

UNIVERSIDADE FEDERAL DE PERNAMBUCO – UFPE  
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA – CCEN  
DEPARTAMENTO DE QUÍMICA FUNDAMENTAL- DQF

TÁSSIA BRENA BARROSO CARNEIRO DA COSTA

METABONÔMICA APLICADA AO DIAGNÓSTICO E ESTADIAMENTO DE  
DOENÇAS HEPÁTICAS

RECIFE

2016

TÁSSIA BRENA BARROSO CARNEIRO DA COSTA

METABONÔMICA APLICADA AO DIAGNÓSTICO E ESTADIAMENTO DE  
DOENÇAS HEPÁTICAS

Dissertação apresentada ao Programa de Pós-Graduação em Química do Departamento de Química Fundamental da Universidade Federal de Pernambuco, como parte das exigências para a obtenção do título de mestre.

Orientador: Prof. Dr. Ricardo Oliveira da Silva

Co-orientador: Prof. Dr. Edmundo Pessoa Lopes

RECIFE

2016

Catálogo na fonte  
Bibliotecária Joana D'Arc Leão Salvador CRB 4-572

C837m Costa, Tássia Brena Barroso Carneiro da.  
Metabonômica aplicada ao diagnóstico e estadiamento de doenças hepáticas / Tássia Brena Barroso Carneiro da Costa. – 2016.  
92 f.: fig., tab.

Orientador: Ricardo Oliveira da Silva.  
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN. Química Fundamental, Recife, 2016.  
Inclui referências e anexos.

1. Ressonância magnética nuclear de Próton. 2. Hepatite B e C. 3. Fibrose hepática. 4. Quimiometria. I. Silva, Ricardo Oliveira da (Orientador). II. Título.

538.362 CDD (22. ed.) UFPE-FQ 2017-03

# **TÁSSIA BRENA BARROSO CARNEIRO DA COSTA**

## **"METABONÔMICA APLICADA AO DIAGNÓSTICO E ESTADIAMENTO DE DOENÇAS HEPÁTICAS"**

Dissertação apresentada ao  
Programa de Pós-Graduação no  
Departamento de Química  
Fundamental da Universidade  
Federal de Pernambuco, como  
requisito parcial para a obtenção do  
título de Mestre em Química.

Aprovado em: 29/03/2016.

### **BANCA EXAMINADORA**

---

**Prof. Ricardo Oliveira da Silva**

Universidade Federal de Pernambuco  
Departamento de Química Fundamental

---

**Prof. Fernando Hallwass**

Universidade Federal de Pernambuco  
Departamento de Química Fundamental

---

**Prof<sup>a</sup>. Maria Fernanda Pimentel Avelar**

Universidade Federal de Pernambuco  
Departamento de Engenharia Química

---

**Prof<sup>a</sup>. Ljubica Tasic**

Universidade Estadual de Campinas  
Instituto de Química

*Dedico a Deus e à minha família, responsáveis pela pessoa que sou hoje e por tudo que consegui conquistar. Meus pais que forneceram todo o ambiente e amor para que eu pudesse me desenvolver como pessoa e como profissional, e meus avós, sinônimo de amor e carinho, que tiveram a serenidade de compreender minha ausência em alguns momentos dessa caminhada.*

*Ao Lucas por todo o amor, carinho, paciência e dedicação que me ofereceu durante essa jornada.*

*Aos meus amigos, por todo o apoio, essencial para que eu conseguisse vencer cada etapa.*

## AGRADECIMENTOS

Ao Prof. Dr. Ricardo Oliveira que me acompanha desde a iniciação científica, me orientando, apoiando e incentivando meu crescimento profissional e pessoal.

Ao Prof. Dr. Edmundo Pessoa Lopes pela orientação e discussões ao longo do desenvolvimento do trabalho.

A todos os colegas de grupo, das equipes de química, de medicina tropical e de urologia, pela parceria no trabalho, no dia-a-dia, e por poder contar com vocês para desenvolver o trabalho. Em especial, à Joelma Cavalcanti, pelas indicações de livros e revisão do texto sobre fígado.

Às professoras Maria Fernanda Pimentel e Claudete Fernandes pelas aulas, cursos e discussões que foram essenciais para a minha aprendizagem de quimiometria. E por terem aberto as portas do LAC para que eu pudesse participar das reuniões e cursos, que me proporcionaram contribuições muito importantes para a formação e ótimas companhias, que contribuíram não só profissionalmente.

Aos meus amigos do dQF pela companhia nos momentos de estresse, contribuições em discussões pelos corredores e pelos momentos de descontração. Em especial, à Carolina Santos, amiga para todas as horas e consultora de quimiometria, e ao Eivson Darlivam, companheiro de sofrimento nas disciplinas, de estudos intermináveis e de conversas descontraídas.

Aos amigos *alops*, por terem feito essa fase passar de forma mais doce, em especial, Juliana Galvão e Igor Brito, pelas revisões e sugestões nos textos.

Ao Laboratório da Central Analítica do Departamento de Química Fundamental; e a toda a sua equipe, por proporcionar a efetivação da pesquisa.

Ao Hospital das Clínicas da Universidade Federal de Pernambuco pela disponibilização de amostras.

Ao CNPq pela bolsa de mestrado. à FACEPE e à CAPES pelo apoio financeiro ao projeto.

## RESUMO

A metabonômica pode ser definida como um conjunto de ferramentas, analíticas e de estatística multivariada, utilizadas para identificar mudanças de concentração dos metabólitos em um dado biofluido, associando-as à perturbação sofrida pelo organismo. Sendo assim, ela seria capaz de identificar qualquer doença no organismo, desde que seja empregado o biofluido adequado e as informações sejam corretamente extraídas. Para isso, a ferramenta mais empregada é a Espectroscopia de Ressonância Magnética Nuclear de Hidrogênio-1 (RMN de  $^1\text{H}$ ), e é necessário o uso de técnicas quimiométricas para extrair as informações do espectro. Neste trabalho, foram construídos modelos metabonômicos para: (1) identificar pacientes portadores de esteatose, e dos vírus da hepatite B (HBV) e da hepatite C (HCV), utilizando amostras de urina; e (2) classificar o grau de fibrose hepática em pacientes com hepatites crônicas por HBV ou HCV, utilizando amostras de soro sanguíneo. O modelo para classificação de pacientes com esteatose, obteve 100% de sensibilidade e de valor preditivo positivo. Para identificar esteatose independentemente de ser um portador de HBV ou HCV, o modelo construído obteve 97,9% de exatidão. Para classificar portadores de HBV e HCV, os modelos apresentaram sensibilidade de 100% e 92,6%, respectivamente. O modelo construído para diferenciar pacientes com diferentes lesões no fígado: esteatose e hepatites virais B ou C, obteve 94% de exatidão. Para classificar pacientes com fibrose significativa; fibrose avançada; e cirrose, alcançamos 98,4; 100; e 96,8% de exatidão, respectivamente. Através da combinação dos resultados dos modelos de fibrose significativa e fibrose avançada, foi possível determinar os pacientes com grau F2, no METAVIR, com percentual de acerto de 96,8%. Nas análises de fibrose, a exatidão observada para os modelos metabonômicos foram superiores aos observados para os métodos não-invasivos normalmente utilizados na prática clínica, APRI (do inglês, *Aspartate aminotransferase Platelet Ratio Index*) e FIB-4. A estratégia metabonômica demonstrou capacidade de avaliar a presença de diferentes doenças hepáticas em uma única análise, não invasiva, e determinar o grau de fibrose hepática, de forma minimamente invasiva.

**Palavras-chave:** RMN de  $^1\text{H}$ . HBV. HCV. Esteatose. Fibrose hepática.

## ABSTRACT

The metabonomics can be defined as a set of analytics and multivariate statistics tools, used to identify the metabolite concentration changes in a certain biofluid, associating them to the disturbance suffered. Therefore, it would be able to identify any disease in the body, if employed the appropriate biofluid and correctly extract the information. The most commonly used tool is Nuclear Magnetic Resonance Spectroscopy for hydrogen-1 ( $^1\text{H}$  NMR), and chemometrics techniques are used to extract the information of the spectrum. In this work we built metabonomics models to: (1) identify patients with steatosis, hepatitis B (HBV) and hepatitis C (HCV), using urine samples; and (2) classify the degree of liver fibrosis in patients with chronic hepatitis, HBV or HCV, using blood serum samples. The classification model for patients with steatosis obtained 100% to sensitivity and positive predictive value. To identify steatosis, without regard the presence of HBV or HCV, the constructed model achieved 97.9% accuracy. To classify carriers of HBV and HCV, the models showed 100 and 92.6% of sensitivity, respectively. The constructed model to differentiate patients with different liver damage: steatosis and viral hepatitis B or C, achieved 94% accuracy. To classify patients with significant fibrosis; advanced fibrosis; and cirrhosis, the models reached 98.4; 100; and 96.8% accuracy, respectively. By combining the results of significant fibrosis models and advanced fibrosis, it determined the patients with F2 in the METAVIR, with 96.8% of accuracy. In fibrosis analysis, the accuracy observed for metabonomics models were higher than those observed for the non-invasive methods commonly used in clinical practice, APRI (*Aspartate aminotransferase Platelet Ratio Index*) and FIB-4. The metabonomics strategy demonstrated ability to assess the presence of different liver diseases in a single non-invasive analysis and determine the degree of liver fibrosis, in a minimally invasive way.

**Keywords:**  $^1\text{H}$  NMR. HBV. HCV. Steatosis. Liver fibrosis.



## LISTA DE FIGURAS

Figura 1: Esquema de objetivos deste trabalho.	19
Figura 2: Ilustração da precessão do spin.	23
Figura 3: Polarização de spin a favor e contra o campo, respectivamente.	24
Figura 4: Resposta do vetor magnetização ao pulso $(\pi/2)_x$ .	26
Figura 5: Populações após aplicação do pulso $(\pi/2)_x$	27
Figura 6: Oscilação do vetor magnetização transversal	27
Figura 7: Decaimento do vetor magnetização transversal.	27
Figura 8: Sequência de pulsos CPMG.	30
Figura 9: Magnetização durante o processo de refocagem.	30
Figura 10: Matriz de dados. Elemento $x_{pn}$ representa o valor da variável $n$ para a amostra $p$ .	33
Figura 11: Rotação das variáveis originais para obtenção das componentes principais.	35
Figura 12: Exemplo de matrizes $\mathbf{X}$ e $\mathbf{Y}$ .	37
Figura 13: Esquema da construção da função linear discriminante	39
Figura 14: Tabela de contingência e parâmetros para avaliação de um teste.	41
Figura 15: Vista frontal do fígado.	42
Figura 16: Espectro de RMN de $^1\text{H}$ ( $\text{Na}_2\text{HPO}_4 / \text{NaH}_2\text{PO}_4 - 0,2 \text{ mol L}^{-1}$ , 300 MHz) de uma amostra de urina.	54
Figura 17: Gráfico de variáveis importantes para a construção do modelo PLS-DA.	57
Figura 18: Subtração do espectro médio controle do espectro médio esteatose	58
Figura 19: Estruturas dos metabólitos identificados e deslocamentos químicos no espectro de RMN de $^1\text{H}$ .	59
Figura 20: Variáveis importantes para o modelo LDA construído a partir das componentes principais	62
Figura 21: Escores do LDA em função do sexo dos indivíduos, no qual 0 indica feminino e 1 masculino.	62
Figura 22: Escores do LDA em função da idade dos indivíduos.	63
Figura 23: Sinais do hipurato no espectro de RMN de $^1\text{H}$ de amostra de urina.	64
Figura 24: Ampliação da região entre 6,8 e 9ppm do espectro COSY de uma amostra de urina.	65
Figura 25: Resumo dos resultados de diagnóstico de hepatites virais e esteatose.	67
Figura 26: Espectro de RMN de $^1\text{H}$ ( $\text{D}_2\text{O}$ , 300 MHz) de uma amostra de soro.	72

Figura 27: Espectro de RMN de $^1\text{H}$ ( $\text{D}_2\text{O}$ , 300 MHz) da amostra de soro utilizando a sequência de pulsos PRESAT.	73
Figura 28: Espectro de RMN de $^1\text{H}$ ( $\text{D}_2\text{O}$ , 300 MHz) da amostra de soro após aplicação da sequência de pulsos CPMG	73
Figura 29: Variáveis importantes para o modelo de fibrose avançada.	77
Figura 30: Curva ROC do APRI para predizer cirrose.	78
Figura 31: Esquema da determinação de amostras F2.	80
Figura 32: Comparação entre modelos metabonômicos e FIB-4 e APRI	81

## LISTA DE TABELAS

Tabela 1 - Propriedades espectroscópicas de alguns núcleos.(PAVIA; LAMPMAN; KRIZ, 2001) .....	24
Tabela 2: Marcadores sorológicos de HBV.....	47
Tabela 3: Número de infectados por HCV em diferentes regiões .....	47
Tabela 4: Resultados de modelo LDA para predição de HBV utilizando as componentes principais .....	55
Tabela 5: Resultados de modelo LDA para predição de HBV, independentemente da presença de esteatose.....	56
Tabela 6: Resultados de modelo LDA para predição de HCV, independentemente da presença de esteatose.....	56
Tabela 7: Resultados de modelo LDA para predição de esteatose usando deslocamentos da subtração de espectros médios.....	58
Tabela 8: Identificação dos principais metabólitos discriminantes, bins correspondentes e níveis urinários relativos em cada grupo. ....	59
Tabela 9: Resultados de modelo LDA para predição de esteatose usando componentes principais. ....	60
Tabela 10: Resultados de modelo LDA para predição de esteatose usando componentes principais. ....	64
Tabela 11: Resultados de modelo LDA para discriminação entre esteatose e hepatites virais .....	66
Tabela 12: Conversão entre Ishak e METAVIR .....	74
Tabela 13: Resultados para predição de fibrose significativa .....	75
Tabela 14: Resultados para predição de fibrose avançada .....	76
Tabela 15: Resultados para predição de cirrose .....	78
Tabela 16: Resultados para predição de F2 por modelos metabonômicos.....	80

## LISTA DE ABREVEATURAS

ALT	Alanina aminotransferase
APRI	APRI, do inglês, <i>Aspartate aminotransferase Platelet Ratio Index</i>
AST	Aspartato aminotransferase
AUROC	Área sob a curva ROC
CHC	Carcinoma Hepatocelular
CPMG	Sequência de pulsos Carr-Pucell-Meiboom-Gill
DHA	Doença Hepática Alcoólica
DHGNA	Doença Hepática Gordurosa Não Alcoólica
DQF	Departamento de Química Fundamental
EHNA	Esteato-hepatite Não Alcoólica
ELISA	Ensaio imunoenzimático, do inglês, <i>Enzyme-Linked Immunosorbent Assay</i>
FID	Decaimento livre de indução, do inglês, <i>Free- Induction Decay</i>
HBV	Vírus da hepatite B
HCV	Vírus da hepatite C
INCA	Instituto Nacional de Câncer José Alencar Gomes da Silva
LDA	Análise Discriminante Linear, do inglês, <i>Linear Discriminant Analysis</i>
LOOCV	Validação Cruzada Total, do inglês, <i>Leave-One-Out Cross Validation</i>
LV	Variáveis Latentes, do inglês, <i>Latent Variable</i>
MS	Espectrometria de Massas, do inglês, <i>Mass Spectroscopy</i>
OMS	Organização Mundial da Saúde
PC	Componente Principal, do inglês, <i>Principal Component</i>
PCA	Análise de Componentes Principais, do inglês, <i>Principal Component Analysis</i>

PCR	Reação em cadeia da polimerase, do inglês, <i>Polymerase Chain Reaction</i>
PD	Doença de Parkinson, do inglês, <i>Parkinson's disease</i>
PLS-DA	Análise Discriminante por Mínimos Quadrados Parciais, do inglês, <i>Partial Least Squares Discriminant Analysis</i>
PRESAT	Sequência de pulsos para saturação do sinal da água
RMN	Ressonância Magnética Nuclear
RMN de $^1\text{H}$	Espectroscopia de Ressonância Magnética Nuclear de Hidrogênio-1
ROC	Do inglês, <i>Receiver Operating Characteristics</i>
RP	Reconhecimento de Padrões
SNV	Do inglês, <i>Standard Normal Variate</i>
$T_2$	Tempo de relaxação transversal
TCLE	Termo de consentimento livre e esclarecido
TMS	Tetrametilsilano
US	Ultrassonografia
VPN	Valor preditivo negativo
VPP	Valor preditivo positivo

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>15</b>
<b>2</b>	<b>OBJETIVOS GERAIS.....</b>	<b>19</b>
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>20</b>
3.1	METABONÔMICA .....	20
3.2	RESSONÂNCIA MAGNÉTICA NUCLEAR .....	22
3.2.1	Princípios Básicos da RMN .....	22
3.2.2	Deslocamento Químico .....	28
3.2.3	Escala e Referências .....	29
3.2.4	PRESAT e CPMG .....	29
3.2.5	Sinal de RMN .....	31
3.2.6	Transformação do Espectro em Matriz de Dados .....	31
3.3	QUIMIOMETRIA .....	32
3.3.1	Análise de Componentes Principais .....	35
3.3.2	Análise Discriminante por Mínimos Quadrados Parciais .....	36
3.3.3	Análise Discriminante Linear .....	38
3.4	FÍGADO.....	42
3.4.1	Esteatose .....	44
3.4.2	Hepatites Virais B e C .....	45
3.4.3	Fibrose Hepática .....	49
<b>4</b>	<b>DIAGNÓSTICO DE HEPATITES VIRAIS, B E C, E ESTEATOSE.....</b>	<b>51</b>
4.1	OBJETIVOS ESPECÍFICOS .....	51
4.2	METODOLOGIA .....	51
4.2.1	Amostras .....	51
4.2.2	Aquisição dos espectros de RMN de $^1\text{H}$ .....	52
4.2.3	Processamento dos Espectros de RMN $^1\text{H}$ .....	52
4.2.4	Análise Estatística Multivariada .....	53
4.3	RESULTADOS E DISCUSSÃO .....	53
4.3.1	HBV versus Controle - puros .....	54
4.3.2	HBV versus Controle – independentemente de esteatose .....	55
4.3.3	HCV versus Controle – independentemente de esteatose .....	56
4.3.4	Esteatose versus Controle - puros .....	57
4.3.5	Esteatose versus Controle – independentemente de vírus .....	63
4.3.6	Esteatose versus Hepatites Virais .....	65
4.4	CONCLUSÃO .....	67
4.5	PERSPECTIVAS .....	68
<b>5</b>	<b>ESTADIAMENTO DE FIBROSE HEPÁTICA EM PACIENTES COM HCV .....</b>	<b>69</b>
5.1	OBJETIVOS ESPECÍFICOS .....	69
5.2	METODOLOGIA .....	69

5.2.1	Amostras .....	69
5.2.2	Aquisição dos espectros de RMN de $^1\text{H}$ .....	70
5.2.3	Processamento dos Espectros de RMN $^1\text{H}$ .....	70
5.2.4	Análise Estatística Multivariada .....	71
5.2.5	APRI e FIB-4 .....	71
5.3	RESULTADOS E DISCUSSÃO .....	72
5.3.1	Obtenção de Espectros e Processamento dos Dados .....	72
5.3.2	Fibrose Significativa .....	74
5.3.3	Fibrose Avançada .....	75
5.3.4	Cirrose .....	77
5.3.5	Pacientes com Classificação F2 no METAVIR .....	79
5.4	CONCLUSÃO .....	81
5.5	PERSPECTIVAS .....	82
<b>6</b>	<b>CONCLUSÕES GERAIS.....</b>	<b>83</b>
<b>7</b>	<b>PERSPECTIVAS GERAIS .....</b>	<b>84</b>
	<b>REFERÊNCIAS .....</b>	<b>85</b>
	<b>ANEXO A – APROVAÇÃO DO COMITÊ DE ÉTICA PARA PROJETO DE DIAGNÓSTICO DE DOENÇAS HEPÁTICAS.....</b>	<b>90</b>
	<b>ANEXO B – APROVAÇÃO DO COMITÊ DE ÉTICA PARA PROJETO DE AVALIAÇÃO DO GRAU DE FIBROSE HEPÁTICA.....</b>	<b>91</b>

## 1 INTRODUÇÃO

Em um seminário assistido durante o primeiro ano do mestrado, ouvi um professor falar sobre quão complicado e demorado é o nosso sistema de saúde. Ele comentou sobre várias idas e vindas, muita espera e pouca resolução dos problemas.

Se refletirmos um pouco, vamos perceber que, em geral, essa é a realidade. Quando alguém está com dor ou percebe algo de incomum acontecendo no seu corpo, vai procurar o serviço de saúde. Espera em uma longa fila para pegar uma ficha de atendimento. Muitas vezes, só consegue a ficha depois de dias de tentativas. Após o atendimento, geralmente em um ambiente precário e com tempo apertado, recebe a solicitação de exames. Para marcar os exames, mais filas e luta por fichas, até conseguir uma data meses depois. Com os exames em mãos, o paciente tem que enfrentar a mesma dificuldade anterior para conseguir voltar ao médico.

Após toda essa jornada, já se passaram meses, e muitas vezes ela só começou, pois é necessária a consulta com um médico de outra especialidade, novos exames, mais filas, mais fichas e bastante tempo.

É muita espera e burocracia para quem sente dor e/ou está sob o risco iminente de complicações sérias de saúde e, em casos extremos, de morte. Frequentemente, pessoas morrem em consequência de uma doença que, se diagnosticada previamente, teria sido tratada e curada. A alta taxa de mortalidade por câncer, por exemplo, é associada ao diagnóstico tardio.(COSTA; MIGLIORATI, 2001; TRUFELLI et al., 2008)

O Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA) atribui o fato de o diagnóstico ser feito em estágios avançados da doença à dificuldade de acesso da população aos serviços e programas de saúde; e à capacidade do Sistema Único de Saúde (SUS) de absorver a demanda que chega às unidades.(PANOBIANCO et al., 2012)

Segundo a Organização Mundial da Saúde (OMS), um atendimento primário eficiente é essencial para a melhoria da saúde da população, uma vez que a maioria dos casos podem ser resolvidos no atendimento de nível primário, mas - quando isso não acontece - evoluem para casos mais complicados. (WORLD; REPORT, 2003) Nos países mais desenvolvidos, o atendimento primário é entendido como o primeiro nível de atenção, porém, em países onde desafios significativos no acesso aos cuidados de



saúde persistem, como no Brasil, é visto mais como uma estratégia de todo o sistema. (WORLD; REPORT, 2003)

Agora, podemos imaginar um sistema em que o atendimento primário seja feito da seguinte forma: ao sentir algum tipo de sintoma, ou simplesmente como medida preventiva, um indivíduo se encaminha a um hospital ou clínica, onde é atendido por um profissional de saúde que, além da anamnese, faz a coleta de amostras de sangue e urina. Durante poucas horas, são realizados testes *screening* de inúmeras doenças e, rapidamente - se for o caso - o paciente já seria encaminhado para o médico especialista e seriam solicitados os exames para a confirmação do diagnóstico.

Esse seria um caminho muito mais curto, rápido e descomplicado. Uma ótima opção para desafogar o sistema de saúde e, ao mesmo tempo, dar o atendimento necessário à população. Mas como é possível uma análise rápida e capaz de identificar diferentes tipos de doenças?

Parece utopia ou ficção científica, mas é uma possibilidade, a partir do estudo e desenvolvimento da metabonômica.

Godoy e colaboradores, em 2010, demonstraram que a metabonômica pode ser uma ferramenta útil para diagnóstico de hepatites, obtendo 94% de sensibilidade e 97% de especificidade em um modelo metabonômico, baseado em RMN de  $^1\text{H}$ , para discriminação entre amostras de urina de pacientes com Hepatite C e voluntários saudáveis. (GODOY et al., 2010) Há estudos que utilizam metabonômica para diferenciar amostras de pacientes com cirrose hepática daqueles com CHC, com sensibilidade de 89,5% e especificidade de 88,9% em população nigeriana infectada pelo HBV, (SHARIFF et al., 2010) e 81% de sensibilidade e 71% de especificidade, em egípcios infectados pelo HCV. (SHARIFF et al., 2011) Em outro estudo, pacientes com cirrose foram diferenciados de pacientes com CHC em uma população chinesa. (GAO et al., 2009) Demonstrando o potencial da estratégia metabonômica na identificação de cirrose.

