosaexaenoico	0,46	3×10^{-4}
Ac. Linoleico	Ac. Araquidônico	0,59	9×10^{-8}
Ac. Linoleico	Ac. Docosaexaenoico	0,56	7×10^{-7}
Ac. Linoleico	Ac. Gamalinoleico	0,70	3×10^{-11}
Ac. Linoleico	Ac. Lignocérico	0,36	6×10^{-3}
Ac. Linoleico	Ac. Linolênico	0,60	5×10^{-8}
Ac. Linolênico	Ac. Araquidônico	0,46	9×10^{-5}
Ac. Linolênico	Ac. Docosaexaenoico	0,38	1×10^{-3}
Ac. Linolênico	Ac. Lignocérico	0,46	4×10^{-4}
Ac. Mirístico	Ac. Araquidônico	0,57	4×10^{-7}
Ac. Mirístico	Ac. Docosaexaenoico	0,56	6×10^{-7}
Ac. Mirístico	Ac. Gamalinoleico	0,65	3×10^{-9}
Ac. Mirístico	Ac. Lignocérico	0,42	1×10^{-3}
Ac. Mirístico	Ac. Linoleico	0,89	0
Ac. Mirístico	Ac. Linolênico	0,65	3×10^{-9}
Ac. Mirístico	Ac. Miristoleico	0,80	9×10^{-16}
Ac. Mirístico	Ac. Oleico	0,82	0
Ac. Miristoleico	Ac. Araquidônico	0,46	8×10^{-5}
Ac. Miristoleico	Ac. Docosaexaenoico	0,38	2×10^{-3}
Ac. Miristoleico	Ac. Gamalinoleico	0,47	9×10^{-5}
Ac. Miristoleico	Ac. Lignocérico	0,44	8×10^{-4}
Ac. Miristoleico	Ac. Linoleico	0,72	1×10^{-11}
Ac. Miristoleico	Ac. Linolênico	0,51	2×10^{-5}
Ac. Miristoleico	Ac. Oleico	0,72	1×10^{-11}
Ac. Oleico	Ac. Araquidônico	0,67	2×10^{-10}
Ac. Oleico	Ac. Docosaexaenoico	0,59	1×10^{-7}
Ac. Oleico	Ac. Gamalinoleico	0,64	5×10^{-9}
Ac. Oleico	Ac. Lignocérico	0,53	2×10^{-5}
Ac. Oleico	Ac. Linoleico	0,79	9×10^{-16}
Ac. Oleico	Ac. Linolênico	0,68	3×10^{-10}
AREG	Ac. Docosaexaenoico	0,25	5×10^{-2}
AREG	ECHDC3	0,23	7×10^{-2}
AREG	IL18R1	0,41	2×10^{-3}
AREG	IRS2	0,45	2×10^{-4}
BCL2A1	Ac. Mirístico	0,23	9×10^{-2}
BCL2A1	Ac. Miristoleico	0,22	1×10^{-1}

Tabela 3.4: Tabela das estimativas da correlação de Pearson que possuem p-valor $< 0,1$

Variáveis		Correlação	P-valor
BCL2A1	Ac. Oleico	0,34	9×10^{-3}
BCL2A1	CA1	-0,33	7×10^{-3}
BCL2A1	iCAM	0,24	6×10^{-2}
BCL2L1	CA1	0,41	6×10^{-4}
BCL2L1	COX7B	0,26	4×10^{-2}
BCL2L1	iCAM	0,31	1×10^{-2}
BCL2L1	IRS2	-0,22	8×10^{-2}
BCL2L1	P. MMP9	-0,33	8×10^{-3}
BCL2L1	MYL4	0,46	1×10^{-4}
CA1	Ac. Linoleico	-0,28	3×10^{-2}
CA1	Ac. Mirístico	-0,22	1×10^{-1}
CA1	Ac. Miristoleico	-0,25	6×10^{-2}
CA1	COX7B	0,36	3×10^{-3}
CA1	MYL4	0,70	1×10^{-10}
COX7B	Ac. Araquidônico	-0,23	9×10^{-2}
COX7B	Ac. Linoleico	-0,27	4×10^{-2}
COX7B	Ac. Linolênico	-0,30	2×10^{-2}
COX7B	Ac. Mirístico	-0,28	4×10^{-2}
COX7B	iCAM	0,25	6×10^{-2}
COX7B	MYL4	0,50	2×10^{-5}
E-Selectina	iCAM	0,28	1×10^{-2}
ECHDC3	P. MMP9	0,21	1×10^{-1}
Framingham	ALOX15	0,38	3×10^{-3}
Framingham	BCL2L1	-0,22	9×10^{-2}
Framingham	Friesinger	0,31	3×10^{-3}
Framingham	vCAM	0,25	3×10^{-2}
Friesinger	ALOX15	0,23	7×10^{-2}
IL18R1	Ac. Mirístico	-0,26	7×10^{-2}
IRS2	IL18R1	0,65	1×10^{-7}
IRS2	KCNE1	0,32	1×10^{-2}
IRS2	MMP9	0,45	2×10^{-4}
KCNE1	Ac. Mirístico	-0,24	9×10^{-2}
KCNE1	IL18R1	0,28	4×10^{-02}
MMP9	Ac. Lignocérico	-0,26	6×10^{-2}
P. MMP9	Ac. Linolênico	0,21	8×10^{-2}
MMP9	IL18R1	0,57	6×10^{-6}
MMP9	P. MMP9	0,31	1×10^{-2}
MYL4	Ac. Linolênico	-0,25	6×10^{-2}
MYL4	Ac. Miristoleico	-0,24	7×10^{-2}
TREML4	Ac. Lignocérico	-0,24	8×10^{-2}
TREML4	CA1	0,27	3×10^{-2}
TREML4	COX7B	0,22	8×10^{-2}
TREML4	IL18R1	0,36	8×10^{-3}
vCAM	Ac. Gamalinoleico	0,24	5×10^{-2}
vCAM	iCAM	0,26	2×10^{-2}

Com base na Figura 3.1 e as Tabelas 3.3 e 3.4 observou-se que associações maiores que 0,20 com a variável escore de Friesinger apresentavam p-valor $< 0,1$. Nota-se grupos com intensidade de correlação mais forte, como os ácidos entre si e expressões gênicas entre si.

Os ácidos são fortemente correlacionados de forma positiva entre si, com exceção do A_8 que possui correlação inversa com os demais. As expressões proteicas possuem baixa correlação com os ácidos. Apenas P_2 com A_5 e P_4 com A_6 foram significantes, ambas positivas. Observou-se correlações positivas e negativas entre vários pares de expressão gênicas. Tais expressões também possuem correlação com os ácidos de forma inversa.

Encontrou-se correlações significantes entre o escore de Friesinger com as variáveis escore de Framingham ($r=0,31$;p-valor $<0,01$) e o gene ALOX15 ($r=0,23$;p-valor $<0,1$). Das variáveis consideradas fatores clássicos de risco temos colesterol LDL (p-valor $<0,1$), idade (p-valor $<0,01$) e PP (p-valor $<0,001$) com as estimativas das correlações de 0,18, 0,24, -0,31 com a variável escore de Friesinger.

Analisando-se os valores das estimativas das correlações que não foram consideradas significativas com p-valor $> 0,1$, decidiu-se utilizar as variáveis de expressão gênica MYL4 ($r=-0,15$;p-valor=0,2), COX7B ($r=-0,15$;p-valor=0,2), BCL2A1 ($r=-0,13$;p-valor=0,3), IL18R1 ($r=-0,13$;p-valor=0,3), MMP9 ($r=-0,10$;p-valor=0,4) e ácido Gama-linoleico ($r=-0,12$;p-valor=0,3) como potenciais fatores preditivos de associação com o escore de Friesinger.

As Figuras 3.2 e 3.3 apresentam os histogramas das variáveis contínuas avaliadas na diagonal com densidade de probabilidade. As estimativas dos coeficientes de correlação de Pearson entre os pares das variáveis acima da diagonal e o diagrama da dispersão bivariada abaixo da diagonal. As elipses entorno dos pontos dentro do diagrama de dispersão são utilizadas para representar o grau de intensidade das correlações.

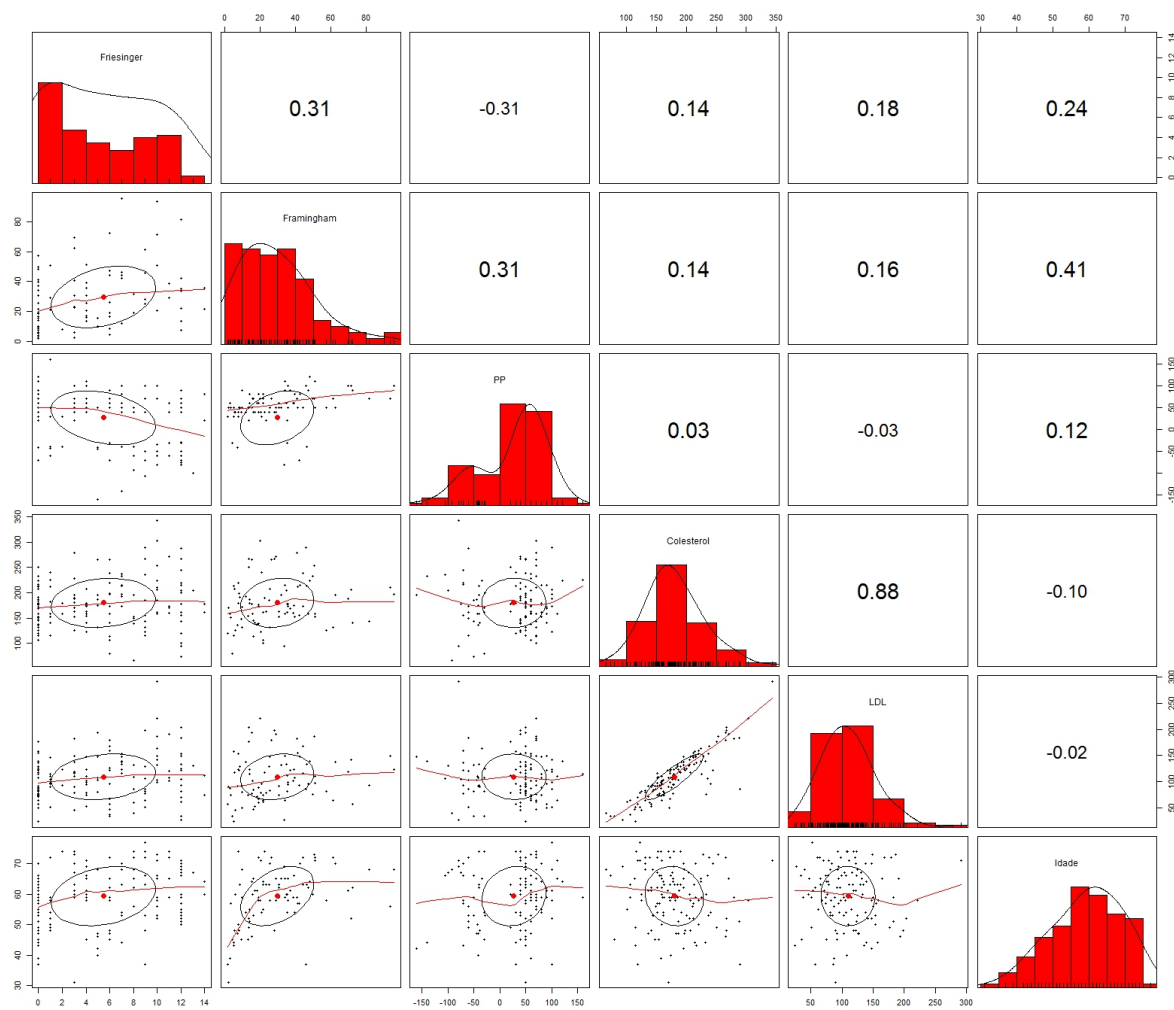


Figura 3.2: Gráfico dos histogramas na diagonal, estimativas dos coeficientes de correlação de Pearson entre os pares acima da diagonal e o diagrama da dispersão bivariada abaixo da diagonal

