



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE QUÍMICA FUNDAMENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

LUÍSA SOUZA ALMEIDA

**INVESTIGAÇÃO DE SIBILÂNCIA RECORRENTE E NÃO RECORRENTE EM
LACTENTES USANDO METABONÔMICA BASEADA EM RMN DE ^1H**

Recife
2021

LUÍSA SOUZA ALMEIDA

**INVESTIGAÇÃO DE SIBILÂNCIA RECORRENTE E NÃO RECORRENTE EM
LACTENTES USANDO METABONÔMICA BASEADA EM RMN DE ^1H**

Dissertação apresentada ao Programa de Pós-Graduação em Química da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de mestre em Química. Área de concentração: Química Analítica

Orientador: Ricardo Oliveira da Silva

Coorientador: José Licarion Pinto Segundo Neto

Recife

2021

Catálogo na fonte
Bibliotecária Nataly Soares Leite Moro, CRB15-861

A447i Almeida, Luísa Souza
Investigação de sibilância recorrente e não recorrente em lactentes usando metabonômica baseada em RMN de ^1H / Luísa Souza Almeida. – 2021.
84 f.: il., fig., tab.

Orientador: Ricardo Oliveira da Silva.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Química, Recife, 2021.
Inclui referências e apêndices.

1. Química analítica. 2. RMN de ^1H . 3. Sibilância. 4. Lactentes. I. Silva, Ricardo Oliveira da (orientador). II. Título.

543 CDD (23. ed.) UFPE- CCEN 2021 - 131

LUÍSA SOUZA ALMEIDA

INVESTIGAÇÃO DE SIBILÂNCIA RECORRENTE E NÃO RECORRENTE EM
LACTENTES USANDO METABONÔMICA BASEADA EM RMN DE ^1H .

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.

Aprovada em: 09/07/2021

BANCA EXAMINADORA

Prof. Ricardo Oliveira da Silva (Orientador)

Universidade Federal de Pernambuco

Profa. Giovannia Araújo de Lima Pereira

Universidade Federal de Pernambuco

Profa. Ana Caroline Cavalcanti Dela Bianca Melo

Universidade Federal de Pernambuco

Dra. Tássia Brena Barroso Carneiro da Costa

Universidade Estadual de Campinas

Dedico este trabalho a todos aqueles que tiveram que fazer ciência em tempos de pandemia e tiveram que se reinventar face aos cortes de orçamento. A todos que tiveram que suportar dores, perdas e lutar contra a ansiedade. E por fim, mas não menos importante, a todas as mulheres que usam sua voz como ferramenta de divulgação científica. Que a ciência permita que abram caminhos para uma sociedade melhor.

AGRADECIMENTOS

Deixo aqui expressos meus agradecimentos a todos que compuseram uma rede de apoio e suporte durante a realização deste mestrado. Minha família que sempre me incentivou nos estudos e leituras mesmo sem entender ao certo o que eu estava pesquisando; meus amigos do IFPE que seguem comigo há mais de uma década; meus amigos que fiz na graduação, sejam do meu curso ou não, e que apesar de seguirmos caminhos diferentes continuamos nos apoiando; minhas amigas do Programa Futuras Cientistas, que durante o mestrado contribuíram com um crescimento pessoal imenso; e meus amigos que fiz no mestrado e que me acolheram tão bem no dQF.

Quero agradecer, também, ao professor Ricardo Oliveira da Silva pela orientação técnica neste trabalho e pela oportunidade concebida em aprender um pouco mais sobre a espectroscopia de ressonância magnética nuclear e a metabonômica. Além disso, por todo apoio, disposição, ensinamentos além da química, explicações e paciência.

Minha gratidão também é direcionada aos amigos do antigo GMEQ Tatiane, Milena, Letícia, Márcio, Ronmilson e Ronaldo e aos amigos do atual LabMeQ, Julieth, Hélder e ao Professor Licarion por todos os momentos descontraídos e de análise juntos, pela parceria e auxílio.

Aos professores do dQF pela contribuição de aprendizado que me foram cedidas cursando as disciplinas do departamento durante estes dois anos. À central analítica do Departamento de Química Fundamental da UFPE por ceder os laboratórios e equipamento para realização deste trabalho.

Ao Hospital das Clínicas e ao Serviço de Alergia e Imunologia Clínica e no Ambulatório de Pediatria e de Puericultura, ao professor Décio Medeiros e a Aline Maranhão pela colaboração e fornecimento das amostras e dados clínicos utilizados na realização deste trabalho.

À CAPES pelo fomento da bolsa de mestrado, cedida em tempos que fazer pesquisa tornou-se um ato de resistência.

A todos, minha gratidão por tornarem esta caminhada mais leve!

“O que importa é que cada um utilize seu intelecto e nunca esteja cem por cento certo de suas convicções. Deve-se sempre deixar espaço para as dúvidas.”.
(BARCELLA e LOPES, 2018, p. 143, apud EBADI).

RESUMO

Estratégias adaptadas às características individuais de pacientes mostram-se promissoras para o diagnóstico de doenças e podem levar a uma melhoria da saúde das populações. A metabonômica faz uso da espectroscopia de Ressonância Magnética Nuclear (RMN) de ^1H para a avaliação de biofluidos, com preparo mínimo de amostra. A incidência de três ou mais episódios de sibilância durante um ano descrevem o quadro de sibilância recorrente, condição associada a níveis mais baixos de função pulmonar que aqueles que não apresentam um quadro de sibilância. Desenvolver ferramentas que permitam o reconhecimento de potenciais fatores associados à sibilância, principalmente na forma recorrente, é de grande importância para pacientes lactentes. Com isso, o presente estudo almejou desenvolver modelos metabonômicos, a partir de dados espectrais de RMN de ^1H de soro, capazes de discriminar lactentes que apresentam quadro de sibilância não recorrente daqueles que tem um quadro de sibilância recorrente. Amostras de soro de lactentes que apresentavam sintomas de sibilância foram analisadas por espectroscopia de RMN de ^1H e foram usadas ferramentas de estatística multivariada para discriminar os pacientes com sibilância recorrente ($n = 25$) (SR) dos que apresentavam sibilância não recorrente ($n = 25$) (SNR). As análises espectroscópicas foram realizadas utilizando um espectrômetro de RMN de ^1H , operando a 400 MHz, aplicando as sequências de pulso HPRESAT para saturação do sinal da água e Carr-Purcell-Meiboom-Gill (CPMG). Os espectros foram processados utilizando o software MestReNova. Para o cálculo dos modelos metabonômicos foram empregadas as técnicas: análise de componentes principais (PCA), para análise exploratória; análise de discriminantes por projeções ortogonais à estrutura latente (OPLS-DA) e análise discriminante linear com o algoritmo genético (GA-LDA), para análise de discriminantes; e, para análise de classe única, utilizou-se o formalismo *Data Driven Soft Independent Modeling of Class Analogy* (DD-SIMCA). Para os cálculos de PCA e OPLS-DA utilizou-se o MetaboAnalyst 5.0 enquanto para as análises de GA-LDA e DD-SIMCA, MATLAB 2010a software. O modelo que apresentou a melhor performance foi o DD-SIMCA, que obteve 78,6% de sensibilidade, 87,5% de especificidade, 81,8% de precisão, 91,7% de valor preditivo positivo e 70,0% de valor preditivo negativo para discriminação dos pacientes com sibilância não recorrente. Nove regiões espectrais foram identificadas

como importantes na separação dos grupos de sibilância recorrente e sibilância não recorrente. A partir dessas regiões espectrais, foi possível identificar sete metabólitos endógenos e duas rotas metabólicas associadas à discriminação observada: (I) metabolismo de alanina, aspartato e glutamato; e (II) metabolismos de glutamina e glutamato. A análise multivariada de dados tornou possível desenvolver um modelo metabonômico com alta sensibilidade e especificidade, útil para o diagnóstico diferencial de sibilância recorrente e sibilância não recorrente.

Palavras-chave: metabolômica; RMN de ^1H ; análise de classe única; sibilância; lactentes.

ABSTRACT

Strategies adapted to patient's individual characteristics are promising to disease diagnosis and could lead to an improvement in the health of population. Metabonomics uses ^1H Nuclear Magnetic Resonance (NMR) spectroscopy for biofluids evaluation, with minimal sample preparation. The incidence of three or more episodes of wheezing in one year characterizes recurrent wheezing, which is a physio pathological condition associated to lower levels of lung function. The development of tools that allow the recognition of potential factors associated with wheezing, especially recurrent wheezing, would be essential for infants. Hence, the present study aimed to develop ^1H NMR-based metabonomics models able to discriminate infants who present recurrent wheezing from non-recurrent wheezing. Serum samples from infants who presented wheezing symptoms were analyzed by ^1H NMR spectroscopy and multivariate statistical tools were used to discriminate recurrent wheezing ($n = 25$) (RW) from Non-Recurrent wheezing ($n = 14$) (NRW). ^1H NMR spectroscopy analysis were performed using a NMR spectrometer operating at 400 MHz, applying HPRESAT and Carr-Purcell-Meiboom-Gill (CPMG) pulse sequences. The spectra were processed using MestReNova software. Metabonomics models were built using different techniques, as follows: Principal Component Analysis (PCA), formalism for exploratory analysis; Orthogonal projection to Latent Structures Discriminant Analysis (OPLS-DA) and Genetic Algorithm – Linear Discriminant Analysis (GA-LDA) for classification by discriminant analysis; and Data Driven Soft Independent Modeling of Class Analogy (DD-SIMCA), for one class analysis. PCA and OPLS-DA were performed using MetaboAnalyst 5.0, while LDA DD-SIMCA metabonomics models were built using MATLAB 2010a software. The DD-SIMCA metabonomics model presented the best performance, as follows: 78.6% sensitivity, 87.5% specificity, 81.8% accuracy, 91.7% positive predictive value and 70.0% negative predictive value to discriminate patients with non-recurrent wheezing. Nine spectral regions shown that are important for discrimination between recurrent and non-recurrent wheezing samples. These spectral regions were assigned to seven endogenous metabolites and the two main metabolites pathways responsible by observed discrimination: (I) alanine, aspartate and glutamate metabolism; and (II) glutamine and glutamate metabolism. Multivariate data analysis allowed to build

metabonomics model with high sensitivity and specificity useful for differential diagnosis of recurrent wheezing and non-recurrent wheezing

Keywords: metabolomics; ^1H NMR; one class analysis; wheezing; infants.