Em 2014, Kimhofer e colaboradores, apresentaram um artigo de revisão que incluiu 22 estudos metabonômicos feitos em plasma, soro ou urina de pacientes com CHC, publicados nos últimos anos. Esse número mostra o esforço empregado para identificar metabólitos específicos detectáveis em fluidos biológicos, tais como sangue ou urina, o que pode servir como biomarcadores para o diagnóstico de CHC nos estágios iniciais. (KIMHOFER et al., 2015)

A revisão desses estudos mostrou que ocorrem mudanças notáveis em vários metabólitos e nos níveis de proteína que se correlacionam com a presença ou progressão doença. Foi destacada a necessidade de padronização nos métodos de construção de modelo, validação e definição dos grupo, assim como de novos estudos para a identificação entre os graus de fibrose, presença de cirrose e progressão do câncer. (KIMHOFER et al., 2015)

Como o sucesso no tratamento do câncer está intrinsicamente relacionado à sua prévia detecção, há trabalhos que empregam a estratégia metabonômica baseada em RMN de  $^1\text{H}$  para identificação de outros tipos de câncer. Através da análise discriminante por mínimos quadrados parciais (PLS-DA) foi identificada a presença dos tumores de Ehrlich e Sarcoma 180 em camundongos, após cinco e dez dias da indução dos tumores, respectivamente. Lipídeos, lipoproteínas e lactato foram os biomarcadores mais importantes. (CAROLINE et al., 2014)

Utilizando PLS-DA e *particle swarm optimization* (PSO), 99% e 85% dos indivíduos foram corretamente classificados nos conjuntos de treinamento e de teste, respectivamente, quanto à presença de câncer de pulmão. Os metabólitos discriminantes identificados foram: lactato, prolina, glicoproteína, glutamato, alanina, treonina, taurina, glicose ( $\alpha$ - e  $\beta$ -), trimetilamina, glutamina, glicina, e mio-inositol. (LI et al., 2014)

A metabonômica também foi utilizada para identificar complicações causadas pela radioterapia no tratamento de pacientes com câncer de colo do útero. Os pacientes com os sintomas intestinais agudos induzidos por radiação (RIAISs) foram identificados aplicando análises quimiométricas aos espectros de extratos de amostras fecais, sendo caracterizados pelas concentrações aumentadas de a-cetobutirato, valina, uracilo, tirosina, N-óxido de trimetilamina, fenilalanina, lisina, isoleucina, glutamina, creatinina, creatina, ácidos biliares, aminohipurato, e alanine; e reduzidas de a-glucose, N-butirato, metilamina, e etanol. (CHAI et al., 2015)

Nos últimos anos, distúrbios neurológicos, fármacos e tratamentos foram objetos de vários estudos metabonômicos. Öhman e Forsgren empregaram análise discriminante ortogonal por mínimos quadrados parciais (OPLS-DA) a espectros de RMN de  $^1\text{H}$  de fluido cerebrospinal de 10 pacientes portadores da doença de Parkinson (PD) e 10 indivíduos controle. Com os grupos satisfatoriamente separados, foram identificados possíveis marcadores para a doença: alanina, creatinina, dimetilamina, glicose, lactato, manose, fenilalanina, ácido 3-hidroxiisobutírico e ácido 3-hidroxiisovalérico, e observou-se uma diminuição generalizada na concentração dos metabólitos dos

pacientes com PD em relação aos indivíduos do grupo controle. (ÖHMAN; FORSGREN, 2015)

A estratégia metabonômica, utilizando cromatografia líquida de ultra-eficiência acoplada à espectrometria de massa (UPLC-MS) e PLS-DA, já havia sido utilizada para confirmar o efeito da administração de *Acanthopanax senticosus Harms* (EAS) em ratos com PD e identificar as diferenças metabólicas antes e depois do tratamento. Foi observado que as perturbações metabólicas causadas pela PD foram restauradas após o tratamento com EAS, identificando biomarcadores para a doença e propondo vias de regulação do metabolismo pelo efeito terapêutico da EAS. (LI et al., 2013)

Outro estudo utilizando UPLC-MS e PLS-DA foi realizado para identificar biomarcadores de esquizofrenia e correlacionar com o efeito terapêutico de olanzapina. Foram coletadas amostras de sangue de 15 pacientes com esquizofrenia, antes e após 4 semanas de tratamento, e de 15 voluntárias saudáveis. Treze metabólitos foram identificados como possíveis biomarcadores da doença, devido a diferença de concentração entre os grupos doentes antes do tratamento e o controle, e tenderam à normalização após o tratamento. (QIAO et al., 2016)

A confirmação do efeito e elucidação de mecanismo terapêutico antidepressivo do medicamento icariin em ratos foi realizada por Gong et al, utilizando RMN de <sup>1</sup>H e OPLS-DA e amostras de soro, urina e extrato de tecido cerebral. Também foram identificados possíveis biomarcadores de depressão. (GONG et al., 2016)

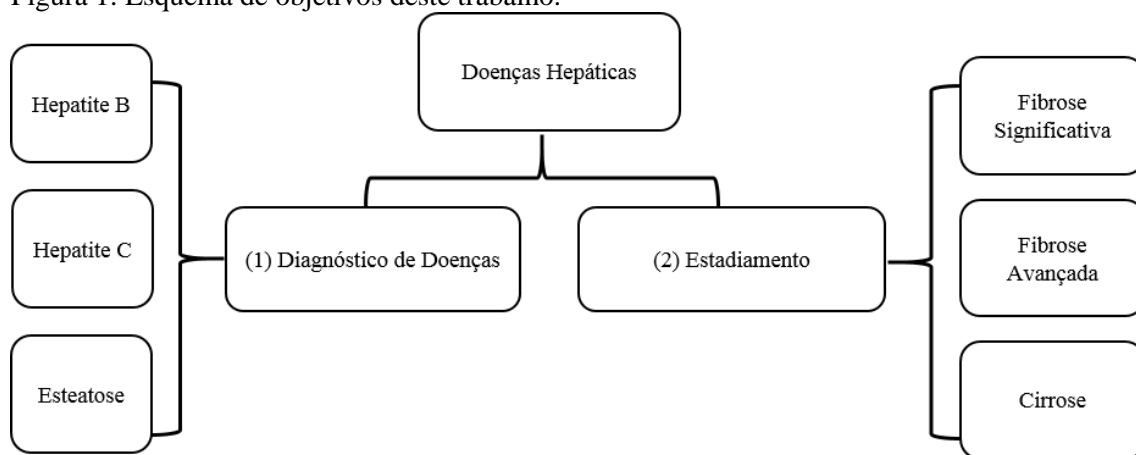
A estratégia metabonômica também pode ser empregada à indústria de alimentos, desde que as diferenças entre as amostras sejam decorrentes de perturbações sofridas pelo organismo vivo que deu origem ao alimento. Como exemplo podemos citar: identificação de grãos de soja submetidos a radiação gama (RIBEIRO et al., 2014); identificação de vinhos produzidos na região do Vale do Rio São Francisco (DA SILVA NETO et al., 2009); e classificação de azeites de diferentes regiões. (SACCHI; MANNINA, 1998)

## 2 OBJETIVOS GERAIS

Construir modelos metabonômicos capazes de: diferenciar amostras de voluntários saudáveis e de pacientes com doenças hepáticas; e classificar o grau de fibrose hepática de pacientes portadores de hepatite C crônica.

Na Figura 1, está ilustrado um esquema dos objetivos deste projeto.

Figura 1: Esquema de objetivos deste trabalho.



Fonte: Autoria própria.

### 3 FUNDAMENTAÇÃO TEÓRICA

#### 3.1 Metabonômica

A metabonômica pode ser definida como um conjunto de ferramentas, analíticas e de estatística multivariada, utilizadas para identificar mudanças de concentração dos metabólitos em um determinado biofluido e associar essas mudanças à perturbação sofrida pelo organismo.

Quando um organismo vivo sofre alguma perturbação, seja por doenças, uso de drogas ou toxinas, ocorre uma série de processos com o objetivo de restaurar o equilíbrio do sistema, causando mudanças nas concentrações de metabólitos endógenos. A metabonômica propõe que essas mudanças possam ser identificadas como uma impressão digital da perturbação.(LINDON; HOLMES; NICHOLSON, 2003; NICHOLSON; LINDON; HOLMES, 1999)

Sendo assim, ela seria capaz de identificar qualquer doença no organismo, desde que seja empregado o biofluido adequado e as informações sejam corretamente extraídas. Para isso, a estratégia metabonômica, em geral, emprega a Espectroscopia de Ressonância Magnética Nuclear de Hidrogênio-1 (RMN de  $^1\text{H}$ ), que é uma técnica rápida, exige mínima preparação de amostras e o espectro obtido reflete um perfil dos metabólitos presentes no biofluido. A duração da análise depende do biofluido a ser analisado, que pode ser, por exemplo: urina, soro sanguíneo ou soro de sêmen.

Como a resposta dessa análise é um espectro muito complexo, devido a quantidade de substâncias presentes nos biofluidos, para observar a diferença entre amostras de doentes e voluntários saudáveis é necessária a aplicação de técnicas quimiométricas à matriz de dados, visando identificar padrões e classificar as amostras.

É comum haver uma confusão entre os termos metabonômica e metabolômica, que são, inclusive, usados como sinônimos em alguns trabalhos, mas, por definição, são técnicas diferentes. Metabonômica (do grego, *meta* -mudança e *nomos* -regras ou leis) foi definida por Jeremy Nicholson, Elaine Holmes e John Lindon, em 1999, como modelos quimiométricos que têm a capacidade de classificar mudanças no metabolismo. (NICHOLSON; LINDON; HOLMES, 1999)

Por outro lado, metabolômica, é um termo com origem não muito bem definida, usado em publicações em 1985 e 1989 como parte da teoria do controle metabólico e análise de fluxo, e não era muito utilizado em publicações até o fim dos anos 90. Em

2011, Fiehn definiu como análise qualitativa e quantitativa de todos os metabólitos de um sistema.(ROBERTSON, 2005)

Então, enquanto a metabolômica visa identificar todos os metabólitos de um sistema, a metabonômica procura a impressão digital de uma perturbação em um sistema vivo e a partir dela, busca identificar a quais metabólitos endógenos está relacionada. Como elas têm objetivos diferentes, são utilizadas técnicas analíticas diferentes.

É comum atribuir essa escolha a fatores históricos: a metabolômica, empregada por cientistas que estudavam plantas, habituados a utilizar Espectrometria de Massas (MS); e metabonômica, empregada pelo grupo do *Imperial College of London*, que quase exclusivamente utilizava espectroscopia de Ressonância Magnética Nuclear (RMN) em seus trabalhos anteriores. Porém, na literatura há casos do uso inverso. Em teoria, qualquer técnica capaz de gerar medidas qualitativas de metabólitos, pode ser utilizada para metabonômica. Em geral, RMN e MS, precedida de uma técnica de separação de misturas, são mais comuns.(ROBERTSON, 2005)

Uma grande vantagem é que a RMN apresenta maior reprodutibilidade, tanto entre amostras de um mesmo trabalho quanto para ser reproduzido por outro grupo, além de apresentar menor tempo de análise. Já a MS apresenta uma grande vantagem em relação à identificação de metabólitos, tornando-a ideal para a análise metabolômica. Como para a análise metabonômica somente alguns metabólitos de interesse são identificados, isso pode ser feito através da espectroscopia de RMN, devido à sua versatilidade, através de experimentos bidimensionais, por exemplo.(ROBERTSON, 2005)

Então, por ser versátil, reprodutível e capaz de obter as informações necessárias para a estratégia metabonômica, a RMN é a técnica mais usada, e foi utilizada neste trabalho.

Para compreender um pouco melhor como funciona a estratégia metabonômica, nas próximas seções serão abordados os princípios de espectroscopia de ressonância magnética nuclear e de quimiometria. Também será apresentada uma explanação sobre fígado e suas principais doenças, objeto desse estudo.

## 3.2 Ressonância Magnética Nuclear

### 3.2.1 Princípios Básicos da RMN

A Ressonância Magnética Nuclear (RMN) é uma técnica espectroscópica, ou seja, se baseia na interação entre a radiação eletromagnética e a matéria.

Os núcleos dos átomos têm spins ( $I$ ) que são o resultado da combinação dos spins das suas partículas nucleares. Quando  $I$  é diferente de zero, o núcleo possui um momento magnético  $\mu$ , proporcional ao spin e à constante magnetogírica ( $\gamma$ ), característica de cada núcleo, dado pela Equação 1 (PAVIA; LAMPMAN; KRIZ, 2001):

Equação 1

$$\mu = \frac{h I \gamma}{2\pi},$$

Sendo  $h$  a constante de Planck.

Quando  $I = 1/2$ , existem dois estados possíveis, normalmente identificados por  $\alpha$  e  $\beta$ . A energia dos dois estados é a mesma, ou seja, são estados degenerados. Para se realizar espectroscopia é necessário quebrar a degenerescência dos estados, isso pode ser feito aplicando um campo magnético externo  $\mathbf{B}_0$ .

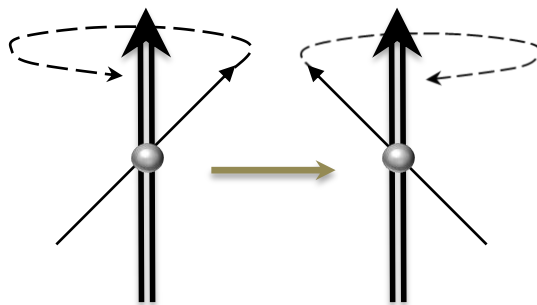
A direção do momento angular de spin é chamada de polarização do spin. Em geral, partículas com spin tem a polarização apontando para todas as possíveis direções. Assim, a partícula de spin  $1/2$  não está restrita aos estados  $\alpha$  e  $\beta$ , pode estar em uma combinação desses estados.

O momento magnético do núcleo aponta para a mesma direção da polarização do spin, para um núcleo com  $\gamma > 0$ , ou na direção oposta da polarização do spin, para um núcleo com  $\gamma < 0$ . Para uma amostra em equilíbrio, e sem ação de campos externos, a distribuição dos momentos magnéticos é isotrópica, todas as direções possíveis estão em proporções iguais.

Quando um campo magnético é aplicado à amostra, a polarização do spin se move em torno do eixo do campo. O momento magnético do spin se move formando um cone, mantendo constante o seu ângulo em relação ao eixo do campo aplicado. Esse movimento é chamado de precessão e está ilustrado na Figura 2.

A seção 2.2 foi baseada nos livros: (1) Spin Dynamics e (2) Fundamentos e Aplicações da Ressonância Magnética Nuclear: Aspectos Quânticos da Ressonância Magnética Nuclear (LEVITT, 2008; VILLAR, 2002).

Figura 2: Ilustração da precessão do spin.



Fonte: Autoria própria.

O estado que estiver na direção do campo terá energia mais baixa do que o estado que estiver alinhado em oposição ao campo, quebrando a degenerescência do sistema.

Esse comportamento, denominado efeito Zeeman, constitui a base do fenômeno de ressonância magnética. Devido à existência de níveis discretos de energia, separados por um quantum de energia igual a  $\hbar\nu$ , é possível observar transições entre tais níveis através da excitação dos spins nucleares por uma radiação com frequência  $\nu$ .

Em termos clássicos, a energia do momento magnético  $\mu$  em um campo magnético  $\mathbf{B}_0$ , é dada pela Equação 2:

Equação 2

$$E = -\vec{\mu} \cdot \vec{B}_0 = -\mu B_0 \cos \theta = -\mu_z B_0$$

Em que  $\theta$  é o ângulo entre  $\mu$  e  $\mathbf{B}_0$ . Substituindo  $\mu_z$ , definido pela mecânica quântica, temos a Equação 3:

Equação 3

$$E = -\frac{(\gamma \hbar B_0 m_l)}{2\pi}$$

E a diferença de energia entre dois estados na Equação 4:

Equação 4

$$\Delta E = -\frac{|\gamma \hbar B_0 \Delta m_l|}{2\pi}$$

Em que  $\Delta m_l$  é a mudança de spin envolvida na transição e  $|\Delta m_l| = 1$ .

Com  $\Delta E = h\nu$  e  $|\Delta m_l| = 1$ , temos a Equação 5:

Equação 5

$$\nu = B_0 \gamma / 2\pi$$



A chamada frequência de Larmor, frequência de ressonância dos diversos núcleos atômicos, em diferentes espécies químicas, é a mesma para núcleos com a mesma constante magnetogírica. Dessa forma, a utilidade da RMN seria extremamente limitada, pois para uma molécula heteropoliatômica obteríamos um espectro de RMN com um sinal para cada classe de elementos presentes.

A Tabela 1 apresenta o spin de alguns núcleos, constante magnetogírica e frequência de ressonância sob a ação de um campo magnético externo de 7,05 T, assim como sua abundância isotópica.

Tabela 1 - Propriedades espectroscópicas de alguns núcleos.(PAVIA; LAMPMAN; KRIZ, 2001)

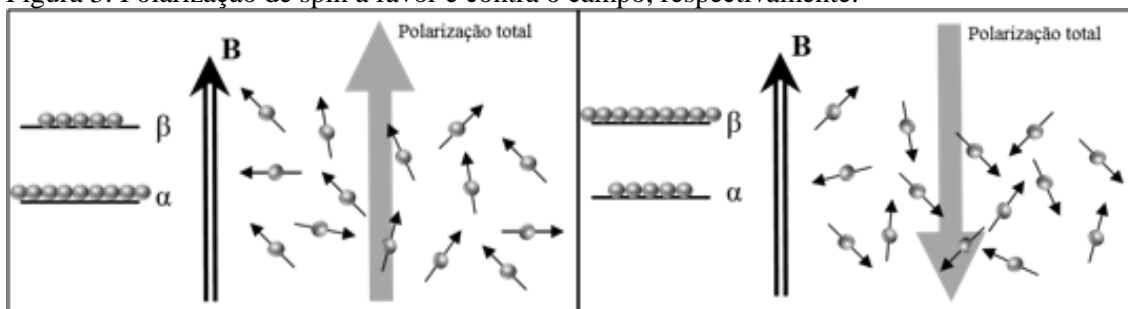
	<b>I</b>	<b><math>\gamma</math></b> ( $10^6 \text{ rad. T}^{-1} \text{ s}^{-1}$ )	<b><math>\nu</math></b> (MHz)	<b>Abundância</b> <b>Isotópica (%)</b>
$^1\text{H}$	1/2	267,53	300	99,98
$^2\text{H}$	1	41,1	46	0,0156
$^{13}\text{C}$	1/2	67,28	75	1,108 *
$^{19}\text{F}$	1/2	251,7	282	100
$^{31}\text{P}$	1/2	108,3	121	100

\* O isótopo mais abundante do carbono,  $^{12}\text{C}$ , tem spin zero. (PAVIA; LAMPMAN; KRIZ, 2001)

Consideramos a soma das populações dos estados como igual a um, então somente a diferença entre populações de dois estados tem significado físico.

A diferença entre as populações dos estados de spin indica a polarização longitudinal de spin, magnetização da amostra na direção do campo. O estado em que a população do estado  $\alpha$  é maior que a do estado  $\beta$ , indica a polarização de spins ao longo da direção do campo externo. Similarmente, o estado no qual a população do estado  $\alpha$  é menor que a do estado  $\beta$ , indica a polarização de spins oposta à direção do campo externo, ver Figura 3.

Figura 3: Polarização de spin a favor e contra o campo, respectivamente.



Fonte própria

Se as populações dos dois estados são iguais, então não existe polarização líquida na direção do campo aplicado. É importante lembrar que a população de um estado não indica a fração de spins que estão naquele estado, pois os spins não estão exatamente no estado  $\alpha$  ou  $\beta$ , estão em combinações desses dois estados de energia, apontando para direções arbitrárias.

Se um sistema de spins não sofre interferência por um longo intervalo de tempo, é esperado que atinja um estado de equilíbrio térmico com o ambiente. Segundo a mecânica estatística, a população dos estados ( $\rho$ ) em equilíbrio térmico segue a distribuição de Boltzmann. Então teremos as populações dos estados  $\alpha$  e  $\beta$ , respectivamente, na Equação 6:

Equação 6

$$\rho_{\alpha} = \frac{e^{\frac{-E_{\alpha}}{T k_B}}}{\sum_n e^{\frac{-E_n}{T k_B}}} \quad \rho_{\beta} = \frac{e^{\frac{-E_{\beta}}{T k_B}}}{\sum_n e^{\frac{-E_n}{T k_B}}}$$

A razão entre as populações está na Equação 7.

Equação 7

$$\frac{\rho_{\alpha}}{\rho_{\beta}} = \frac{e^{\frac{-E_{\alpha}}{T k_B}}}{e^{\frac{-E_{\beta}}{T k_B}}} = e^{\frac{(E_{\beta} - E_{\alpha})}{T k_B}} = e^{\frac{\Delta E}{T k_B}}$$

Em que T é a temperatura,  $E_i$  é a energia do estado i e  $k_B$  é a constante de Boltzmann.

A diferença entre as populações de dois estados em equilíbrio térmico depende da diferença de energia entre eles, comparada com a energia térmica disponibilizada pela temperatura da amostra.

Como a frequência associada à diferença de energia entre os estados de spin está na região de radiofrequência, então a diferença de energia,  $\Delta E$ , é cerca de quatro ordens de grandeza menor do que a energia térmica à temperatura ambiente. Consequentemente, a diferença de população entre os dois autoestados de Zeeman é muito pequena.

Isso significa que, a temperatura ambiente, a energia térmica da molécula é muito maior do que a diferença de energia entre os estados de spin nuclear, fazendo com que o estado de mais alta energia seja preenchido, mas não totalmente.

Para  $\gamma$  positivo, o estado de mais baixa energia,  $\alpha$  terá uma população ligeiramente maior que a do estado de alta energia,  $\beta$ . Para  $\gamma$  negativo a energia dos estados se invertem.

O estado de um único spin  $1/2$  é, comumente, representado por uma seta, indicando a direção do momento angular de spin. A resposta do spin ao campo magnético pode ser ilustrada pela rotação desta seta em torno de diferentes eixos no espaço tridimensional.

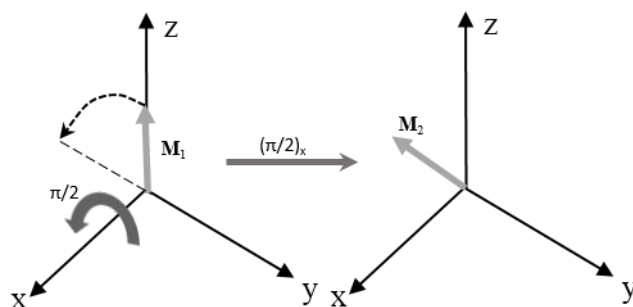
Similarmente, é possível representar um conjunto de spins  $1/2$  como um vetor magnetização  $\mathbf{M}$ , indicando a magnitude e direção da magnetização.

A magnetização longitudinal de spin nuclear é quase indetectável. A espectroscopia de ressonância magnética nuclear utiliza uma diferente abordagem: ao invés de medir a magnetização de spin nuclear ao longo do eixo do campo, mede a magnetização perpendicular a esse campo.

Suponha que o sistema de spins atingiu o equilíbrio térmico, após ser submetido a um campo magnético externo. A magnetização ao longo da direção do campo aplicado, magnetização longitudinal, é  $M_z$ . Na situação de equilíbrio, todos os spins quimicamente equivalentes precessam em torno do campo magnético com a mesma frequência,  $\nu$ . Não há magnetização perpendicular ao campo, pois, em equilíbrio térmico, a distribuição de magnetização é cilindricamente simétrica em torno do eixo  $z$ , portanto as componentes em  $x$  e  $y$  se cancelam.

Não há transição espontânea entre diferentes estados de spin. É necessário um estímulo externo para que isso ocorra. Se aplicarmos um pulso de radiofrequência  $\pi/2$  radianos em torno de  $x$ ,  $(\pi/2)_x$ , o vetor magnetização é rotacionado de  $z$  para  $-y$ , como na Figura 4.

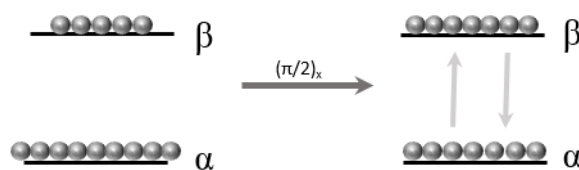
Figura 4: Resposta do vetor magnetização ao pulso  $(\pi/2)_x$ .



Fonte: Autoria própria.

Imediatamente após o pulso, observamos que o vetor magnetização não tem componente no eixo  $z$ , isso ocorre porque o pulso iguala as populações dos dois estados, ou seja, os spins individuais são rotacionados de forma que ao somarmos a componente de  $\alpha$  e de  $\beta$ , de todos, obtém-se o mesmo valor para a soma de cada estado, ilustrado na Figura 5.

Figura 5: Populações após aplicação do pulso  $(\pi/2)_x$

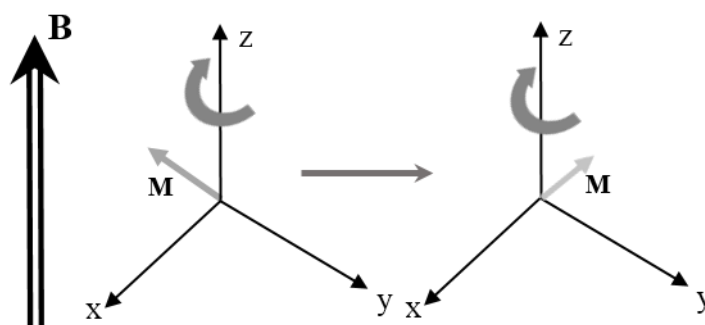


Fonte: Autoria própria.

Esse momento magnético perpendicular ao campo magnético é chamado magnetização transversal.

Quando o pulso é desligado, os spins retornam ao seu movimento de precessão. Microscopicamente, cada spin individual precessa no seu cone. Em escala macroscópica, o estado imediatamente após o pulso corresponde a uma polarização ao longo de  $-y$ , que precessa em torno de  $z$ , rotacionando no plano  $xy$ , com a frequência de Larmor, como na Figura 6.

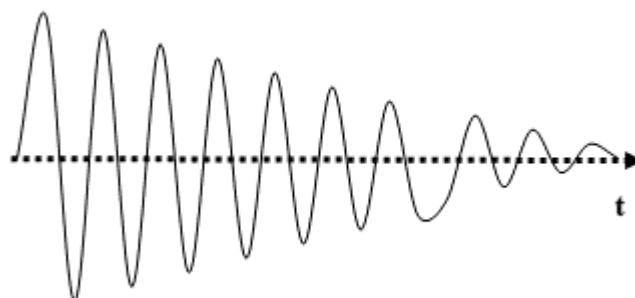
Figura 6: Oscilação do vetor magnetização transversal



Fonte: Autoria própria.

A magnetização em  $y$ ,  $M_y$ , e em  $x$ ,  $M_x$ , oscilam como uma função senoidal com a frequência de Larmor. Porém, a intensidade em  $x$  e  $y$  decai com o tempo, como pode ser observado na Figura 7.

Figura 7: Decaimento do vetor magnetização transversal.



Fonte: Autoria própria.

Isso acontece porque é impossível manter a sincronia entre os spins individuais e, aos poucos, a posição relativa entre eles vai mudando. Dessa forma, a soma vetorial entre os componentes em x e y vai diminuindo até alcançar o zero, enquanto a componente em z aumenta até voltar ao estado inicial.

### 3.2.2 Deslocamento Químico

A equação da frequência de Larmor descreve o comportamento de um único núcleo atômico isolado em um campo magnético  $\mathbf{B}_0$ . Porém, os sistemas moleculares são formados por átomos, e não simplesmente núcleos, ligados entre si através de densidades eletrônicas e de interações eletrostáticas. É necessário incluir os efeitos da presença de densidades eletrônicas e de outros núcleos atômicos na frequência de ressonância dos núcleos.

Quando colocado em um campo magnético  $\mathbf{B}_0$ , o elétron tende a se movimentar em órbita circular perpendicular ao campo. O movimento do elétron implica na geração de corrente elétrica, que gera um campo magnético induzido  $\mathbf{B}_i$ , proporcional a  $\mathbf{B}_0$ . Então o campo sentido pelo núcleo,  $B_T$ , seria menor, devido à blindagem eletrônica, Equação 8.

Equação 8

$$B_T = B_0 - B_i = B_0 - \sigma B_0 = B_0 (1 - \sigma)$$

Sendo  $\sigma$  a constante de proporcionalidade entre o campo externo e o induzido, chamada de constante de blindagem.

Substituindo o  $\mathbf{B}_0$  na Equação 5 pelo encontrado na Equação 8, temos a frequência de ressonância dos núcleos descrita na Equação 9:

Equação 9

$$\nu = B_0 (1 - \sigma) (\gamma / 2\pi)$$

Essa equação leva em consideração o efeito da densidade eletrônica na frequência de RMN. Como esta densidade depende do ambiente químico onde o núcleo se encontra, podemos utilizar a frequência de ressonância para estudar o ambiente químico onde se encontra os diversos átomos em uma molécula.