3.1. ANÁLISE DESCRITIVA

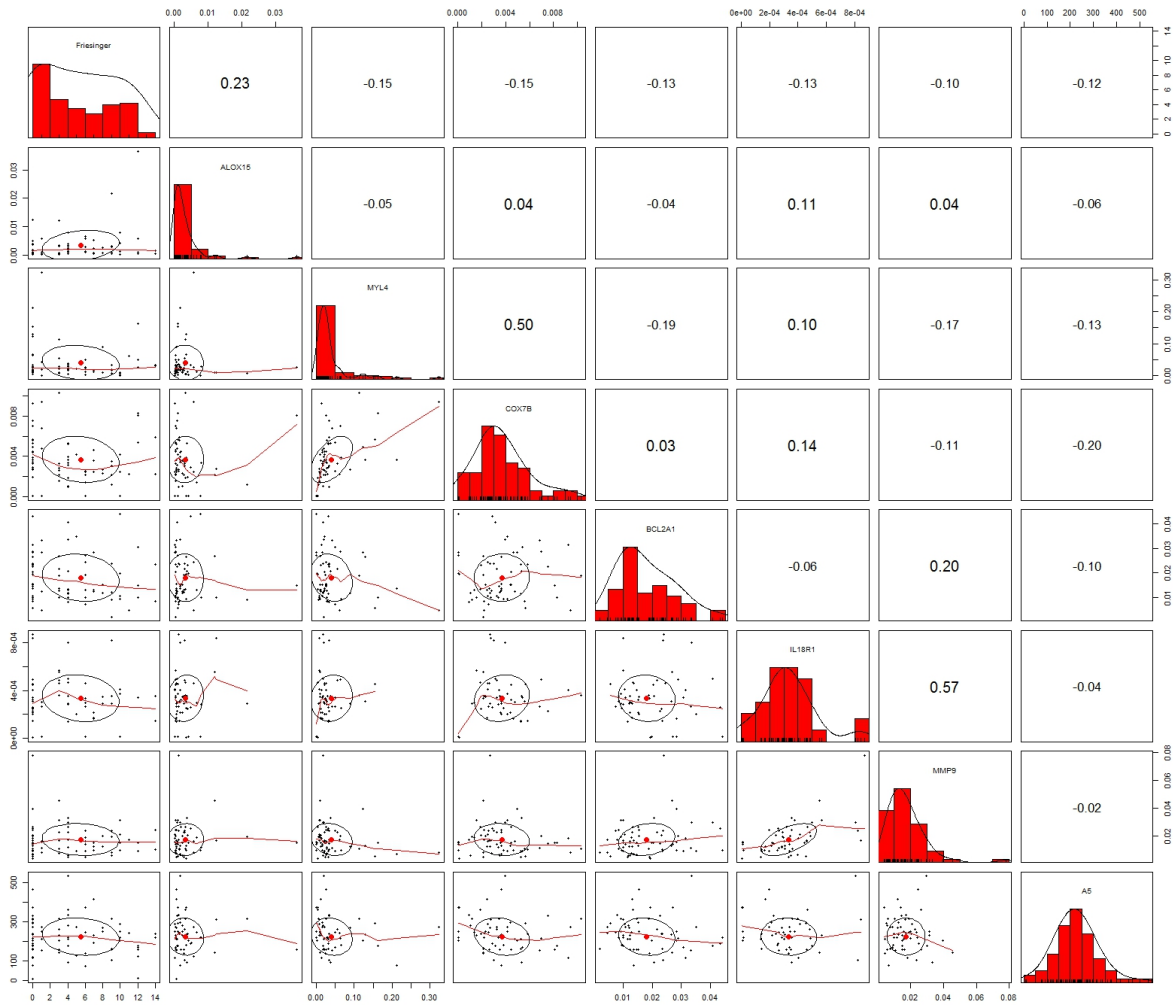


Figura 3.3: Gráfico dos histogramas na diagonal, estimativas dos coeficientes de correlação de Pearson entre os pares de expressão gênica e ácidos acima da diagonal e o diagrama da dispersão bivariada abaixo da diagonal

3.2 Regressão Linear múltipla

Um modelo com os fatores de risco que apresentaram significância no estudo foram o escore de Framingham, colesterol LDL e PP. Levando em consideração os valores faltantes encontrados, 44 observações foram excluídas dessa análise por não apresentarem valores de alguma das variáveis explicativas. O primeiro modelo obteve p -valor $< 0,001$, $R^2 = 0,20$ e $R_a^2 = 0,17$.

Tabela 3.5: Modelo para Friesinger com alguns fatores de risco conhecidos

Parâmetro	Estimativa	Erro padrão	P-valor
Intercepto	1,92	1,32	0,15
Framingham	0,08	0,02	$< 0,001$
PP	-0,03	0,01	$< 0,01$
LDL	0,02	0,01	$< 0,05$

Notas: PP-pressão de pulso e LDL-Lipoproteínas de baixa densidade.

A Tabela 3.5 apresenta os coeficientes estimados e seus respectivos p -valores. Para tal modelo, pela estimativa do coeficiente, observa-se um aumento em um ponto no escore de Framingham temos um aumento em 0,08 pontos no escore de Friesinger. Levando em consideração que Friesinger é uma variável quantitativa discreta, em média, se aumentarmos em 12,5 o valor do escore de Framingham, o escore de Friesinger aumentaria em 1 ponto. A mesma ideia ocorre para colesterol LDL, sendo a variação de 0,02. Para a PP, temos que para o aumento de uma unidade, quanto maior for a diferença entre as pressões sistólica e diastólica, ocorre um decréscimo de 0,03 no escore de Friesinger.

Ao sugerirmos um modelo que incluía as variáveis expressões gênicas, expressões proteicas e níveis de ácidos graxos observou-se o problema de observações deletadas. As expressões e os ácidos tornavam-se não significativos e afetavam substancialmente as demais variáveis preditoras, que também, deixavam de ser significativas para explicar a variação da variável escore de Friesinger.

Portanto, um possível modelo para Friesinger, com a ressalva de 65 observações deletadas, $R^2 = 0,13$, $R_a^2 = 0,10$ e p -valor $< 0,01$ é apresentado na Tabela 3.6 e pode ser considerado um possível modelo para futuras análises mais apuradas.

Tabela 3.6: Possível Modelo para Friesinger com gene e LDL

Parâmetro	Estimativa	Erro padrão	P-valor
Intercepto	0,52	1,53	0,74
ALOX15	205,84	92,79	$< 0,04$
LDL	0,03	0,01	$< 0,03$

Notas: PP-pressão de pulso e LDL-Lipoproteínas de baixa densidade.

3.2.1 Adequação do modelo

O gráfico dos resíduos avalia a adequação do modelo proposto e as suposições utilizadas para a construção do modelo.

Na Figura 3.4, o primeiro gráfico apresenta a dispersão dos resíduos com a variável preditora e considera-se um perfil aleatório. O segundo gráfico é de probabilidade normal em escala transformada e apresenta observações que se desviam da reta nas caldas. O teste de hipótese para a normalidade dos resíduos apresentou um p-valor $< 0,05$. Rejeitando-se a hipótese nula (H_0) de não normalidade dos resíduos. Para o terceiro, temos uma forma alternativa que leva em consideração a variância. Os resíduos studentizados apresentam uma curvatura nos dados e sugere a não homogeneidade nas variâncias. O último gráfico mostra a distância de Cook para a influência dos dados na estimação dos parâmetros. Destaca-se na Figura 3.5 três observações de maior valor mas ambas ainda dentro do limite 1.