LISTA DE ILUSTRAÇÕES

Quadro 1 –	Representação da Multidisciplinaridade da Metabonômica	20
Figura 1 –	Esquema das vias aéreas normais e durante a asma. Pode-se observar a broncoconstricção (1), produção de muco (2) e edema (3)	23
Figura 2 –	“Roda da urina” publicada no livro <i>Epiphanie Medicorum</i> em 1506 por Ulrich Pinder	26
Figura 3 –	Objetos de pesquisa das ciências ômicas numa abordagem que parte da informação genética até a manifestação metabólica de biocomponentes	27
Gráfico 1 –	Resultados de produção de trabalhos na área de Metabonomica/Metabolômica por ano no período de 2000 a 2020	28
Figura 4 –	Divisão do momento magnético de spin (m_l) com a aplicação de um campo magnético (B_0)	31
Figura 5 –	Deslocamentos químicos de ^1H (δ_{H}) presentes em grupos funcionais orgânicos ao longo do eixo de um espectro de RMN.	33
Figura 6 –	Representação do fenômeno da RMN num sistema de coordenadas girantes. O campo magnético oscilante B_1 aplicado perpendicularmente a B_0 é utilizado para excitar a amostra	34
Figura 7 –	Sequência de pulsos no experimento de PRESAT para suprimir o sinal da água	35
Figura 8 –	Sequência de pulsos CPMG com sinal HPRESAT. Em preto tempo em segundos, em azul tempo em microssegundos, em marrom tempo em milissegundos, em ciano número de ciclos	36
Figura 9 –	Esquematização da sequência de pulsos Carr-Purcell-Meiboom-Gill	36
Figura 10 –	Esquematização da sequência de passos realizada para extrair informação dos dados coletados numa análise	39

metabonômica

- Figura 11 – Esquema de uma matriz hipotética $X_{(i,j)}$ no cálculo de PCA 40
- Figura 12 – Representação das matrizes utilizadas no PLS-DA. X 41
contendo os dados originais e Y, os índices de classe 0
(pertencente a classe analisada) e 1 (não pertencente) do
exemplo, i representa as amostras e j as variáveis
- Figura 13 – Esquema gráfico do cálculo da OPLS-DA, indicando os 43
escores (p'), escores ortogonais (p_{orto}'), autovetores (w'),
autovetores ortogonais (w_{orto}'), pesos (t), pesos ortogonais
(t_{orto}') para a matriz X, u as colunas da matriz Y e c' as
classes do modelo
- Quadro 2 – Métricas da matriz de confusão, seu cálculo e significância 46
estatística para modelos metabonômicos
- Figura 14 – Típico espectro de RMN de 1H (400 MHz, D_2O) de soro 51
obtido nesse estudo de um paciente do grupo de Sibilância
não Recorrente. Alguns sinais foram atribuídos a
metabólitos
- Figura 15 – Análise de componentes principais do conjunto de dados 52
analisado. Em vermelho amostras de Sibilância Recorrente
e em verde amostras de Sibilância não Recorrente
- Figura 16 – Modelo de PLS-DA. Em vermelho Sibilância Recorrente e 53
em verde Sibilância não Recorrente
- Figura 17 – Figuras de mérito do PLS-DA. A) Teste de permutação. B) 54
Valores de performance do modelo
- Figura 18 – Modelo metabonômica de OPLS-DA calculado para 55
discriminação entre os grupos de Sibilância Recorrente e
não Recorrente (A). Resultado do teste de permutação do
OPLS-DA (B)
- Figura 19 – Modelo DDSIMCA para o grupo de treinamento (A) e o 57
grupo teste (B). Em preto as amostras da classe de
treinamento (SR), em cinza as amostras *outliers* preditas
pelo modelo. A linha verde indica o intervalo de confiança
do grupo de amostras modeladas no formalismo

Figura 20 – (A) Balanço da conversão do glutamato em glutamina e a 60
produção de intermediários de vias energéticas. (B) Vias
metabólicas de biossíntese do glutamato

Figura 21 – Identificação das vias metabólicas mais relevantes para a 62
discriminação. (a) Metabolismo da alanina, aspartato e
glutamato (p -value 0.0045); (b) Metabolismo da D-glutamina
e D-glutamato (p -value 0.0230)

LISTA DE TABELAS

Tabela 1 –	Atividade de alguns núcleos observados em RMN e sua abundância natural	34
Tabela 2 –	Exemplo de matriz de confusão em formalismos quimiométricos	45
Tabela 3 –	Características demográficas, clínicas e laboratoriais dos voluntários por grupo	50
Tabela 4 –	Matriz de confusão das classes preditas usando o Índice Preditivo de Asma modificado (mAPI)	51
Tabela 5 –	Matriz de confusão dos modelos de GA-LDA para treinamento e teste	56
Tabela 6 –	Matriz de Confusão do modelo metabonômico de DD-SIMCA utilizando o espectro de RMN de ^1H do soro de lactentes para discriminação entre Sibilância Recorrente e Sibilância não Recorrente	58
Tabela 7 –	Metabólitos endógenos mais importantes para discriminação entre Sibilância Recorrente e Sibilância não Recorrente e sua variação de concentração nestes grupos de acordo com o modelo metabonômico de GA-LDA	59

LISTA DE ABREVIATURAS E SIGLAS

API	do inglês, <i>Asthma Predictive Index</i>
CPMG	Sequência Carr-Purcell-Meiboon-Gill
CLAE	Cromatografia líquida de alta eficiência
DD-SIMCA	do inglês, <i>Data Driven Soft Independent Modeling of Class Analogy</i>
EBC	do inglês, <i>Exhaled Breath Condensate</i>
EISL	do espanhol, <i>Estudio Internacional de Sibilancias em Lactentes</i>
FN	Falsa negativa
FP	Falsa positiva
GA	do inglês, <i>Genetic Algorithm</i>
HPRESAT	Sequência de pulsos com pulso seletivo para supressão do sinal da água
IA	Inteligência Artificial
IgE	Imunoglobulina E
IV	Espectroscopia de Infravermelho
KNN	do inglês, <i>K-Nearest Neighbor</i>
LDA	Análise Discriminante Linear, do inglês <i>Linear Discriminant Analysis</i>
LDL	do inglês <i>Low Density Lipid</i>
mAPI	do inglês, <i>Modified Asthma Predictive Index</i>
OPLS-DA	Análise discriminante por projeções ortogonais à estrutura latente, do inglês <i>Orthogonal Projection to Latent Structures Discriminant Analysis</i>
PC	Componente principal
PCA	Análise de Componentes Principais, do inglês <i>Principal Component Analysis</i>
PIAMA	do inglês, <i>Prevention and Incidence of Asthma and Mite Allergy</i>
PLS-DA	Análise discriminantes por Mínimos Quadrados Parciais, do inglês <i>Partial Least Square Discriminant Analysis</i>
RF	Radiofrequência
RMN de ¹ H	Ressonância Magnética Nuclear de hidrogênio
SIMCA	do inglês, <i>Soft Independent Modeling of Class Analogy</i>
SNR	Sibilância não recorrente
SNV	do inglês, <i>Standard Normal Variation</i>
SR	Sibilância recorrente

TCLE	Termo de consentimento livre e esclarecido
TMS	Tetrametilsilano
VLDL	do inglês, <i>Very Low Density Lipid</i>
VN	Verdadeira negativa
VP	Verdadeira positiva
VPN	Valor preditivo positivo
VPP	Valor preditivo negativo

SUMÁRIO

1	INTRODUÇÃO	19
2	OBJETIVOS	22
2.1	GERAL	22
2.2	ESPECÍFICOS	22
3	FUNDAMENTAÇÃO TEÓRICA	23
3.1	ASMA E SIBILÂNCIA: CARACTERÍSTICAS, CONDIÇÕES CLÍNICAS E SAÚDE PÚBLICA	23
3.2	METABONÔMICA: O ESTUDO DO PERFIL METABÓLICO	25
3.2.1	Estado da Arte	28
3.3	RESSONÂNCIA MAGNÉTICA NUCLEAR	29
3.3.1	Sequência de Pulsos	33
3.3.2	Processamento das Informações Espectrais	37
3.4	FORMALISMOS QUIMIOMÉTRICOS	39
3.4.1	Figuras de Mérito	45
4	METODOLOGIA	47
4.1	SELEÇÃO DE PACIENTES E PROCEDIMENTOS ÉTICOS	47
4.2	ENSAIOS METABONÔMICOS	48
5	RESULTADOS E DISCUSSÃO	50
5.1	DADOS DOS PACIENTES	50
5.2	FORMALISMOS QUIMIOMÉTRICOS	51
5.3	VIAS METABÓLICAS	59
6	PERSPECTIVAS	63
7	CONCLUSÃO	64
	REFERÊNCIAS	66
	APÊNDICE A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE) DO ESTUDO	77
	APÊNDICE B – FICHA CADASTRAL DE PACIENTES DO AMBULATÓRIO DE PEDIATRIA E DE PUERICULTURA DO HOSPITAL DAS CLÍNICAS DE PERNAMBUCO	79
	APÊNDICE C – QUESTIONÁRIO PADRÃO DE COLETA DE DADOS DE PACIENTES DO AMBULATÓRIO DE PEDIATRIA	80

E DE PUERICULTURA DO HOSPITAL DAS CLÍNICAS DE PERNAMBUCO, BASEADO NO EISL	
APÊNDICE D – MODELOS DE TREINAMENTO E TESTE, RESPECTIVAMENTE, DA MODELAGEM COM O GRUPO DE SIBILÂNCIA RECORRENTE COMO GRUPO ALVO E MATRIZ DE CONTIGÊNCIA	81
APÊNDICE E – RESULTADO DO ÍNDICE PREDITOR DE ASMA MODIFICADO DOS VOLUNTÁRIOS DO ESTUDO	82
APÊNDICE F – CLASSIFICAÇÃO DOS PACIENTES VOLUNTÁRIOS DO ESTUDO SEGUNDO O mAPI, GA-LDA E DDSIMCA	83
ANEXO A – VIAS METABÓLICAS DA ALANINA-ASPARTATO-GLUTAMATO E D-GLUTAMINA E D-GLUTAMATO	84

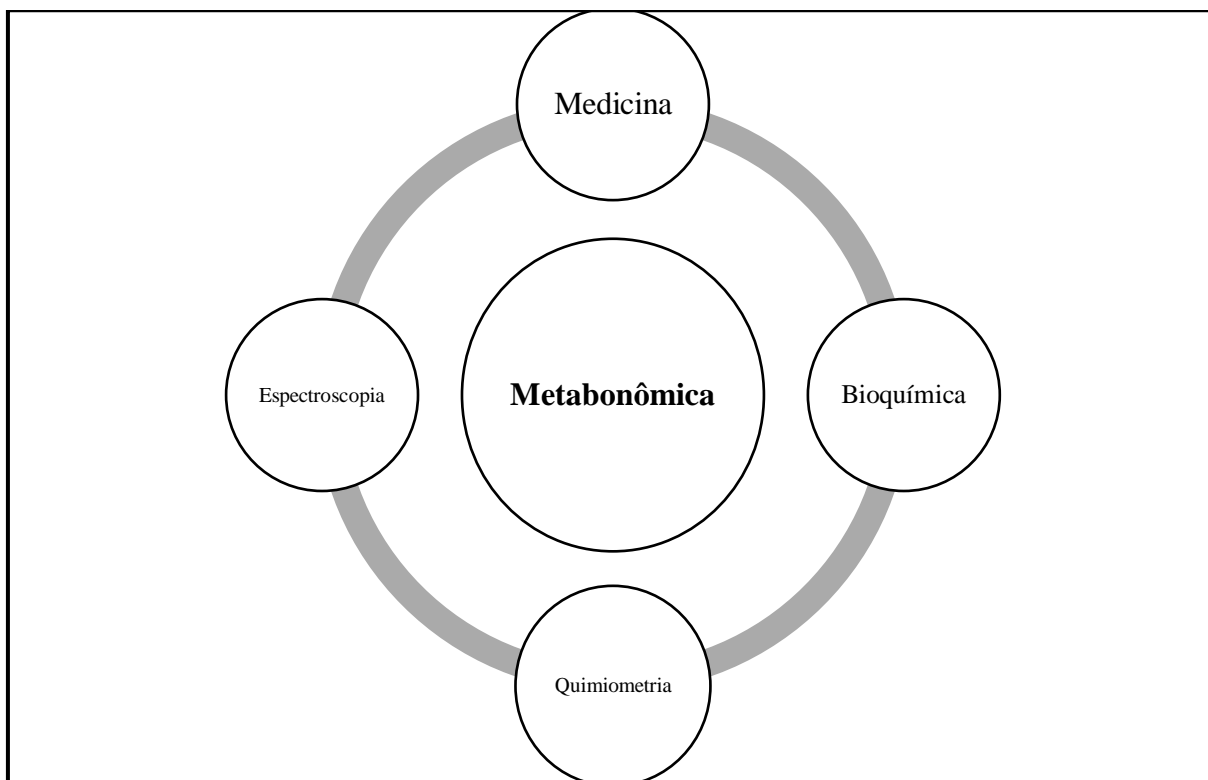
1 INTRODUÇÃO

Grandes mudanças nas ciências são esporadicamente experimentadas diante de inovações tecnológicas, modificação de protocolos e perfis profissionais que as implementem (ÁLVAREZ-FERNÁNDEZ e NÚÑEZ-REIZ, 2016). Estratégias adaptadas às características individuais de pacientes mostram-se promissoras para o diagnóstico de doenças e podem levar a uma melhoria da saúde da população, otimizando os resultados para cada paciente individualmente (SAVOIA *et al.*, 2017).