A constante de blindagem deve incluir todos os efeitos no deslocamento químico do núcleo atômico que sejam de origem eletrônica, sendo então expressa como um somatório de diferentes efeitos eletrônicos:

$$\sigma = \sigma_D + \sigma_P + \sigma_A + \sigma_E + \sigma_S \dots$$

Que correspondem, respectivamente, às blindagens diamagnética e paramagnética, aos efeitos anisotrópicos, aos efeitos de campos elétricos e aos efeitos do solvente.

### 3.2.3 Escala e Referências

Como a frequência ( $\nu$ ) depende do campo aplicado, seria necessário indicar em qual campo o espectro foi obtido para identificar o núcleo. Para solucionar esse problema foi criada uma escala independente do campo aplicado,  $\delta$ .

Essa escala é definida utilizando um sinal como referência. Assim, para um núcleo com constante magnetogírica  $\gamma$ , temos a Equação 10:

Equação 10

$$\delta = \frac{\nu_A - \nu_R}{\nu_0}$$

Sendo  $\nu_A$ ,  $\nu_R$  e  $\nu_0$  as frequências do núcleo: na amostra, na referência e a do espectrômetro, respectivamente, dadas por:

Equação 11

$$\nu_A = B_0(1 - \sigma_A) \left( \frac{\gamma}{2\pi} \right) , \quad \nu_R = B_0(1 - \sigma_R) \left( \frac{\gamma}{2\pi} \right) \text{ e } \nu_0 = B_0(\gamma/2\pi)$$

Como a diferença no numerador está em Hz e a frequência do denominador está em MHz, então o  $\delta$  é em partes por milhão, ppm.

A referência mais utilizada é o tetrametilsilano, TMS. Por definição, os spins do TMS têm  $\delta = 0$  ppm. É um composto inerte, com um intenso e definido sinal de RMN.

A escala  $\delta$  sempre aumenta da direita para esquerda do espectro, uma convenção em RMN.

### 3.2.4 PRESAT e CPMG

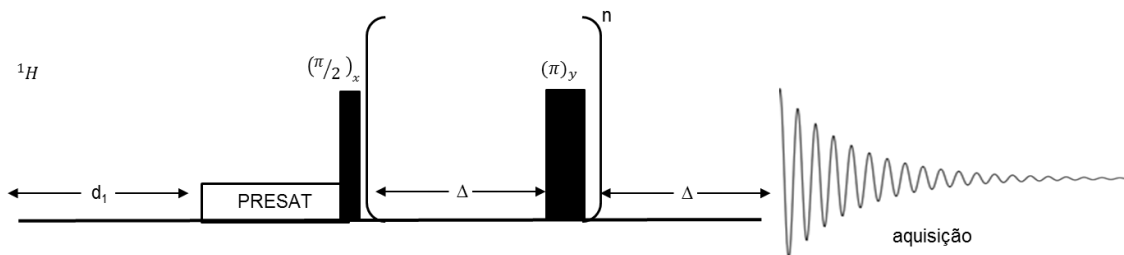
Há sequências de pulsos utilizadas para suprimir seletivamente alguns sinais, como por exemplo: PRESAT e CPMG (Carr-Purcell-Meiboom-Gill).

PRESAT consiste na aplicação de um pulso de baixa potência, em um tempo relativamente longo, numa frequência pré-determinada, tipicamente a frequência atribuída ao sinal da água. Este pulso iguala as populações dos dois estados e, conseqüentemente, sinais não são observados nessa frequência. Esta sequência de pulsos é particularmente útil para as amostras aquosas ou que tem um sinal muito intenso, que dificulte a observação dos sinais de interesse.

A sequência de pulsos CPMG, quando utilizada como filtro de  $T_2$  (tempo de relaxação transversal), tem o objetivo de eliminar os sinais provenientes de moléculas com  $T_2$  curto e possibilitar a análise dos sinais de moléculas maiores. Para isso, como observado na Figura 8: é aplicado um pulso de  $90^\circ$ , enviando a magnetização para o plano  $xy$ . Em seguida, são aplicados sucessivos pulsos de  $180^\circ$  entre intervalos de tempo  $\tau$ , até que os spins com valores menores de  $T_2$  tenham relaxado, e o espectro é adquirido contendo apenas as ressonâncias de interesse.

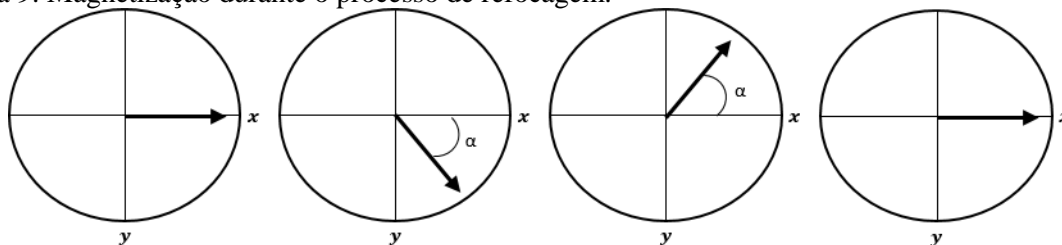
A combinação do pulso de  $180^\circ$  entre os intervalos de tempo  $\tau$ , servem para refocar a magnetização, pois durante o tempo  $\tau$  os spins precessam e a magnetização gira em um ângulo  $\alpha$  no sentido horário, por exemplo. Ao aplicar o pulso, a magnetização inverte em relação ao eixo, e com o segundo intervalo  $\tau$ , a magnetização gira no mesmo sentido e com o mesmo ângulo, voltando à posição inicial, como na Figura 9. Porém, núcleos com  $T_2$  curto relaxam durante os intervalos  $\tau$ , até que seus spins voltem totalmente ao eixo  $z$  e não apareçam seus sinais.

Figura 8: Sequência de pulsos CPMG.



Fonte: Autoria própria.

Figura 9: Magnetização durante o processo de refocagem.



Fonte: Autoria própria.

### 3.2.5 Sinal de RMN

A magnetização transversal, obtida pela rotação da magnetização longitudinal através da aplicação de um pulso de radiofrequência –  $90^\circ$  ou  $45^\circ$  em x ou y –, é muito pequena, apesar disso, é detectável, porque oscila a uma frequência bem definida.

Um momento magnético oscilante gera um campo magnético oscilante que pode gerar corrente elétrica oscilante em uma bobina. Essa corrente elétrica gerada pode ser detectada por um sensor de radiofrequência.

Os spins nucleares em uma amostra são detectados utilizando o decaimento livre de indução (FID, do inglês, *free- induction decay*), a detecção da corrente elétrica oscilante induzida pela precessão nuclear da magnetização transversal.

O decaimento da magnetização transversal é uma função do tempo, como vimos na Figura 7. Aplicando a transformada de Fourier, é possível visualizar os componentes individuais do sinal e obter o gráfico das intensidades em função das frequências, ou seja, o espectro de RMN.

### 3.2.6 Transformação do Espectro em Matriz de Dados

Cada espectro é dividido em regiões de mesma largura em ppm, denominadas bins, e é convertido em uma matriz linha, na qual os elementos são preenchidos com a integral do bin correspondente a cada coluna. É construída então uma matriz com cada caso (amostra) correspondendo a uma linha, e as colunas correspondendo às faixas de deslocamento químico pré-selecionadas, os bins.

É determinada a janela espectral (limites de deslocamento químico de interesse para a análise) baseada no conhecimento do espectro do biofluido, como sua composição e localização de sinais correspondentes a determinados metabólitos, podendo ser incluídos ou excluídos regiões e sinais específicos.(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007)



### 3.3 Quimiometria

Quimiometria é uma área da química que utiliza métodos matemáticos e estatísticos para: planejar ou selecionar condições ótimas para medidas e experimentos; ou extrair o máximo de informações relevantes a partir de análise de dados químicos.(BRUNS; FAIGLE, 1984)

Em sistemas com poucas variáveis é fácil encontrar a solução do problema, por exemplo: em uma dissolução de sal em água, a pressão e temperatura ambiente, podemos facilmente relacionar a quantidade de sal dissolvido à quantidade de água. Mas em um sistema em que se tenham dois sais sendo dissolvidos, com variação de temperatura e de pressão, essa relação não é tão direta, fazendo com que a quimiometria seja muito útil, tanto na escolha de condições ótimas para o procedimento, quanto para obter a relação final entre a dissolução do sal e os fatores que interferem.

Sistemas dessa natureza, multivariados, são muito comuns na química, como: reações que dependem de temperatura, pH, solvente, concentração dos reagentes; e utilização de espectros de Infravermelho (IR), ressonância magnética nuclear (RMN), espectrometria de massas (MS), e outras técnicas espectrométricas, para a determinação de concentração de um analito ou classificação de amostras a partir da sua impressão digital.(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007; BRUNS; FAIGLE, 1984)

Como os dados químicos utilizados pela estratégia metabonômica são espectros de RMN das amostras, que são dados multivariados, se faz necessário o uso de tratamentos quimiométricos para extrair as informações relacionadas à impressão digital da perturbação. Dessa forma, o conjunto de dados é organizado na forma de uma matriz de dados. As linhas correspondentes aos espectros de cada amostra são agrupadas em uma única matriz, contendo as amostras nas linhas (os casos) e os bins nas colunas (as variáveis). A matriz de dados está ilustrada na Figura 10.

Figura 10: Matriz de dados. Elemento  $x_{pn}$  representa o valor da variável  $n$  para a amostra  $p$ .

		Variáveis				
		1	2	3	...	n
Casos	1	$x_{11}$	$x_{12}$	$x_{13}$		$x_{1n}$
	2	$x_{21}$	...	$\vdots$		$\vdots$
	3	$x_{31}$	...	$x_{33}$		$\vdots$
	4	$x_{41}$	...	$\vdots$		$\vdots$
	$\vdots$					
	p	$x_{p1}$	...	$x_{3p}$		$x_{pn}$

Fonte: Autoria própria.

Para extrair as informações dessas matrizes de dados são utilizadas as técnicas de Reconhecimento de Padrões (RP), ferramentas que buscam identificar, caracterizar e avaliar similaridades e diferenças entre amostras de um conjunto de dados. As técnicas de RP são divididas em supervisionadas e não supervisionadas.(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007; BRUNS; FAIGLE, 1984)

As técnicas não supervisionadas recebem esse nome por não utilizarem as informações de classe na etapa de construção do modelo. São normalmente usadas para Análise Exploratória dos Dados (EDA), que visa obter um conhecimento do grupo de amostras e simplificação dos dados, removendo as informações redundantes e ruídos sem perder as informações importantes. Um exemplo de técnica de RP não supervisionada é a Análise de Componentes Principais (PCA).(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007; BRUNS; FAIGLE, 1984)

As técnicas de RP supervisionadas utilizam a informação de classe das amostras para construir modelos de classificação, capazes de prever a qual classe pertence uma amostra desconhecida. Neste trabalho utilizamos duas técnicas de RP supervisionadas: Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA) e Análise Discriminante Linear (LDA).

Antes de aplicar as técnicas de RP à matriz de dados, é necessário pré-processar os dados para corrigir alguns fatores que podem fazer com que algumas amostras ou

variáveis influenciem mais no resultado da análise, sem ter essa real importância, como variáveis em ordens de grandeza diferentes e amostras em diferentes concentrações.

Como as amostras usadas são de biofluidos, há uma diferença de diluição, isso acontece principalmente na urina. Para evitar que amostras mais concentradas pesem mais na construção dos modelos, há vários tipos de normalizações, como normalização pela soma e *Standard Normal Variate* (SNV).

A normalização pela soma consiste em dividir cada bin ( $x_i$ ) pela soma de todos os bins da amostra, Equação 12. SNV subtrai de cada bin ( $x_i$ ) o valor da média ( $\bar{x}$ ) de todos os bins da amostra e divide pelo desvio padrão ( $\sigma$ ), Equação 13.

Nas variáveis, é utilizado o autoescalonamento, necessário quando há um metabólito, ou um grupo de metabólitos, que aparece em grande concentração no biofluido, para todas as amostras. Pois, quando há variáveis com valores que variam em ordens de grandeza diferentes, a comparação direta entre elas atribui uma ponderação maior às variáveis com maior valor numérico, dessa forma os sinais de maior intensidade dariam contribuições muito maiores para o modelo criado. Porém, têm que se tomar cuidado para não dar a mesma importância para sinais de metabólitos e ruídos. O autoescalonamento utiliza a mesma equação que o SNV, porém, é aplicado às variáveis.

Equação 12: Normalização pela soma.

$$x'_i = \frac{x_i}{\sum x}$$

Equação 13: SNV para as amostras e autoescalonamento para as variáveis.

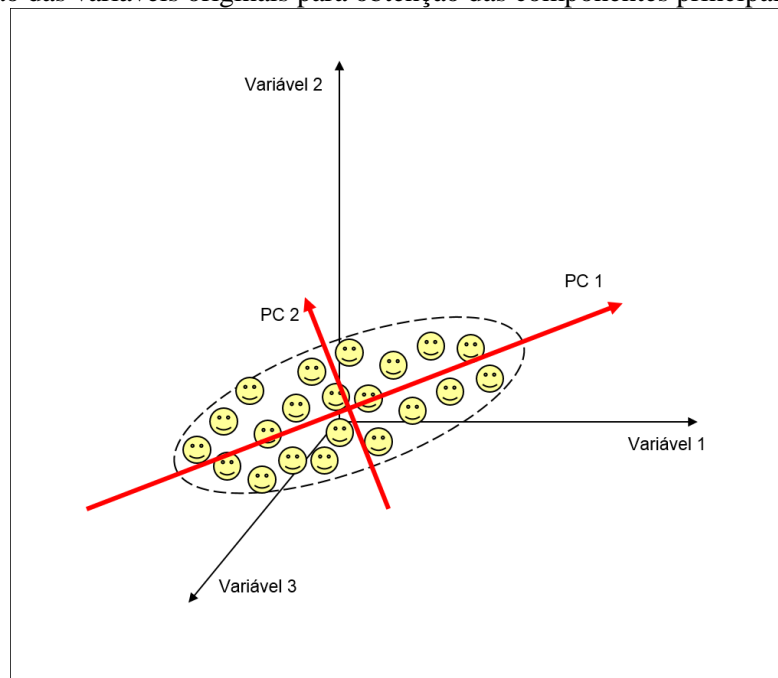
$$x'_i = \frac{x_i - \bar{x}}{\sigma}$$

Em que  $x'_i$  é o valor da amostra na variável  $i$  após a normalização,  $x_i$  é o valor da amostra na variável  $i$  antes da normalização,  $\bar{x}$  é a média dos valores de todas as variáveis naquela amostra e  $\sigma$  é o desvio padrão, no caso do SNV. Para o autoescalonamento,  $x'_i$  é o valor da variável na amostra  $i$  após a normalização,  $x_i$  é o valor da variável na amostra  $i$  antes da normalização,  $\bar{x}$  é a média e  $\sigma$  é o desvio padrão dos valores de todas as amostras naquela variável.

### 3.3.1 Análise de Componentes Principais

A análise de componentes principais (PCA) consiste em transformar as variáveis originais em um novo grupo de variáveis não correlacionadas, chamadas de Componentes Principais (PC). As componentes principais são eixos ortogonais construídos a partir de combinações lineares das variáveis originais e que representam as direções de maior variância dos dados. A primeira componente principal (PC1) explica a maior variância dos dados, a segunda componente principal (PC2) explica o máximo da variância residual, e assim sucessivamente até que toda a variância dos dados seja explicada. (BERRUETA; ALONSO-SALCES; HÉBERGER, 2007; NETO; MIOTA, 1998) As duas primeiras componentes principais de um grupo de dados hipotético estão ilustradas na Figura 11.

Figura 11: Rotação das variáveis originais para obtenção das componentes principais.



Fonte: Autoria própria.

Cada componente principal é definida pela variância explicada, o conjunto de pesos das variáveis e o conjunto de escores das amostras. Os pesos das variáveis, ou *loadings*, são definidos a partir da matriz de covariância e correspondem ao cosseno do ângulo entre o eixo da PC e das variáveis originais, dando origem à matriz de pesos, **L**.

O escore é a coordenada da amostra no eixo da PC. É obtido pelo somatório do valor do peso de cada variável, naquela componente principal, multiplicado pelo valor da variável original para aquela amostra, Equação 14.

Equação 14

$$t_{ih} = \sum_j x_{ij} l_{jh}$$

Sendo  $t_{ih}$  o escore da amostra  $i$  na componente  $h$ ,  $x_{ij}$  o valor da variável  $j$  na amostra  $i$  e  $l_{jh}$  o peso da variável  $j$  na componente  $h$ .(DE SOUZA; POPPI, 2012) Os escores de todas as amostras nas componentes principais construídas compõem a matriz de escores,  $\mathbf{T}$ .

Podemos então representar o modelo PCA na sua forma matricial a partir da Equação 15.

Equação 15

$$\mathbf{X} = \mathbf{T} \mathbf{L}^t + \mathbf{E}$$

Em que  $\mathbf{X}$  é a matriz de dados originais,  $\mathbf{T}$  é a matriz de escores,  $\mathbf{L}$  é a matriz de pesos e  $\mathbf{E}$  é a matriz de erro, que corresponde aos resíduos do modelo.(DE SOUZA; POPPI, 2012; LEE; GILMORE, 2006)

A PCA permite visualizar a estrutura dos dados e encontrar a similaridade entre amostras, através da observação de agrupamentos das amostras quando projetadas nas componentes principais.

Também é utilizada para detecção de amostras anômalas, ou *outliers*, que são amostras que se apresentam fora do padrão ou agrupamento mostrado pela grande maioria das amostras. É importante que essas amostras sejam excluídas pois estarão adicionando, aos modelos construídos, informações que não correspondem ao grupo estudado. Mas, para ser considerada uma amostra anômala, é necessário avaliar o resíduo, distância da amostra para o modelo, e a sua influência na construção da componente principal.

A PCA permite a redução da dimensionalidade dos dados, uma vez que remove as informações redundantes e os ruídos. Além disso, como as componentes são construídas de forma decrescente em variância explicada, a maior parte das informações significativas estarão nas primeiras componentes.(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007)

### 3.3.2 Análise Discriminante por Mínimos Quadrados Parciais

A PLS-DA é uma técnica supervisionada, portanto, utiliza uma matriz que contém as informações correspondentes às classes das amostras na etapa de construção do modelo ( $\mathbf{Y}$ ). O número de colunas de  $\mathbf{Y}$  é igual a quantidade de classes das amostras, os

casos pertencentes àquela classe recebem o valor 1 e os que não pertencem, recebem o valor 0. A Figura 12 mostra um exemplo das matrizes **X** e **Y**, para duas classes A e B.

A PLS-DA se baseia em encontrar componentes que melhor expliquem as variações das variáveis em **X** que sejam relacionadas aos valores de **Y**, dando menos pesos às variações não correlacionadas a **Y** e aos ruídos.

Figura 12: Exemplo de matrizes **X** e **Y**.

		X					Y		
		Variáveis					Classes		
		1	2	3	...	n	A	B	
Casos	1	$x_{11}$	$x_{12}$	$x_{13}$		$x_{1n}$	1	0	amostras da classe A
	2	$x_{21}$	...	$\vdots$		$\vdots$	1	0	
	3	$x_{31}$	...	$x_{33}$		$\vdots$	1	0	
	4	$x_{41}$	...	$\vdots$		$\vdots$	0	1	
	$\vdots$						$\vdots$	$\vdots$	amostras da classe B
	p	$x_{p1}$	...	$x_{3p}$		$x_{pn}$	0	1	

Fonte: Autoria própria.

Na PCA, as componentes principais são calculadas com a mínima correlação entre si e que maximizam a variância em **X**. Já a PLS-DA modela **X** e **Y** simultaneamente, procurando Variáveis Latentes (LV), combinações lineares das variáveis originais assim como as componentes principais, que expliquem a máxima correlação entre as matrizes. Geometricamente, as variáveis latentes podem ser interpretadas como uma rotação das componentes principais que visa aumentar a correlação entre **X** e **Y**.(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007; WOLD; SJÖSTRÖM; ERIKSSON, 2001)

Uma limitação desse método é que inclui erros de ambas as matrizes, portanto é mais indicado para conjuntos de dados com pouco ruído. Uma forma de evitar esses erros é não utilizar um grande número de variáveis latentes para construir o modelo de classificação.(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007)

Na construção do modelo, as duas matrizes são decompostas simultaneamente em matrizes de escores e *loadings*, encontrando pesos que determinem as mesmas projeções, matriz de escores **T**, para **X** e **Y**. As equações usadas na construção das variáveis latentes são Equação 16 e Equação 17.(LEE; GILMORE, 2006)

Equação 16

$$\mathbf{X} = \mathbf{T} \mathbf{P}^t + \mathbf{E}$$

Equação 17

$$Y = T Q^t + F$$

Em que **X** é a matriz das variáveis independentes, **Y** é a matriz das variáveis dependentes (classes), **T** é a matriz de escores, **P** é a matriz de pesos e **E** é a matriz de erro, ambas de **X**, enquanto, **Q** é a matriz de pesos e **F** é a matriz de erro, de **Y**.(LEE; GILMORE, 2006)

Para classificar amostras desconhecidas, as equações 16 e 17 são utilizadas para encontrar o valor de **Y** e um ponto de corte é definido para classificar a amostra. Normalmente é utilizado 0,5, a partir desse valor em **Y**, a amostra é considerada pertencente à classe.

Ao se analisar **P**, é possível identificar variáveis importantes para a classificação. No caso da metabonômica, essas variáveis indicam regiões do espectro que são mais intensas em alguma das categorias, caracterizando metabólitos presentes em concentrações diferentes devido à perturbação estudada.

### 3.3.3 Análise Discriminante Linear

Análise discriminante linear (LDA) se baseia em procurar uma função linear que permita a discriminação entre pontos relativos a amostras de duas diferentes categorias. Para isso, as classes devem seguir uma distribuição normal e serem linearmente separáveis.(BRUNS; FAIGLE, 1984)

Assim como a PCA, a LDA projeta dados de grandes dimensões em planos de dimensões menores. No entanto, enquanto a PCA seleciona a direção que retém a máxima variância dos dados em uma dimensão inferior, LDA seleciona a direção que alcança a máxima separação entre classes e minimiza a separação intra - classes.(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007)

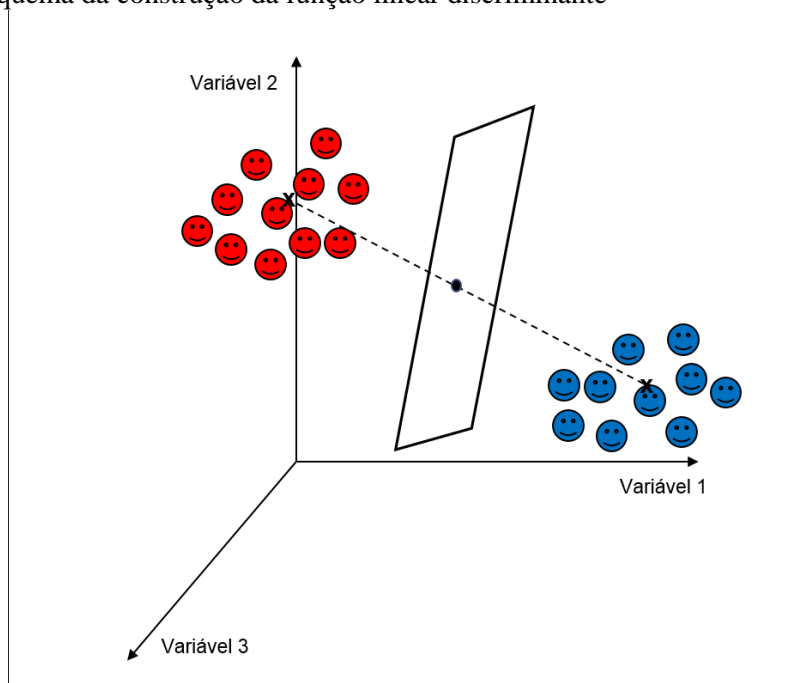
Para se obter uma função discriminante linear, como está ilustrado na Figura 13, determina-se o centro de gravidade para cada classe e, então, uma linha reta é traçada entre eles. Perpendicular a essa reta, equidistante dos centros de gravidade, é traçado um plano de decisão, a função discriminante. O espaço fica, então, dividido em duas regiões. Projetando-se uma amostra de categoria desconhecida, é possível classificá-la de acordo com a região na qual aparece. Em um espaço tridimensional, como na Figura 13, temos um plano discriminante, mas em espaços de quatro ou mais dimensões, a

função linear será uma superfície, um hiperplano, de dimensão  $m-1$ , em que  $m$  é o número de variáveis usadas. (BRUNS; FAIGLE, 1984)

A LDA tem uma limitação: o número de variáveis não pode ser maior que o número de amostras. Mas esse problema pode ser facilmente resolvido aplicando-se um método de seleção de variáveis ou de redução de dimensionalidade.

Métodos de seleção de variáveis são muito utilizados, além do caso em que o número de objetos é relativamente pequeno e o número de variáveis é grande, quando há variáveis que contêm informações redundantes ou ruído, para que o modelo não perca sua capacidade de generalização. (BERRUETA; ALONSO-SALCES; HÉBERGER, 2007)

Figura 13: Esquema da construção da função linear discriminante



Fonte: Autoria própria.

Para análise discriminante, um tipo de seleção de variáveis busca, entre todos os possíveis subconjuntos de variáveis, o que melhor discriminar as classes segundo um determinado critério, como por exemplo, o Lambda de Wilks.

O lambda de Wilks é um parâmetro similar ao teste F, um teste de variância só que para dados multivariados. O valor do Lambda de Wilks é calculado para todo os subconjuntos de variáveis para testar a hipótese de que as duas (ou mais) classes são diferentes. (CRICHTON, 2000)

Quando existem muitas variáveis e diversos tamanhos de subconjuntos, o número total de subconjuntos possíveis e o tempo de processamento podem se tornar



excessivamente grandes. Nesses casos, é indicado utilizar um método de redução de dimensionalidade e determinar um tamanho específico para os subconjuntos testados.

Uma opção é utilizar a PCA como técnica de redução da dimensão das variáveis, excluindo as informações redundantes e ruídos, antes de submeter ao método de seleção de variáveis para a construção do modelo LDA.