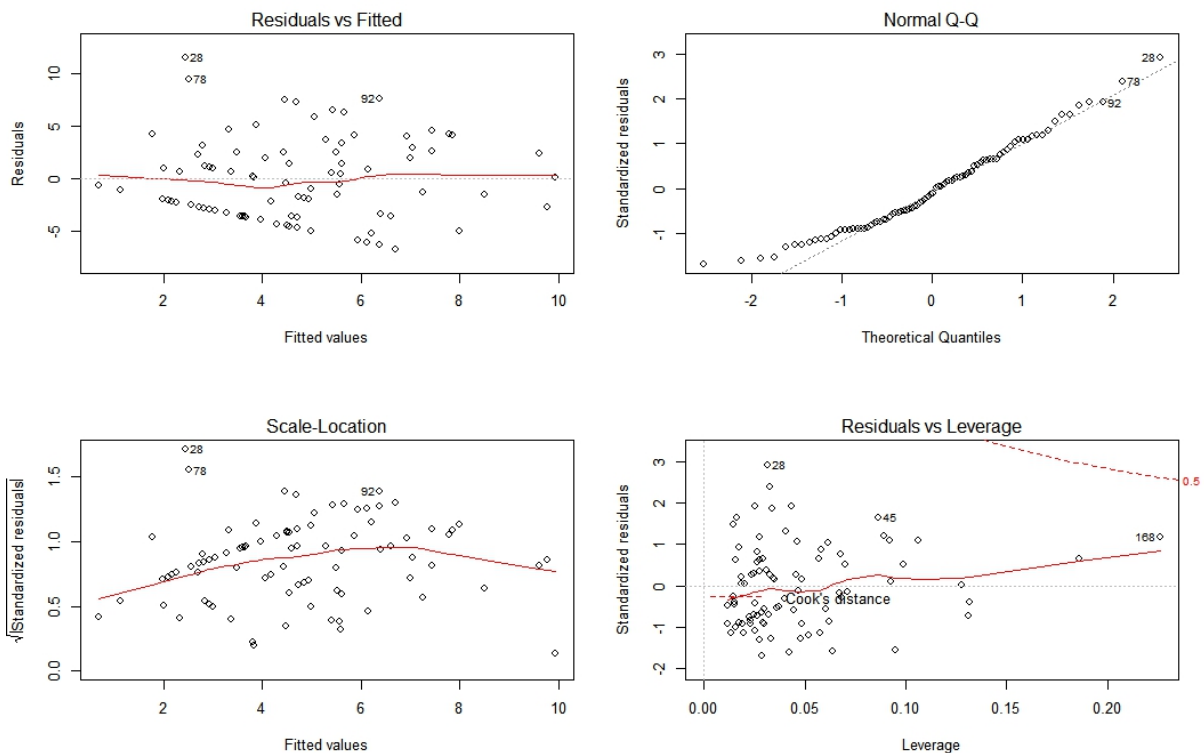


Figura 3.4: Gráfico dos resíduos para o modelo com os fatores clássicos de risco

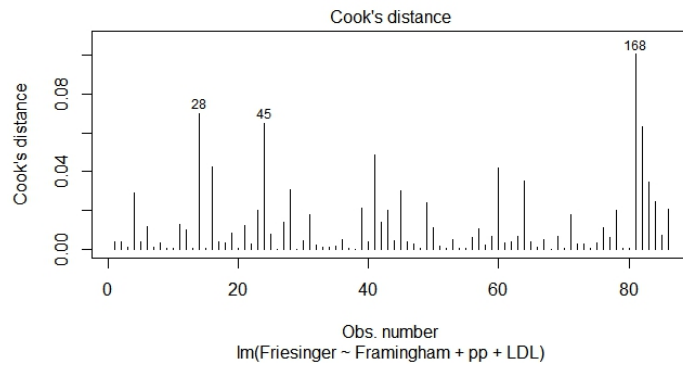


Figura 3.5: Gráfico da distância de Cook

Para o segundo modelo, também não encontramos um bom ajuste aos conjunto de dados. O primeiro gráfico de dispersão apresenta valores concentrados à esquerda. O gráfico do quantil da probabilidade normal em escala transformada, também, observações que se desviam da reta na calda. O teste de hipótese de normalidade apresentou um $p\text{-valor} < 0,01$ e confirma a não normalidade dos resíduos. O gráfico com os resíduos studentizados observa-se uma concentração à esquerda e um valor aparentemente discrepante. A distância de Cook para a influência dos dados não apresentou nenhuma observação fora dos limites. A Figura 3.7 destaca três observações com maior distância.

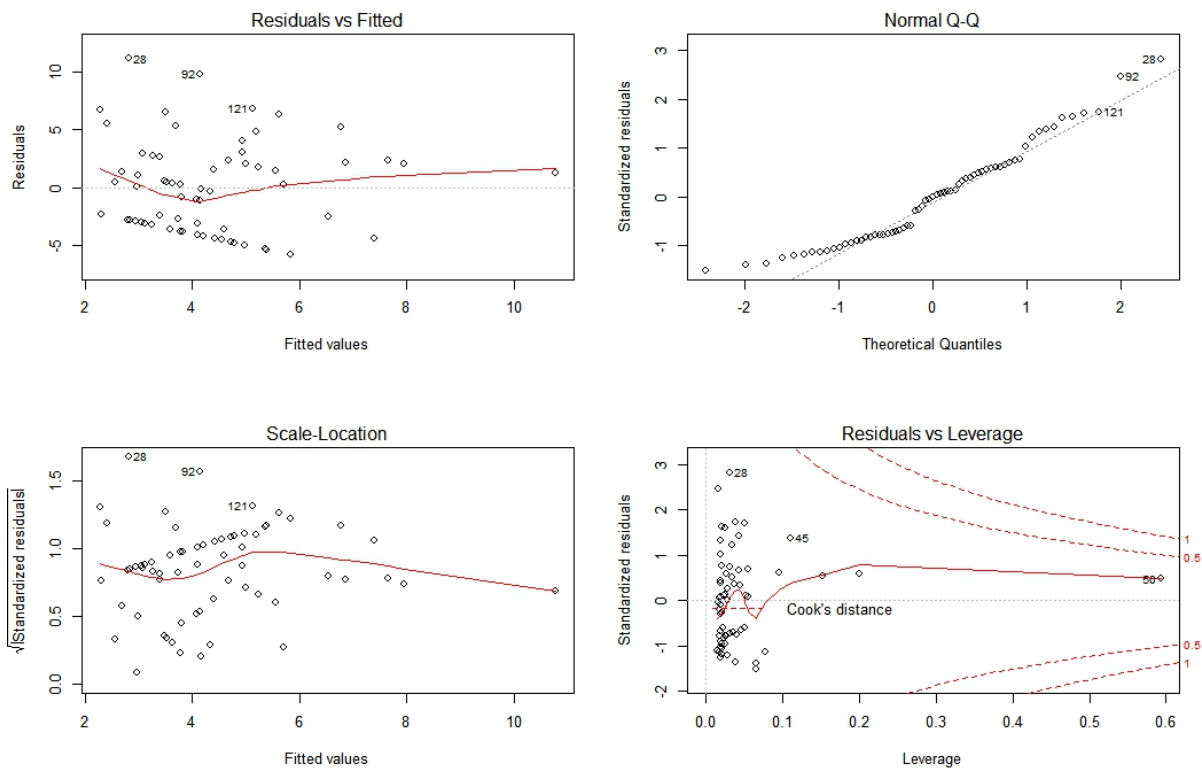


Figura 3.6: Gráfico dos resíduos para o modelo com ALOX15 e LDL

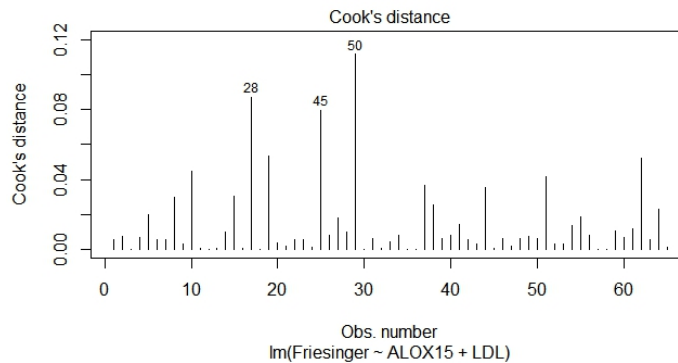


Figura 3.7: Gráfico da distância de Cook

Mediante a violação das suposições, até mesmo o modelo que apresenta variáveis explicativas significantes já conhecidas por outros estudos e, por não apresentarem um bom ajuste nos dados, sugere-se novas análises e novo modelo.

3.3 Modelo linear generalizado

Tendo em vista que as análise dos resíduos dos modelos de regressão linear múltipla sugeridos busca-se um melhor modelo para a variação da variável Friesinger. Utilizou-se um modelo linear generalizado com distribuição de Poisson.

A Tabela 3.7 contém os valores das estimativas, erros padrão e p-valores para o modelo linear generalizado de Poisson. Três das variáveis com estimativas de correlação consideradas potenciais para a correlação com Friesinger apresentaram-se significativas para o modelo generalizado. O gene BCL2A1 ($\beta = -20,7$), COX7B ($\beta = -70,9$) e ácido gamalinoleico ($\beta = -0,0031$) todos com valores negativos indicam um relação inversa para o aumento no escore de Friesinger. LDL, idade e Framingham aparecem no modelo com os coeficientes iguais a 0,01 indicando que, quando as variáveis fatores de risco aumentam, o escore também sofre aumento. Apenas o intercepto se manteve não significativo com p-valor = 0,12. Setenta e oito observações foram deletadas por falta de informações e o resíduo *Deviance* foi de 197,76 com 45 graus de liberdade sendo necessárias 6 interações. O critério de seleção do modelo AIC foi de 336,52.

Tabela 3.7: Modelo para Friesinger com alguns fatores de risco conhecidos

Parâmetro	Estimativa	Erro padrão	P-valor
Intercepto	0,90	0,6	0,12
Framingham	0,01	0,0	<0,0001
BCL2A1	-20,7	6,6	<0,001
COX7B	-70,9	30,1	<0,01
AC. Gamalinoleico	-0,0031	0,0	<0,001
LDL	0,01	0,0	<0,0001
Idade	0,01	0,01	<0,1

Notas: AC - ácido, LDL-Lipoproteínas de baixa densidade.

A análise de resíduos e diagnóstico do ajuste para os modelos lineares generalizados é semelhante à análise de resíduos para o modelo linear múltiplo, com uma ou outra adaptação devido a estrutura dos modelos generalizados.

Para o modelo generalizado o gráfico de dispersão dos resíduos e dos resíduos studentizados apresenta valores mais dispersos em torno do valor zero mas com algumas observações mais relevantes, principalmente quando se avalia o valor padronizado. O gráfico do quantil da probabilidade apresenta observações que se desviam da reta nas caldas. A distância de Cook identificou pontos fora dos limites, o que indica influência da observação na estimação dos parâmetros do modelo.

A Figura 3.10 evidencia um ajuste não muito adequado. Esse fato pode ser ocasionado pela presença de zeros no escore de Friesinger.