Avanços rápidos em capacidade computacional, criação de bancos de dados vinculados e novas terapias direcionadas estão tornando cada vez mais possível a prevenção ou o tratamento de doenças com base nas características individuais (WEIL, 2018). A prática de caracterizar um paciente em termos de genoma ou proteoma, pode lançar luz sobre sua resposta fisiológica a uma intervenção e é uma necessidade, dado o fato de que a variação interindividual clinicamente significativa existe (GOETZ e SCHORK, 2018).

Trabalhos que visam desenvolver um modelo diagnóstico, com uma abordagem focada nos metabólitos produzidos por um tipo de doença, são alvo de investigação das ciências ômicas – Metabonômica e Metabolômica. Definida em 1999, por Nicholson, Lindon e Holmes, a Metabonômica é a medida quantitativa da resposta metabólica multiparamétrica e dinâmica de sistemas vivos a estímulos fisiopatológicos ou modificação genética. Diferentemente da Metabolômica, que visa caracterizar todo o metaboloma de um sistema, a Metabonômica visa identificar padrões de alterações presentes mediante estímulos ou modificações no sistema em estudo. A Metabonômica é uma área de pesquisa multifacetada, implicando a interação de diferentes profissionais e expertises. Em algumas abordagens, tem-se uma integração entre a **Medicina**, que avalia uma fisiopatologia ou uma resposta homeostática à ação de um agente externo; a **Química Analítica**, que produz uma medida quantitativa da matriz biológica estudada e avalia, de maneira multiparamétrica, os dados coletados. O Quadro 1 traz uma representação geral dessa multidisciplinaridade.

Quadro 1 – Representação da Multidisciplinaridade da Metabonômica



Fonte: Autoria própria (2021).

Doenças respiratórias crônicas apresentam-se como um problema de saúde pública há anos, tendo registrado aproximadamente 3,91 milhões de mortes em 2017, sendo 7% de todas as mortes do mundo (LI *et al.*, 2020). Dentre estas, a asma merece destaque por ser a doença crônica mais comum presente em crianças (SOLÉ *et al.*, 2017). Sua prevalência vem aumentando ao redor do mundo, e esses indicadores de crescimento parecem estar conectados às taxas de industrialização e urbanização (RODRIGUES, 2007). Estudos longitudinais apontam que a asma é uma das causas de sibilância de repetição, por conseguinte, a identificação de possíveis fatores de risco pode possibilitar a instalação de condutas terapêuticas mais específicas (DELA BIANCA *et al.*, 2007).

A incidência de três ou mais episódios de sibilância, durante um ano descreve o quadro de sibilância recorrente (MALLOL *et al.*, 2010), sendo uma condição clínica associada aqueles pacientes que apresentam chiado no peito, ou sibilos. Bebês com doenças respiratórias associadas à sibilância no primeiro ano de vida apresentam níveis mais baixos de função pulmonar que aqueles que não apresentam um quadro de sibilo (MARTINEZ *et al.*, 1995). Desenvolver ferramentas que permitam o

reconhecimento de potenciais fatores associados à sibilância, principalmente na forma recorrente, seriam imprescindíveis para pacientes lactentes (ARANDA *et al.*, 2015).

A estratégia metabonômica faz uso da espectroscopia de Ressonância Magnética Nuclear (RMN) de ^1H para a avaliação de biofluidos, com preparo mínimo de amostra, sendo também uma técnica não destrutiva e rápida (DE GRAAF *et al.*, 2015; FAN *et al.*, 2018). Os espectros de RMN de ^1H do plasma e soro sanguíneos mostram bandas largas de sinais de proteínas e lipoproteínas, e outras macromoléculas que compõem o metaboloma de um indivíduo, considerando que as intensidades dos sinais observados podem ser editadas com base nos coeficientes de difusão molecular ou nos tempos de relaxação da RMN (LINDON e NICHOLSON, 2008). O metaboloma é a coleção de pequenas moléculas endógenas que marcam impressões digitais específicas da bioquímica celular (TURI *et al.*, 2018).

Os dados espectrais obtidos em estudos metabonômicos são tratados por métodos de estatística multivariada, buscando o reconhecimento de padrão associado ao *status* bioquímico dos pacientes e/ou voluntários. A Inteligência Artificial (IA) é uma poderosa ferramenta de reconhecimento de padrão e tem atraído a atenção de pesquisadores em diversas áreas (LIU, 2018). Aprendizagem de máquinas é uma forma de IA na qual um modelo aprende com base em exemplos e não regras pré-programadas (SHILLAN, 2019). Com isso, pode-se obter modelos precisos, exatos, reproduzíveis e de baixo custo de implementação, sendo possível ainda analisar dados à distância, possibilitando a cooperação entre pares.

No que diz respeito a modelos metabonômicos para diagnóstico de asma em lactentes que apresentam sibilos, ainda não existem trabalhos na literatura, como aponta Kelly e colaboradores (2017) e nossa busca na literatura. A metabonômica propicia uma abordagem eficiente e de baixo custo, sendo assim um método promissor para a compreensão da sibilância pelo desenvolvimento de testes de diagnóstico robustos, sensíveis e reproduzíveis (ZHANG, 2012).

2 OBJETIVOS

2.1 GERAL

Desenvolver modelos metabonômicos, a partir de dados espectrais de RMN de ^1H de soro, que sejam capazes de discriminar lactentes que apresentam quadro de sibilância recorrente daqueles que não apresentam sibilância recorrente.

2.2 ESPECÍFICOS

Obter os espectros de RMN de ^1H de amostras de soro dos lactentes com sibilância recorrente e com sibilância não recorrente;

Empregar técnicas de aprendizagem de máquinas para o reconhecimento de padrões, identificando padrões metabonômicos, discriminando lactentes que apresentam sibilância recorrente dos que apresentam sibilância não recorrente;

Validar os modelos metabonômicos construídos;

Identificar as regiões espectrais responsáveis pela discriminação e atribuí-las aos metabólitos endógenos produzidos pelos indivíduos do estudo;

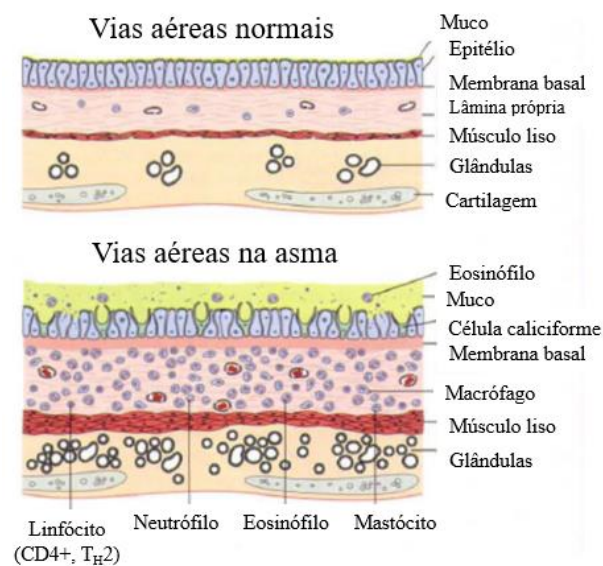
Identificar as vias metabólicas associadas ao quadro fisiopatológico dos indivíduos participantes do estudo.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 ASMA E SIBILÂNCIA: CARACTERÍSTICAS, CONDIÇÕES CLÍNICAS E SAÚDE PÚBLICA

A asma é um distúrbio inflamatório crônico das vias aéreas que causa episódios recorrentes de sibilos, falta de ar e opressão torácica (KUMAR, 2010). É causada por fatores genéticos e ambientais, e se manifesta com a contração reversível do músculo liso das vias aéreas e a hiper-reatividade brônquica (MATTIUZZI e LIPPI, 2020; GASTON, 2011) característica de processos inflamatórios (Figura 1).

Figura 1 – Esquema representativo das vias aéreas normais (A) e durante a asma (B) e estruturas morfológicas e celulares em cada situação



Fonte: Kumar (2010).

A sibilância de repetição é a característica principal da asma de início precoce (DELIU *et al.*, 2017). Um estudo internacional evidenciou que 45,2% dos lactentes desenvolvem pelo menos um episódio de sibilância no primeiro ano de vida e que 20,3% apresentam sibilância recorrente (MALLOL *et al.*, 2007).

A má percepção da gravidade da doença pelo paciente, familiares e/ou equipe médica, assim como as exacerbações, contribui para o aumento do risco de mortalidade (PITCHON *et al.*, 2018). Além disso, a exacerbação do quadro de asma é um dos fatores responsáveis pelos seus altos níveis de mortalidade e morbidade (DE BENEDICTIS e ATTANASI, 2016).

A ocorrência de três ou mais episódios de sibilância no último ano caracteriza uma sibilância recorrente (SOLÉ, 2008). Apesar de sibilos, também conhecidos como chiado no peito, serem sintomas comuns nos pacientes asmáticos, nem todos os lactentes sibilantes serão asmáticos no futuro (BACHARIER *et al.*, 2008).

Identificar precocemente quais dessas crianças sibilantes serão asmáticas no futuro é imprescindível devido a ansiedade da família, a qual deseja saber o diagnóstico e, principalmente, o prognóstico da doença na criança (FIRMIDA, 2013). Além do mais, já foi demonstrado que as crianças que desenvolviam sibilância durante os primeiros três anos de vida, e cujo sintomas persistiram até seis anos de idade, apresentavam níveis significativamente mais baixos da função pulmonar aos seis anos de idade, em comparação com crianças cujo sintomas de chiado começaram após os três anos de idade (MARTINEZ *et al.*, 1995). Sendo assim, o diagnóstico precoce permitiria a otimização do tratamento com objetivo de preservar a função pulmonar.

Embora existam ferramentas diagnósticas para a asma, ela segue subdiagnosticada nos países do MERCOSUL (SOLÉ *et al.*, 2017). Para países como Brasil, Bolívia e Venezuela que apresentam sistemas públicos de saúde (PEREIRA *et al.*, 2012), um diagnóstico precoce de asma pode levar a uma diminuição dos custos associados a essa doença.