A PLS-DA tem a capacidade de construir modelos satisfatórios com um número de variáveis maior que o número de casos. Contudo, é preciso estar atento ao sobreajuste do modelo PLS-DA, quando o número de variáveis latentes usadas for maior que o necessário, e ao sub-ajuste, quando é menor que o necessário.(BERRUETA; ALONSO-SALCES; HÉBERGER, 2007)

As duas técnicas de RP supervisionadas apresentadas (PLS-DA e LDA) não contemplam a possibilidade de a amostra pertencer a uma categoria ainda não definida, ou seja, a amostra sempre será classificada como pertencente a uma das categorias modeladas.(BRUNS; FAIGLE, 1984)

Após a construção do modelo é feita a validação. Essa etapa tem como objetivos avaliar: o carácter representativo dos dados utilizados para produzir o modelo; o número de variáveis necessárias para caracterizar o conjunto de dados; e a capacidade do modelo de prever amostras desconhecidas.

A validação externa é usada quando há amostras suficientes para criar grupos de treinamento, usado para construção do modelo, e de teste, usado para avaliar o modelo, cada um contendo amostras representativas de cada classe. Mas é comum a quantidade de amostras disponíveis não ser suficiente para dividir em dois grupos representativos, nesses casos, a validação cruzada é aplicada.

Um exemplo de validação cruzada muito utilizado é a Validação Cruzada Total (LOOCV, do inglês, *Leave-One-Out Cross Validation*). A LOOCV consiste em retirar uma amostra do conjunto de treinamento, construindo um novo modelo sem a amostra retirada, e usá-la para validá-lo. Em seguida, o mesmo procedimento é realizado com a próxima amostra, até que todas as amostras sejam retiradas do conjunto de treinamento e usadas como conjunto de teste. Há outros tipos de validação cruzada, nas quais se tira uma quantidade diferente de amostras do conjunto de treinamento, para constituir o conjunto de teste. A exatidão (porcentagem de acertos) é calculada para cada modelo e a performance na validação cruzada é a média das exatidões obtidas em todos os modelos.

Quando as categorias às quais as amostras pertencem são relacionadas a doenças, é comum representar os resultados por meio da tabela de contingência e, a partir dela,

calcular sensibilidade, especificidade, Valor Preditivo Positivo (VPP), Valor Preditivo Negativo (VPN) e exatidão. Sensibilidade expressa a proporção de doentes, identificados pelo padrão ouro para o diagnóstico de determinada doença, que foram corretamente classificados pelo novo teste. Especificidade indica a proporção de saudáveis, também comparando com padrão ouro, corretamente classificados pelo novo teste. Valor preditivo positivo indica qual a chance de acerto quando o novo teste der positivo, e valor preditivo negativo indica a chance de acerto quando o novo teste der negativo.

Na Figura 14 está ilustrada uma tabela de contingência genérica e como calcular os parâmetros de avaliação do modelo.

Figura 14: Tabela de contingência e parâmetros para avaliação de um teste.

		Padrão Ouro	
		Presença da doença	Saudável
Novo teste	Presença da doença	a	b
	Saudável	c	d
Sensibilidade (%) = $\frac{a}{a+c} \times 100$		Valor preditivo positivo (%) = $\frac{a}{a+b} \times 100$	
Especificidade (%) = $\frac{d}{b+d} \times 100$		Valor preditivo negativo (%) = $\frac{d}{c+d} \times 100$	
Exatidão (%) = $\frac{a+d}{a+b+c+d} \times 100$			

Fonte: Autoria própria.

Em que a é a quantidade de doentes classificados pelo novo teste como doentes, b é a quantidade de saudáveis classificados como doentes, c é a quantidade de doentes classificados como saudáveis e d é a quantidade de saudáveis classificados como saudáveis.

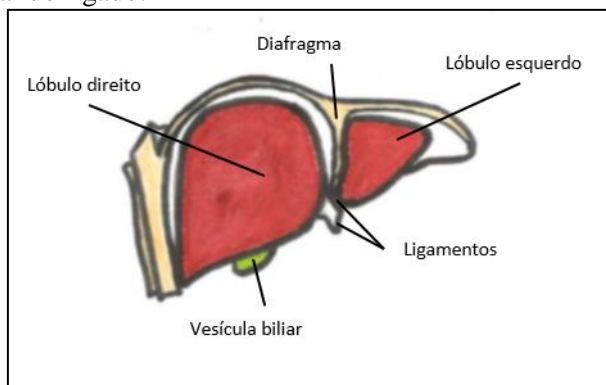
### 3.4 Fígado

O fígado é o maior órgão interno do corpo humano, em um adulto normal pesa de 1200 a 1600 g, constituindo aproximadamente 2,5% do peso corporal.(KUMAR et al., 2006) (SHERLOCK; DOOLEY, 1997)

Este órgão possui duas vias de entrada de sangue: 60 a 70% do sangue é levado do intestino e baço pela veia porta, e é rico em produtos bacterianos, toxinas e antígenos alimentares; os outros 30 a 40% são fornecidos por vascularização da artéria hepática. No interior do fígado, a veia porta, a artéria hepática e os ductos biliares se ramificam paralelamente em tratos portais, de até 17 a 20 ordens de ramos.(KUMAR et al., 2006) (SHERLOCK; DOOLEY, 1997)

De acordo com o modelo lobular, o fígado é constituído pelos lóbulos esquerdo e direito, separados por ligamentos, sendo o direito cerca de seis vezes maior que o esquerdo, conforme ilustrados na Figura 15. O lóbulo direito tem dois seguimentos menores: lóbulos quadrado e caudado. (SHERLOCK; DOOLEY, 1997) No centro dos lóbulos há terminais da veia hepática e, nas periferias, há tratos portais.(KUMAR et al., 2006)

Figura 15: Vista frontal do fígado.



Fonte: Autoria própria.

Cerca de 80% do volume do fígado é formado por hepatócitos, entre 60 a 70% das células desse órgão. Estão organizados em placas, que vão das veias hepáticas terminais até os tratos portais. Entre essas placas estão os sinusoides, canais por onde passa o sangue. Os hepatócitos vizinhos à veia hepática terminal são chamados de centrolobulares e os situados próximos ao trato portal são chamados periportais. Os outros 20% são formados por células não parênquimas, como endoteliais, células de Kupffer (macrófagos capazes de fagocitar bactérias e outras matérias estranhas do sangue) e linfócitos.(SHERLOCK; DOOLEY, 1997)

O fígado tem a capacidade de se regenerar. Acredita-se que durante a regeneração hepática, os hepatócitos se repliquem uma ou duas vezes até atingir o tamanho e volume hepático normal, voltando ao estado de repouso. (KUMAR et al., 2006)

Os hepatócitos são responsáveis por suprir as necessidades de metabolização e desintoxicação do organismo. (HALL; GUYTON, 2011; KUMAR et al., 2006)

O fígado tem muitas funções diferentes, o que fica evidente quando ocorre uma disfunção neste órgão, pois percebe-se várias alterações no organismo ocorrendo simultaneamente. Elas podem ser resumidas em:

**Reservatório de sangue:** O volume sanguíneo normal do fígado é cerca de 450 mL, ou 10% do volume sanguíneo corporal. Porém, como é um órgão expansível, tem a capacidade de armazenar em suas veias e sinusoides de 0,5 a 1 L de sangue extra.

**Filtração de sangue:** O sangue que flui pelo trato digestivo recolhe muitas bactérias. Elas entram no fígado pelas veias porta e à medida que passam pelos sinusoides, são fagocitados pelas células Kupffer, permanecendo alojadas até serem digeridas. Estima-se que menos de 1% das bactérias que entram no sangue porta consiga atravessar o fígado e atingir a circulação sistêmica.

**Armazenamento de vitaminas e ferro:** É capaz de armazenar vitamina A suficiente para suprir a necessidade do corpo por 10 meses; vitamina D para 3 ou 4 meses; e vitamina B<sub>12</sub> para, pelo menos, um ano. Além disso, armazena a maior proporção de ferro do organismo, depois da hemoglobina sanguínea, sob a forma de ferritina.

**Função metabólica do fígado:** Este órgão é um grande agrupamento celular quimicamente reativo, que processa e sintetiza muitas substâncias que são transportadas para outras partes do corpo. Suas principais atividades metabólicas são: metabolismo de carboidratos (importante na manutenção do nível de glicose no sangue), de gorduras (importante para a produção de energia e sais biliares) e de proteínas (forma cerca de 90% das proteínas plasmáticas, assim como alguns aminoácidos e outras substâncias a partir deles).

**Formação de fatores de coagulação:** É responsável pela sintetização de várias substâncias sanguíneas utilizadas na coagulação.

**Formação de bile e excreção de hormônios e substâncias estranhas ao organismo:** É conhecido pela sua capacidade de desintoxicar o organismo. Metaboliza e/ou excreta fármacos, hormônios e outras substâncias. A excreção dessas substâncias é feita através da bile, passando para o intestino e sendo liberadas nas fezes. (HALL; GUYTON, 2011)

Geralmente as doenças hepáticas têm um processo insidioso, no qual os sintomas só aparecem muito tempo após ocorrer a lesão, quando o quadro já está avançado. Por isso, a maioria dos pacientes são encaminhados para hepatologistas quando apresentam doença hepática crônica.(KUMAR et al., 2006)

As principais doenças do fígado são as hepatites virais, a Doença Hepática Alcoólica (DHA), a Doença Hepática Gordurosa Não Alcoólica (DHGNA) e o Carcinoma Hepatocelular (CHC). Porém, a lesão hepática também pode ocorrer devido a algumas outras doenças, como descompensação cardíaca, câncer disseminado e infecções.(KUMAR et al., 2006)

### 3.4.1 Esteatose

A DHGNA é considerada a forma mais comum do grupo de doenças hepáticas atribuídas a distúrbios do metabolismo, adquiridos ou hereditários, e à má nutrição. É a causa mais frequente de distúrbio hepático em muitos países desenvolvidos, e a principal causa de doença hepática crônica nos Estados Unidos afetando, em suas várias formas, mais de 30% da população. Mas essa estimativa pode estar subestimando a incidência e prevalência desta doença, que muitas vezes é assintomática.(KUMAR et al., 2006; SHERLOCK; DOOLEY, 1997)

Na Europa, a prevalência da DHGNA é estimada em 20 a 30% da população geral, e 2,6 a 10% na população pediátrica.(LABRECQUE et al., 2014)

A DHGNA pode se manifestar como esteatose simples ou acompanhada por uma inflamação mínima, mas, geralmente, não apresenta problemas clínicos significativos. Porém, quando há o excesso do acúmulo de gordura no fígado subsequente de inflamação hepática, tem-se a esteato-hepatite Não Alcoólica (EHNA), na qual ocorre lesão de hepatócitos. É uma das causas mais comuns de cirrose, evoluindo em 10 a 20% dos casos.(KUMAR et al., 2006)

Apesar de não apresentar sintomas e grandes problemas clínicos, a esteatose pode evoluir para problemas mais graves, como esteato-hepatite e, posteriormente, cirrose ou CHC. Além disso, as diferentes formas de DHGNA contribuem para a progressão de outras doenças hepáticas, como infecção por vírus da hepatite C (HCV) e CHC. Portanto é importante que seja feito o diagnóstico no início da doença.(KUMAR et al., 2006)

O acúmulo de gordura no fígado pode ser observado por exames de imagem, mas o padrão ouro para seu diagnóstico é a biópsia hepática, que identifica também a extensão da lesão, a presença de EHNA e o grau de fibrose.(KUMAR et al., 2006)

### 3.4.2 Hepatites Virais B e C

O termo “hepatites virais” é aplicado às infecções hepáticas causadas por um grupo de vírus que possuem uma afinidade particular pelo fígado, conhecidos como vírus hepatotrópicos. Estão incluídos os vírus das hepatites A, B, C, D e E.(KUMAR et al., 2006) Vamos tratar especificamente das hepatites virais B e C.

O vírus da hepatite B (HBV) pode se manifestar de várias formas: hepatite aguda com recuperação e eliminação do vírus; hepatite crônica não progressiva; doença crônica progressiva terminando em cirrose; ou um estado de portador assintomático.(KUMAR et al., 2006; SHERLOCK; DOOLEY, 1997)

A doença hepática decorrente da infecção pelo HBV constitui um enorme problema de saúde pública global, pois cerca de um terço da população mundial, 2 bilhões de pessoas, está infectada pelo HBV.(KUMAR et al., 2006)

Na maioria das pessoas, é assintomática durante a fase aguda, no entanto, algumas apresentam sintomas que podem durar várias semanas, como icterícia, fadiga extrema, náuseas, vômitos e dor abdominal. Estima-se que mais de 90% dos adultos saudáveis infectados por HBV se recuperam naturalmente dentro do primeiro ano, e menos de 5% desenvolvem infecção crônica. A estimativa é de que haja 240 milhões de pessoas cronicamente infectadas por HBV, no mundo.(ORGANIZATION, 2015a)

A hepatite crônica é definida como a evidência sintomática, bioquímica ou sorológica de doença hepática contínua ou recorrente por mais de 6 meses.(KUMAR et al., 2006). O principal objetivo do tratamento da hepatite B crônica consiste em reduzir o risco de progressão da doença hepática e de seus desfechos primários, especificamente cirrose, hepatocarcinoma e, conseqüentemente, o óbito. (MINISTÉRIO DA SAÚDE; SECRETARIA DE VIGILÂNCIA EM SAÚDE; DEPARTAMENTO DE DST, 2011)

Em cerca de 20 a 30% dos indivíduos infectados pelo HBV, a hepatite crônica pode evoluir para cirrose e CHC. Entre 1990 e 2010, mais de 780 mil pessoas morriam a cada dia, devido às complicações de hepatite B. (ORGANIZATION, 2015a)

O modo de transmissão do HBV varia, em áreas com alta prevalência, a transmissão perinatal - durante o parto - representa 90% dos casos. Em áreas de baixa

prevalência, os principais modos de disseminação são a relação sexual sem proteção e o abuso de drogas intravenosas.(KUMAR et al., 2006)

O HBV tem um período de incubação prolongado, de 4 a 26 semanas, e permanece no sangue durante a hepatite aguda e crônica. Esse vírus consiste em uma molécula de DNA circular de fita parcialmente dupla, que apresenta aproximadamente 3.200 nucleotídeos, e é membro de uma família de vírus DNA que causa hepatite em várias espécies animais, *Hepadnaviridae*. Foram identificadas dez variações genotípicas do HBV (CROAGH; DESMOND; BELL, 2015) com distribuição geográfica ao redor do mundo.(KUMAR et al., 2006; SHERLOCK; DOOLEY, 1997)

O diagnóstico da infecção pelo HBV é baseado no exame laboratorial sorológico através das técnicas de ensaio imunoenzimático (ELISA, do inglês, *enzyme-linked immunosorbent assay*), ensaio enzimático de micropartículas (MEIA, do inglês, *microparticle enzyme immunoassay*), ou de quimioluminescência que analisa a presença dos antígenos HBsAg e HBeAg, e dos anticorpos anti-HBc, anti-HBs e anti-HBe. Também são utilizados testes de biologia molecular específicos como a reação em cadeia da polimerase (PCR, do inglês, *polymerase chain reaction*), método que amplifica as sequências genéticas específicas de modo que uma única molécula de DNA seja detectada dentre milhões, para detectar o HBV-DNA.(BRANDÃO et al., 2001; LIANG, 2009) Os marcadores serológicos estão resumidos na Tabela 2. (KUMAR et al., 2006; SHERLOCK; DOOLEY, 1997)

Também são medidos os níveis séricos de alanina e aspartato aminotransferase (ALT, AST), que começam a subir após a infecção.(LIANG, 2009)

Há uma vacina de prevenção a HBV, que consiste em HBsAg purificado e induz a produção de anticorpos anti-HBs em mais de 95% dos casos.(KUMAR et al., 2006; ORGANIZATION, 2015a)

A OMS estima que cerca de 3% da população mundial está infectada por HCV e que cerca de 170 milhões de portadores crônicos estão em risco de desenvolver cirrose ou CHC. (ORGANIZATION, 2015b) A incidência de HCV em algumas regiões do mundo está disposta na Tabela 3. Além disso, cerca de 60-70% das infecções agudas de HCV são assintomáticas, dificultando a obtenção de dados mais exatos. (ALTER, 2007; LAVANCHY, 2011) Aproximadamente 4,1 milhões de norte-americanos, ou 1,6% da população, apresentam infecção crônica por HCV, a infecção transmitida pelo sangue mais comum e responsável por quase metade de todos os indivíduos com doença hepática crônica nos EUA.(KUMAR et al., 2006)

Tabela 2: Marcadores sorológicos de HBV.

Marcador	Detalhes	Significância
HBsAg	Aparece antes do início dos sintomas e diminui até níveis indetectáveis em 3 a 6 meses.	Presença de hepatite B aguda ou crônica
IgM anti-HBc	Detectável no soro pouco tempo antes do início dos sintomas.	Hepatite B aguda
IgG anti-HBc	Após alguns meses IgM anti-HBc é substituído por IgG anti-HBc.	Exposição a hepatite B (sem HBsAg) Hepatite B crônica (com HBsAg)
HBV DNA	Permanece no sangue durante hepatite aguda e crônica.	Infecção continuada
HBeAg	Aparece no soro logo após HBsAg e indica replicação viral ativa.	Persistência indica provável hepatite B crônica
Anti-HBe	Aparece quando a hepatite aguda atingiu seu pico e está declinando.	Convalescença
Anti-HBs	Detectável após o desaparecimento do HBsAg e pode persistir por toda a vida.	Imune a hepatite B

Fonte: KUMAR et al., 2006; SHERLOCK; DOOLEY, 1997

Tabela 3: Número de infectados por HCV em diferentes regiões

Região	Número de infectados por HCV
África	28.100.000
Américas	14.000.000
Ásia	83.000.000
Austrália e Oceania	400.000
Europa	17.500.000
Oriente Médio	16.000.000
Total	159.000.000

Fonte: (LAVANCHY, 2011)



Em aproximadamente 85% dos indivíduos, o curso clínico da infecção aguda é assintomático e facilmente ignorado. Em contraste com o HBV, 80% dos casos de hepatite C aguda progridem para infecção crônica; 10 a 20% desenvolvem cirrose, que pode desenvolver-se ao longo de 5 a 20 anos após a infecção; 1 a 5% apresentam CHC; e 7 a 8% evoluem para óbito. De mais de 500 mil casos de câncer de fígado que ocorrem por ano, aproximadamente, 22% são atribuídos à infecção por HCV. (KUMAR et al., 2006; LAVANCHY, 2011; NIEDERAU et al., 1998)

De acordo com dados de 2008 dos Centros para Controle de Doenças nos EUA, as formas mais comuns de infecção por HCV são: abuso de drogas intravenosas; múltiplos parceiros sexuais ; realização de cirurgia nos últimos 6 meses ; ferimento por picada de agulha; contatos múltiplos com uma pessoa infectada por HCV; e atividade profissional na área médica ou odontológica.(KUMAR et al., 2006)

O HCV é um vírus de RNA pequeno, envelopado, de fita única, com um genoma de 9,6 kb, membro da família *Flaviviridae*. O HCV tem múltiplos genótipos e subtipos, em qualquer paciente, o vírus circula como uma população de variantes conhecidas como quase-espécies. Com o tempo, dezenas de quase-espécies podem ser detectadas em um indivíduo e mapeadas como derivadas da HCV original que infectou o paciente. Isso dificulta seriamente os esforços para desenvolver uma vacina contra o HCV. O período de incubação da hepatite por HCV varia de 2 a 26 semanas, com uma média de 6 a 12 semanas.(KUMAR et al., 2006)

Um aspecto característico da infecção por HCV consiste em surtos repetidos de lesão hepática, resultantes da reativação de uma infecção preexistente ou do aparecimento de uma variação nova do vírus, resultado de uma mutação recente. HCV-RNA é detectável no sangue após 1 a 3 semanas da infecção, coincidindo com as elevações de ALT e AST. Em mais de 90% dos pacientes com infecção crônica, o HCV-RNA persiste apesar da presença de anticorpos neutralizantes.(KUMAR et al., 2006)

O diagnóstico da infecção pelo HCV é frequentemente realizado pela detecção dos anticorpos e antígenos específicos do vírus. Para isso, são usados testes de rastreamento como o ELISA, que utiliza proteínas recombinantes ou peptídeos sintéticos para a captação do anti-HCV. O teste ELISA está na quarta geração e apresenta uma sensibilidade de 97%.(BRANDÃO et al., 2001)

Contudo, resultados positivos no ELISA podem ser devido a uma infecção já curada, visto que os anticorpos podem permanecer por anos após a cura da doença. Por

isso, o padrão ouro no diagnóstico de hepatite C é a determinação do RNA do vírus através da PCR.(BRANDÃO et al., 2001)

Espera-se que o número de casos de cirrose e CHC aumentem significativamente na próxima década, pois estudos mostram que os índices de prevalência das hepatites B e C aumentaram desde a última estimativa,(LAVANCHY, 2011) e doenças do fígado podem se desenvolver anos depois da infecção, além dos casos não diagnosticados e não tratados.(ALTER, 2007) Portanto, infecção por HBV ou HCV caracteriza um grande problema de saúde pública em todo o mundo.

É importante que haja a disponibilidade de diferentes testes de diagnóstico, o que viabiliza o diagnóstico precoce, minimizando o potencial para disseminação da infecção e proporcionando a possibilidade de controle e tratamento da doença, evitando que ocorra o desenvolvimento de complicações mais graves como cirrose e CHC.

### 3.4.3 Fibrose Hepática

Como a hepatite crônica causa lesão nas células hepáticas, começam a ocorrer eventos de regeneração e reparo dessas células. A regeneração resulta na restituição completa do tecido perdido ou lesado. O reparo pode restaurar algumas estruturas originais, mas pode causar desarranjos estruturais, pois consiste em uma combinação de regeneração das células originais e formação de cicatriz pela deposição de colágeno.(KUMAR et al., 2006)

Os hepatócitos têm a capacidade de se regenerar rapidamente, porém, quando há presença de inflamação, por causa da hepatite crônica, o processo de formação de cicatriz fica mais forte, e as células parenquimais são substituídas por colágeno e outros componentes da matriz celular, formando o que é chamado de fibrose.(KUMAR et al., 2006) O grau de fibrose hepática caracteriza o nível de comprometimento do fígado e é, geralmente, classificado pelo índice METAVIR, que vai de F0, sem fibrose, a F4, cirrose.

O tratamento das hepatites crônicas é recomendado em pacientes que apresentam fibrose hepática mais severa, METAVIR F3 ou F4, ou que tenham diagnóstico de F2 há mais de 3 anos. Os demais pacientes devem ser monitorados periodicamente até que se configure uma indicação de terapia. (CALVARUSO; CRAXÌ, 2012; ORGANIZATION, 2015b) Portanto, o conhecimento do grau de fibrose hepática, decorrente da infecção crônica por HBV ou HVC, é de fundamental importância.

Segundo a OMS e o Ministério da Saúde, o padrão ouro para definição do grau de acometimento hepático é biópsia hepática. A biópsia hepática é muito útil no diagnóstico de outras doenças hepáticas concomitantes, como a doença gordurosa, que impactam de maneira significativa a evolução dos casos. Entretanto, é um procedimento invasivo que requer uma estrutura, nem sempre disponível. Quando disponível, a principal limitação da biópsia é o erro de amostragem muitas vezes relacionado ao tamanho do fragmento e ao local em que foi coletado. (“Protocolo Clínico e Diretrizes Terapêuticas para Hepatite C e Coinfecções”, 2015)

Além disso, existe a dificuldade no acesso ao serviço de saúde e aos procedimentos de média e alta complexidade, como a biópsia. Essa dificuldade é muito comum em locais remotos ou até mesmo devido à precariedade do sistema de saúde. Portanto, é importante que haja opções não invasivas de avaliação do grau de fibrose, principalmente por causa de pacientes que não preenchem os critérios necessários ou que apresentam contraindicações.

Quando a biópsia é considerada de inviável realização, seja por questões de custo-efetividade, indisponibilidade, falta de treinamento para o método ou ainda contraindicação clínica, são indicados, métodos não invasivos como a elastografia hepática ou os índices APRI (do inglês, *Aspartate aminotransferase Platelet Ratio Index*) e FIB-4. Os índices podem ser utilizados tanto para identificar a fibrose avançada e a cirrose, quanto para deferir o tratamento da infecção após avaliação médica, na ausência de métodos preferenciais com biópsia. (“Protocolo Clínico e Diretrizes Terapêuticas para Hepatite C e Coinfecções”, 2015)

O tratamento também é indicado para pacientes com sinais clínicos de cirrose hepática. Esses pacientes não necessitam de biópsia hepática ou outro método para iniciar o tratamento. (CALVARUSO; CRAXÌ, 2012; ORGANIZATION, 2015b)

A cirrose é responsável pela maioria das mortes relacionadas ao fígado. As causas mais importantes de cirrose no mundo todo são o abuso de álcool, a hepatite viral e EHNA. (KUMAR et al., 2006)

A cirrose hepática se caracteriza pela destruição das células parenquimatosas hepáticas, sendo substituídas por tecido fibroso, que, eventualmente, se contrai em torno dos vasos sanguíneos, impedindo assim o fluxo de sangue pelo fígado. Aproximadamente 40% dos indivíduos com cirrose são assintomáticos até um ponto tardio na evolução da doença. (KUMAR et al., 2006) Portanto, é necessário que existam exames para diagnosticar precocemente a cirrose.

## 4 DIAGNÓSTICO DE HEPATITES VIRAIS, B E C, E ESTEATOSE

Em 2009, nosso grupo demonstrou que a metabonômica era capaz de identificar portadores de HCV, com altas sensibilidade e especificidade, utilizando amostras de urina.(GODOY et al., 2010) Visto que a obtenção de amostras de urina é não-invasiva, decidimos investigar a capacidade dos modelos metabonômicos em identificar portadores de HBV, que se caracterizaria como um método prático e não invasivo de diagnóstico. Como tínhamos alguns pacientes com esteatose e também portadores de HBV ou HCV, decidimos investigar também se é possível discriminar amostras de urina de portadores de HBV ou HCV de não portadores, independentemente de terem esteatose ou não. Posteriormente, foram construídos modelos para identificar a presença de esteatose, independentemente de os indivíduos serem portadores de HBV ou HCV.

### 4.1 Objetivos Específicos

Construir modelos metabonômicos a partir da análise de RMN de  $^1\text{H}$  de amostras de urina com a finalidade de diagnosticar as hepatites virais B e C, e esteatose; assim como diferenciar amostras de portadores de hepatites virais, B ou C, de amostras de pacientes com esteatose.

### 4.2 Metodologia

#### 4.2.1 Amostras

Foram utilizadas amostras de urina de pacientes com diagnóstico de esteatose e hepatite crônica ou cirrose por HBV e HCV, do ambulatório de hepatologia do Hospital das Clínicas da UFPE, com idade acima de 18 anos, de ambos os sexos, e que concordaram em participar do estudo, após leitura e assinatura do termo de consentimento livre e esclarecido (TCLE). Pacientes com doença hepática esquistossomótica, esteatohepatite, síndrome de imunodeficiência adquirida, hepatite auto-imune, hepatopatia metabólica, neoplasia e alcoolismo (consumo de etanol > 210g/semana em homens e 140 g/semana em mulheres) foram excluídos do estudo.