3.3. MODELO LINEAR GENERALIZADO

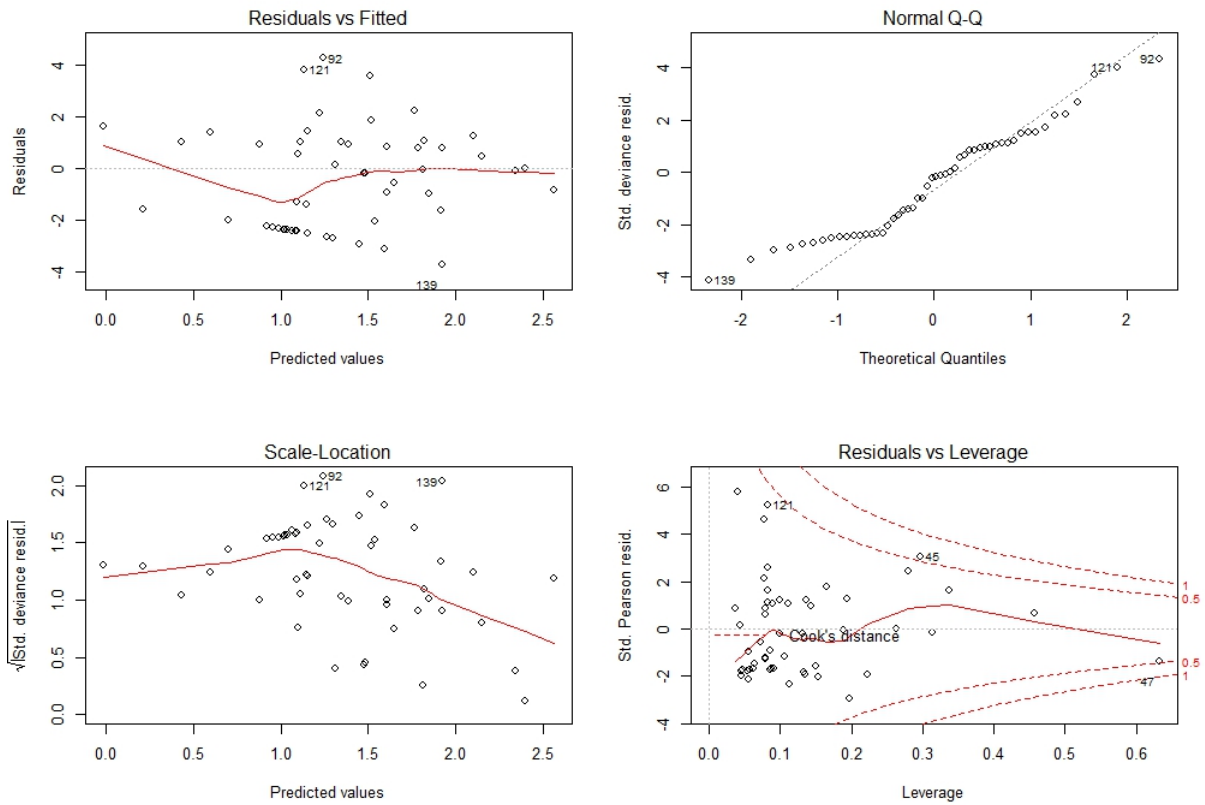


Figura 3.8: Gráfico dos resíduos para o modelo linear generalizado com Distribuição de Poisson

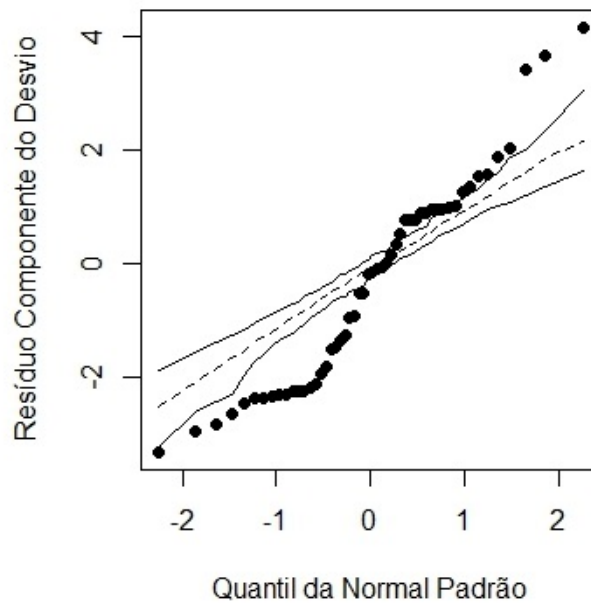


Figura 3.9: Envelope simulado para os componentes do desvio no modelo de Poisson

Capítulo 4

Considerações Finais

Neste trabalho foram avaliadas as relações de expressões gênicas, expressões proteicas e ácidos graxos com o escore de Friesinger. Com a finalidade de identificar a relação das covariáveis com os diferentes níveis de lesão aterosclerótica. Avaliou-se os treze genes, quatro proteínas e dez ácidos, bem com a relação entre o escore de Framingham e os fatores clássicos de risco.

O escore de Framingham apresentou uma correlação moderada porém significativa com o escore de Friesinger. Destacou-se como sendo uma das variáveis mais significantes apresentadas nos modelos de regressão. Não se encontrou um modelo com bom ajuste e sendo o escore de Framingham usado vastamente pelos cardiologistas como uma estimativa do risco de desenvolver DCVs, percebe-se que ainda faz-se necessário buscar melhores formas de estimar a aterosclerose e outras DCVs.

Observou-se correlação moderada de Friesinger positiva com os genes ALOX15, MYL4, COX7B, BCL2A1, IL18R1 e MMP9 e com o ácido Gamalinoleico, sendo significativa para o modelo de regressão múltipla proposto apenas o gene ALOX15. Estudos anteriores reportaram a associação entre o gene lipoxigenase 12/15 e a aterosclerose em camundongos [21]. Em lesões ateroscleróticas humanas avançadas, grandes quantidades de produtos de oxigenação foram detectadas cuja estrutura sugere uma origem não enzimática. Estes dados sugerem que a *arachidonate 15-lipoxigenase*, é um fisiopatológico importante durante as fases iniciais da aterosclerose. Em fases posteriores da placa com peroxidação de lipídios não enzimáticos o desenvolvimento torna-se mais relevante [22].

Em estudo realizado pela Universidade do Rio Grande do Norte no Laboratório de Bioquímica Clínica e Biologia Molecular, também não se observou nenhuma associação entre os genes AREG, TREML4, BCL2A1, BCL2L1, IRS2, KCNE1, MMP9, IL18R1, ALOX15 com o escore de Friesinger. Porém, encontrou significância para os genes CA1, COX7B, ECHDC3 e MYL4. Portanto, MYL4 e COX7B. Podendo vir a serem alvos de novos estudos pela relação encontrada em ambos os estudos.

Destaca-se o ácido Gamalinoleico em tal estudo pois foi o ácido que apresentou menor p-valor na correlação, semelhante ao presente trabalho. Os demais ácidos também não apresentaram associação.

No presente estudo as variáveis sexo, IMC e triglicérides não se mostraram significantes. Fato curioso, já que tais variáveis são considerados como fatores de risco de DCV estabelecidos na literatura. Isso pode ser um indicativo do perfil homogêneo do grupo dos indivíduos selecionados em relação aos fatores clássicos de risco considerados nas análises. Os participantes apresentavam características plausíveis para a realização do cateterismo perdendo a referência de predição das DCVs. Com perfil menos severo para aterosclerose, os fatores clássicos de risco, antes descritos, deixam de ser significantes e o nível de lesão acaba por não depender dos fatores.

Sabe-se que o gene ALOX15, o escore de Framingham, PP, idade e colesterol LDL tiveram correlações consideradas significativas com o escore de Friesinger. Porém, ao se utilizar um modelo de regressão linear múltipla com as variáveis gene, proteína, ácidos e fatores de risco observou-se que Framingham, idade e PP deixaram de ser significativas.

Na tentativa de buscar um melhor modelo após as análises de resíduos do modelo linear múltiplo, encontrou-se no modelo generalizado outras potenciais variáveis preditoras mas sem sucesso na adequação dos dados ao novo modelo. Destaca-se no modelo linear generalizado as variáveis gene BCL2A1, COX7B, ácido Gama-linoleico, Framingham, LDL e Idade como sendo significativas para explicar a variabilidade do escore de Friesinger. Um fator importante que justificaria o ajuste não tão adequado são as 78 observações deletadas por falta de informação. A partir dessas informações, modelos alternativos devem ser considerados a fim de melhorar a variabilidade presente no conjunto de dados.

Após diversas análises, foi concluído que quando as variáveis de expressão gênica, proteica ou ácidos graxos eram incluídas nos modelos, devido ao número muito elevado de observações não coletadas, perdiam-se informações relevantes. Devido a esse número reduzido de informações das expressões e ácidos os resultados do presente trabalho não mostraram-se favoráveis, fazendo com que mais da metade das observações fossem deletadas.

Vários estudos evidenciam a importância de se conhecer novos marcadores para DCV precoce e marcadores de aterosclerose menos invasivos. Novos estudos são necessários para elucidar o papel das expressões gênicas, proteicas e ácidos graxos no desenvolvimento da aterosclerose.

Referências Bibliográficas

- [1] BUSSAB, W.O.; MORETTIN, P.A. **ESTATÍSTICA BÁSICA**. 7 ed. São Paulo, EDITORA, 2012.
- [2] CHARNET, R., et al. **Análise de Regressão Linear: com aplicações**. 2 ed. Campinas, SP: Editora da Unicamp, 2008.
- [3] Marciano, F. W. P., **Principais tipos de resíduos utilizados na análise de diagnóstico em MLG com aplicações para os modelos: Poisson, ZIP e ZINB**. Fortaleza-CE, Setembro, 2009. 14
- [4] Kutner, Michael H., et al. **Applied linear statistical models**. 5 ed. McGraw-Hill Irwin series Operations and decision sciences, 2004. 10, 13
- [5] Kleinbaum, David G., et al. **Applied regression Analysis and other multivariable methods**. 4 ed. Emory University, Brooks/Cole Cengage Learning 2008. 11
- [6] Weisberg, S. **Applied linear Regression**. 2 ed. University Minnesota, St. Paul, Minnesota 1985.
- [7] MOTTA, V.T. **BIOESTATÍSTICA**. 2 ed. Livro da Universidade de Caxias do Sul, EDITORA, 2000.
- [8] Ralph B. D'Agostino, Sr, PhD; Ramachandran S. Vasan, MD; Michael J. Pencina, PhD; Philip A. Wolf, MD; Mark Cobain, PhD; Joseph M. Massaro, PhD; William B. Kannel, MD **Prediction of coronary heart disease using risk factor categories**. *Circulation*. 2008; 117: 743-753. January 22, 2008 6
- [9] WILSON PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. **Prediction of coronary heart disease using risk factor categories**. *Circulation*. 1998;97:1837-1847. 6
- [10] LOTUFO, P.A., **O Escore de Risco de Framingham para doenças cardiovasculares**. *Revista de Medicina*. São Paulo, out.-dez.;87(4):232-7, 2008. 2, 6
- [11] Deepesh Agarwal, Ramanand P. Sinha, Ram Mohan Jaiswal, Nikhil Pursnani, et al., Link between Friesinger Index and Lipid Profile in Patients Undergoing Diagnostic Coronary

Angiography. International Journal of Clinical Medicine, 2015, 6, 538-546. Published Online August 2015 in SciRes. <http://www.scirp.org/journal/ijcm> 7