Atualmente, para o diagnóstico de asma em crianças, é importante a confirmação da limitação das vias respiratórias, que pode ser avaliada pelo exame de espirometria (GINA, 2019) e por ferramentas preditivas tais como o Índice Preditivo de Asma (API, do inglês *Asthma Predictive Index*) e a Prevenção e Incidência de Asma e Alergia a Ácaro (PIAMA, do inglês *Prevention and Incidence of Asthma and Mite Allergy*) (COLICINO *et al.*, 2019). No que diz respeito às ferramentas preditivas, estas apresentam uma razão de verossimilhança negativa relativamente baixa, o que o torna uma ferramenta ineficiente para prever a asma

(LI *et al.*, 2019; CASTRO-RODRIGUEZ *et al.*, 2019). Essa razão relaciona a proporção da probabilidade do resultado de um teste em pessoas que têm a doença com a probabilidade de pessoas que não a têm (DEEKS e ALTMAN, 2004). No que diz respeito ao teste de espirometria, este exige técnicas de expiração forçada, que só poderiam ser executadas com segurança a partir de crianças de 5 anos (BACHARIER *et al.*, 2008).

O Estudo Internacional de Sibilância em Lactentes (EISL) multicêntrico, realizado entre 2005 e 2007, teve como objetivo determinar a prevalência de sibilância recorrente durante o primeiro ano de vida (DELA BIANCA *et al.*, 2007). O EISL identificou que 45,2% dos 30.093 lactentes avaliados apresentaram pelo menos 1 episódio de sibilância no primeiro ano de vida e 20,3% apresentaram sibilância recorrente (MALLOL *et al.*, 2010). Este questionário por sua vez apresenta sensibilidade, especificidade, valor preditivo positivo e valor preditivo negativo de 86%, 91,8%, 76,8% e 95,4%, respectivamente (CHONG NETO *et al.*, 2007). Entretanto, apesar das variáveis clínicas apresentadas pelo EISL, fatores genéticos ou ambientais, quer façam parte da anamnese, do exame físico ou de meios complementares de diagnóstico, não apresentam poder suficiente para serem usados como marcador único na identificação de asma no futuro em pacientes lactentes (MEDEIROS *et al.*, 2011). Com isso, a busca de biomarcadores para pacientes lactentes que apresentem quadro de sibilância recorrente ou não recorrente constituem uma importante área a ser explorada.

3.2 METABONÔMICA: O ESTUDO DO PERFIL METABÓLICO

A ideia de que vários estados de doença podem ser refletidos por mudanças nas concentrações de metabólitos é fundamental para compreender a metabonômica e remete aos povos da Grécia Antiga (MADSEN *et al.*, 2010). Esse pensamento também permeou pensadores da Idade Média como Ulrich Pinder (LEHMAN-MCKEEMAN, 2005), que associava cores, cheiros e sabores a urinas em diferentes patologias (Figura 2).

Figura 2 - “Roda da urina” publicada no livro *Epiphanie Medicorum* em 1506 por Ulrich Pinder



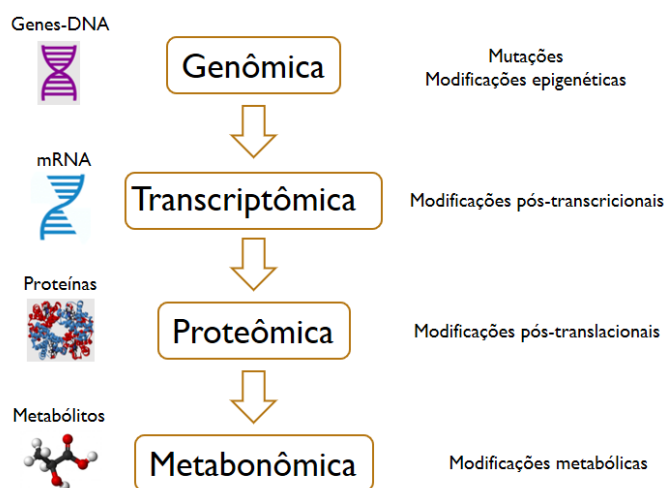
Fonte: Lehman-McKeeman (2005).

Ao longo das décadas de 1970 e 1980, conforme os espectrômetros de RMN melhoraram e as intensidades dos campos magnéticos aumentaram, a avaliação de perfis metabólicos empregando essa técnica começou a atrair a atenção de outros cientistas (WISHART, 2019). Lindon, Nicholson e Holmes (1991, p.1881) definiram o termo “metabonômica” da seguinte maneira: “(...) é a medida quantitativa da resposta metabólica, multiparamétrica, dependente do tempo, de sistemas vivos a estímulos fisiopatológicos ou modificações genéticas”.

Enquanto isso, o termo metabolômica foi introduzido em 2001, por Oliver Fiehn, como sendo a análise abrangente e quantitativa do metaboloma de um sistema biológico após estímulos fisiopatológicos ou modificações genéticas (CANUTO *et al.*, 2018). Atualmente, metabonômica e metabolômica são termos usados como sinônimos (MADSEN *et al.*, 2010), mas há uma diferença metodológica importante: a análise metabolômica exige a utilização de uma ferramenta de separação de misturas, normalmente a Cromatografia Líquida de Alta Eficiência (CLAE), ou Cromatografia em fase Gasosa (CG), associada à espectrometria de massas; na metabonômica, a análise é feita com mínima intervenção na amostra, normalmente usando RMN (DUNN e ELIS, 2005).

Juntamente com a genômica, transcriptômica e proteômica, a metabonômica fornece informações específicas que permitem a compreensão de vias metabólicas em processos patológicos (ZHANG, 2012) (Figura 3).

Figura 3 - Objetos de pesquisa das ciências ômicas numa abordagem que parte da informação genética até a manifestação metabólica de biocomponentes



Fonte: Autoria própria (2020).

Para desenvolver um estudo metabonômico, um problema biológico a ser estudado deve ser previamente definido, através da elaboração de uma ou mais perguntas a serem elucidadas ao final do estudo (CANUTO *et al.*, 2018). As etapas de uma investigação do perfil metabólico envolvem a coleta da amostra biológica, análise instrumental, processamento dos espectros, pré-processamento das amostras, análise multivariada de dados e a identificação dos metabólitos.

O uso de amostras de sangue tem como vantagem, considerando outros biofluidos, uma menor alteração diária de componentes (GHINI *et al.*, 2019), o que não ocorreria com amostras de urina. Além disso, existe a conveniência do uso de pequenos volumes para as amostras, o que pode ser uma vantagem se empregado na prática clínica.

A espectroscopia de RMN de ^1H e a espectrometria de massas são as ferramentas mais comumente usadas para medir o metaboloma, mas as técnicas de espectroscopia óptica, como espectroscopia Raman e espectroscopia na região do infravermelho (IV) também já foram relatadas na literatura (MADSEN, 2010).

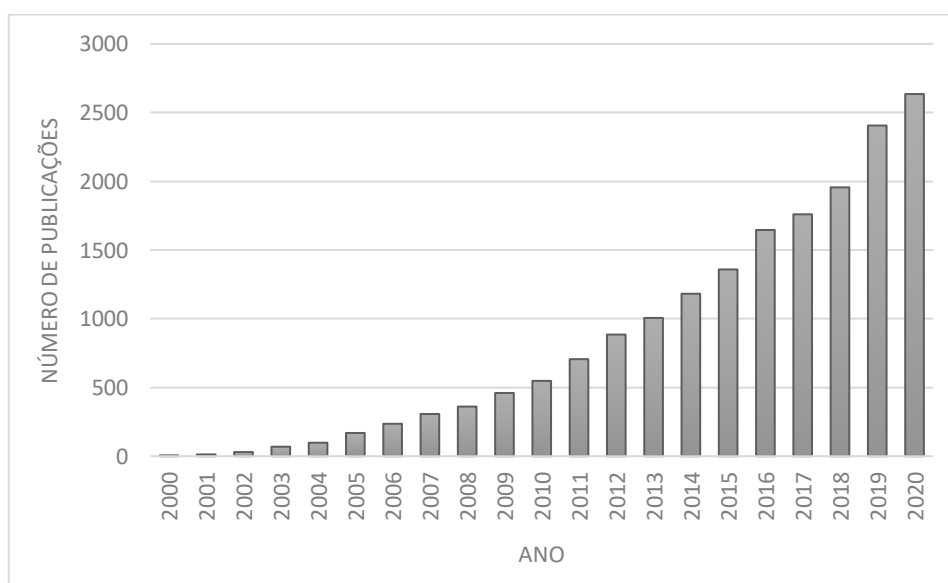
Geralmente, na metabonômica baseada em espectroscopia de RMN de ^1H , poucas etapas (ou nenhuma) de pré-processamento da amostra são requeridas, podendo ser empregadas amostras sólidas ou líquidas, além de tratar-se de uma técnica não destrutiva, quando comparada com a espectrometria de massas (CANUTO *et al.*, 2018).

Todos os estudos metabonômicos resultam em conjuntos de dados multivariados complexos que são então analisados a partir de ferramentas de reconhecimento de padrão que permitem a identificação de “impressões digitais” bioquímicas e a classificação de amostras (LINDON e NICHOLSON, 2008; BECKONERT *et al.*, 2007).

3.2.1 Estado da Arte

A metabonômica é um campo de pesquisa multifacetado, integrando Ciências Biológicas, Ciências Médicas, Química e *Data Science*, possuindo um enorme potencial para que grupos de pesquisa trabalhem em cooperação. Uma busca realizada na Web of Science, com os termos “*metabonomics or metabonomic or metabolomics or metabolomic*” no título dos trabalhos, nos anos de 2000 a 2020, resultou em 17837 documentos (artigos, resumos de conferências ou revisões). Esses dados são apresentados no Gráfico 1.

Gráfico 1 – Números de artigos publicados por ano contendo o termo *Metabonomic(s)* ou *Metabolomic(s)* no título. Período considerado: de 2000 a 2020



Fonte: Autoria própria a partir de dados do Web of Science (2021).

Os modelos diagnósticos disponíveis na literatura abrangem diversas doenças, tais como hepatite C viral (GODOY *et al.*, 2010), hepatite B crônica (GILANY *et al.*, 2019), câncer geniturinário (MANZI *et al.*, 2020), fibrose cística (ZANG *et al.*, 2019) e diagnóstico diferencial de infertilidade na presença de varicocele (LIRA NETO *et al.*, 2020). Nos estudos metabonômicos, a avaliação da capacidade preditiva de modelos de classificação baseia-se nos parâmetros de exatidão, sensibilidade e especificidade (SOARES, 2016).

Modelos metabonômicos para o tratamento de asma podem utilizar como biofluido o condensado do ar exalado (EBC, do inglês *Exhaled Breath Condensate*), urina ou o soro de pacientes. Um estudo de revisão sistemática realizado por Kelly e colaboradores (2017) comparou estudos metabonômicos que tinham a asma como objeto de pesquisa. Dentre os artigos avaliados, num total de vinte e um, onze autores utilizaram como biofluido o EBC, enquanto quatro utilizaram urina, dois utilizaram soro e quatro utilizaram plasma. Utilizando os valores de exatidão, os autores da revisão sistemática classificaram os trabalhos como bons ($\geq 80\%$) ou muito bons ($\geq 85\%$).