Após a consulta, foram colhidos 50 mL de urina, em jejum, e coletados 10 mL de sangue, em veia periférica, no Laboratório Central do HC-UFPE, onde, em seguida, foi separado o soro.

Na amostra de soro, foram determinadas as dosagens dos níveis séricos de: alanina aminotransferase, aspartato aminotransferase, fosfatase alcalina, gama-glutamil transferase, bilirrubina, albumina, plaquetas e INR, nos setores de hematologia e bioquímica; e HBsAg, anti-HBs, anti-HBc e anti-HCV através de ensaio imunoenzimático (Abbott®), no setor de sorologia. A pesquisa do HCV-RNA e genoma dos pacientes com o anti-HCV positivo e do HBV-DNA em pacientes HBsAg positivos foi realizada através da técnica PCR, no Laboratório Central da Secretaria de Saúde de Pernambuco (LACEN-PE).

O diagnóstico de DHGNA foi feito por Ultrassonografia (US), realizada no Centro de Imagens (Diagimagem) do Hospital das Clínicas – UFPE, com o aparelho GE – *Healthcare*, Logic P6.

Uma alíquota de 2 mL de urina foi enviada à Central Analítica do Departamento de Química Fundamental (DQF) da UFPE para obtenção dos espectros de RMN de  $^1\text{H}$  e construção dos modelos metabonômicos.

#### 4.2.2 Aquisição dos espectros de RMN de $^1\text{H}$

Na Central Analítica do DQF - UFPE, obtivemos espectros de RMN de  $^1\text{H}$  de 400  $\mu\text{L}$  de urina em 200  $\mu\text{L}$  de solução-tampão ( $\text{Na}_2\text{HPO}_4$  /  $\text{NaH}_2\text{PO}_4$  – 0,2 mol  $\text{L}^{-1}$ ), utilizando um aparelho de RMN VARIAN *Unity Plus 300*.

Utilizamos a sequência de pulsos PRESAT, para supressão do sinal da água. Foram utilizadas 128 repetições, line broadening (lb) igual a 0,3 Hz e tempo de espera, *delay*, igual a 2 segundos.

#### 4.2.3 Processamento dos Espectros de RMN $^1\text{H}$

Com auxílio do software *MestreNova*, realizamos o processamento dos espectros:

- ☐ Correção manual de linha de base e fase;
- ☐ Referência de deslocamento químico, utilizando o simpleto atribuído ao grupo metil da creatinina em 3,06ppm;

- Divisão dos espectros, de 0 a 10ppm, em bins de 0,05ppm, formando uma matriz linha para cada amostra, excluindo a região onde se encontram os sinais atribuídos à água e à ureia (intervalo entre 4,2 e 6,4ppm).

Para cada investigação, agrupamos as matrizes linhas das amostras em uma única matriz, contendo as amostras (casos) nas linhas e os bins nas colunas (variáveis), e adicionamos uma coluna com a classificação das amostras (variável de classe). Em seguida, submetemos as matrizes às análises quimiométricas.

#### 4.2.4 Análise Estatística Multivariada

As análises de componentes principais foram realizadas utilizando a plataforma online *MetaboAnalyst* ou o *software Statistica*, para observar se as amostras se agrupavam nas classes de interesse; para detecção de amostras anômalas; e, em alguns casos, como método de redução de dimensionalidade.

Com o objetivo de classificar as amostras, construímos modelos de PLS-DA e/ou LDA, na plataforma *MetaboAnalyst* ou no *software Statistica*, respectivamente. Para a construção dos modelos LDA, utilizamos o Lambda de Wilks como método de seleção de variáveis.

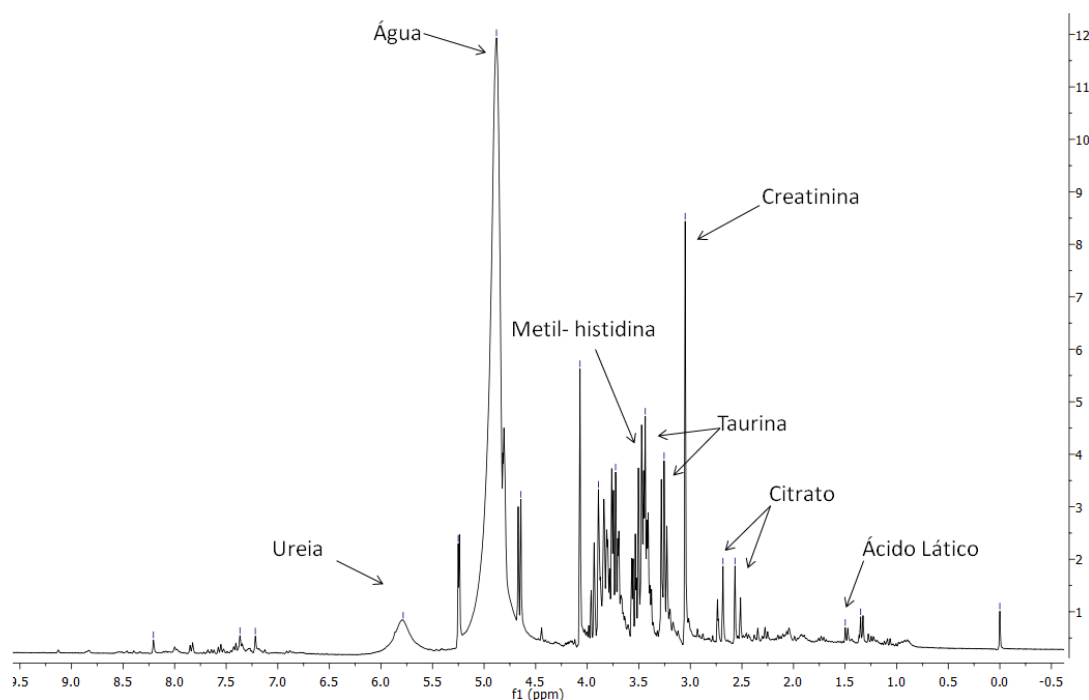
Como o número de amostras é reduzido, devido à dificuldade de obtenção das amostras de biofluidos e perda de material, tanto por falta de exames essenciais para o diagnóstico quanto por falhas no processo de coleta das amostras, armazenamento e encaminhamento para a análise de RMN  $^1\text{H}$ , não temos amostras suficientes para uma validação externa, portanto, utilizamos a LOOCV.

### 4.3 Resultados e Discussão

Os resultados aqui apresentados são os modelos construídos com amostras de urina para o diagnóstico de hepatites virais e esteatose. O mérito desta abordagem está associado ao fato de, a partir de uma única análise espectroscópica (RMN de  $^1\text{H}$ ), obter informações relativas à presença - ou não - de esteatose e infecção viral por HBV ou HCV.

Na Figura 16 podemos ver um exemplo de espectro de urina, com sinais de alguns metabólitos indicados.

Figura 16: Espectro de RMN de  $^1\text{H}$  ( $\text{Na}_2\text{HPO}_4 / \text{NaH}_2\text{PO}_4 - 0,2 \text{ mol L}^{-1}$ , 300 MHz) de uma amostra de urina.



#### 4.3.1 HBV *versus* Controle - puros

A matriz de dados continha os espectros de RMN de  $^1\text{H}$  de urina de 28 portadores de HBV, foram excluídos desta análise indivíduos que apresentavam esteatose, e 10 voluntários saudáveis, totalizando 38 amostras e 155 variáveis.

Foram utilizados os pré-processamentos SNV e normalização pela soma, nas amostras, e autoescalamento nas variáveis.

A PCA não demonstrou agrupamento das amostras nas classes de interesse, também não foram identificadas amostras anômalas. Mas foi utilizada como método de redução de variáveis, para a construção de um modelo de classificação com LDA. Os modelos foram construídos com os escores de cinco componentes principais, escolhidas pelo método de seleção de variáveis Lambda de Wilks. A PLS-DA não conseguiu construir modelos de classificação com resultados satisfatórios.

O modelo que apresentou a melhor performance foi utilizando SNV como pré-processamento, com 92,9% de sensibilidade na LOOCV. Na Tabela 4, estão apresentados os resultados da validação do modelo.

Tabela 4: Resultados de modelo LDA para predição de HBV utilizando as componentes principais

		HBV	Controle	Sensibilidade	Especificidade	VPP	VPN	Exatidão
				(%)	(%)	(%)	(%)	(%)
Modelo Metabonômico	HBV	26	2	92,9	80,0	92,9	80,0	89,5
	Controle	2	8					

VPP: Valor preditivo positivo, VPN: valor preditivo negativo.

Este resultado demonstra que a estratégia metabonômica tem potencial para identificar portadores de HBV de voluntários saudáveis, a partir de amostras de urina.

#### 4.3.2 HBV *versus* Controle – independentemente de esteatose

Nesta análise, consideramos pacientes que apresentaram esteatose, tanto no grupo HBV, quanto no grupo controle. À matriz de dados utilizada no modelo só com pacientes HBV puros (sem esteatose), foram acrescentados espectros de RMN de  $^1\text{H}$  de urina de: quatro portadores de HBV e oito voluntários saudáveis, totalizando 32 amostras no grupo doença e 18 no grupo controle.

Testamos normalização pela soma, SNV e autoescalonamento, como pré-processamentos para a matriz de dados originais.

Não observamos agrupamento nas classes de interesse, nem amostras anômalas, na PCA. Os modelos PLS-DA construídos não conseguiram classificar satisfatoriamente as amostras.

Utilizamos a PCA como método de redução de dimensionalidade para a construção de modelos LDA com os escores das cinco componentes principais que melhor discriminaram os grupos, escolhidas por Lambda de Wilks.

O modelo de classificação que demonstrou o melhor desempenho, foi construído com LDA a partir dos escores das componentes principais obtidas pela PCA da matriz de dados pré-processada com SNV. Os resultados da LOOCV estão dispostos na Tabela 5.

O modelo foi capaz de classificar corretamente todos os pacientes com HBV. Apesar da especificidade não ter sido muito alta, o resultado é favorável para uma técnica de *screening*, muito útil na ideia de com uma única análise, se investigar várias doenças. Principalmente pela sensibilidade ter sido 100%, pode ser usada como uma avaliação inicial da infecção por HBV, assim como a avaliação de outras doenças com um único



exame, e no caso de resultado positivo, ser feito outro exame, mais invasivo, para confirmação.

Tabela 5: Resultados de modelo LDA para predição de HBV, independentemente da presença de esteatose.

Modelo			Sensibilidade	Especificidade	VPP	VPN	Exatidão
			(%)	(%)	(%)	(%)	(%)
	HBV	Controle					
Metabonômico	HBV	32	4	100	77,8	88,9	100
	Controle	0	14				
							92,0

VPP: Valor preditivo positivo, VPN: valor preditivo negativo.

#### 4.3.3 HCV *versus* Controle – independentemente de esteatose

A análise incluiu amostras de urina de 27 portadores de HCV, grupo HCV, e de 18 voluntários saudáveis, grupo controle, ou seja, 45 casos e 155 variáveis.

Aplicamos SNV, normalização pela soma e autoescalonamento à matriz original, como pré-processamento.

Não observamos agrupamentos das amostras nos grupos HCV e controle, nem amostras anômalas na PCA. A PLS-DA não foi capaz de construir modelos que classificassem satisfatoriamente as amostras. Os escores das componentes principais foram usados para construção de modelos de classificação LDA, e o melhor modelo encontrado foi com cinco componentes principais, escolhidas por Lambda de Wilks, utilizando normalização pela soma como pré-processamento. Após a validação, por LOOCV, o modelo apresentou 92,6% de sensibilidade. Mais detalhes do resultado são apresentados na Tabela 6.

Tabela 6: Resultados de modelo LDA para predição de HCV, independentemente da presença de esteatose.

Modelo			Sensibilidade	Especificidade	VPP	VPN	Exatidão
			(%)	(%)	(%)	(%)	(%)
	HCV	Controle					
Metabonômico	HCV	25	4	92,6	77,8	86,2	87,5
	Controle	2	14				
							86,7

VPP: Valor preditivo positivo, VPN: valor preditivo negativo.

#### 4.3.4 Esteatose *versus* Controle - puros

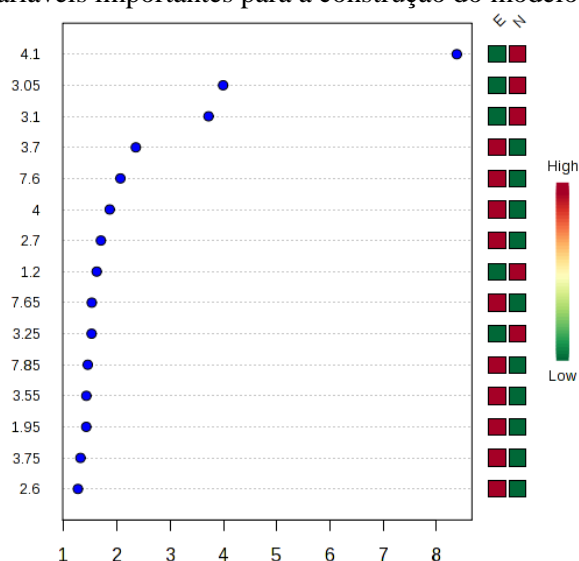
A matriz de dados continha os espectros de RMN de  $^1\text{H}$  de urina de 10 voluntários saudáveis e 20 pacientes com esteatose, foram excluídos desta análise indivíduos com infecção por HBV ou HCV, totalizando 30 amostras e 155 variáveis. Para corrigir a diferença de concentração das amostras, foi feita a normalização pela soma.

Na PCA não foi observada separação entre as classes de interesse e também não foram identificadas amostras anômalas.

Foi construído um modelo de classificação PLS-DA. O melhor resultado utilizou três variáveis latentes, porém não foi satisfatório, pois a exatidão foi de 73,33%. Apesar disso, achamos importante avaliar o gráfico das variáveis mais influentes para a construção do modelo, gráfico VIP, na Figura 17, pois mesmo a quantidade de acertos não sendo suficiente para um teste de diagnóstico, ele ainda conseguiu acertar mais de 73% das amostras, portanto, os bins indicados, devem estar relacionados à informação química em comum entre amostras de mesma classe.

No eixo das ordenadas estão os deslocamentos químicos e no eixo das abscissas está o peso correspondente. Na escala de cores ao lado, podemos observar em que classe aquela região é mais intensa, por exemplo: 4,1ppm, 3,05ppm e 3,1ppm têm sinais mais intensos nas amostras do grupo controle; enquanto 3,7ppm, 7,6ppm e 4,0ppm, são mais intensos no grupo esteatose.

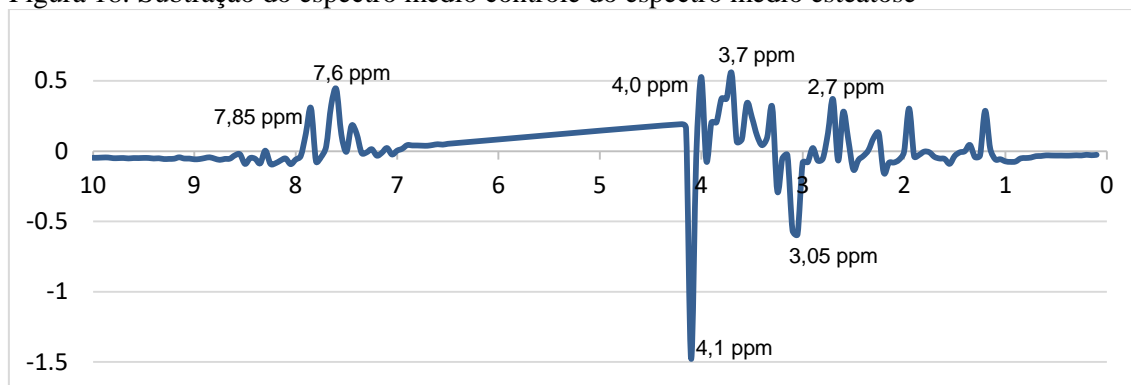
Figura 17: Gráfico de variáveis importantes para a construção do modelo PLS-DA.



Fonte: *MetaboAnalyst*

Paralelamente, obtivemos o espectro médio das amostras de cada classe, após aplicação de SNV, e então subtraímos a média do grupo controle da média do grupo esteatose, como exposto na Figura 18. É possível perceber que o resultado concorda com o observado no gráfico VIP do PLS-DA: 7,6ppm, 7,65ppm 7,85ppm, 4,0ppm, 3,7ppm, 2,7ppm, 2,6ppm e 1,95ppm são mais intensos no grupo esteatose; enquanto 4,1ppm, 3,25ppm, 3,1ppm e 3,05ppm são mais intensos no grupo controle.

Figura 18: Subtração do espectro médio controle do espectro médio esteatose



Fonte: Autoria própria.

Estas regiões foram selecionadas e usadas como variáveis para a construção de um modelo LDA usando cinco variáveis, selecionadas usando o lambda de Wilks. Aplicamos LOOCV ao modelo construído, que classificou incorretamente três amostras do grupo controle e uma do grupo esteatose, como pode ser visto na Tabela 7.

Tabela 7: Resultados de modelo LDA para predição de esteatose usando deslocamentos da subtração de espectros médios

		Esteatose	Controle	Sensibilidade	Especificidade	VPP	VPN	Exatidão
				(%)	(%)	(%)	(%)	(%)
Modelo Metabonômico	Esteatose	19	3	95,5	70,0	86,4	87,5	86,7
	Controle	1	7					

VPP: Valor preditivo positivo, VPN: valor preditivo negativo.

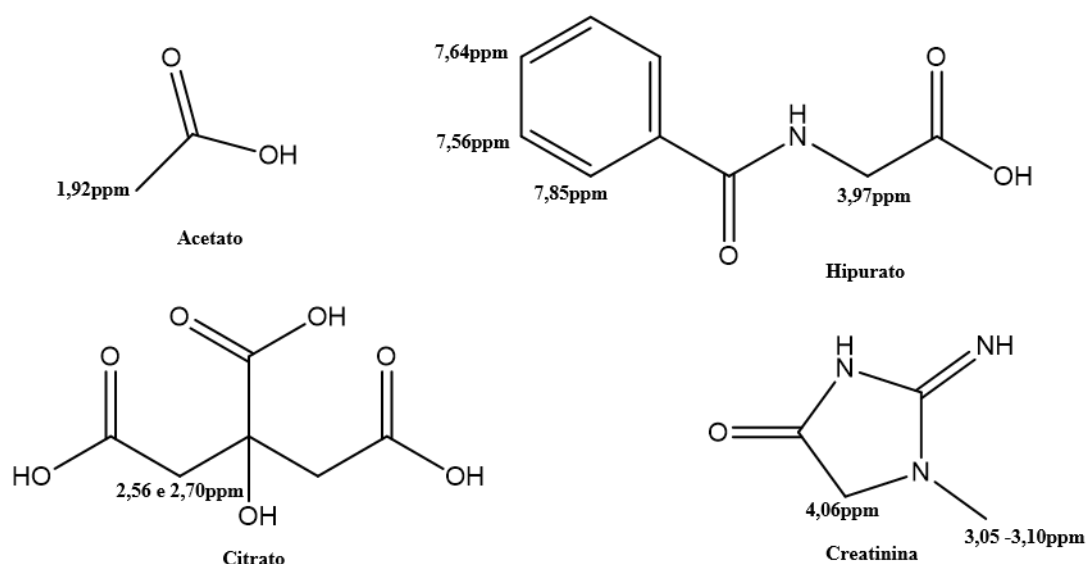
Foi realizada a atribuição dos sinais das regiões usadas no LDA, resumidas na Tabela 8. As estruturas dos metabólitos e respectivos sinais identificados estão dispostos na Figura 19. (MIAO et al., 2015)

Tabela 8: Identificação dos principais metabólitos discriminantes, bins correspondentes e níveis urinários relativos em cada grupo.

Metabólitos	$\delta$ (ppm)	Esteatose	Controle
Acetato	1,95	+	–
Citrato	2,60 e 2,70	+	–
Creatinina	3,05 - 3,10 e 4,10	–	+
Hipurato	4,00, 7,60, 7,65 e 7,85	+	–

+ mais concentrado no grupo; – menos concentrado no grupo;  $\delta$  deslocamento químico.

Figura 19: Estruturas dos metabólitos identificados e deslocamentos químicos no espectro de RMN de  $^1\text{H}$ .



Fonte: Autoria própria.

Os sinais em 3,05ppm e 4,10ppm são atribuídos à creatinina (MIAO et al., 2015), e estão mais intensos no grupo controle, o que indica que a creatinina tem uma concentração menor na urina dos pacientes com esteatose.

Existem trabalhos que relatam alterações nos níveis séricos de creatinina na presença de doenças hepáticas: Soga e colaboradores identificaram alterações nos níveis séricos de creatinina quando investigados pacientes com infecção por HBV ou HCV, lesão hepática induzida por drogas, cirrose, hepatocarcinoma, esteato-hepatite não alcoólica e esteatose simples (SOGA et al., 2011); Chen e colaboradores observaram redução nos níveis séricos de creatinina de pacientes com hepatocarcinoma (CHEN et al., 2011); e, Cocchetto e colaboradores, relataram redução na produção de creatinina em pacientes com doença hepática. (COCCHETTO; TSCHANZ; BJORNSSON, 1983)

Os sinais em 4,00, 7,65, 7,60 e 7,85ppm, atribuídos ao hipurato, estão mais intensos nas amostras de urina do grupo esteatose. Há trabalhos que observaram que pacientes com diabetes têm níveis urinários maiores de hipurato, associando ao metabolismo alterado da glicose (CALVANI et al., 2010), e de acetato, que também se

apresenta mais intenso nos pacientes com esteatose.(MESSANA et al., 1998) Visto isso, observamos que os 20 pacientes com esteatose do nosso estudo têm níveis séricos de glicose maiores que os do grupo controle (esteatose:  $120,15 \pm 40,7$  mg/dL e controle:  $72,68 \pm 13,14$  mg/dL,  $p < 0,001$ ), portanto, os maiores níveis de hipurato e acetato encontrados podem ser associados ao metabolismo alterado da glicose.

Os sinais em 2,60 e 2,70ppm, são atribuídos ao citrato, mais intensos no grupo esteatose. O grupo esteatose apresentou um nível sérico de triglicerídeos maior que o grupo controle (esteatose:  $231,64 \pm 221,64$  mg/dL e controle:  $104,20 \pm 48,22$  mg/dL,  $p = 0,008$ ). Podemos associar a maior concentração de citrato à de triglicerídeos, já que este é metabolizado formando glicerol, que por sua vez forma piruvato, que é convertido a citrato no ciclo de Krebs. (KIM et al., 2009)

Com o intuito de buscar um modelo de classificação que tivesse uma performance melhor, usamos a PCA como método de redução de dimensionalidade dos dados, pré-processados com SNV, para construir um modelo LDA. O modelo foi construído com as cinco componentes principais que melhor separavam as classes (1, 3, 5, 6 e 8), escolhidas usando Lambda de Wilks, e alcançou 100% de sensibilidade e de valor preditivo negativo, após a aplicação da LOOCV. Este resultado está descrito na Tabela 9.

Tabela 9: Resultados de modelo LDA para predição de esteatose usando componentes principais.

		Esteatose	Controle	Sensibilidade	Especificidade	VPP	VPN	Exatidão
				(%)	(%)	(%)	(%)	(%)
Modelo Metabonômico	Esteatose	20	2	100	80,0	90,9	100	93,3
	Controle	0	8					

VPP: Valor preditivo positivo, VPN: valor preditivo negativo.

Combinando os pesos das componentes principais utilizadas na construção do modelo e os modelo e os coeficientes da função canônica discriminante, encontramos as regiões do espectro espectro mais importantes para a discriminação (

Figura 20). Comparando com as regiões que aparecem no VIP do PLS-DA e na subtração dos espectros médios, podemos observar que há mais regiões consideradas como importante, por isso o modelo teve um maior número de acertos.

Há duas opções a se considerar: há outros metabólitos capazes de discriminar os grupos esteatose e controle, que não foram considerados no modelo anterior; ou o

segundo modelo está superajustado aos dados, incluindo informações relacionadas a amostras individuais, fazendo com que sejam corretamente classificadas.

Devido ao pequeno número de amostras, é muito provável que a PCA esteja incluindo informações de amostras individuais, que estão sendo posteriormente utilizadas pelo LDA, já que este busca a combinação de componentes que apresente a melhor performance, ou seja, classifique corretamente um maior número de amostras. Já na subtração dos espectros médios, é mais difícil ser incluída informação de amostras individuais, portanto, acreditamos que a segunda opção é a mais provável.

Observamos que havia diferença de média de idade e sexo entre os grupos. A média de idade do grupo doença era de 55, enquanto do grupo controle era 37 anos. No grupo esteatose, 30% era do sexo masculino, e 80% no grupo de voluntários saudáveis. Levantamos, então, a hipótese da LDA estar usando informações de sexo e idade, além da doença, para separar os grupos.

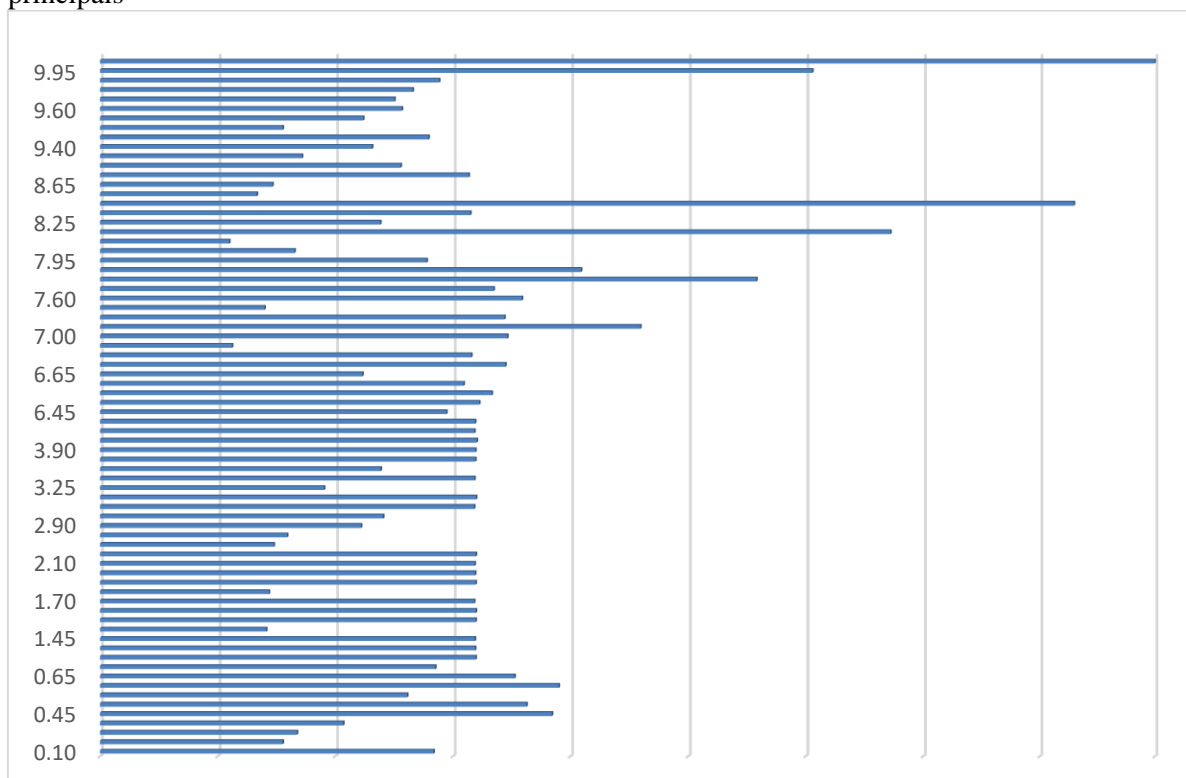
Para avaliar esta hipótese, inicialmente, fizemos os gráficos de escores das componentes principais usadas, indicando as amostras de acordo com a idade e com o sexo, separadamente. Não foi observada separação das amostras por sexo ou idade, nas componentes utilizadas. Em seguida, construímos o gráfico dos escores das amostras no modelo LDA construído, em função de cada uma das classes, sexo e idade, para identificar se tinha algum ponto de corte que separasse as amostras.