- [12] DUARTE, V.H.R., AVALIAÇÃO DA EXPRESSÃO E GENOTIPAGEM DOS GENES TREML4 E ECHDC3 COMO POTENCIAIS BIOMARCADORES DAS DOENÇAS CARDIOVASCULARES. Rio grande do Norte, 2015. 1, 2, 3, 5, 7, 8, 9
- [13] White MY, Tchen AS, McCarron HC, Hambly BD, Jeremy RW, Cordwell SJ., **Proteomics of ischemia and reperfusion injuries in rabbit myocardium with and without intervention by an oxygen-free radical scavenger.** Proteomics. 2006 Dec;6(23):6221-33. 3
- [14] Glatz J.F., G.J. van der Vusse, Stam H.C., Reneman R.S., **Fatty acid homeostasis in the normoxic and ischemic heart.** Physiological Reviews Published 1 October 1992 Vol. 72 no. 4, 881-940] 3
- [15] TOTH PP., **Making a case for quantitative assessment of cardiovascular risk.** J Clin Lipidol. 2007 Aug;1(4):234?241. 8
- [16] RIDKER, P. M., et al. (2000), **C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women.**, N Engl J Med, 342 (12), 836-43. 8
- [17] PAULA, G.A. **Modelos de Regressão com apoio computacional.** Instituto de Matemática e Estatística. São Paulo, 2013.
- [18] III Diretrizes brasileiras sobre dislipidemias e diretriz de prevenção de aterosclerose do Departamento de Aterosclerose da Sociedade Brasileira de Cardiologia. Arq Bras Cardiol 2001; 77 Supl 3:S1-S48, São Paulo, Nov. 2001. 1
- [19] AZAMBUJA, M.I.R. , et al. **Impacto Econômico dos Casos de Doença Cardiovascular Grave no Brasil: Uma estimativa baseada em dados secundários.** Arq. Bras. Cardiol. [online]. 2008, vol.91, n.3, pp. 163-171. ISSN 0066-782X. 1
- [20] Javed F., Aziz E. F., Nadkarni N., Khan S. A., et al. **The association of the Friesinger score and pulse pressure in an urban South Asian patient population: pulse pressure, an independent predictor of coronary artery disease.** Journal of the American Society of Hypertension. Volume 4, Issue 3, May?June 2010, Pages 142?147 6
- [21] Cyrus, T., Witztum, J.L., Rader, D.J., Tangirala, R., Fazio, S., Linton, M.F., Funk, C.D. J. Clin. Invest. **Disruption of the 12/15-lipoxygenase gene diminishes atherosclerosis in apo E-deficient.** J Clin Invest. 1999 Jun 1; 103(11): 1597-1604. 31
- [22] H Kühn, J Belkner, S Zaiss, T Fährenklempfer, and S Wohlfeil. **Involvement of 15-lipoxygenase in early stages of atherogenesis.** Published June 1, 1994. JEM vol. 179 no. 6 1903-1911. The Rockefeller University Press, doi: 10.1084/jem.179.6.1903. 31
- [23] Brasil, Ministério da Saúde. Banco de dados do Sistema Único de Saúde-DATASUS. Disponível em <http://www.datasus.gov.br> [Acessado em 3 de março de 2008]. 1

- [24] GOMES, R.C. Portal Brasil. Publicado em: 06/09/2011. Última modificação: 28/07/2014. **Doenças Cardiovasculares Causam Quase 30% das Mortes no País**. Ministério da Saúde. Disponível em: <<http://www.brasil.gov.br/saude/2011/09/doencas-cardiovasculares-causam-quase-30-das-mortes-no-pais>>. Acesso em:16/05/2015. 1, 2
- [25] R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <<http://www.R-project.org/>>.

Apêndice A

Gráficos e Tabela adicionais

A.1 Estimativas de Coeficientes de Correlação de Pearson entre expressões gênicas e expressões proteicas

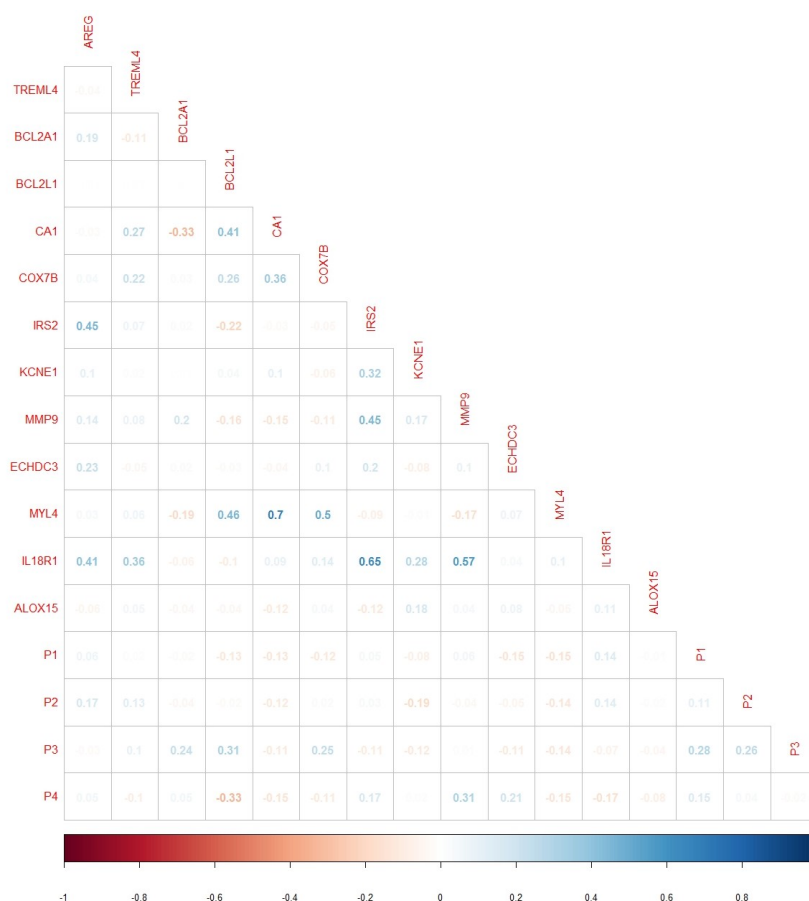


Figura A.1: Gráfico das correlações

AREG: *Amphiregulin*; TREML4: *Triggering Receptor Expressed On Myeloid Cells-Like 4*; BCL2A1: *BCL2-Related Protein A1*; BCL2L1: *BCL2-Like 1*; CA1: *Carbonic Anhydrase I*; COX7B: *Cytochrome C Oxidase Subunit VIIb*; IRS2: *Insulin Receptor Substrate 2*; KCNE1: *Potassium Channel, Voltage Gated Subfamily E Regulatory Beta Subunit 1*; MMP9: *Matrix Metalloproteinase 9*; ECHDC3: *Enoyl CoA Hydratase Domain Containing 3*; MYL4: *Myosin, Light Chain 4*; IL18R1: *Interleukin 18 Receptor 1*; ALOX15: *Arachidonate 15-Lipoxygenase*; P₁:E-Selectina Luminex; P₂:vCAM Luminex; P₃:iCAM Luminex e P₄:MMP9 Luminex.

A.2 Estimativas de Coeficientes de Correlação de Pearson entre níveis de ácidos graxos

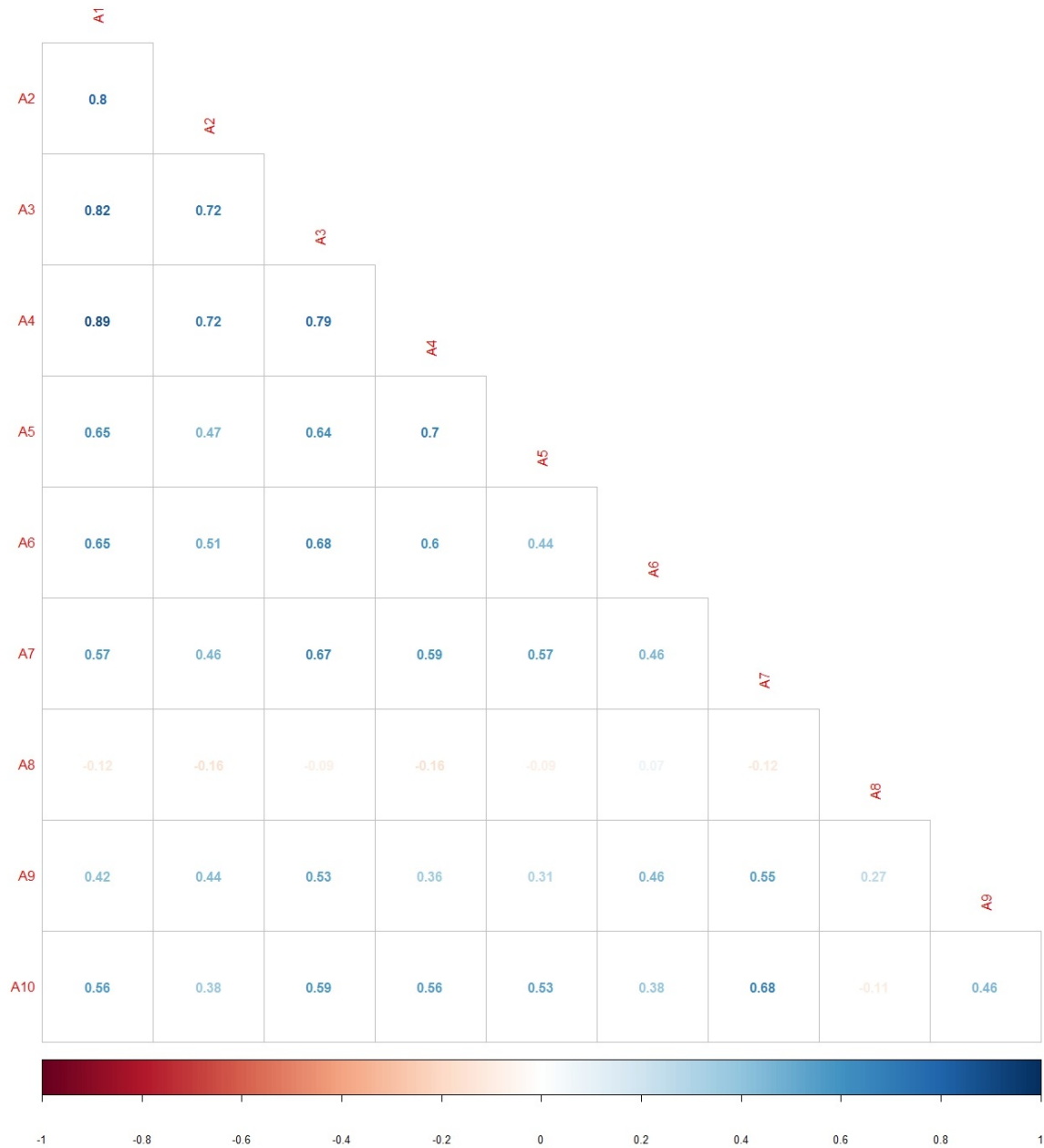


Figura A.2: Gráfico das correlações

A₁-Ácido mirfístico, A₂-Ácido miristoleico, A₃-Ácido oleico, A₄-Ácido linoleico, A₅-Ácido gamalinoleico, A₆-Ácido linolênico, A₇-Ácido araquidônico, A₈-Ácido eicosapentaenoico, A₉-Ácido lignocérico e A₁₀-Ácido docosaenoico.