A escolha do soro do paciente pode otimizar o diagnóstico na prática clínica, uma vez que geralmente são pedidos exames de dosagem de IgE para pacientes asmáticos (BACHARIER *et al.*, 2008), também avaliados no soro. Além disso, uma vez que a espectroscopia de RMN é uma técnica reprodutível, adequa-se bem como ferramenta diagnóstica num ambiente clínico (MADSEN, 2010).

3.3 RESSONÂNCIA MAGNÉTICA NUCLEAR

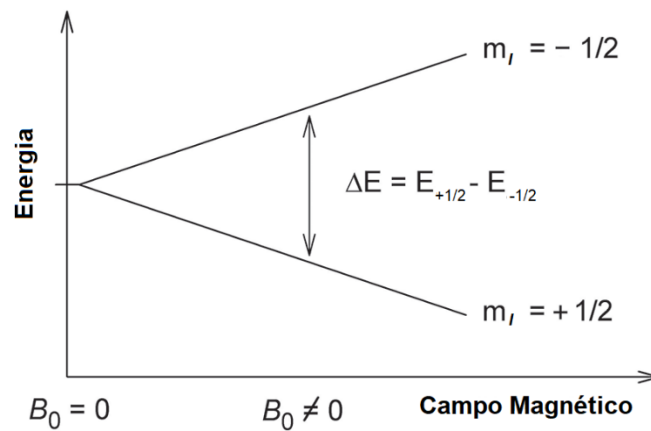
Em 1896, Pieter Zeeman demonstrou uma nova relação entre luz e magnetismo, ao provar que um campo magnético externo provoca a modificação da emissão e absorção de radiação luminosa (RHODES, 2017; INGLE e CROUCH, 1998). Décadas mais tarde, Isidor Rabi e seus colaboradores observaram que um campo magnético oscilante poderia provocar a reorientação do spin nuclear e do momento magnético (RABI *et al.*, 1938). Bloembergen, Purcell e Pound, em 1947, publicaram o primeiro estudo que aplicava a RMN à amostras líquidas, abrindo assim espaço para maiores aplicações dessa técnica. Este estudo empregou em seus cálculos a constante magnetogírica, que é definida como uma propriedade

nuclear intrínseca, referente a razão entre os momentos magnético e angular do núcleo (Bloembergen, Purcell e Pound, 1947; MLYNÁRIK, 2017).

Tais descobertas contribuíram para traçar as bases teóricas da espectroscopia de Ressonância Magnética Nuclear. Nesses estudos, provou-se a fundamental importância do spin nuclear para o fenômeno da RMN (CLARIDGE, 2009). Para cada núcleo com spin diferente de zero, o número de estados de spin permitidos que se pode quantizar é determinado por seu número quântico de spin nuclear (I). Para cada número I existem $(2I+1)$ estados, podendo variar de $+I$ a $-I$ (BLOEMBERGEN, *et al*, 1947; PAVIA, 2009). Em átomos que não apresentem número atômico e número de massa pares, os números quânticos de momento angular orbital e do spin dão origem a momentos magnéticos (μ), o que significa que, de certo modo, eles se comportam de forma semelhante a pequenos ímãs (ATKINS, 2008). Deste modo, somente núcleos com spin nuclear (I) diferente de zero possuem atividade na RMN (ROCHA, 2013).

Na ausência de um campo magnético externo, os estados de spin são degenerados, perdendo essa degenerescência na presença de um campo magnético externo (SILVA, 2017). No caso do núcleo de hidrogênio-1, os estados de spin possíveis são $+1/2$ e $-1/2$, sendo um desses estados correspondente à orientação paralela ao campo magnético, e o outro à orientação antiparalela (MLYNÁRIK, 2017). Tal orientação de momentos magnéticos nucleares num campo magnético estático B_0 dá origem a um vetor de magnetização, que é a soma vetorial de momentos magnéticos nucleares individuais (MLYNÁRIK, 2017). Ao aplicar um pulso de radiação eletromagnética, com frequência correspondente à diferença de energia entre ambos estados, é promovida a transição de spins do estado de menor energia para o de maior energia (HEMMINGA, 1992) (Figura 4). A diferença de energia aplicada é proporcional à força do campo magnético externo e à constante magnetogírica do núcleo (HEMMINGA, 1992).

Figura 4 - Diagrama de energia apresentando os estados de spin (m_I), como função do campo magnético externo (B_0)



Fonte: Autoria própria (2020).

Uma vez que a absorção de energia é um processo quantizado, a energia absorvida deve ser igual à diferença de energia entre os dois estados envolvidos (PAVIA, 2009). As equações 1 e 2 demonstram como ocorre a quantização dessa energia em núcleos de spin 1/2.

$$\Delta E = E_\beta - E_\alpha = h\nu \quad (1)$$

$$\Delta E = \left(-\frac{1}{2}\gamma\hbar B_0\right) - \left(+\frac{1}{2}\gamma\hbar B_0\right) \therefore |\Delta E| = \gamma\hbar B_0 \quad (2)$$

Sendo \hbar a unidade de quantização do momento angular do núcleo, ($h/2\pi$), e h a constante de Planck. Outro termo que aparece na Equação 1 é a constante magnetogírica (γ), que é uma propriedade intrínseca do núcleo atômico e determina a dependência de energia com o campo magnético (PAVIA, 2009), como descrita na Equação 3.

$$\gamma = \frac{2\pi\mu}{hI} \quad (3)$$

A condição de ressonância é atingida quando a frequência da radiação eletromagnética se iguala a frequência de Larmor (ATKINS, 2008), como descrito nas equações abaixo.

$$\Delta E = h\nu \quad (4)$$

$$\Delta E = \frac{h}{2\pi} \gamma B_0 \quad (5)$$

$$\nu = \frac{\gamma B_0}{2\pi} \quad (6)$$

A Equação 6 é a equação fundamental da Ressonância Magnética Nuclear. Na presença de um campo magnético, a distribuição (população) não é igualitária para os estados possíveis de energia, ficando o estado de menor energia ligeiramente em excesso (BARROS, 2017). Com isso, as populações obedecem a distribuição de Boltzmann (Equação 7).

$$\ln\left(\frac{N_\alpha}{N_\beta}\right) = \frac{-\Delta E}{k_B T} \quad (7)$$

Sendo N_α e N_β as populações nos estados α e β , respectivamente, k_B a constante de Boltzmann, T é a temperatura absoluta (K) e ΔE a diferença de energia entre os estados. A diferença de população entre os estados de spin impacta diretamente na intensidade do sinal de RMN, sendo, portanto, diretamente proporcional às concentrações das espécies na amostra.

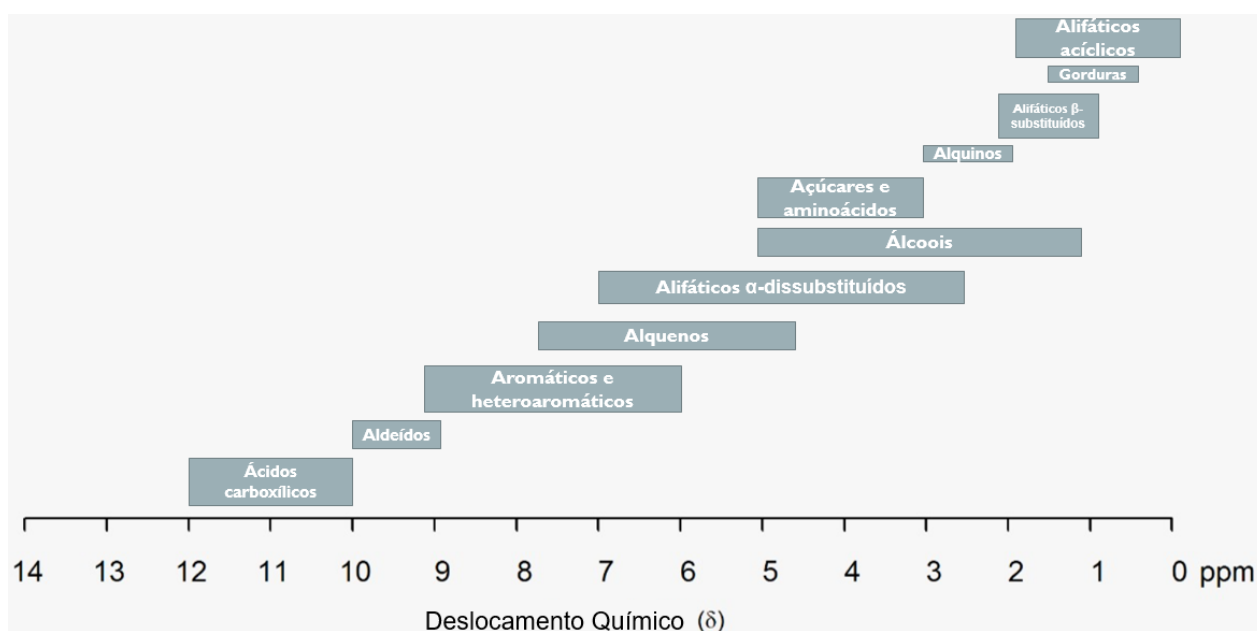
Sistemas moleculares são formados por átomos, e não simplesmente núcleos, ligados entre si através de densidades eletrônicas e de interações eletrostáticas (COSTA, 2016). Sendo assim, é necessário considerar os efeitos de densidades eletrônicas e de outros núcleos atômicos na frequência de ressonância de cada núcleo. Quanto maior a densidade eletrônica em torno de um núcleo, maior será o campo contrário induzido e se opondo ao campo aplicado. Desta forma, este campo contrário que protege o núcleo diminui a intensidade do campo magnético aplicado ao núcleo observado. Consequentemente, núcleos em diferentes ambientes químicos apresentam frequências diferentes. Esse fenômeno é chamado de deslocamento químico, sendo representado pela letra grega delta (δ) (HEMMINGA, 1992). Usualmente, o δ é expresso em ppm (partes por milhão) e, no caso dos núcleos de ^1H e ^{13}C , o sinal atribuído às metilas do tetrametilsilano (TMS) é usado como referência interna de deslocamento químico (MLYNÁRIK, 2017). A

Equação 8 demonstra como a medida do deslocamento químico é obtida (PAVIA, 2009).

$$\delta = \frac{\nu_{observado} - \nu_{referencia}}{\nu_{referencia}} \times 10^6 \quad (8)$$

Logo, núcleos em diferentes ambientes químicos possuem diferentes deslocamentos químicos, e, por conseguinte, os que se apresentam em um mesmo ambiente químico possuem o mesmo deslocamento químico (GOUVEIA, 2017). A Figura 5 traz os deslocamentos químicos observados para alguns grupos funcionais orgânicos.

Figura 5 - Deslocamentos químicos de ^1H (δ_{H}) presentes em grupos funcionais orgânicos ao longo do eixo de um espectro de RMN.



Fonte: Adaptado de Pavia (2009).

3.3.1 Sequência de pulsos

O núcleo mais importante para análise de biomoléculas com atividade na RMN é o de ^1H , sendo relatados alguns estudos com ^{13}C , ^{15}N , porém estes núcleos apresentam baixa abundância natural e o núcleo de ^{31}P apesar da alta abundância natural (Tabela 1), os sinais da maioria dos compostos fosforilados se sobrepõem (MARKLEY *et al.*, 2017). Se os picos de um espectro de RMN apresentarem boa

resolução e uma relação sinal ruído aceitável, suas intensidades podem se correlacionar linearmente com sua concentração relativa (MARKLEY *et al.*, 2017).