Na Figura 21, podemos observar que as amostras de indivíduos do sexo feminino (valor zero no eixo das abscissas) estão distribuídas de 0 a 1,04 nos escores do LDA (eixo das ordenadas), assim como as amostras do sexo masculino (valor um no eixo das abscissas), estão distribuídas de 0 a 1,3, não podendo ser separadas por essa função discriminante linear em nenhum ponto de corte.

Na Figura 22, podemos observar que as amostras das diferentes idades estão aleatoriamente distribuídas nos escores de 0 a 1,3, não podendo ser separadas por essa função discriminante.

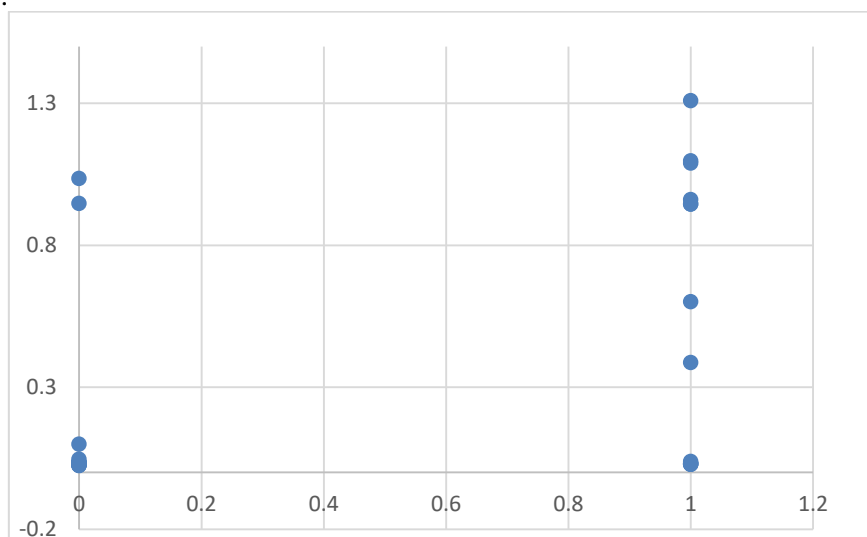
Conseguimos construir um modelo capaz de identificar satisfatoriamente pacientes com esteatose e voluntários saudáveis, assim como identificar os metabólitos envolvidos. Porém, muito dos pacientes com esteatose são portadores de hepatites B ou C, decidimos então construir um modelo para separar indivíduos com esteatose de indivíduos sem esteatose, independentemente de serem portadores de HBV ou HCV.

Figura 20: Variáveis importantes para o modelo LDA construído a partir das componentes principais



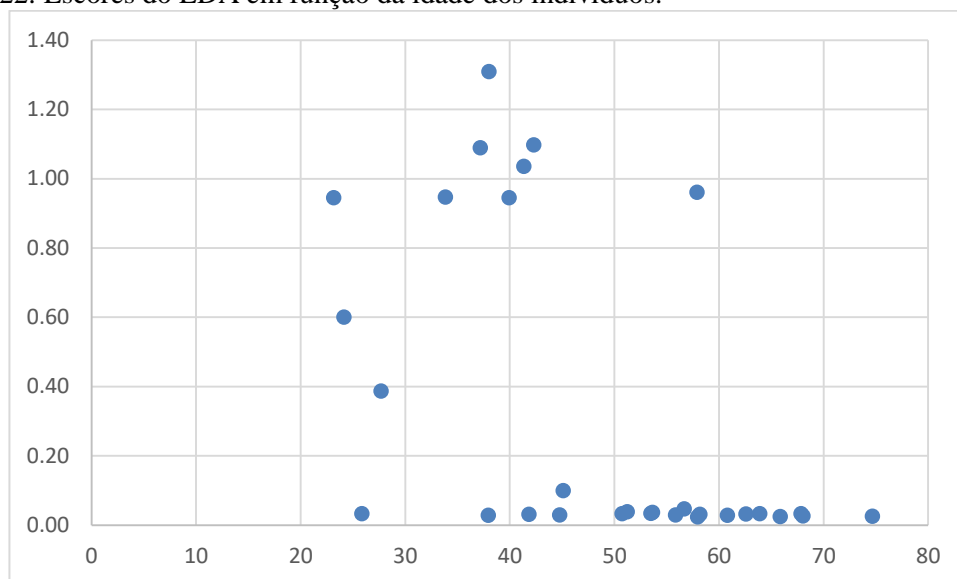
Fonte: Autoria própria.

Figura 21: Escores do LDA em função do sexo dos indivíduos, no qual 0 indica feminino e 1 masculino.



Fonte: Autoria própria.

Figura 22: Escores do LDA em função da idade dos indivíduos.



Fonte: Autoria própria.

#### 4.3.5 Esteatose *versus* Controle – independentemente de vírus

O grupo esteatose incluía 40 pacientes com esteatose, portadores ou não de HBV ou HCV, e o grupo controle incluía 57 indivíduos sem esteatose, portadores ou não de HBV ou HCV. A matriz de dados originais continha o espectro de RMN de  $^1\text{H}$  de amostras de urina dos 97 casos e 155 variáveis.

Foram testados os pré-processamentos: normalização pela soma e SNV, nas linhas; e autoescalamento nas colunas. Em nenhum dos casos observamos separação das amostras nas classes esteatose e controle, na PCA, e não obtivemos resultados satisfatórios construindo modelos de classificação PLS-DA, que alcançou, no máximo, 61% de exatidão.

Foi então usada a PCA como método de redução de dimensionalidade da matriz de dados, pré-processada com SNV, para aplicação da LDA. O modelo foi construído com as três componentes principais que melhor discriminaram as amostras de diferentes classes, 5, 6 e 7, escolhidas utilizando Lambda de Wilks. Aplicada a LOOCV, o modelo apresentou 100% de sensibilidade e de valor preditivo negativo, pois classificou corretamente todas as amostras de pacientes com esteatose. Os resultados estão descritos na Tabela 10.

Mostramos, então que os modelos metabonômicos são capazes de classificar pacientes com e sem esteatose, independentemente da presença outras lesões no fígado.



Tabela 10: Resultados de modelo LDA para predição de esteatose usando componentes principais.

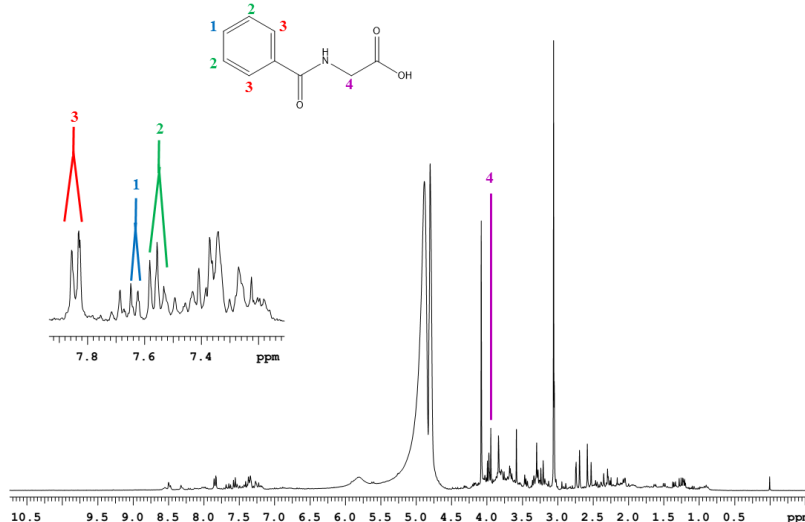
Modelo Metabonômico		Esteatose	Controle	Sensibilidade	Especificidade	VPP	VPN	Exatidão
				(%)	(%)	(%)	(%)	(%)
	<b>Esteatose</b>	40	2	100	96,5	95,2	100	97,9
	<b>Controle</b>	0	55					

VPP: Valor preditivo positivo, VPN: valor preditivo negativo.

É de se esperar que alguns metabólitos sejam diferentes dos encontrados no modelo anterior, pois parte deles indicava lesão no fígado, independentemente de ser por esteatose ou hepatites virais. Neste modelo, os metabólitos característicos de lesões no fígado sem levar em conta a sua causa não devem ser tidos como importantes, ganhando destaque os relacionados diretamente à agressão específica pela esteatose. Então avaliamos, neste modelo, o peso dos bins relacionados aos metabólitos destacados no modelo anterior.

Dos cinco metabólitos anteriormente identificados, somente os bins associados aos sinais do hipurato foram importantes para a construção do modelo de classificação de pacientes com esteatose, independentemente de serem portadores de HBV ou HCV. Na Figura 23 estão identificados os sinais correspondentes aos hidrogênios do hipurato em um espectro de urina.

Figura 23: Sinais do hipurato no espectro de RMN de  $^1\text{H}$  de amostra de urina.

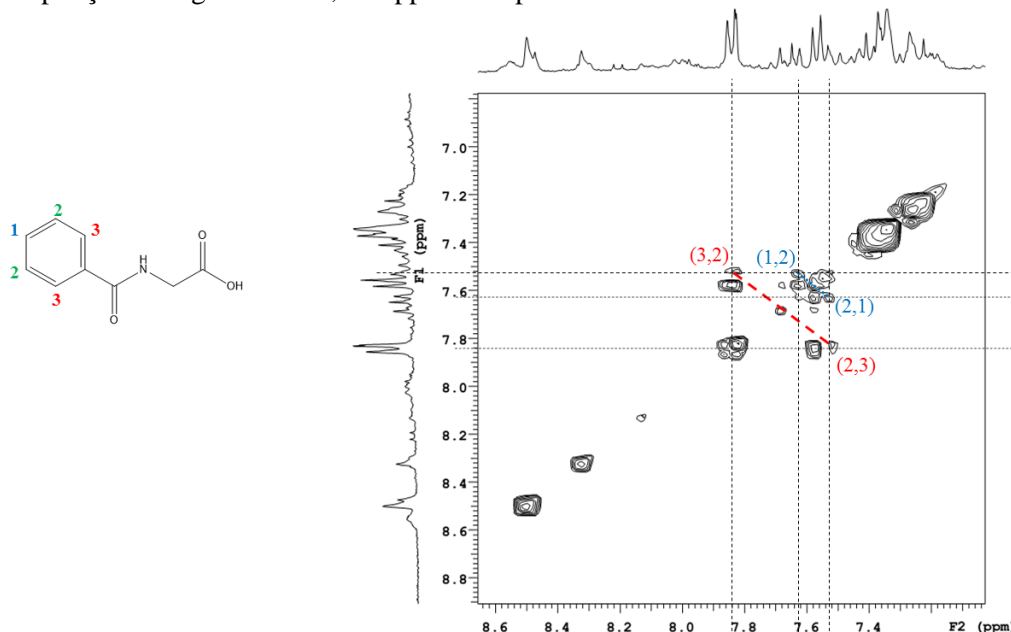


Fonte: Autoria própria.

Para confirmar que esses sinais são do hipurato, fizemos um experimento COSY, que identifica as correlações entre hidrogênios acoplados por  $^3J_{\text{HH}}$ , já que nesse

metabólito há hidrogênios que acoplam a três ligações. Conseguimos observar as correlações esperadas entre os hidrogênios 1 e 2, e entre os hidrogênios 2 e 3, como mostrado na Figura 24.

Figura 24: Ampliação da região entre 6,8 e 9ppm do espectro COSY de uma amostra de urina.



Fonte: Autoria própria.

Portanto, podemos inferir que a concentração relativa do hipurato entre os grupos, maior no grupo de doentes, teve um papel importante na identificação de pacientes com esteatose. Além disso, devido à presença de portadores de HBV ou HCV em ambos os grupos, acreditamos que o hipurato esteja diretamente relacionado à esteatose, e não a doenças hepáticas no geral.

A partir disso, surgiu o questionamento se modelos metabonômicos seriam capazes de diferenciar esteatose de hepatites virais.

#### 4.3.6 Esteatose *versus* Hepatites Virais

A matriz de dados foi construída com espectros de RMN de  $^1\text{H}$  de amostras de urina de 20 pacientes com esteatose e 47 portadores de HBV ou HCV, totalizando 155 variáveis e 67 casos. Foram testados os pré-processamentos SNV e normalização pela soma amostras, e autoescalonamento nas variáveis.

Na PCA não foram observados agrupamentos das amostras nas classes esteatose e hepatites virais, e também não foram identificadas amostras anômalas. A PLS-DA não obteve resultados satisfatórios com nenhum dos pré-processamentos testados.

Foram construídos modelos LDA a partir da seleção de componentes principais construídas pela PCA a partir das matrizes com os pré-processamentos citados. O modelo LDA construído com cinco componentes principais, selecionadas pelo método do Lambda de Wilks, a partir da matriz de dados normalizada pela soma e autoescalada, obteve 94% de exatidão. Analisando a Tabela 11, constatamos que a estratégia metabonômica é capaz de diferenciar as lesões hepáticas consequente de acúmulo de gordura no fígado das causadas por infecção por vírus HBV ou HCV.

Tabela 11: Resultados de modelo LDA para discriminação entre esteatose e hepatites virais

		<b>Esteatose</b>	<b>Hepatites Virais</b>	<b>Sensibilidade</b>	<b>Especificidade</b>	<b>VPP</b>	<b>VPN</b>	<b>Exatidão</b>
				(%)	(%)	(%)	(%)	(%)
<b>Modelo Metabonômico</b>	<b>Esteatose</b>	18	2	90,0	95,7	90,0	95,7	94,0
	<b>Hepatites Virais</b>	2	45					

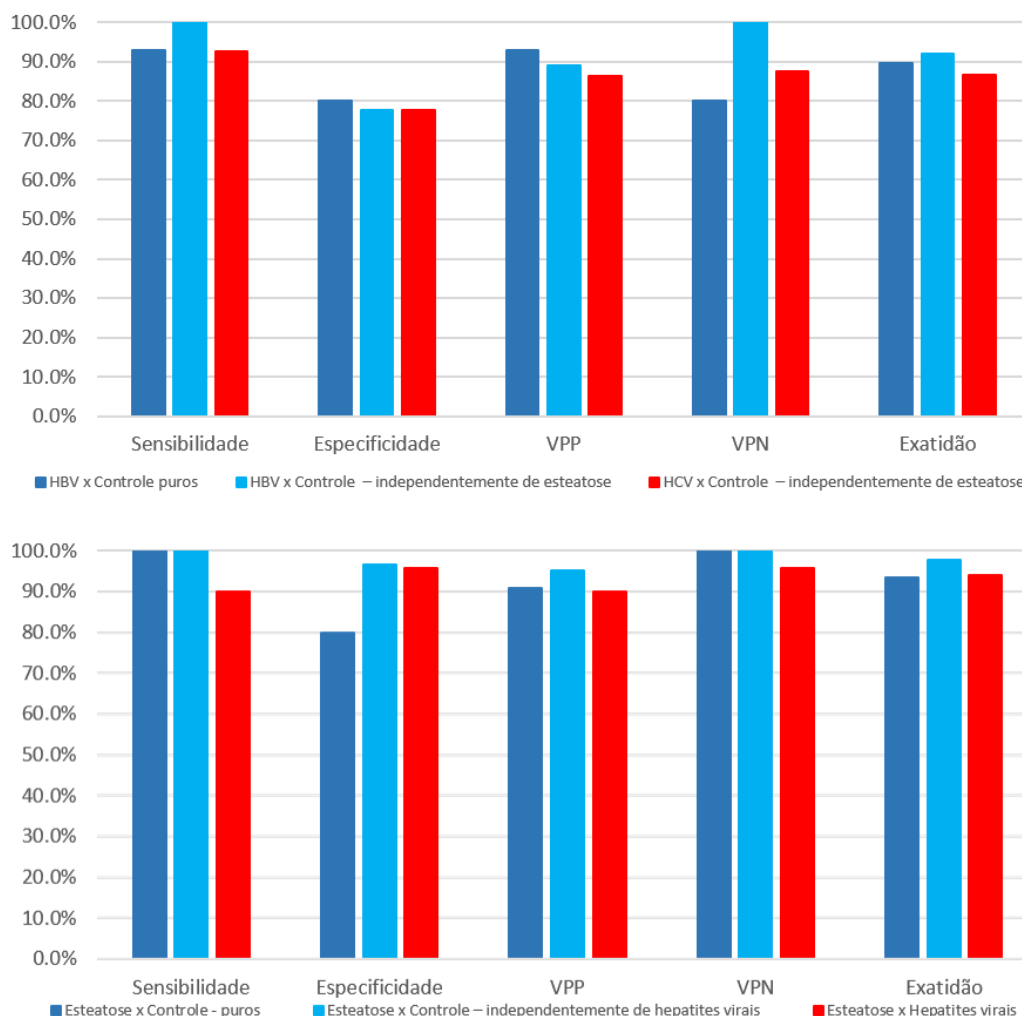
VPP: Valor preditivo positivo, VPN: valor preditivo negativo.

#### 4.4 Conclusão

Observando a Figura 25, percebemos que estratégia metabonômica aplicada a amostras de urina obteve excelentes resultados para, em um único exame: identificar se indivíduos são portadores de HBV e HCV, independentemente de terem outro tipo de lesão no fígado, como esteatose; reconhecer pacientes com esteatose, tanto com a presença de hepatites virais HBV ou HCV, quanto em pacientes sem hepatites virais; e diferenciar pacientes com esteatose de pacientes com hepatites virais dos tipos B ou C.

Podemos então inferir que a estratégia metabonômica demonstrou capacidade de avaliar diferentes doenças hepáticas em uma única análise, identificando simultaneamente a presença de esteatose e dos vírus HBV e HCV, caracterizando-se como uma eficaz técnica de *screening* não invasiva para doenças no fígado.

Figura 25: Resumo dos resultados de diagnóstico de hepatites virais e esteatose.



Fonte: Autoria própria.

#### 4.5 Perspectivas

- ☐ Obter um número maior de amostras para fazer validação externa dos modelos para discriminação entre pacientes com esteatose e voluntários saudáveis;
- ☐ A partir do resultado da validação externa, decidir se o modelo está sobreajustado;
- ☐ Fazer um estudo com esteato-hepatite;
- ☐ Determinar quais são os metabólitos responsáveis pela classificação e associar às rotas metabólicas do organismo.

## 5 ESTADIAMENTO DE FIBROSE HEPÁTICA EM PACIENTES COM HCV

Baseados na necessidade da prática clínica, construímos modelos para fibrose significativa, avançada e cirrose. Comparamos a performance dos modelos metabonômicos a testes de diagnóstico não invasivos que são recomendados, pela Organização Mundial da Saúde e Ministério da Saúde, como alternativas para identificação do grau de fibrose, APRI e FIB-4. (“Protocolo Clínico e Diretrizes Terapêuticas para Hepatite C e Coinfecções”, 2015)

### 5.1 Objetivos Específicos

Construir modelos metabonômicos, a partir da análise de RMN de  $^1\text{H}$  de amostras de soro de pacientes com hepatite crônica por HCV, com a finalidade de avaliar o grau de fibrose hepática.

### 5.2 Metodologia

#### 5.2.1 Amostras

Foram utilizadas amostras de soro de pacientes com diagnóstico de hepatite crônica ou cirrose por HCV, do ambulatório de hepatologia do Hospital das Clínicas da UFPE, com idade acima de 18 anos, de ambos os sexos, e que concordaram em participar do estudo, após leitura e assinatura do termo de consentimento livre e esclarecido (TCLE). Pacientes com doença hepática esquistossomótica, esteatohepatite, síndrome de imunodeficiência adquirida, hepatite auto-imune, hepatopatia metabólica, neoplasia e alcoolismo (consumo de etanol  $> 210\text{g/semana}$  em homens e  $140\text{ g/semana}$  em mulheres) foram excluídos do estudo.

Após a consulta, foram coletados 10 mL de sangue, em veia periférica, no Laboratório Central do HC-UFPE, onde, em seguida, foi separado o soro, que foi dividido em duas alíquotas. Na primeira alíquota de soro foram realizadas as análises já descritas no item 4.2.1 do Capítulo 2.

Os pacientes cirróticos foram caracterizados pelo quadro clínico e por alterações sugestivas de doença hepática crônica e/ou hipertensão portal à US, como: textura hepática heterogênea; ecogenicidade aumentada; diminuição do tamanho do fígado;

aumento do calibre da veia porta; presença de circulação colateral e/ou esplenomegalia; e alterações sugestivas de hipertensão portal à endoscopia digestiva alta (EDA), como a presença de varizes esofageanas, varizes gástricas e/ou gastropatia hipertensiva.

O grau de fibrose hepática em pacientes com hepatites crônicas por HCV foi determinado por biópsias hepáticas, realizadas no Serviço de Gastroenterologia do Hospital das Clínicas/UFPE, conforme indicação clínica. Os fragmentos de tecido hepático foram processados, analisados e classificados utilizando o sistema METAVIR, no Departamento de Patologia Clínica do HC/UFPE, por um patologista com experiência no diagnóstico das doenças hepáticas.

A segunda alíquota de soro, das 63 amostras obtidas, foi enviada para a Central Analítica do Departamento de Química Fundamental (DQF) da UFPE para obtenção dos espectros de RMN de  $^1\text{H}$  e construção dos modelos metabonômicos.

#### 5.2.2 Aquisição dos espectros de RMN de $^1\text{H}$

Na Central Analítica do DQF - UFPE, obtivemos espectros de RMN de  $^1\text{H}$  de 400  $\mu\text{L}$  de soro sanguíneo diluído em 200  $\mu\text{L}$  de água deuterada utilizando um aparelho de RMN VARIAN *Unity Plus 300*.

Utilizamos as sequências de pulsos: PRESAT, para supressão do sinal da água; e CPMG, filtro de Tempo de Relaxação Transversal ( $T_2$ ) para eliminar os sinais de moléculas de alto peso molecular, com o objetivo de otimizar os tratamentos quimiométricos. Foram utilizadas 128 repetições, line broadening (lb) igual a 0,3 Hz e tempo de espera, *delay*, igual a 2 segundos. A sequência de refocagem foi repetida 88 vezes, com o valor de  $\tau$  igual a 361,3  $\mu\text{s}$ .

#### 5.2.3 Processamento dos Espectros de RMN $^1\text{H}$

Com auxílio do software *MestreNova*, realizamos o processamento dos espectros:

- ☐ Correção manual de linha de base e fase;
- ☐ Referência com um sinal conhecido, duplete atribuído ao grupo metil do lactato em 1,33ppm;
- ☐ Divisão dos espectros de 0 a 8ppm em bins de 0,04ppm, formando uma matriz linha para cada amostra, excluindo a região onde se encontra o sinal atribuído à água, entre 4,5 e 5,2ppm.

Para cada investigação, agrupamos as matrizes linhas das amostras em uma única matriz, contendo as amostras (casos) nas linhas e os bins nas colunas (variáveis), e adicionamos uma coluna com a classificação das amostras (variável de classe). Em seguida, submetemos as matrizes às análises quimiométricas.

#### 5.2.4 Análise Estatística Multivariada

Fizemos PCA das amostras, na plataforma *MetaboAnalyst* ou no *software Statistica*, para observar se as amostras se agrupavam nas classes de interesse, para detecção de amostras anômalas e, em alguns casos, como método de redução de variáveis.

Com o objetivo de construir modelos de classificação das amostras, construímos modelos de PLS-DA e/ou LDA, na plataforma *MetaboAnalyst* ou no *software Statistica*, respectivamente. Para a construção dos modelos LDA, utilizamos o Lambda de Wilks como método de seleção de variáveis.

Como o número de amostras é reduzido, devido à dificuldade de obtenção das amostras de biofluidos e perda de material, tanto por falta de exames essenciais para o diagnóstico quanto por falhas no processo de coleta das amostras, armazenamento e encaminhamento para a análise de RMN  $^1\text{H}$ , não temos amostras suficientes para uma validação externa, portanto, utilizamos a LOOCV.

#### 5.2.5 APRI e FIB-4

O APRI é calculado a partir da Equação 18. (WAI et al., 2003)

Equação 18

$$APRI = \frac{AST}{Plaquetas} \times 100$$

E o FIB-4 é calculado de acordo com a Equação 19. (STERLING et al., 2006)

Equação 19

$$FIB - 4 = \frac{idade \times AST}{Plaquetas \times \sqrt[2]{ALT}}$$

Sendo, AST, nível sérico de aspartato aminotransferases em unidades normalizadas por litro, Plaquetas em  $10^9$  unidades por litro, e ALT, nível sérico de alanina aminotransferases, em unidades normalizadas por litro.