A.3 Gráfico de diagnóstico para não linearidade das variáveis dos modelos propostos

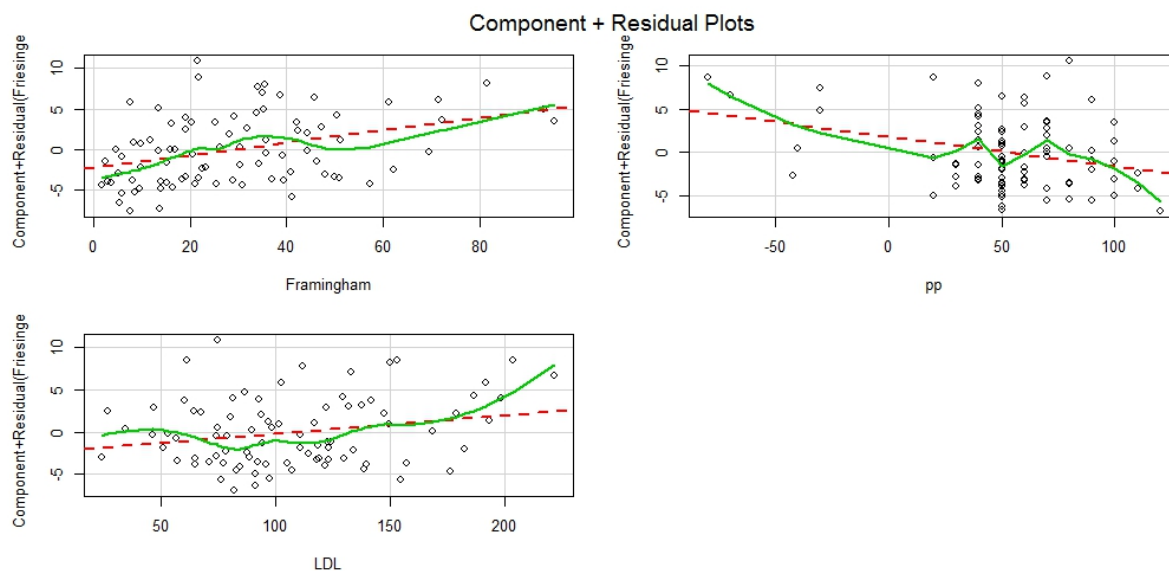


Figura A.3: Gráfico dos resíduos parciais e as variáveis do modelo linear com os fatores clássicos de risco

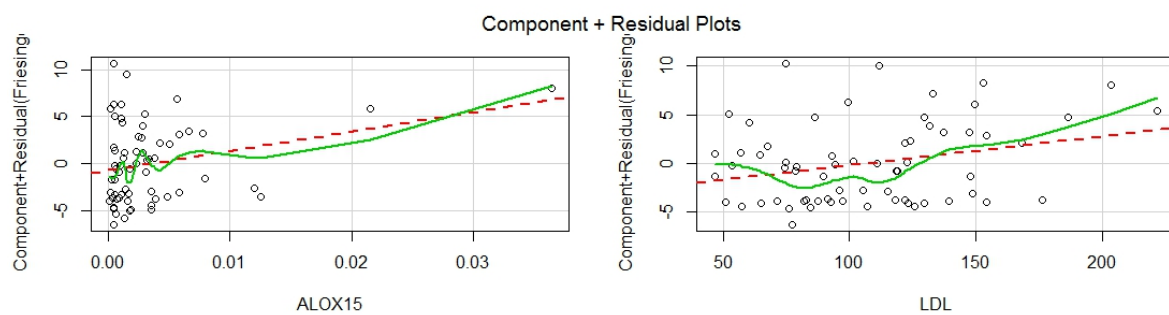


Figura A.4: Gráfico dos resíduos parciais e as variáveis do modelo linear com a variável gene ALOX15

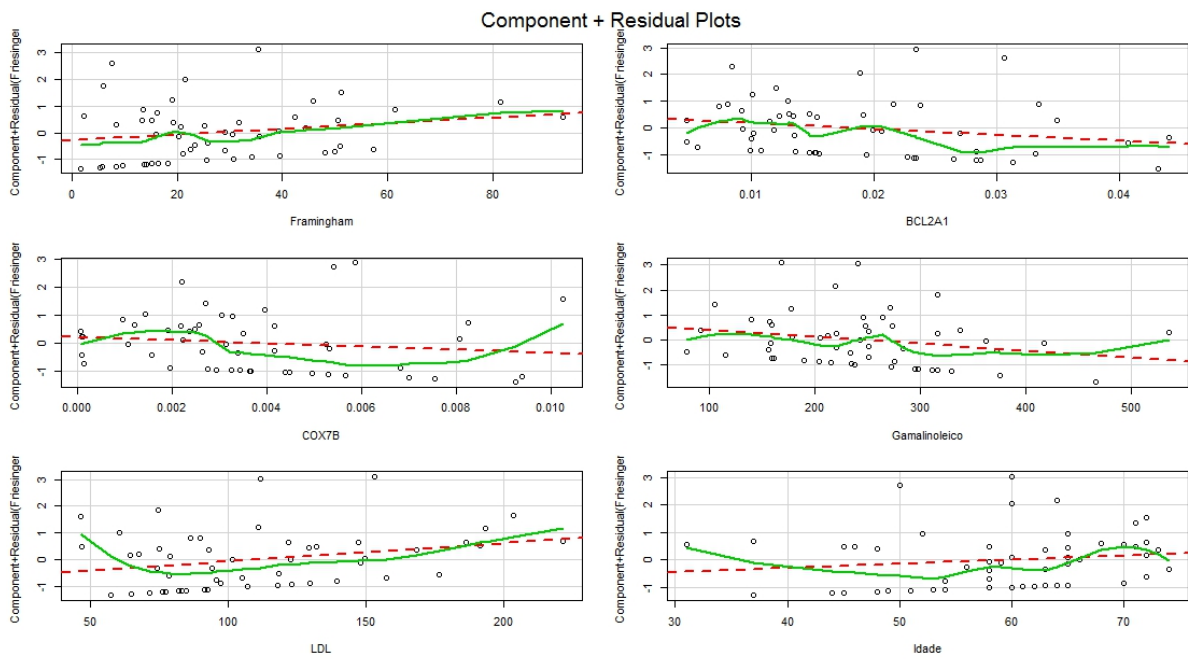


Figura A.5: Gráfico dos resíduos parciais e as variáveis do modelo linear generalizado

A.4 Tabela completa das estimativas de Coeficientes de Correlação de Pearson e P-valores

Tabela A.1: Tabela das estimativas da correlação de Pearson

Variáveis		Correlação	P-valor
Ac. Araquidônico	Ac. Docosaexaenoico	0.68	0.0000
Ac. Araquidônico	Ac. Lignocérico	0.55	0.0000
Ac. Eicosapentaenoico	Ac. Lignocérico	0.27	0.0492
Ac. Gamalinoleico	Ac. Araquidônico	0.57	0.0000
Ac. Gamalinoleico	Ac. Docosaexaenoico	0.53	0.0000
Ac. Gamalinoleico	Ac. Lignocérico	0.31	0.0203
Ac. Gamalinoleico	Ac. Linolênico	0.44	0.0002
Ac. Lignocérico	Ac. Docosaexaenoico	0.46	0.0003
Ac. Linoleico	Ac. Araquidônico	0.59	0.0000
Ac. Linoleico	Ac. Docosaexaenoico	0.56	0.0000
Ac. Linoleico	Ac. Gamalinoleico	0.70	0.0000
Ac. Linoleico	Ac. Lignocérico	0.36	0.0058
Ac. Linoleico	Ac. Linolênico	0.60	0.0000
Ac. Linolênico	Ac. Araquidônico	0.46	0.0001
Ac. Linolênico	Ac. Docosaexaenoico	0.38	0.0013
Ac. Linolênico	Ac. Lignocérico	0.46	0.0004
Ac. Mirístico	Ac. Araquidônico	0.57	0.0000
Ac. Mirístico	Ac. Docosaexaenoico	0.56	0.0000
Ac. Mirístico	Ac. Gamalinoleico	0.65	0.0000
Ac. Mirístico	Ac. Lignocérico	0.42	0.0010
Ac. Mirístico	Ac. Linoleico	0.89	-
Ac. Mirístico	Ac. Linolênico	0.65	0.0000
Ac. Mirístico	Ac. Miristoleico	0.80	0.0000
Ac. Mirístico	Ac. Oleico	0.82	-
Ac. Miristoleico	Ac. Araquidônico	0.46	0.0001
Ac. Miristoleico	Ac. Docosaexaenoico	0.38	0.0018
Ac. Miristoleico	Ac. Gamalinoleico	0.47	0.0001
Ac. Miristoleico	Ac. Lignocérico	0.44	0.0008
Ac. Miristoleico	Ac. Linoleico	0.72	0.0000
Ac. Miristoleico	Ac. Linolênico	0.51	0.0000
Ac. Miristoleico	Ac. Oleico	0.72	0.0000
Ac. Oleico	Ac. Araquidônico	0.67	0.0000
Ac. Oleico	Ac. Docosaexaenoico	0.59	0.0000
Ac. Oleico	Ac. Gamalinoleico	0.64	0.0000
Ac. Oleico	Ac. Lignocérico	0.53	0.0000
Ac. Oleico	Ac. Linoleico	0.79	0.0000
Ac. Oleico	Ac. Linolênico	0.68	0.0000
Ac. Araquidônico	Ac. Eicosapentaenoico	-0.12	0.3566
Ac. Eicosapentaenoico	Ac. docosaexaenoico	-0.11	0.4032
Ac. Gamalinoleico	Ac. Eicosapentaenoico	-0.09	0.4954
Ac. Linoleico	Ac. Eicosapentaenoico	-0.16	0.2087
Ac. Linolênico	Ac. Eicosapentaenoico	0.07	0.5850
Ac. Mirístico	Ac. Eicosapentaenoico	-0.12	0.3496
Ac. Miristoleico	Ac. Eicosapentaenoico	-0.16	0.2245
Ac. Oleico	Ac. Eicosapentaenoico	-0.09	0.4977