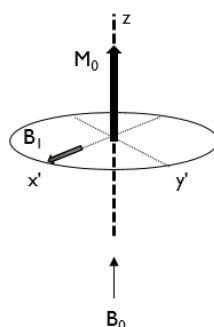
Tabela 1 – Características de alguns núcleos estudados por espectroscopia de RMN.

Núcleo	^1H	^2H	^{13}C	^{14}N	^{15}N	^{31}P
Constante magnetogírica γ (radianos/Tesla)	267,53	41,1	67,28	19,3	-27,1	108,3
Abundância natural	99,985	0,015	1,11	99,63	0,37	100,00

Fonte: adaptado de Rocha (2013) e Pavia (2009).

Para que seja possível a realização de experimentos via RMN, é necessário que haja um mecanismo de interação entre a radiação eletromagnética e o sistema em estudo (MONARETTO, 2019). No experimento de RMN, um campo magnético oscilante B_1 na frequência de Larmor é aplicado perpendicularmente a B_0 na forma de um pulso de radiofrequência. O pulso faz com que a magnetização M_0 , que estava originalmente na direção do campo B_0 , no eixo z , gire num ângulo θ , gerando assim magnetização no plano $x'y'$ (COLNAGO e ANDRADE, 2016) (Figura 6).

Figura 6 – Representação do fenômeno da RMN num sistema de coordenadas girantes. O campo magnético oscilante B_1 aplicado perpendicularmente a B_0 é utilizado para excitar a amostra



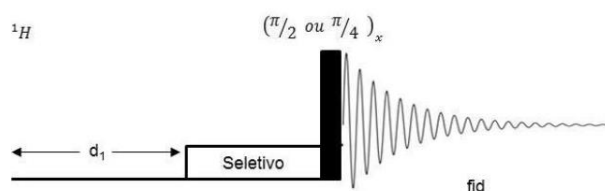
Fonte: adaptado de Colnago e Andrade (2016).

A Equação 9 permite determinar o ângulo θ entre o vetor magnetização e o eixo z após aplicação de um pulso de radiofrequência (RF) pelo tempo t_p .

$$\theta = \gamma B_1 t_p \quad (9)$$

A sequência de pulsos muda temporariamente o comportamento evolutivo do sistema de spin nuclear, às vezes imitando a remoção de algumas das interações de spin nuclear (LEVITT, 2008). A sequência de pulsos HPRESAT consiste em um pulso relativamente longo e de baixa potência para saturar seletivamente uma frequência específica, geralmente a frequência atribuída ao sinal da água, e outro pulso não seletivo de 45° ou 90° para excitar as frequências de interesse (GOUVEIA, 2017) (Figura 7).

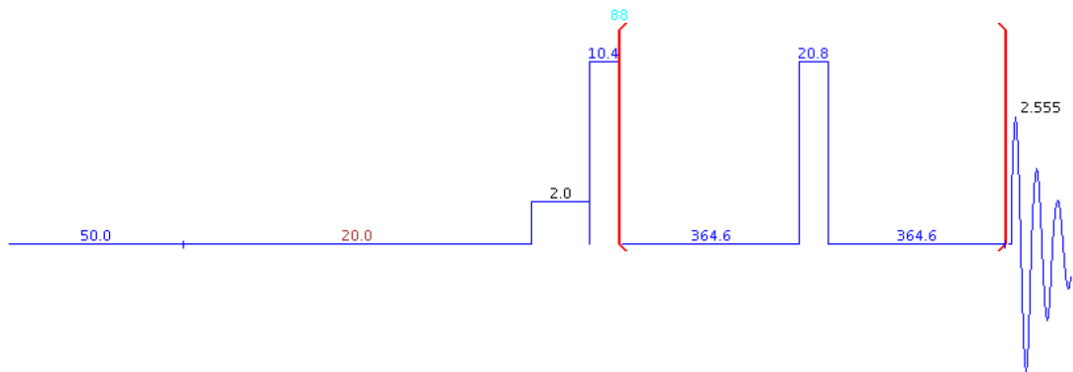
Figura 7 – Sequência de pulsos no experimento de HPRESAT para suprimir o sinal da água



Fonte: Costa (2016).

Um problema secundário à abundante quantidade de água na análise de biofluidos é a presença de macromoléculas, tais como lipídeos e proteínas (VAN *et al.*, 2003). A presença desses biocompostos numa amostra submetida à espectroscopia de RMN de ^1H pode resultar em espectros com sinais alargados, o que leva, muitas vezes, a uma maior dificuldade para a visualização de sinais menos intensos (GOUVEIA, 2017). Tal adversidade pode ser contornada aplicando-se a sequência de pulsos desenvolvida por Carr-Purcell-Meiboon-Gill (CPMG), que explora a taxa de relaxação transversal (T_2) dessas macromoléculas, funcionando como um filtro de T_2 (Figura 8).

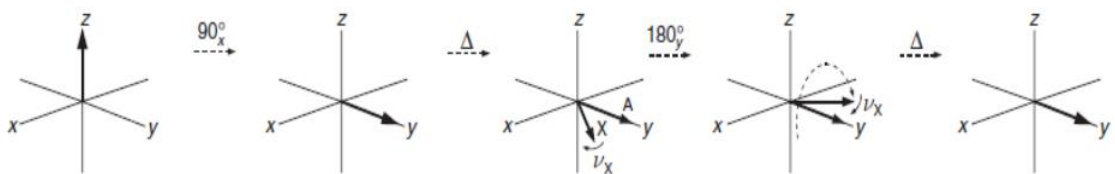
Figura 8 – Sequência de pulsos CPMG com sinal HPRESAT. Em preto tempo em segundos, em azul tempo em microssegundos, em marrom tempo em milissegundos, em ciano número de ciclos



Fonte: Autoria própria (2021).

Assim como na sequência HPRESAT, ao aplicar-se um pulso de RF de 90° , ao longo do eixo x, o vetor magnetização tem sua orientação alterada e passa a ser analisado considerando as componentes no eixo z e no plano xy (SILVA, 2017). Em seguida, são aplicados sucessivos pulsos de 180° com intervalos de tempo τ , até que os spins com valores menores de T_2 tenham relaxado, e o espectro é adquirido contendo apenas as ressonâncias de interesse (COSTA, 2016). Na sequência apresentada na Figura 8, a supressão do sinal da água é realizada aplicando um pulso seletivo por 2,0 s, enquanto espécies com T_2 menor que 400 ms não são observados no espectro porque o sinal de RMN é “filtrado”, pois as magnetizações dos núcleos de menor T_2 são perdidas durante os ciclos de eco de spin. A Figura 9 apresenta um esquema com a dinâmica associada ao vetor magnetização, ao aplicar a sequência de pulsos CPMG como filtro de T_2 .

Figura 9 – Esquematização da sequência de pulsos Carr-Purcell-Meiboom-Gill



Fonte: Silva (2017).

Inicialmente, aplica-se um pulso de RF de 90° , mudando a orientação do vetor magnetização, que agora está no plano xy. Durante o tempo Δ (ou T), as componentes no plano xy evoluem, de forma que o vetor resultante no plano xy diminua ao longo do tempo. Segue-se a aplicação de um pulso de RF de 180° , gerando o “eco de spin”, pois a evolução das componentes no plano xy fazem com que o vetor resultante no plano xy aumente de intensidade com o tempo. Após o tempo Δ , a intensidade do sinal começa a decair e segue-se a aquisição ou aplica-se novo pulso de RF de 180° . A definição do tempo Δ é muito importante, pois espécies que tenham um T_2 menor que o valor de Δ perderão a magnetização no plano xy antes da aplicação do pulso de RF de 180° . Com isso, os sinais atribuídos a essas espécies não serão observados no espectro.

3.3.2 Processamento das Informações Espectrais

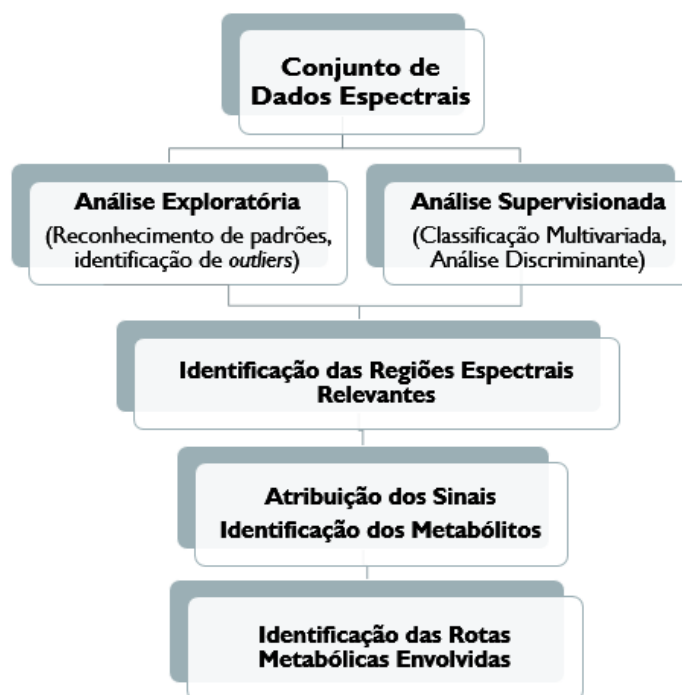
Comumente, dados espectrais extraídos da RMN para uma análise metabonômica são analisados nas seguintes etapas: processamento de dados pós-instrumental, tais como a transformada de Fourier, ajuste de fase e correção da linha de base; quantificação de características espectrais comumente implementadas via *binning*; normalização e escalonamento; e modelagem quimiométrica multivariada estatística de dados (ANDERSON *et al.*, 2011). A transformada de Fourier é responsável por recuperar as frequências de ressonância, presentes no domínio do tempo, num decaimento livre de indução (ATKINS, 2008). No entanto, se a magnetização for invertida no plano xy, e não perfeitamente alinhada ao longo do eixo x é necessário recorrer ao ajuste de fase (IZQUIERDO-GARCÍA *et al.*, 2011). Na técnica de *binning*, os espectros são divididos em janelas uniformemente espaçadas, ou bin, com base em conhecimento prévio ou de maneira estocástica utilizando um algoritmo (EMWAS *et al.*, 2018), permitindo que seja realizada uma redução de dados agrupando respostas espectrais. As intensidades dentro de cada bin são somadas, de modo que a área sob cada região espectral é usada em vez de intensidades individuais (SOUZA *et al.*, 2013). Os dados são reunidos numa matriz, de forma que cada linha contenha os dados espectrais de uma amostra; enquanto cada coluna contém a intensidade do sinal observada para cada bin. As intensidades dos sinais presentes no espectro podem variar em ordens de magnitude, levando a um viés ou a uma influência indevida de metabólitos altamente

concentrados nos resultados de um estudo metabonômico (EMWAS *et al.*, 2018). As variações podem ocorrer de maneira intra- e inter-amstral. Para contornar esse problema, realiza-se uma normalização por escalamento dos dados com base na fórmula da Equação 10, na qual x é o valor da variável que é subtraído da média (da linha ou coluna) e então dividido pelo desvio padrão (da linha ou da coluna).

$$\tilde{x}_{ij} = \frac{x_i - \bar{x}}{s} \quad (10)$$

A escolha de ser na linha ou na coluna está relacionada se a variação que dificulta ou insere um erro à modelagem se dá na amostra ou no conjunto de amostras, respectivamente. O autoescalamento pelas linhas da matriz de dados é chamado também de Padronização Normal de Sinal (SNV, do inglês *Standard Normal Variate*) (DA-COL *et al.*, 2017). Após a etapa de normalização, os dados podem ser analisados a partir da matriz de correlação. Técnicas de estatística multivariada são utilizados no conjunto de dados espectrais, devido à vasta dimensionalidade, visando identificar padrões naturais de agrupamentos. Isso pode ser feito a partir de análises exploratórias ou de reconhecimento de padrão e/ou análises classificatórias supervisionadas, que são usadas para extrair a informação relevante dos dados de entidades ou metabólitos responsáveis por diferenciar os grupos de amostras (CANUTO *et al.*, 2018). Uma vez identificado padrões ou construídos modelos supervisionados para discriminar as amostras, segue-se para a identificação das regiões espectrais mais relevantes para a discriminação e, depois, para a atribuição dessas regiões aos metabólitos (Figura 10).