Todos os resultados de APRI e FIB-4 têm um paciente a menos em relação aos resultados dos modelos metabonômicos, apresentando 62 amostras analisadas, pois



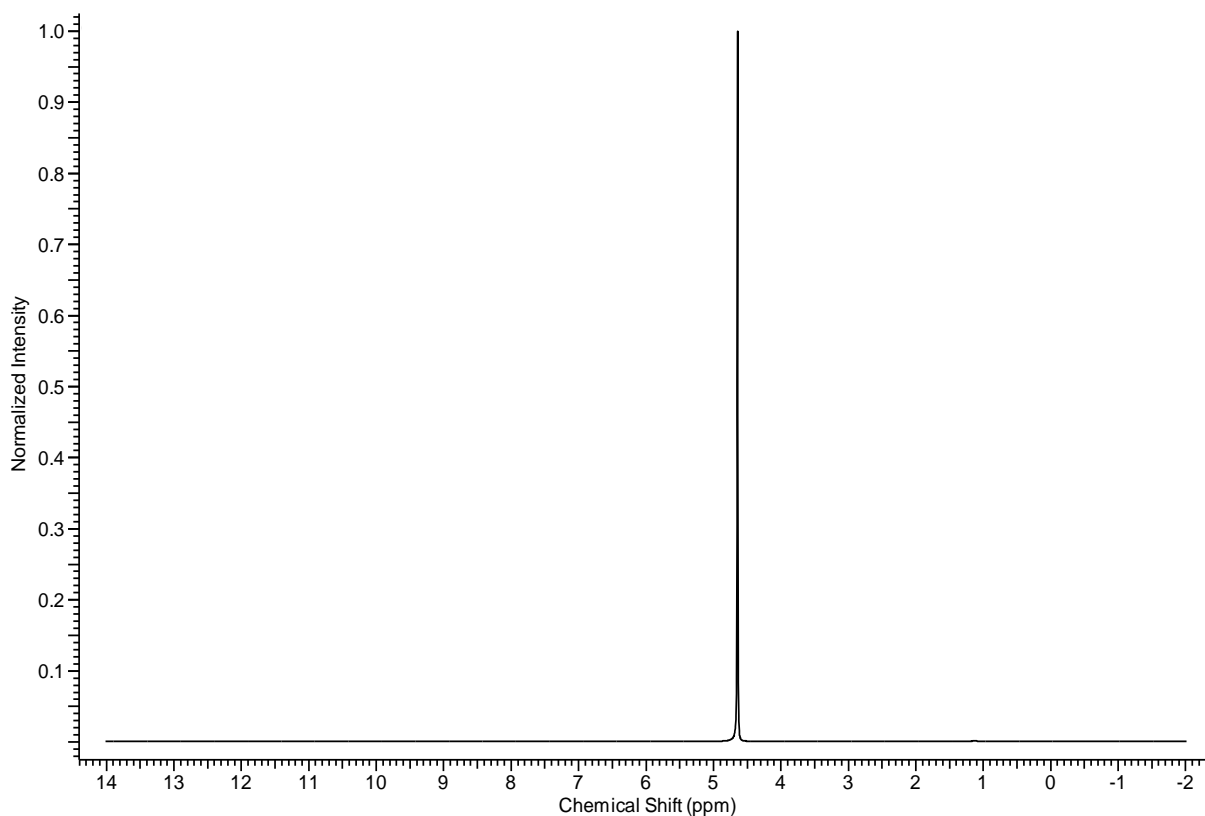
ambos dependem do valor de plaquetas e não tínhamos essa informação de um dos pacientes.

### 5.3 Resultados e Discussão

#### 5.3.1 Obtenção de Espectros e Processamento dos Dados

Como a quantidade de água nos biofluidos é muito grande em relação aos demais metabólitos, ao realizar a análise de RMN de  $^1\text{H}$  das amostras só é observado o sinal atribuído a água,  $\delta$  4,70ppm, como podemos observar na Figura 26. Portanto, é necessária a utilização da sequência de pulsos de supressão do sinal da água, PRESAT.

Figura 26: Espectro de RMN de  $^1\text{H}$  ( $\text{D}_2\text{O}$ , 300 MHz) de uma amostra de soro.



Após a utilização do PRESAT podemos observar outros sinais, ver Figura 27. Há sinais muito alargados e sobrepostos na região entre 0,50 e 4,00ppm, atribuídos às moléculas de grande massa molar, como os lipídios, que se apresentam nessa região espectral e tem  $T_2$  curto. Para aumentar a resolução do espectro e viabilizar os tratamentos estatísticos, buscamos suprimir esses sinais através da utilização da sequência de pulsos CPMG acoplada à sequência PRESAT, obtendo o espectro da Figura 28.

Figura 27: Espectro de RMN de  $^1\text{H}$  ( $\text{D}_2\text{O}$ , 300 MHz) da amostra de soro utilizando a sequência de pulsos PRESAT.

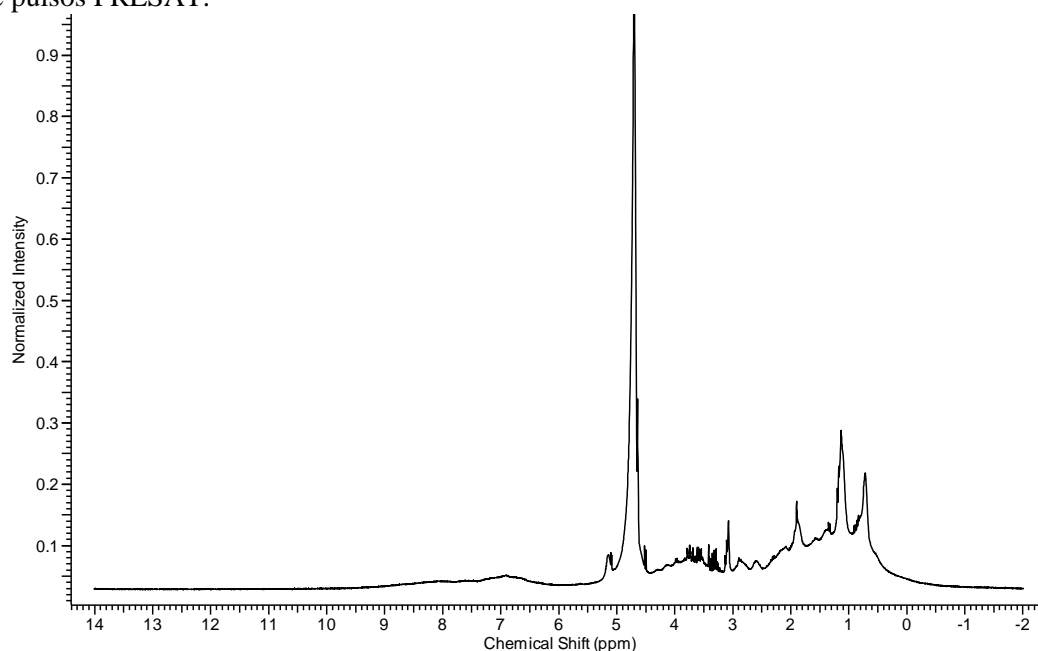
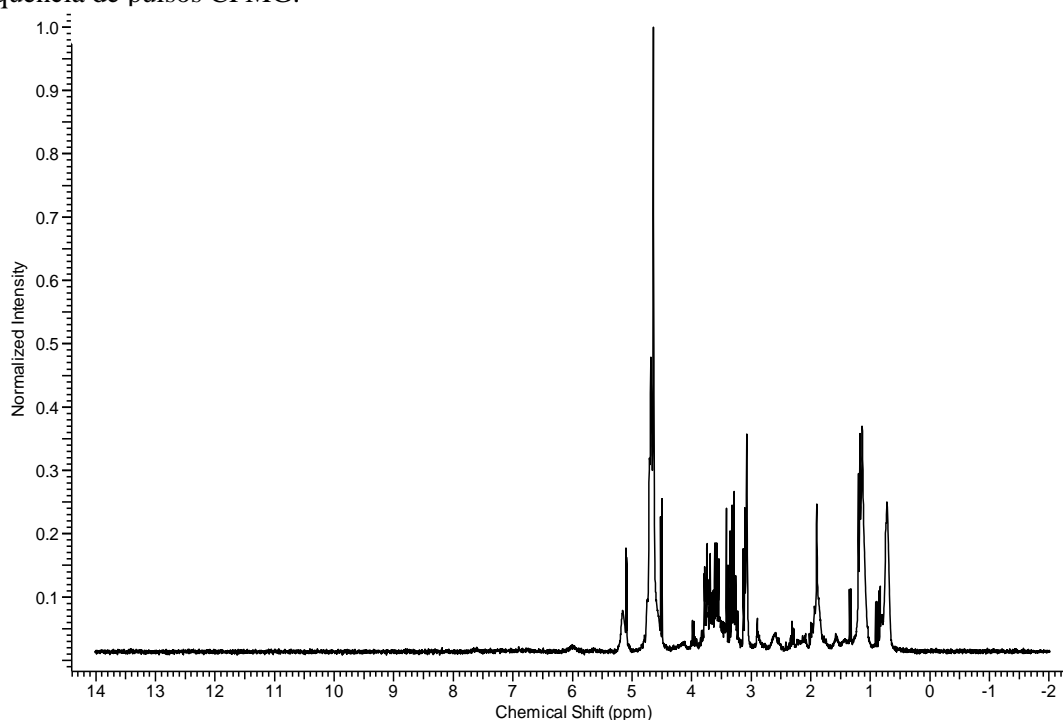


Figura 28: Espectro de RMN de  $^1\text{H}$  ( $\text{D}_2\text{O}$ , 300 MHz) da amostra de soro após aplicação da sequência de pulsos CPMG.



Uma vantagem do CPMG em relação à forma de supressão de sinal do PRESAT é que os sinais das espécies de alta massa molar foram suprimidos, porém os demais sinais na mesma região não foram afetados, fazendo com que não fossem perdidas mais informações. No caso da PRESAT, todos os sinais com o deslocamento químico na região da supressão também são perdidos.

### 5.3.2 Fibrose Significativa

Originalmente, o APRI foi desenvolvido utilizando a classificação Ishak e somente para pacientes com hepatite C. (WAI et al., 2003) Como as biopsias foram realizadas considerando a escala METAVIR, foi então necessário fazer a conversão entre os escores Ishak e METAVIR, para determinar os grupos. Na Tabela 12 estão as equivalências entre os dois escores.

Tabela 12: Conversão entre Ishak e METAVIR

Ishak	METAVIR
0	0
1 - 2	1
3	2
4 - 5	3
6	4

Fonte: (GAMAL; ZALATA, 2011)

Fibrose significativa foi considerada pelos autores como Ishak maior ou igual a 3, usando os pontos de corte: APRI  $< 0,5$ , sem fibrose ou fibrose leve; e APRI  $> 1,5$ , fibrose significativa. (WAI et al., 2003) Consideramos então, no nosso grupo de amostras, fibrose significativa quando a classificação METAVIR for maior ou igual a 2.

A matriz foi construída com espectros de RMN de  $^1\text{H}$  de amostras de soro de 63 pacientes portadores de HCV, sendo 39 com fibrose significativa, escore METAVIR  $\geq 2$ , e 24 com fibrose não significativa, escore METAVIR  $\leq 1$ , e 184 variáveis. Foi utilizado o pré-processamento SNV na matriz de dados originais.

Na PCA, não observamos agrupamento natural das amostras nos grupos de interesse e também não foram identificadas amostras anômalas. Mas foi utilizada como técnica de redução de variáveis, para aplicação da LDA. A LDA utilizou as cinco componentes principais que melhor discriminaram as duas classes, selecionadas utilizando o Lambda de Wilks, para construir um modelo de classificação. Aplicamos a LOOCV ao modelo construído, obtendo os resultados apresentados na Tabela 13.

O índice APRI foi calculado para cada paciente e, segundo os pontos de corte anteriormente definidos, foram calculados os valores de: sensibilidade, especificidade, VPP, VPN e exatidão do APRI no nosso grupo de estudo.

Na Tabela 13 podemos observar que a performance do modelo metabonômico foi muito superior à do APRI, enquanto o primeiro teve seu menor resultado acima de 95%, o outro não chegou a 88%.

O trabalho original obteve como maior valor de: sensibilidade, 91%; especificidade, 85%; VPP, 88%; e VPN, 86%, nos pontos de corte de fibrose significativa. (WAI et al., 2003) Portanto, o modelo metabonômico construído teve um poder de classificação maior que o de um teste largamente utilizado na prática médica.

Tabela 13: Resultados para predição de fibrose significativa

	<b>F2- F4</b>	<b>F0- F1</b>	<b>Sensibilidade</b>	<b>Especificidade</b>	<b>VPP</b>	<b>VPN</b>	<b>Exatidão</b>
<b>APRI</b>			(%)	(%)	(%)	(%)	(%)
<b>≤0.5</b>	34	10	87,2	56,5	77,3	72,2	75,8
<b>&gt;0.5</b>	5	13					
<b>≤1.5</b>	15	3	38,5	87,0	83,3	45,5	56,5
<b>&gt;1.5</b>	24	20					
<b>Modelo Metabonômico</b>							
<b>Significativa</b>	39	1	100	95,8	97,5	100	98,4
<b>Não significativa</b>	0	23					

VPP: Valor preditivo positivo, VPN: valor preditivo negativo

### 5.3.3 Fibrose Avançada

O FIB-4 foi construído e validado para identificar pacientes com fibrose hepática, coinfectados com HCV e HIV, utilizando a classificação Ishak. (STERLING et al., 2006) Posteriormente, foi validado para pacientes monoinfectados com o vírus HCV, utilizando a classificação METAVIR, convertida do Ishak segundo a Tabela 12, com os mesmos pontos de corte do trabalho original: escore FIB-4 < 1,45, fibrose leve ou moderada (escore METAVIR ≤ 2); e escore FIB-4 > 3,25, fibrose avançada (escore METAVIR ≥ 3). (VALLET-PICHARD et al., 2007)

Um estudo mostra que a exatidão do FIB-4, no ponto de corte 3,25, diminui com o aumento da idade. (KUROSAKI; IZUMI, 2008) Como nossa população era mais velha que a dos trabalhos originais e o FIB-4 depende da idade, comparamos os acertos do FIB-4 no nosso grupo e no grupo estudado por Kurosaki e colaboradores. Obtivemos 87% de acerto para pacientes com idades ≤ 50 anos; 82% de acerto entre 51 e 60 anos; e 77% de acerto para > 60 anos, enquanto o estudo apresentou, 68%, 48% e 49%, respectivamente. Nossos resultados foram muito acima dos resultados desse trabalho, e tivemos 85% de acertos no total, o mesmo valor da área sob a curva ROC (do inglês,

*Receiver Operating Characteristics*) do trabalho de validação. (VALLET-PICHARD et al., 2007) Com isso, demonstramos que a média de idade superior na população que investigamos não apresentou significativa diminuição da performance do FIB-4.

Para a construção dos modelos metabonômicos, os 63 pacientes foram classificados como: 22 com fibrose avançada e 41 com fibrose leve ou moderada. A matriz utilizada foi a mesma do modelo construído para fibrose significativa, mudando somente as classes das amostras. Também foi utilizado SNV na matriz de dados.

Pela PCA, não observamos tendência de agrupamento de acordo com a fibrose avançada ou leve/moderada, mas usamos as componentes construídas como método de redução da dimensionalidade, para construir um modelo LDA. A validação foi feita por LOOCV, na qual o modelo não errou nenhuma amostra, apesar de ter sido construído utilizando apenas quatro componentes principais.

O modelo metabonômico apresentou uma performance melhor em todos os quesitos, como podemos atestar na Tabela 14.

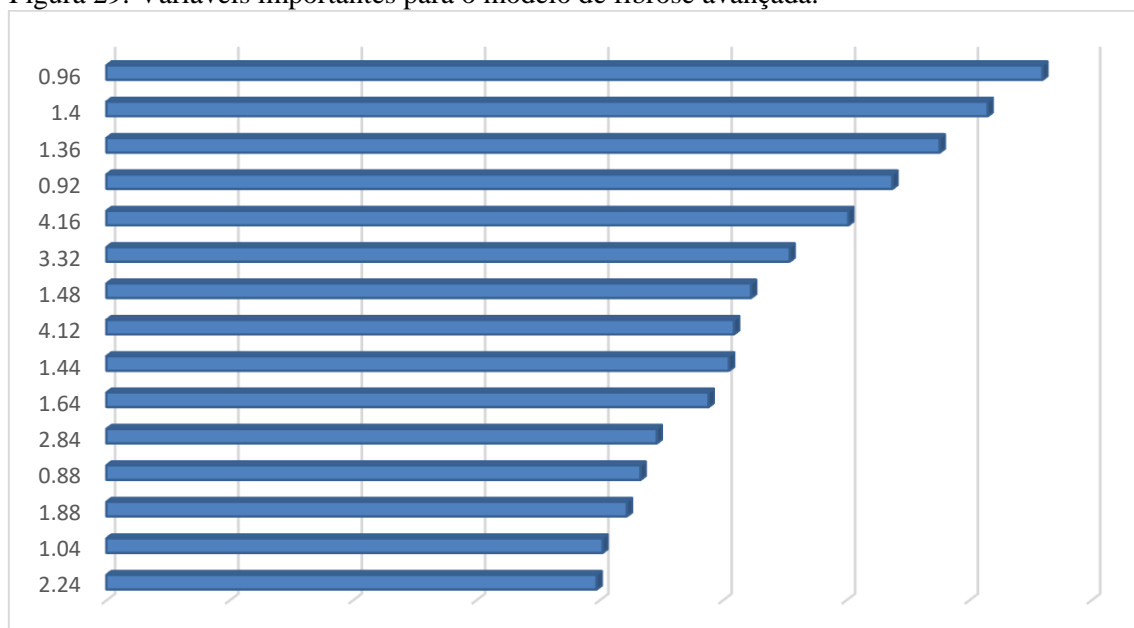
Combinando os pesos das componentes principais utilizadas na construção do modelo e os coeficientes da função canônica discriminante, encontramos as regiões do espectro mais importantes para a discriminação. Na Figura 29, estão dispostos os 15 bins que tiveram maior influência na construção do modelo de classificação das amostras em fibrose avançada ou leve/moderada.

Tabela 14: Resultados para predição de fibrose avançada

	F3-F4	F0-F2	Sensibilidade	Especificidade	VPP	VPN	Exatidão
FIB-4			(%)	(%)	(%)	(%)	(%)
≤1.45	20	25	90,9	37,5	44,4	88,2	56,5
>1.45	2	15					
≤3.25	17	4	77,3	90,0	81,0	87,8	85,5
>3.25	5	36					
Modelo Metabonômico							
Avançada	22	0	100	100	100	100	100
Não Avançada	0	41					

VPP: Valor preditivo positivo, VPN: valor preditivo negativo

Figura 29: Variáveis importantes para o modelo de fibrose avançada.



Fonte: Autoria própria.

#### 5.3.4 Cirrose

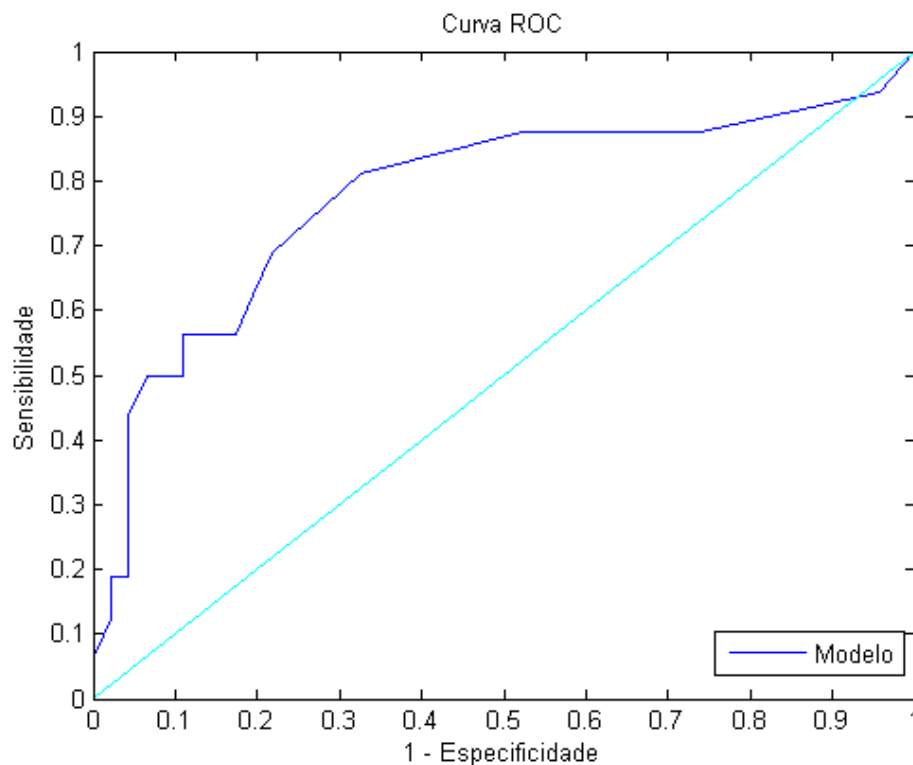
O artigo original no FIB-4 também classifica cirrose, como Ishak maior ou igual a 5. (STERLING et al., 2006) Como 4 e 5 no Ishak são correspondentes a 3 no METAVIR, não pudemos usar esse método para comparar com o resultado de cirrose do METAVIR.

Há estudos que validaram o APRI para determinação de cirrose, escore igual a 4 no METAVIR. Porém, haviam discordâncias entre os valores de ponto de corte, como mostrada numa meta-análise de 2011. (LIN et al., 2011) Para assegurar a melhor performance do APRI, construímos uma curva ROC, Figura 30, para encontrar os pontos de corte que melhor se ajustassem à população em estudo. A área sob a curva foi de 0.81, e os pontos escolhidos foram:  $APRI < 1,0$ , sem cirrose; e  $APRI > 2,00$ , com cirrose. As 63 amostras de soro da matriz de dados foram divididas em: 16 pacientes com cirrose, clinicamente atestada ou com escore METAVIR igual 4; e 47 sem cirrose, escore METAVIR  $\leq 3$ . Foi utilizado o pré-processamento SNV na matriz de dados originais.

Assim como para as duas classificações anteriores, não foi identificado agrupamento das amostras nas classes de interesse, mas a PCA foi utilizada para redução de variáveis. O LDA foi construído utilizando apenas três componentes principais. Avaliado por

LOOCV, o modelo de classificação apresentou: 94% de sensibilidade; 98% de especificidade; e 97% de exatidão.

Figura 30: Curva ROC do APRI para prever cirrose.



Fonte: Autoria própria.

Ao avaliar a Tabela 15, percebemos que o modelo metabonômico construído para identificar pacientes com cirrose é bastante superior aos melhores resultados obtidos pelo APRI.

Tabela 15: Resultados para predição de cirrose

	F4	F0-F3	Sensibilidade	Especificidade	VPP	VPN	Exatidão
APRI			(%)	(%)	(%)	(%)	(%)
≤1.0	11	10	68,8	78,3	52,4	87,8	75,8
>1.0	5	36					
≤2.0	9	5	56,3	89,1	64,3	85,4	80,6
>2.0	7	41					
Modelo Metabonômico							
Cirrose	15	1	93,8	97,9	93,8	97,9	96,8
Sem cirrose	1	46					

VPP: Valor preditivo positivo, VPN: valor preditivo negativo

### 5.3.5 Pacientes com Classificação F2 no METAVIR

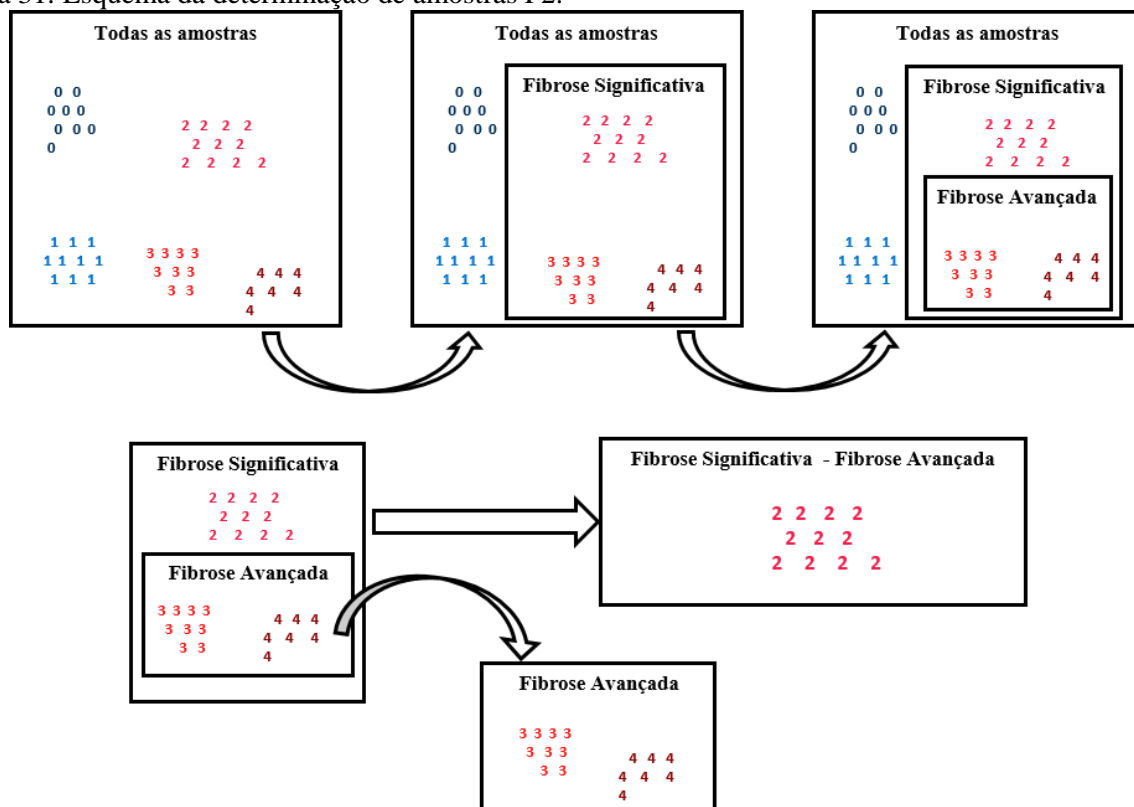
Devido à importância da classificação desse estágio, construímos modelos de classificação utilizando PLS-DA e LDA, com o objetivo de separar as amostras com grau F2 das demais. Porém, não obtivemos sucesso. Acreditamos que o problema estava em reunir graus de fibrose F0 e F1 com F3 e F4, que devem ter perfis bioquímicos muito distintos. Pois como as técnicas utilizadas tentam construir um modelo que maximize a diferença interclasses e minimize as diferenças intraclasses, eles buscaram dar ênfase a semelhanças entre amostras muito diferentes, como um paciente sem fibrose (F0) e um com cirrose (F4). Buscamos, então, uma estratégia diferente: utilizar os modelos de fibrose significativa e avançada para classificar as amostras F2.

As amostras estão classificadas em: F0, F1, F2, F3 e F4. O grupo fibrose significativa engloba F2, F3 e F4. Já o grupo fibrose avançada, inclui F3 e F4, podendo ser interpretado como um subgrupo do fibrose significativa. Dessa forma, se retirarmos os elementos F3 e F4, grupo fibrose avançada, do fibrose significativa, só restarão os elementos F2, como ilustrado na Figura 31.