Tabela A.2: Tabela das estimativas da correlação de Pearson

Variáveis		Correlação	P-valor
ALOX15	Ac. Araquidônico	0.06	0.6667
ALOX15	Ac. docosaexaenoico	0.08	0.5424
ALOX15	Ac. Eicosapentaenoico	-0.18	0.1834
ALOX15	Ac. Gamalinoleico	-0.06	0.6420
ALOX15	Ac. Lignocerico	-0.01	0.9689
ALOX15	Ac. Linoleico	0.04	0.7517
ALOX15	Ac. Linolênico	0.06	0.6681
ALOX15	Ac. Mirístico	0.03	0.8241
ALOX15	Ac. Miristoleico	0.00	0.9844
ALOX15	Ac. Oleico	-0.03	0.8302
ALOX15	E-Selectina	-0.01	0.9213
ALOX15	iCAM	-0.04	0.7263
ALOX15	P. MMP9	-0.08	0.5259
ALOX15	vCAM	-0.02	0.8954
AREG	Ac. Docosaexaenoico	0.25	0.0533
AREG	Ac. Araquidônico	0.00	0.9927
AREG	Ac. Eicosapentaenoico	0.03	0.8094
AREG	Ac. Gamalinoleico	-0.07	0.5878
AREG	Ac. Lignocerico	0.14	0.3176
AREG	Ac. Linoleico	-0.03	0.8051
AREG	Ac. Linolênico	-0.02	0.8632
AREG	Ac. Mirístico	-0.03	0.8251
AREG	Ac. Miristoleico	-0.09	0.4898
AREG	Ac. Oleico	0.01	0.9149
AREG	ALOX15	-0.06	0.6702
AREG	BCLA1	0.19	0.1359
AREG	BCLL1	-0.01	0.9663
AREG	CA1	-0.03	0.8166
AREG	COX7B	0.04	0.7331
AREG	E-Selectina	0.06	0.6418
AREG	ECHDC3	0.23	0.0648
AREG	iCAM	-0.03	0.8254
AREG	IL18R1	0.41	0.0020
AREG	IRS	0.45	0.0002
AREG	KCNE1	0.10	0.4490
AREG	MMP9	0.14	0.2517
AREG	P. MMP9	0.05	0.6735
AREG	MYL4	0.03	0.7858
AREG	TREML4	-0.04	0.7748
AREG	vCAM	0.17	0.1835
BCL2A1	Ac. Mirístico	0.23	0.0890
BCL2A1	Ac. Miristoleico	0.22	0.0971
BCL2A1	Ac. Oleico	0.34	0.0093
BCL2A1	Ac. Araquidônico	0.10	0.4724
BCL2A1	Ac. docosaexaenoico	0.17	0.1940
BCL2A1	Ac. Eicosapentaenoico	-0.07	0.5975
BCL2A1	Ac. Gamalinoleico	-0.10	0.4433
BCL2A1	Ac. Lignocerico	0.08	0.5582
BCL2A1	Ac. Linoleico	0.18	0.1681
BCL2A1	Ac. Linolênico	0.13	0.3191
BCL2A1	ALOX15	-0.04	0.7288
BCL2A1	BCLL1	-0.00	0.9750
BCL2A1	CA1	-0.33	0.0068
BCL2A1	COX7B	0.03	0.8293
BCL2A1	E-Selectina	-0.02	0.8463

Tabela A.3: Tabela das estimativas da correlação de Pearson

Variáveis	Correlação	P-valor
BCL2A1	ECHDC3	0.02 0.8597
BCL2A1	iCAM	0.24 0.0613
BCL2A1	IL18R1	-0.06 0.6541
BCL2A1	IRS	0.02 0.8551
BCL2A1	KCNE1	0.01 0.9699
BCL2A1	MMP9	0.20 0.1140
BCL2A1	P. MMP9	0.05 0.6757
BCL2A1	MYL4	-0.19 0.1201
BCL2A1	vCAM	-0.04 0.7433
BCL2L1	Ac. Araquidônico	0.08 0.5651
BCL2L1	Ac. docosaexaenoico	0.02 0.8785
BCL2L1	Ac. Eicosapentaenoico	0.03 0.8010
BCL2L1	Ac. Gamalinoleico	0.04 0.7403
BCL2L1	Ac. Lignocerico	0.01 0.9314
BCL2L1	Ac. Linoleico	-0.06 0.6469
BCL2L1	Ac. Linolênico	-0.06 0.6350
BCL2L1	Ac. Mirístico	-0.04 0.7797
BCL2L1	Ac. Miristoleico	-0.18 0.1698
BCL2L1	Ac. Oleico	0.03 0.8190
BCL2L1	ALOX15	-0.04 0.7401
BCL2L1	CA1	0.41 0.0006
BCL2L1	COX7B	0.26 0.0401
BCL2L1	E-Selectina	-0.13 0.3165
BCL2L1	ECHDC3	-0.03 0.8348
BCL2L1	iCAM	0.31 0.0129
BCL2L1	IL18R1	-0.10 0.4458
BCL2L1	IRS	-0.22 0.0825
BCL2L1	KCNE1	0.04 0.7524
BCL2L1	P. MMP9	-0.33 0.0078
BCL2L1	MMP9	-0.16 0.2119
BCL2L1	MYL4	0.46 0.0001
BCL2L1	vCAM	-0.02 0.8739
CA1	Ac. Linoleico	-0.28 0.0324
CA1	Ac. Mirístico	-0.22 0.0985
CA1	Ac. Miristoleico	-0.25 0.0611
CA1	Ac. Araquidônico	0.04 0.7688
CA1	Ac. docosaexaenoico	-0.03 0.8442
CA1	Ac. Eicosapentaenoico	-0.06 0.6431
CA1	Ac. Gamalinoleico	-0.10 0.4502
CA1	Ac. lignocerico	0.06 0.6459
CA1	Ac. Linolênico	-0.16 0.2211
CA1	Ac. Oleico	-0.11 0.4057
CA1	ALOX15	-0.12 0.3649
CA1	COX7B	0.36 0.0033
CA1	E-Selectina	-0.13 0.3044
CA1	ECHDC3	-0.04 0.7309
CA1	iCAM	-0.11 0.3922
CA1	IL18R1	0.09 0.4966
CA1	IRS	-0.03 0.8304
CA1	KCNE1	0.10 0.4455
CA1	MMP9	-0.15 0.2410
CA1	P. MMP9	-0.15 0.2551
CA1	MYL4	0.70 0.0000
CA1	vCAM	-0.12 0.3575

Tabela A.4: Tabela das estimativas da correlação de Pearson

Variáveis		Correlação	P-valor
COX7B	Ac. Araquidônico	-0.23	0.0865
COX7B	Ac. Linoleico	-0.27	0.0381
COX7B	Ac. Linolênico	-0.30	0.0221
COX7B	Ac. Mirístico	-0.28	0.0362
COX7B	Ac. Docosaexaenoico	-0.05	0.6859
COX7B	Ac. Eicosapentaenoico	-0.11	0.4191
COX7B	Ac. Gamalinoleico	-0.20	0.1399
COX7B	Ac. Lignocerico	-0.14	0.3320
COX7B	Ac. Miristoleico	-0.15	0.2665
COX7B	Ac. Oleico	-0.13	0.3428
COX7B	ALOX15	0.04	0.7495
COX7B	E-Selectina	-0.12	0.3413
COX7B	ECHDC3	0.10	0.4428
COX7B	iCAM	0.25	0.0549
COX7B	IL18R1	0.14	0.3214
COX7B	IRS	-0.05	0.7176
COX7B	KCNE1	-0.06	0.6515
COX7B	MMP9	-0.11	0.3621
COX7B	P. MMP9	-0.11	0.3939
COX7B	MYL4	0.50	0.0000
COX7B	vCAM	0.02	0.8546
E-Selectina	iCAM	0.28	0.0095
E-Selectina	Ac. Araquidônico	0.20	0.1060
E-Selectina	Ac. docosaexaenoico	0.15	0.2310
E-Selectina	Ac. Eicosapentaenoico	0.02	0.8921
E-Selectina	Ac. Gamalinoleico	0.08	0.5021
E-Selectina	Ac. lignocerico	-0.08	0.5381
E-Selectina	Ac. Linoleico	0.15	0.2310
E-Selectina	Ac. Linolênico	-0.05	0.6860
E-Selectina	Ac. miristico	0.09	0.4684
E-Selectina	Ac. mirist Oleico	0.10	0.4317
E-Selectina	Ac. Oleico	0.07	0.5420
E-Selectina	P. MMP9	0.15	0.1622
E-Selectina	vCAM	0.11	0.3317
ECHDC3	Ac. Araquidônico	-0.03	0.8231
ECHDC3	Ac. docosaexaenoico	0.02	0.8829
ECHDC3	Ac. Eicosapentaenoico	-0.04	0.7850
ECHDC3	Ac. Gamalinoleico	0.15	0.2658
ECHDC3	Ac. lignocerico	-0.13	0.3507
ECHDC3	Ac. Linoleico	-0.07	0.6148
ECHDC3	Ac. Linolênico	-0.07	0.5806
ECHDC3	Ac. miristico	-0.08	0.5620
ECHDC3	Ac. mirist Oleico	-0.12	0.3702
ECHDC3	Ac. Oleico	-0.06	0.6290
ECHDC3	ALOX15	0.08	0.5492
ECHDC3	E-Selectina	-0.15	0.2561
ECHDC3	iCAM	-0.11	0.3723
ECHDC3	IL18R1	0.04	0.7725
ECHDC3	MMP9	0.21	0.0999
ECHDC3	MYL4	0.07	0.5711
ECHDC3	vCAM	-0.05	0.6866