Figura 10 – Sequência básica para extração de dados em ensaios metabonômicos



Fonte: Autoria própria (2021).

3.4 FORMALISMOS QUIMIOMÉTRICOS

As últimas duas décadas testemunharam uma explosão de dados decorrentes do desenvolvimento e maturação gradual de tecnologias ômicas e bioinformática (BOJA *et al.*, 2014). Definida em 1974 por Svante Wold como “A arte de extração de informações quimicamente relevantes de dados produzido em experimentos químicos” (WOLD, 1995), a quimiometria emergiu da necessidade de extrair informação química da avalanche de dados produzidos pela moderna instrumentação (FERREIRA, 2015). Atualmente, abrange a análise de dados químicos, para calibração e classificação de dados, e o planejamento de experimentos químicos e simulações (IUPAC, 2021). Devido à variabilidade e riqueza de dados multivariados, é quase impossível reconhecer padrões no conjunto de dados em estudo de maneira direta (LIU *et al.*, 2018).

A aprendizagem de máquinas ou *machine learning* (ML) é um ramo bem estabelecido da Inteligência Artificial (IA) (LAPADULA *et al.*, 2020). Trata-se de uma forma de IA na qual um modelo aprende com base em exemplos e não regras pré-

programadas (SHILLAN *et al.*, 2019). Dentre os métodos de *machine learning*, geralmente os mais empregados para a construção de modelos quimiométricos são o de Análise de Componentes Principais (PCA), Análise Discriminantes por Mínimos Quadrados Parciais (PLS-DA), Análise discriminante por projeções ortogonais à estrutura latente (OPLS-DA) e a Análise Discriminante Linear (LDA).

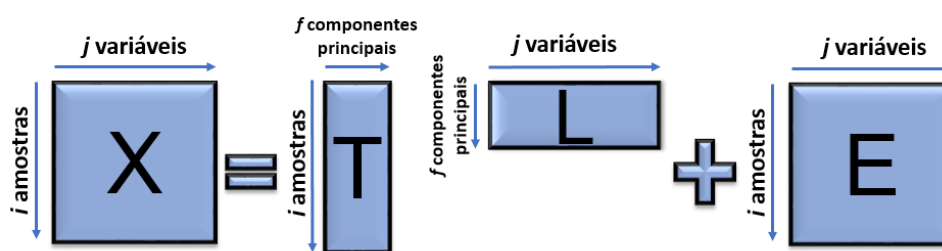
A análise de componentes principais é uma ferramenta quimiométrica de análise multivariada utilizada para mitigar redundâncias e reduzir a dimensionalidade do conjunto de variáveis, por meio da criação de uma nova base, cujas componentes são linearmente independentes e em menor número, comparando em relação ao conjunto inicial de dimensões (ROSSI, 2017).

Como uma ferramenta de análise exploratória, a PCA permite revelar a existência ou não de amostras anômalas, de relações entre as variáveis medidas e de relações ou agrupamentos entre amostras (LYRA, 2010). A realização da PCA consiste em fatorar a matriz de dados \mathbf{X} , como demonstrada na Equação 8. Nesta matriz, \mathbf{X} é a matriz de dados originais de entrada (*input*), o número de linhas é igual ao número de casos (amostras) e o número de colunas é igual ao número de variáveis (espectrais).

$$\mathbf{X} = \mathbf{T}\mathbf{L}^T + \mathbf{E} \quad (8)$$

Em que \mathbf{L} é a matriz dos pesos, \mathbf{T} a matriz dos escores e \mathbf{E} a matriz dos resíduos e \mathbf{T} sobrescrito corresponde a transposição da matriz (LYRA *et al.*, 2010) (Figura 11).

Figura 11 – Esquema de uma matriz hipotética $X_{(i,j)}$ no cálculo de PCA

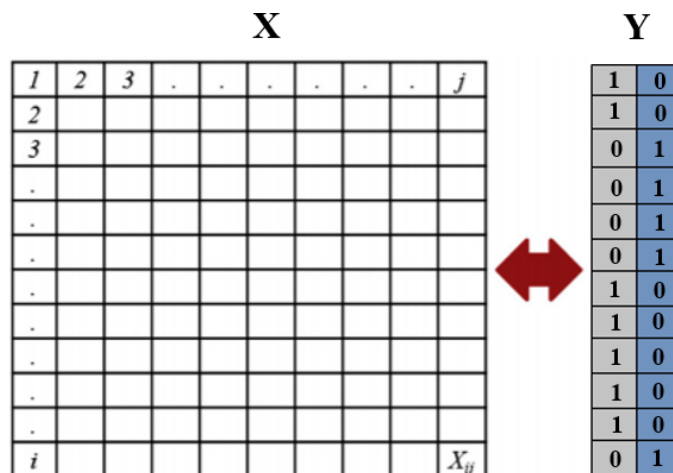


Fonte: Autoria própria (2021).

A PCA, entretanto, trata-se de um método exploratório não supervisionado. Para a construção de modelos de classificação são necessários métodos supervisionados. Nos métodos supervisionados, os modelos são construídos utilizando amostras com características conhecidas e, subsequentemente, o modelo é utilizado para prever a classe de amostras desconhecidas (SANTANA, 2020). São exemplos de métodos supervisionados o *Soft Independent Modeling of Class Analogy* (SIMCA), *K-Nearest Neighbor* (KNN), PLS-DA, OPLS-DA e LDA.

O método de classificação PLS-DA utiliza a técnica de regressão multivariada por mínimos quadrados parciais (PLS) (SANTANA *et al.*, 2020). Para a construção do modelo por PLS, os dados são organizados em dois blocos: o conjunto de dados é representado numa matriz \mathbf{X}_{ij} em que as i linhas representam as amostras e as j colunas as variáveis instrumentais e a matriz \mathbf{Y}_{ixn} que contém os valores para a propriedade de interesse que está sendo avaliada (LOPES, 2015) e sua quantidade de colunas depende da quantidade de classes em estudo. Tanto a matriz \mathbf{X} quanto a matriz \mathbf{Y} são decompostos em novas variáveis, denominadas variáveis latentes (VL) (Figura 12).

Figura 12 – Representação das matrizes utilizadas no PLS-DA. X contendo os dados originais e Y, os índice de classe 0 (pertencente a classe analisada) e 1 (não pertencente) do exemplo, i representa as amostras e j as variáveis



Fonte: Adaptado de Groski *et al.* (2015).

Como o método de PLS maximiza a relação entre a variável dependente e os escores, pode-se afirmar que as variáveis latentes representam as direções que melhor discriminam as classes em estudo (FERREIRA, 2015). Essa decomposição em variáveis latentes acontece da seguinte maneira (Equação 9):

$$\mathbf{X} = \mathbf{T}_A \mathbf{P}_A^T + \mathbf{E} \text{ e } \mathbf{Y} = \mathbf{T}_A \mathbf{q}_A^T + \mathbf{F} \quad (9)$$

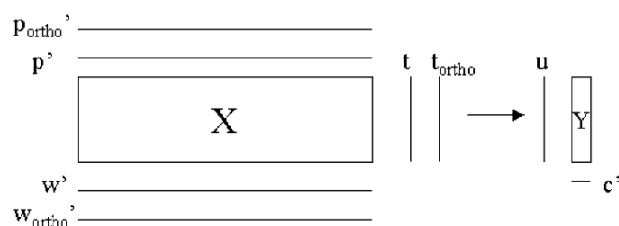
Nesta equação, A é o número de variáveis latentes. A matriz \mathbf{P}_A e o vetor \mathbf{q}_A são denominados pesos (do inglês *loadings*) da matriz \mathbf{X} e \mathbf{Y} , respectivamente (LOPES, 2015), sendo o T sobrescrito correspondente às suas transposições. A matriz \mathbf{T}_A é a matriz de escores e \mathbf{E} e \mathbf{F} são os resíduos.

A incorporação de uma etapa de filtração ortogonal no método de PLS-DA, o OPLS-DA, pode alcançar a separação da variação sistemática de interesse eliminando variações derivadas de fontes de confusão (HOLMES *et al.*, 2006). Com isso, informações não correlacionadas da matriz \mathbf{X} em relação a matriz \mathbf{Y} são removidas (BYLESJÖ *et al.*, 2006) o que é uma propriedade útil para a identificação de fontes desconhecidas da variação dos dados bem como para a identificação das variáveis preditivas do modelo (BYLESJÖ e RANTALAINEN, 2009). O método proposto por Trygg e Wold (2001) para cálculo da OPLS-DA é abordado na Equação 10.

$$\mathbf{X} = \mathbf{t}_p \mathbf{p}_p^T + \mathbf{T}_o \mathbf{P}_{y_o}^T + \mathbf{E} \quad (10)$$

Sendo \mathbf{t}_p o valor de escores preditivos de uma amostra, \mathbf{p}_p o vetor de pesos preditivo, \mathbf{T}_o a matriz de escores ortogonais, \mathbf{P}_{y_o} a matriz de pesos ortogonais e \mathbf{E} a matriz residual de \mathbf{X} . Os escores ortogonais são calculados pela decomposição dos autovalores e usando os autovetores (w) como vetores de peso (w') (TRYGG e Wold, 2001). A Figura 13 esquematiza a equação 10.

Figura 13 – Esquema gráfico do cálculo da OPLS-DA, indicando os escores (\mathbf{p}'), escores ortogonais (\mathbf{p}_{ortho}'), autovetores (\mathbf{w}'), autovetores ortogonais (\mathbf{w}_{ortho}'), pesos (\mathbf{t}), pesos ortogonais (\mathbf{t}_{ortho}') para a matriz \mathbf{X} , \mathbf{u} as colunas da matriz \mathbf{Y} e \mathbf{c}' as classes do modelo



Fonte: Trygg e Wold (2001).

No caso de sobreposição de classes numa direção preditiva, a variação Y -ortogonal pode ser empregada para intensificar as taxas de classificação (BYLESJÖ *et al.*, 2006), o que pode levar o OPLS-DA a uma maior facilidade na visualização dos *loadings* usados na discriminação.