Seguindo esse raciocínio, selecionamos as amostras classificadas pelo modelo metabonômico como fibrose significativa e excluimos as classificadas como fibrose avançada, obtendo as amostras classificadas como grau F2. Construímos a tabela de contingência do teste (Tabela 16) e encontramos 100% de sensibilidade e valor preditivo negativo.



Figura 31: Esquema da determinação de amostras F2.



Fonte: Autoria própria.

Tabela 16: Resultados para predição de F2 por modelos metabonômicos

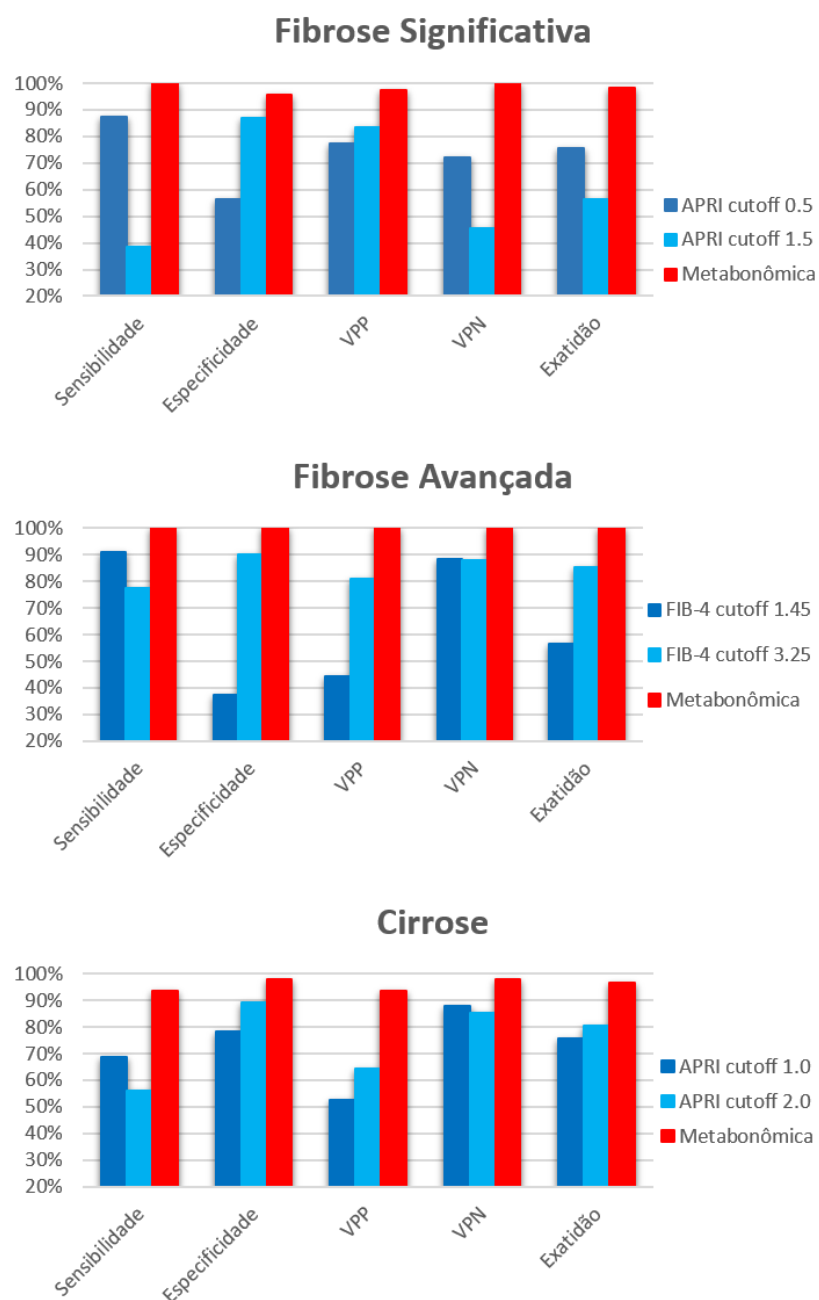
Modelos Metabonômicos	F2		Outros	Sensibilidade	Especificidade	VPP	VPN	Exatidão
				(%)	(%)	(%)	(%)	(%)
	F2	17	2					
	Outros	0	44	100	95,6	89,5	100	96,8

VPP: Valor preditivo positivo, VPN: valor preditivo negativo. Outros: F0, F1, F3 e F4.

## 5.4 Conclusão

Na Figura 32, é possível notar como os modelos metabonômicos apresentaram uma performance melhor que o APRI e o FIB-4 na população de estudo. Outra vantagem dos modelos metabonômicos em relação ao APRI e FIB-4, é que todos os pacientes são classificados, enquanto os outros dois métodos têm uma região duvidosa, entre os dois pontos de corte escolhidos, na qual não se pode atribuir classificação à fibrose hepática do paciente, fazendo-se necessária a utilização de outro método de diagnóstico.

Figura 32: Comparação entre modelos metabonômicos e FIB-4 e APRI



Fonte: Autoria própria.

## 5.5 Perspectivas

- ☐ Validação externa dos modelos construídos – conseguimos mais amostras de soro de pacientes com grau de fibrose hepática definido por biópsia, então esta etapa já está em andamento;
- ☐ Estudar os deslocamentos químicos mais importantes na construção dos modelos, para determinar quais são os metabólitos associados aos diferentes graus de fibrose hepática e associar às rotas metabólicas do organismo.

## 6 CONCLUSÕES GERAIS

A metabonômica se mostrou uma ferramenta com potencial para identificar, simultaneamente, diferentes lesões no fígado: hepatites virais por HBV ou HCV; e esteatose, utilizando amostras de urina, caracterizando-se como uma análise rápida e não invasiva.

Além disso, utilizando amostras de soro sanguíneo, se apresentou como uma alternativa, minimamente invasiva, à biópsia hepática, para classificação do grau de fibrose hepática, importante para determinação do estadiamento da doença e indicação para tratamento. Foram construídos modelos para classificar fibrose significativa, fibrose avançada e cirrose, alcançando 98,4, 100 e 96,8% de exatidão, respectivamente.

Os resultados obtidos foram superiores a outros testes alternativos para estadiamento da fibrose hepática, largamente empregados na prática médica, APRI e FIB-4, empregados ao mesmo conjunto de dados. Além disso, o modelo metabonômico permitiu identificar o estágio F2 (no METAVIR), que não é possível com esses testes mais simples.

## 7 PERSPECTIVAS GERAIS

- ☐ Obter amostras suficientes para serem divididas em um grupo de treinamento e um grupo de teste com amostras representativas de cada classe;
- ☐ Fazer estudos dos metabólitos responsáveis pelas classificações e associar às rotas metabólicas do organismo;
- ☐ Utilizar a estratégia metabonômica para o diagnóstico e acompanhamento de outras doenças.

## REFERÊNCIAS

- ALTER, M. J. Epidemiology of hepatitis C virus infection. **World journal of gastroenterology : WJG**, v. 13, n. 17, p. 2436–41, 7 maio 2007.
- BERRUETA, L. A; ALONSO-SALCES, R. M.; HÉBERGER, K. Supervised pattern recognition in food analysis. **Journal of chromatography. A**, v. 1158, n. 1–2, p. 196–214, 27 jul. 2007.
- BRANDÃO, A. B. DE M. et al. Diagnóstico da hepatite C na prática médica: revisão da literatura. **Revista Panamericana de Salud Pública**, v. 9, n. 3, p. 161–168, mar. 2001.
- BRUNS, R. E.; FAIGLE, J. F. G. Quimiometria. **Química Nova**, p. 84–99, 1984.
- CALVANI, R. et al. Gut microbiome-derived metabolites characterize a peculiar obese urinary metabotype. **International journal of obesity (2005)**, v. 34, n. 6, p. 1095–1098, 2010.
- CALVARUSO, V.; CRAXÌ, A. 2011 European Association of the Study of the Liver hepatitis C virus clinical practice guidelines. **Liver International**, v. 32, p. 2–8, fev. 2012.
- CAROLINE, C. W. P. et al. Ehrlich and sarcoma 180 tumour characterisation and early detection by <sup>1</sup>H NMR-based metabonomics of mice serum. **Journal of the Brazilian Chemical Society**, v. 25, n. 5, p. 841–852, 2014.
- CHAI, Y. et al. Application of <sup>1</sup>H NMR spectroscopy-based metabonomics to feces of cervical cancer patients with radiation-induced acute intestinal symptoms. **Radiotherapy and Oncology**, v. 117, n. 2, p. 294–301, 2015.
- CHEN, T. et al. Serum and urine metabolite profiling reveals potential biomarkers of human hepatocellular carcinoma. **Molecular & cellular proteomics : MCP**, v. 10, n. 7, p. M110.004945, 2011.
- COCCHETTO, D. M.; TSCHANZ, C.; BJORNSSON, T. D. Decreased rate of creatinine production in patients with hepatic disease: implications for estimation of creatinine clearance. **Therapeutic drug monitoring**, v. 5, n. 2, p. 161–168, 1983.
- COSTA, E. G.; MIGLIORATI, C. A. Câncer Bucal : Avaliação do tempo decorrente entre a detecção da lesão e o início do tratamento. **Revista Brasileira de Cancerologia**, v. 47, n. 3, p. 283–9, 2001.
- CRICHTON, N. Information point: Wilks' lambda. **Journal of clinical nursing**, v. 9, n. 3, p. 381, 2000.
- CROAGH, C. M.; DESMOND, P. V.; BELL, S. J. Genotypes and viral variants in chronic hepatitis B: A review of epidemiology and clinical relevance. **World Journal of Hepatology**, v. 7, n. 3, p. 289–303, 2015.
- DA SILVA NETO, H. G. et al. Determination of metabolite profiles in tropical wines by <sup>1</sup>H NMR spectroscopy and chemometrics. **Magnetic Resonance in Chemistry**, v. 47, n. S1, p. S127–S129, dez. 2009.
- DE SOUZA, A. M.; POPPI, R. J. Experimento didático de quimiometria para análise

exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: Um tutorial, parte I. **Química Nova**, v. 35, n. 1, p. 223–229, 2012.

GAMAL, S.; ZALATA, K. Ishak versus METAVIR: Terminology, Convertibility and Correlation with Laboratory Changes in Chronic Hepatitis C. **Liver Biopsy, Dr Hirokazu Takahashi (Ed.)**, 2011.

GAO, H. et al. Application of <sup>1</sup>H NMR-based metabonomics in the study of metabolic profiling of human hepatocellular carcinoma and liver cirrhosis. **Cancer Science**, v. 100, n. 4, p. 782–785, 2009.

GODOY, M. M. G. et al. Hepatitis C virus infection diagnosis using metabonomics. **Journal of viral hepatitis**, v. 17, n. 12, p. 854–8, dez. 2010.

GONG, M. et al. Icaritin reverses corticosterone-induced depression-like behavior, decrease in hippocampal brain-derived neurotrophic factor (BDNF) and metabolic network disturbances revealed by NMR-based metabonomics in rats. **Journal of Pharmaceutical and Biomedical Analysis**, v. 123, p. 63–73, maio 2016.

GUYTON, Arthur Clifton; HALL, John E.; GUYTON, Arthur C. **Tratado de fisiologia médica**. Elsevier Brasil, 2006.

KIM, S.-H. et al. <sup>1</sup>H-nuclear magnetic resonance spectroscopy-based metabolic assessment in a rat model of obesity induced by a high-fat diet. **Analytical and bioanalytical chemistry**, v. 395, n. 4, p. 1117–24, out. 2009.

KIMHOFER, T. et al. Proteomic and metabonomic biomarkers for hepatocellular carcinoma: a comprehensive review. **British journal of cancer**, v. 112, n. 7, p. 1141–56, 2015.

KUMAR, V. et al. **Fundamentos de Robbins & Cotran - Patologia: Bases Patológicas das Doenças**. Oitava ed. Rio de Janeiro: [s.n.], v. 41

KUROSAKI, M.; IZUMI, N. External validation of FIB-4: Diagnostic accuracy is limited in elderly populations [1]. **Hepatology**, v. 47, n. 1, p. 352, 2008.

LABRECQUE, D. R. et al. World Gastroenterology Organisation Global Guidelines. **Journal of Clinical Gastroenterology**, p. 1, jun. 2014.

LAVANCHY, D. Evolving epidemiology of hepatitis C virus. **Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases**, v. 17, n. 2, p. 107–15, fev. 2011.

LEE, J.; GILMORE, I. **A guide to the practical use of Chemometrics - with applications for Static SIMS** Teddington National Physical Laboratory, , 2006.

LEVITT, M. H. Spin Dynamics. **JOHN WILEY & Sons, Ltd**, v. second edi, p. 693, 2008.

LI, X. Z. et al. Cerebral metabonomics study on Parkinson's disease mice treated with extract of *Acanthopanax senticosus* harms. **Phytomedicine**, v. 20, n. 13, p. 1219–1229, 2013.

LI, Y.-Q. et al. Particle Swarm Optimization-Based Protocol for Partial Least- Squares Discriminant Analysis: Application to  $^1\text{H}$  Nuclear Magnetic Resonance Analysis of Lung Cancer Metabonomics. **Chemometrics and Intelligent Laboratory Systems**, v. 135, p. 192–200, 2014.

LIANG, T. J. Hepatitis B: the virus and disease. **Hepatology (Baltimore, Md.)**, v. 49, n. 5 Suppl, p. S13-21, maio 2009.

LIN, Z. H. et al. Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: An updated meta-analysis. **Hepatology**, v. 53, n. 3, p. 726–736, 2011.

LINDON, J. C.; HOLMES, E.; NICHOLSON, J. K. Peer Reviewed: So What's the Deal with Metabonomics? **Analytical Chemistry**, v. 75, n. 17, p. 384 A-391 A, set. 2003.

MESSANA, I. et al. Proton nuclear magnetic resonance spectral profiles of urine in type II diabetic patients. **Clinical chemistry**, v. 44, n. 7, p. 1529–1534, 1998.

MIAO, Z. et al. The application of HPLC and microprobe NMR spectroscopy in the identification of metabolites in complex biological matrices. **Analytical and bioanalytical chemistry**, v. 407, n. 12, p. 3405–16, 2015.

MINISTÉRIO DA SAÚDE; SECRETARIA DE VIGILÂNCIA EM SAÚDE; DEPARTAMENTO DE DST, A. E H. V. **Protocolo Clínico e Diretrizes Terapêuticas para o Tratamento da Hepatite Viral Crônica B e Coinfecções**. [s.l.: s.n.].

NETO, J. M. M.; MIOTA, G. C. Uma Introdução à Análise Exploratória de Dados Multivariados. **Química Nova**, v. 21, n. 4, p. 467–469, 1998.

NICHOLSON, J. K.; LINDON, J. C.; HOLMES, E. “Metabonomics”: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. **Xenobiotica; the fate of foreign compounds in biological systems**, v. 29, n. 11, p. 1181–9, nov. 1999.

NIEDERAU, C. et al. Prognosis of chronic hepatitis C: results of a large, prospective cohort study. **Hepatology (Baltimore, Md.)**, v. 28, n. 6, p. 1687–95, dez. 1998.

ÖHMAN, A.; FORSGREN, L. NMR metabonomics of cerebrospinal fluid distinguishes between Parkinson's disease and controls. **Neuroscience Letters**, v. 594, p. 36–39, 2015.

WORLD HEALTH ORGANIZATION. **Hepatitis B**. Disponível em: <<http://www.who.int/mediacentre/factsheets/fs204/en/>>.

WORLD HEALTH ORGANIZATION. **Hepatitis C**. Disponível em: <<http://www.who.int/csr/disease/hepatitis/whocdscsrlyo2003/en/index3.html>>.

PANOBIANCO, M. S. et al. Mulheres com Diagnóstico Avançado do Câncer do Colo do Útero: Enfrentando a Doença e o Tratamento. **Revista Brasileira de Cancerologia**, v. 58, n. 3, p. 517–523, 2012.

PAVIA, D. L.; LAMPMAN, G. M.; KRIZ, G. S. **Introduction to Spectroscopy**. Third ed. [s.l.] Thomson Learning, 2001.



- RIBEIRO, J. E. et al. **Protocolo clínico e diretrizes terapêuticas para hepatite viral C e coinfeções**. Ministério da Saúde. Série A. Normas e Manuais Técnicos, Brasília, 2011.
- QIAO, Y. et al. Plasma metabonomics study of first-Episode schizophrenia treated with olanzapine in female patients. **Neuroscience Letters**, v. 617, p. 270–276, 2016.
- RIBEIRO, A. S. et al. Discriminating gamma-irradiated soybean seeds by <sup>1</sup>H NMR-based metabonomics. **Food Control**, v. 36, p. 266–272, 2014.
- ROBERTSON, D. G. Metabonomics in toxicology: a review. **Toxicological sciences : an official journal of the Society of Toxicology**, v. 85, n. 2, p. 809–22, jun. 2005.
- SACCHI, R.; MANNINA, L. Characterization of Italian extra virgin olive oils using <sup>1</sup>H-NMR spectroscopy. **Journal of Agricultural ...**, p. 3947–3951, 1998.
- SHARIFF, M. I. F. et al. Characterization of Urinary Biomarkers of Hepatocellular Carcinoma Using Magnetic Resonance Spectroscopy in a Nigerian Population. **Journal of Proteome Research**, v. 9, n. 2, p. 1096–1103, 5 fev. 2010.
- SHARIFF, M. I. F. et al. Urinary metabolic biomarkers of hepatocellular carcinoma in an Egyptian population: A validation study. **Journal of Proteome Research**, v. 10, n. 4, p. 1828–1836, 2011.
- SHERLOCK, S.; DOOLEY, J. **Diseases of the Liver and Biliary System**. Tenth ed. Oxford: Blackwell Science, 1997.
- SOGA, T. et al. Serum metabolomics reveals  $\gamma$ -glutamyl dipeptides as biomarkers for discrimination among different forms of liver disease. **Journal of Hepatology**, v. 55, n. 4, p. 896–905, 2011.
- STERLING, R. K. et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. **Hepatology**, v. 43, n. 6, p. 1317–1325, 2006.
- TRUFELLI, D. et al. Análise do atraso no diagnóstico e tratamento do câncer de mama em um hospital público. **Revista da Associação Médica Brasileira**, v. 54, n. 1, p. 72–6, 2008.
- VALLET-PICHARD, A. et al. FIB-4: An inexpensive and accurate marker of fibrosis in HCV infection. Comparison with liver biopsy and FibroTest. **Hepatology**, v. 46, n. 1, p. 32–36, 2007.
- VILLAR, J. D. F. **Fundamentos e Aplicações da Ressonância Magnética Nuclear: Aspectos Quânticos da Ressonância Magnética Nuclear**. 2<sup>a</sup> ed. Rio de Janeiro: AUREMN, 2002.
- WAI, C. T. et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. **Hepatology**, v. 38, n. 2, p. 518–526, 2003.
- WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, v. 58, n. 2, p. 109–130, 2001.

WORLD HEALTH ORGANIZATION. **The world health report 2003: shaping the future.** World Health Organization, 193.p, 2003.

ANEXO A – APROVAÇÃO DO COMITÊ DE ÉTICA PARA PROJETO DE  
DIAGNÓSTICO DE DOENÇAS HEPÁTICAS



SERVIÇO PÚBLICO FEDERAL  
UNIVERSIDADE FEDERAL DE PERNAMBUCO  
Comitê de Ética em Pesquisa

Av. da Engenharia, s/n – 1º Andar, Cid. Universitária, CEP 50740-600, Recife - PE,  
Tel/fax: 81 2126 8588 - [www.ufpe.br/ccs](http://www.ufpe.br/ccs); e-mail: [cepccs@ufpe.br](mailto:cepccs@ufpe.br)

Of. Nº. 335/2012 - CEP/CCS

Recife, 04 de setembro de 2012

A  
Profª. Michele Maria de Godoy  
Hospital das Clínicas - UFPE

Registro do SISNEP FR - 448995  
CAAE – 0382.0.172.000-11  
Registro CEP/CCS/UFPE Nº 396/11  
Título: Emprego da metabonômica no diagnóstico das doenças hepáticas em amostras de urina.  
Pesquisador Responsável: Michele Maria de Godoy

Senhor (a) Pesquisador (a):

Informamos que o Comitê de Ética em Pesquisa Envolvendo Seres Humanos do Centro de Ciências da Saúde da Universidade Federal de Pernambuco (CEP/CCS/UFPE) registrou e analisou de acordo com a Resolução N.º 196/96 do Conselho Nacional de Saúde, o protocolo de pesquisa em epígrafe, liberando-o para início da coleta de dados em 03 de setembro de 2012.

Ressaltamos que a aprovação definitiva do projeto será dada após a entrega do relatório final, conforme as seguintes orientações:

- a) Projetos com, no máximo, 06 (seis) meses para conclusão: o pesquisador deverá enviar apenas um relatório final;
- b) Projetos com períodos maiores de 06 (seis) meses: o pesquisador deverá enviar relatórios semestrais.

Dessa forma, o ofício de aprovação somente será entregue após a análise do relatório final.

Atenciosamente

Prof. Geraldo Bosco Lindoso Couto  
Coordenador do CEP/CCS / UFPE

## ANEXO B – APROVAÇÃO DO COMITÊ DE ÉTICA PARA PROJETO DE AVALIAÇÃO DO GRAU DE FIBROSE HEPÁTICA

**UNIVERSIDADE FEDERAL DE  
PERNAMBUCO CENTRO DE  
CIÊNCIAS DA SAÚDE / UFPE-**



### PROJETO DE PESQUISA

**Título:** AVALIAÇÃO DO GRAU DE FIBROSE HEPÁTICA EM PACIENTES COM DOENÇA HEPÁTICA CRÔNICA PELOS VÍRUS DAS HEPATITES B E C UTILIZANDO

**Área Temática:**

**Versão:** 1

**CAAE:** 06572612.3.0000.5208

**Pesquisador:** Andrea Dória

**Instituição:** CENTRO DE CIÊNCIAS DA SAÚDE

### PARECER CONSUBSTANCIADO DO CEP

**Número do Parecer:** 93.127

**Data da Relatoria:** 05/09/2012

#### **Apresentação do Projeto:**

O estudo intitulado AVALIAÇÃO DO GRAU DE FIBROSE HEPÁTICA EM PACIENTES COM DOENÇA HEPÁTICA CRÔNICA PELOS VÍRUS DAS HEPATITES B E C UTILIZANDO METABONÔMICA, é um projeto de Doutorado do Programa de Pós-Graduação em Medicina Tropical UFPE, que tem como pesquisadora responsável Andrea Dória Batista, sob orientação do Prof. Dr. Edmundo Pessoa de Almeida Lopes e do Prof. Dr. Ricardo Oliveira da Silva.

O conhecimento do grau de fibrose hepática, decorrente da infecção crônica pelo vírus da hepatite B (HBV) e vírus da hepatite C (HVC), é de fundamental importância no manejo clínico do paciente, sendo utilizada no estadiamento da doença, na avaliação da necessidade de tratamento e na predição de resposta à terapêutica. O padrão-ouro na avaliação da fibrose hepática é o exame histopatológico, obtido através de biópsia, geralmente por agulha fina. Na prática também é utilizado marcadores bioquímicos diretos da fibrose hepática e os marcadores indiretos, que refletem hipertensão portal, atividade inflamatória e/ou função hepática, no entanto podem ser avaliados isoladamente ou, como ocorre mais frequentemente, uma pontuação (escore) é determinada a partir da combinação deles. A acurácia para identificar pacientes com fibrose intermediária é pequena por este método. A metabonômica é um novo método diagnóstico que utiliza análise estatística de espectro de fluido biológico ou tecido. Perturbação bioquímica, causada por doença, por exemplo, leva a mecanismos de ajuste, que são expressos como uma impressão digital desta perturbação. Biofluidos são geralmente analisados por espectroscopia de ressonância nuclear magnética de prótons (1H NMR), visto que o hidrogênio é abundante nestes materiais. Estudos clínicos em doenças hepáticas têm demonstrado utilidade da metabonômica no diagnóstico da hepatotoxicidade por droga, hepatite alcoólica, hepatites crônicas B e C, encefalopatia hepática e carcinoma hepatocelular. Desta forma o presente estudo pretende avaliar se a metabonômica, método minimamente invasivo, é capaz de diferenciar os graus de fibrose hepática em pacientes com doença hepática crônica pelos vírus das hepatites B e C, atendidos em um hospital universitário de Recife/PE. Os resultados poderão contribuir para diminuição da necessidade de realização de biópsia hepática e uma melhor avaliação prognóstica e terapêutica destes pacientes.

#### **Objetivo da Pesquisa:**

##### **OBJETIVO PRIMÁRIO:**

Diferenciar o grau de fibrose hepática em pacientes cronicamente infectados pelo vírus das hepatites B e C, a partir de modelo metabonômico, obtido por espectroscopia por RNM de amostras de soro.

##### **OBJETIVOS SECUNDÁRIOS:**

Obter espectros de RMN 1H de amostras de soro de pacientes com doença hepática crônica pelos vírus das hepatites B e C.

Desenvolver modelo metabonômico, a partir dos espectros de RMN 1H de soro, capaz de diferenciar

**Endereço:** Av. da Engenharia s/nº - 1º andar, sala 4, Prédio do CCS

**Bairro:** Cidade Universitária

**CEP:** 50.740-600

**UF:** PE

**Município:** RECIFE

**Telefone:** (81)2126-8588

**Fax:** (81)2126-8588

**E-mail:** cepccs@ufpe.br

UNIVERSIDADE FEDERAL DE  
PERNAMBUCO CENTRO DE  
CIÊNCIAS DA SAÚDE / UFPE-



**Conclusões ou Pendências e Lista de Inadequações:**

O projeto é relevante, pode ter eventualmente impacto no diagnóstico dos pacientes com fibrose. Desta forma, pelo exposto somos de parecer que o projeto seja aprovado.

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

**Considerações Finais a critério do CEP:**

O Colegiado aprova o parecer do protocolo em questão para início da coleta de dados.  
Projeto foi avaliado e sua APROVAÇÃO definitiva será dada, por meio de ofício impresso, após a entrega do relatório final ao Comitê de Ética em Pesquisa/UFPE.

RECIFE, 10 de Setembro de 2012

---

Assinado por:  
GERALDO BOSCO LINDOSO COUTO

**Endereço:** Av. da Engenharia s/nº - 1º andar, sala 4, Prédio do CCS  
**Bairro:** Cidade Universitária **CEP:** 50.740-600  
**UF:** PE **Município:** RECIFE  
**Telefone:** (81)2126-8588 **Fax:** (81)2126-8588 **E-mail:** cepccs@ufpe.br