Tabela A.5: Tabela das estimativas da correlação de Pearson

Variáveis		Correlação	P-valor
Framingham	Ac. Araquidônico	0.05	0.6897
Framingham	Ac. docosaexaenoico	0.14	0.2648
Framingham	Ac. Eicosapentaenoico	-0.11	0.4226
Framingham	Ac. Gamalinoleico	0.16	0.2328
Framingham	Ac. lignocérico	0.07	0.6433
Framingham	Ac. Linoleico	0.15	0.2384
Framingham	Ac. Linolênico	0.01	0.9688
Framingham	Ac. mirístico	0.08	0.5189
Framingham	Ac. miristOleico	0.18	0.1675
Framingham	Ac. Oleico	0.12	0.3679
Framingham	ALOX15	0.38	0.0033
Framingham	AREG	-0.04	0.7510
Framingham	BCLA1	0.10	0.4409
Framingham	BCLL1	-0.22	0.0888
Framingham	CA1	-0.16	0.2298
Framingham	COX7B	0.03	0.8177
Framingham	E-Selectina	0.01	0.9000
Framingham	ECHDC3	-0.01	0.9564
Framingham	Friesinger	0.31	0.0034
Framingham	iCAM	0.13	0.2831
Framingham	IL18R1	0.13	0.3826
Framingham	IRS	0.10	0.4492
Framingham	KCNE1	0.02	0.9066
Framingham	MMP9	0.17	0.2045
Framingham	P. MMP9	0.12	0.2952
Framingham	MYL4	-0.11	0.4031
Framingham	TREML4	0.01	0.9372
Framingham	vCAM	0.25	0.0290
Friesinger	Ac. Araquidônico	-0.07	0.5687
Friesinger	Ac. docosaexaenoico	-0.02	0.8387
Friesinger	Ac. Eicosapentaenoico	-0.06	0.6233
Friesinger	Ac. Gamalinoleico	-0.12	0.3195
Friesinger	Ac. lignocérico	-0.08	0.5395
Friesinger	Ac. Linoleico	-0.04	0.7390
Friesinger	Ac. Linolênico	-0.07	0.5789
Friesinger	Ac. mirístico	-0.06	0.6228
Friesinger	Ac. miristOleico	-0.05	0.7182
Friesinger	Ac. Oleico	-0.02	0.8520
Friesinger	ALOX15	0.23	0.0660
Friesinger	AREG	-0.02	0.8925
Friesinger	BCLA1	-0.13	0.2839
Friesinger	BCLL1	-0.02	0.9002
Friesinger	CA1	0.07	0.5774
Friesinger	COX7B	-0.15	0.2430
Friesinger	E-Selectina	-0.08	0.4928
Friesinger	ECHDC3	-0.03	0.7809
Friesinger	iCAM	-0.02	0.8316
Friesinger	IL18R1	-0.13	0.3336
Friesinger	IRS	0.01	0.9115
Friesinger	KCNE1	-0.04	0.7364
Friesinger	MMP9	-0.10	0.4456
Friesinger	P. MMP9	-0.02	0.8564
Friesinger	MYL4	-0.15	0.2251
Friesinger	TREML4	0.06	0.6219
Friesinger	vCAM	0.06	0.5682

Tabela A.6: Tabela das estimativas da correlação de Pearson

Variáveis		Correlação	P-valor
iCAM	Ac. Araquidônico	0.07	0.5451
iCAM	Ac. docosaexaenoico	-0.02	0.8404
iCAM	Ac. Eicosapentaenoico	0.20	0.1255
iCAM	Ac. Gamalinoleico	0.18	0.1368
iCAM	Ac. lignocérico	-0.13	0.3229
iCAM	Ac. Linoleico	0.10	0.4075
iCAM	Ac. Linolênico	-0.08	0.5133
iCAM	Ac. mirístico	0.07	0.5489
iCAM	Ac. miristOleico	-0.00	0.9876
iCAM	Ac. Oleico	0.09	0.4831
iCAM	P. MMP9	-0.02	0.8567
IL18R1	Ac. Mirístico	-0.26	0.0672
IL18R1	Ac. Araquidônico	-0.13	0.3739
IL18R1	Ac. docosaexaenoico	-0.13	0.3823
IL18R1	Ac. Eicosapentaenoico	-0.15	0.3194
IL18R1	Ac. Gamalinoleico	-0.04	0.7988
IL18R1	Ac. lignocérico	-0.03	0.8572
IL18R1	Ac. Linoleico	-0.18	0.2053
IL18R1	Ac. Linolênico	-0.21	0.1387
IL18R1	Ac. miristOleico	-0.17	0.2384
IL18R1	Ac. Oleico	-0.14	0.3444
IL18R1	ALOX15	0.11	0.4325
IL18R1	E-Selectina	0.14	0.3219
IL18R1	iCAM	-0.07	0.6271
IL18R1	P. MMP9	-0.17	0.2332
IL18R1	vCAM	0.14	0.3378
IRS2	Ac. Araquidônico	0.01	0.9288
IRS2	Ac. docosaexaenoico	0.03	0.8110
IRS2	Ac. Eicosapentaenoico	0.00	0.9957
IRS2	Ac. Gamalinoleico	0.00	0.9822
IRS2	Ac. lignocérico	-0.06	0.6471
IRS2	Ac. Linoleico	-0.15	0.2729
IRS2	Ac. Linolênico	-0.15	0.2725
IRS2	Ac. mirístico	-0.20	0.1354
IRS2	Ac. miristOleico	-0.19	0.1623
IRS2	Ac. Oleico	-0.02	0.8999
IRS2	ALOX15	-0.12	0.3412
IRS2	E-Selectina	0.05	0.7101
IRS2	ECHDC3	0.20	0.1160
IRS2	iCAM	-0.11	0.3974
IRS2	IL18R1	0.65	0.0000
IRS2	KCNE1	0.32	0.0128
IRS2	MMP9	0.45	0.0002
IRS2	P. MMP9	0.17	0.1938
IRS2	MYL4	-0.09	0.4828
IRS2	vCAM	0.03	0.7948
KCNE1	Ac. Mirístico	-0.24	0.0858
KCNE1	Ac. Araquidônico	0.10	0.4969
KCNE1	Ac. docosaexaenoico	0.09	0.5378
KCNE1	Ac. Eicosapentaenoico	-0.11	0.4292
KCNE1	Ac. Gamalinoleico	0.15	0.2883
KCNE1	Ac. lignocérico	0.04	0.8049
KCNE1	Ac. Linoleico	-0.10	0.4821
KCNE1	Ac. Linolênico	-0.04	0.7738

Tabela A.7: Tabela das estimativas da correlação de Pearson

Variáveis		Correlação	P-valor
KCNE1	Ac. miristOleico	-0.23	0.1071
KCNE1	Ac. Oleico	0.04	0.7789
KCNE1	ALOX15	0.18	0.1832
KCNE1	E-Selectina	-0.08	0.5498
KCNE1	ECHDC3	-0.08	0.5324
KCNE1	iCAM	-0.12	0.3900
KCNE1	IL18R1	0.28	0.0411
KCNE1	MMP9	0.17	0.2107
KCNE1	P. MMP9	0.02	0.8838
KCNE1	MYL4	-0.01	0.9124
KCNE1	vCAM	-0.19	0.1638
MMP9	Ac. Lignocérico	-0.26	0.0616
MMP9	Ac. Linolênico	0.21	0.0816
MMP9	Ac. Araquidônico	-0.02	0.8862
MMP9	Ac. docosaexaenoico	-0.07	0.5839
MMP9	Ac. Eicosapentaenoico	-0.18	0.1815
MMP9	Ac. Gamalinoleico	-0.02	0.8644
MMP9	Ac. Linoleico	0.01	0.9420
MMP9	Ac. Linolênico	-0.07	0.6223
MMP9	Ac. miristico	-0.06	0.6446
MMP9	Ac. miristOleico	0.04	0.7756
MMP9	Ac. Oleico	-0.03	0.8049
MMP9	ALOX15	0.04	0.7304
MMP9	E-Selectina	0.06	0.6435
MMP9	ECHDC3	0.10	0.4315
MMP9	iCAM	0.01	0.9230
MMP9	IL18R1	0.57	0.0000
MMP9	MMP9	0.31	0.0123
MMP9	MYL4	-0.17	0.1711
MMP9	vCAM	-0.04	0.7658
P. MMP9	Ac. Araquidônico	0.10	0.4229
P. MMP9	Ac. Docosaexaenoico	0.11	0.3594
P. MMP9	Ac. Eicosapentaenoico	0.02	0.9080
P. MMP9	Ac. Gamalinoleico	0.16	0.2061
P. MMP9	Ac. Lignoceric	-0.01	0.9353
P. MMP9	Ac. Linoleico	0.14	0.2391
P. MMP9	Ac. Miristico	0.06	0.6442
P. MMP9	Ac. Miristoleico	0.11	0.3842
P. MMP9	Ac. Oleico	0.05	0.6901
MYL4	Ac. Linolênico	-0.25	0.0547
MYL4	Ac. Miristoleico	-0.24	0.0700
MYL4	Ac. Araquidônico	-0.08	0.5437
MYL4	Ac. docosaexaenoico	0.07	0.6208
MYL4	Ac. Eicosapentaenoico	-0.13	0.3509
MYL4	Ac. Gamalinoleico	-0.13	0.3263
MYL4	Ac. lignoceric	0.06	0.6757
MYL4	Ac. Linoleico	-0.18	0.1638
MYL4	Ac. Miristico	-0.14	0.2827
MYL4	Ac. Oleico	-0.15	0.2634
MYL4	ALOX15	-0.05	0.7137
MYL4	E-Selectina	-0.15	0.2426
MYL4	iCAM	-0.14	0.2823
MYL4	IL18R1	0.10	0.4655
MYL4	P. MMP9	-0.15	0.2447
MYL4	vCAM	-0.14	0.2934

Tabela A.8: Tabela das estimativas da correlação de Pearson

Variáveis		Correlação	P-valor
TREML4	Ac. Lignocérico	-0.24	0.0801
TREML4	Ac. Araquidônico	-0.13	0.3016
TREML4	Ac. Docosaexaenoico	-0.13	0.3156
TREML4	Ac. Eicosapentaenoico	-0.03	0.8237
TREML4	Ac. Gamalinoleico	0.03	0.8192
TREML4	Ac. Linoleico	-0.09	0.4758
TREML4	Ac. Linolênico	-0.21	0.1025
TREML4	Ac. Mirístico	-0.10	0.4520
TREML4	Ac. miristoleico	-0.03	0.8161
TREML4	Ac. Oleico	-0.18	0.1536
TREML4	ALOX15	0.05	0.6726
TREML4	BCLA1	-0.11	0.3802
TREML4	BCLL1	0.01	0.9648
TREML4	CA1	0.27	0.0286
TREML4	COX7B	0.22	0.0781
TREML4	E-Selectina	0.02	0.8867
TREML4	ECHDC3	-0.05	0.7143
TREML4	iCAM	0.10	0.4246
TREML4	IL18R1	0.36	0.0077
TREML4	IRS	0.07	0.5633
TREML4	KCNE1	0.02	0.9024
TREML4	MMP9	0.08	0.5231
TREML4	P. MMP9	-0.10	0.4366
TREML4	MYL4	0.06	0.6288
TREML4	vCAM	0.13	0.3008
vCAM	Ac. Gamalinoleico	0.24	0.0452
vCAM	iCAM	0.26	0.0170
vCAM	Ac. Araquidônico	0.06	0.6366
vCAM	Ac. Docosaexaenoico	0.15	0.2341
vCAM	Ac. Eicosapentaenoico	-0.08	0.5498
vCAM	Ac. Lignocérico	-0.01	0.9268
vCAM	Ac. Linoleico	0.15	0.2169
vCAM	Ac. Linolênico	-0.02	0.8928
vCAM	Ac. mirístico	0.09	0.4592
vCAM	Ac. miristoleico	0.09	0.4800
vCAM	Ac. Oleico	0.08	0.5209
vCAM	P.MMP9	0.04	0.7402