A análise discriminante linear (do inglês, *Linear Discriminant Analysis*), LDA, é uma técnica supervisionada de análise multivariada e supervisionada cujo principal intuito é projetar os dados em um subespaço menor que possa resultar em uma melhor separação entre as classes (SILVA, 2017). A classificação por LDA emprega a distância de Mahalanobis e estima uma única covariância agrupada matriz S , em vez de usar uma estimativa separada para cada classe (MOREIRA *et al.*, 2009). O GA foi proposto por John Holland e é uma técnica que simula matematicamente os mecanismos de seleção natural e a teoria da evolução das espécies (SOARES, 2016), possibilitando a seleção de variáveis presentes num vasto conjunto de dados. O algoritmo genético básico envolve cinco passos: codificação das variáveis, criação da população inicial, avaliação da resposta, cruzamento e mutação (COSTA FILHO, 1998). Com a seleção de variáveis é possível proceder com a LDA, uma vez que esta não pode construir um modelo com mais variáveis que amostras (FERREIRA, 2015).

Os formalismos de análise de discriminantes citados até então classificam seus indivíduos com base em informação prévia fornecida ao modelo. Contudo, cabe citar também os formalismos modelativos ou de classe única, que têm como objetivo distinguir objetos de uma classe particular, ou classe alvo, de outros objetos e classes (ZONTOV *et al.*, 2017). O método de *Soft Independent Modeling of Class*

Analogy (SIMCA) é um destes, assumindo que os valores medidos para um dado grupo de amostras similares terão uma distribuição uniforme e modelável supondo que haja uma distribuição probabilística, permitindo estimar o grau de certeza na classificação (FERREIRA, 2015). Versões mais recentes do cálculo do SIMCA, o *Data Driven SIMCA* (DD-SIMCA) permitem calcular os erros de classificação do modelo (ZONTOV *et al.*, 2017). Assim, a área de aceitação é construída para um valor especificado α de um erro tipo I e levando em consideração a presença/ausência de *outliers* (POMERANSTEV *et al.*, 2014).

Na metodologia empregada para cálculo do SIMCA os limites de cada classe são definidos por um modelo de PCA ajustado a cada classe de treinamento (FERREIRA, 2015). Para a tomada de decisões, o modelo se baseia em dois fatores que são provenientes da análise de componente principal de cada grupo: *Hotelling* (H), valor que representa o quadrado da distância de Mahalanobis do centro de um modelo a uma amostra projetada dentro do subespaço da PCA; e erro quadrado de predição (Q), valor da distância Euclidiana de uma amostra ao subespaço da PCA (POMERANSTEV *et al.*, 2020). Assim sendo, uma amostra deve obedecer a ambos critérios para ser classificada num determinado grupo:

$$\begin{pmatrix} Q < Q_{\alpha} \\ H < H_{\alpha} \end{pmatrix}$$

Sendo Q_{α} e H_{α} limites críticos decididos visando diminuir o erro tipo I. Já no DD-SIMCA, por ser um modelo que pode seguir um critério de componentes principais robusto, a amostra deve seguir o critério da distância completa (D) descrito nas equações 11 e 12 e ser aceita na condição $D < D_{\alpha}$ (POMERANSTEV *et al.*, 2020).

$$D = L_H \frac{H}{H_0} + L_Q \frac{Q}{Q_0} \propto X^2 (L_H + L_Q) \quad (11)$$

$$D_{\alpha} = X^{-2}(\alpha, L_D), L_D = L_H + L_Q \quad (12)$$

Assume-se que a distribuição de H e Q sigam o modelo de chi-quadrado. L representa graus de liberdade que possam ser assumidos por H ou Q, χ^2 é $(1-\alpha)$ do quantil da distribuição nos graus de liberdade (POMERANSTEV *et al.*, 2020). Modelos robustos têm sido empregados em conjuntos de dados que apresentam grande complexidade nas suas informações e são uma alternativa que apresenta

insensibilidade a pequenos desvios de suposições e, ao mesmo tempo, não são alterados devido a desvios maiores (POMERANSTEV *et al.*, 2014)

O emprego de métodos multivariados é de especial importância para a metabonômica pois geralmente um único biomarcador não é responsável para descrever uma condição clínica individualmente (MADSEN, 2010).

3.4.1 Figuras de Mérito

Apesar de determinar um limite para aceitação de erros, os modelos podem ainda apresentar classificação errônea de amostras. Para calcular a taxa dos erros cometidos pelo classificador, os resultados são organizados na forma de tabela de contingência ou matriz de confusão (FERREIRA, 2015) (Tabela 2). Numa matriz de confusão, se uma amostra, classificada como positiva pelo padrão-ouro, for classificada pelo modelo corretamente, é contada como verdadeira positiva (VP). Porém, se for classificado como negativo, é considerado falso negativo (FN), ou erro Tipo II. Se a amostra, classificada como negativa pelo padrão-ouro, for classificada como negativa é considerada verdadeiro negativo (VN). Entretanto, se for classificada como positiva, é contado como falso positivo (FP), ou erro Tipo I (THARWAT, 2018).

Tabela 2 – Exemplo de matriz de confusão em formalismos quimiométricos

		Classe Verdadeira	
		A	B
Classe Predita	A	VP	FP
	B	FN	VN

Fonte: autoria própria (2021).

As métricas de exatidão, sensibilidade, especificidade, valor preditivo positivo (VPP) e valor preditivo negativo (VPN) são calculadas como mostra o Quadro 2.

Quadro 2 – Métricas da matriz de confusão, seu cálculo e significância estatística para modelos metabonômicos

Métrica	Cálculo
Exatidão	$\frac{VP + VN}{n \text{ de amostras}}$
Sensibilidade	$\frac{VP}{VP + FN}$
Especificidade	$\frac{VN}{FP + VN}$
Valor Preditivo Positivo (VPP)	$\frac{VP}{VP + FP}$
Valor Preditivo Negativo (VPN)	$\frac{VN}{FN + VN}$
Razão de verossimilhança negativa	$\frac{1 - \text{sensibilidade}}{\text{especificidade}}$
Intervalo de confiança a 95%*	$1,96 * \sqrt{\left[\frac{(P)(1 - P)}{n} \right]}$

*O P indicado na fórmula é a proporção para a qual se pretende calcular o intervalo de confiança.

Fonte: Adaptado de Luque *et al.* (2019).

Exatidão trata-se da fração de predições corretas; especificidade mede a fração das amostras negativas preditas corretamente, enquanto sensibilidade é a proporção das amostras positivas medidas corretamente (LEVER *et al.*, 2016). O valor preditivo positivo (VPP) representa a métrica de correta classificação das amostras para o número total de amostras positivas, enquanto o valor preditivo negativo (VPN) é expresso como a razão entre as amostras negativas classificadas corretamente e o número total de amostras classificadas como negativas (THARWAT, 2018).

4 METODOLOGIA

4.1 SELEÇÃO DE PACIENTES E PROCEDIMENTOS ÉTICOS

Os pacientes selecionados tinham entre seis e trinta e seis meses de idade. Tratavam-se de pacientes do Serviço de Alergia e Imunologia Clínica e no Ambulatório de Pediatria e de Puericultura do Hospital das Clínicas de Pernambuco que apresentavam sintomas anteriores de sibilância. As crianças participantes tiveram o termo de consentimento livre e esclarecido (TCLE) (APÊNDICE A) assinado pelos pais e/ou responsável legal. Posteriormente, foi aplicado o questionário padrão de coleta de dados de pacientes do Ambulatório de Pediatria e de Puericultura do Hospital das Clínicas de Pernambuco, baseado no EISL (APÊNDICE B e C) para registro de dados epidemiológicos. No questionário dos pacientes também houve o registro de achados laboratoriais como teste cutâneo (com testagem para *Dermatophagoides pteronyssinus*, *Blomia tropicalis*, barata mix e epitélio de cão) e dosagem de eosinófilos absoluta e relativa. Os pacientes foram classificados, usando os critérios deste questionário, como positivos ou negativos e, para este estudo, esses grupos foram classificados como Sibilância Recorrente (SR), para os pacientes positivos, e Sibilância não Recorrente (SNR), para os negativos. No universo de 81 pacientes atendidos, foram selecionados 14 voluntários no grupo positivo e 25 no grupo negativo.

Os critérios de inclusão para os voluntários selecionados foram: lactentes que apresentaram três ou mais episódios de sibilância nos últimos seis meses; ausência de outras doenças respiratórias, como infecções respiratórias recentes, fibrose cística, broncodisplasia ou bronquiectasia. Os critérios de exclusão foram: o uso de medicamentos para o tratamento de sibilância ou asma, como corticoide inalatório, antileucotrienos e/ou broncodilatadores; uso de medicação para tratamento de crise aguda em menos de quinze dias. Para cada paciente, foi empregado o Índice Preditivo de Asma modificado, conforme descrito por Guilbert *et al.* (2004). O protocolo de preparação das amostras incluiu centrifugação (4000 RPM por 10 minutos) do sangue total dos participantes do estudo e amostragem apenas do soro.

O presente estudo foi aprovado pelo Comitê de Ética em Pesquisa do Centro de Ciências da Saúde da Universidade Federal de Pernambuco (número de aprovação 2.337.325).

4.2 ENSAIOS METABONÔMICOS

Os espectros de RMN de ^1H foram adquiridos usando um espectrômetro VNMRs400, operando a 400 MHz. Foram utilizados 400 μL de soro e 200 μL de D_2O num tubo de RMN de 5 mm. A sequência de pulso Carr-Purcell-Meiboom-Gill (CPMG) com pré-saturação do sinal de água foi usada com os seguintes parâmetros: janela espectral igual a 6,4 kHz, tempo de aquisição igual a 2,55 s, 128 transientes, spin echo delay igual a 400 microssegundos, 88 ciclos, dando um tempo total de eco igual a 70,4 milissegundos e tempo de aplicação do pulso seletivo para saturação do sinal da água igual a 2,0 s, a 27° C (BECKONERT, 2007). O sinal atribuído ao grupo metil do lactato (δ 1,33 ppm) foi utilizado como referência de deslocamento químico (BATISTA, 2018). A utilização do software MestReNova possibilitou o processamento dos espectros e da correção da linha de base, pelo Polinômio de Bernstein e ajuste manual de fase. Em seguida, foi selecionada a região entre δ 0,65 e 4,5 ppm dos espectros e utilizou-se um bin de 0,004 ppm/bin. Os dados espectrais foram coletados em uma matriz com 39 casos (linhas) e 963 variáveis (coluna). O conjunto de dados foi pré-processado usando normalização por média (nas linhas e, em seguida, nas colunas, nessa ordem).

Empregou-se como análise exploratória a PCA e para análise classificatória empregou-se PLS-DA e OPLS-DA, utilizando a plataforma online MetaboAnalyst. Para validar os modelos obtido por PLS-DA e OPLS-DA, foi utilizado o *Leaving-One-Out Cross Validation* (LOOCV) e teste de permutação, usando 2000 permutações (PONTES, 2019). O modelo de LDA foi realizado usando o software MATLAB 2010a. Para o modelo LDA, as amostras foram divididas em dois grupos: treinamento e teste na proporção de 70:30 de cada grupo classificado. O Algoritmo Kennard-Stone foi utilizado para fazer a seleção das amostras. O Algoritmo Genético (GA) (população = 500, gerações = 50) foi executado para promover uma seleção de variáveis para o modelo LDA. Utilizou-se também o MATLAB 2010a para a análise DD-SIMCA, utilizando 6 componentes principais. Os dados foram centrados na

média, utilizou-se como parâmetros $\alpha = 0.05$, *outlier significance* de 0.01 com chi-quadrado como tipo de área de aceitação num modelo robusto. Para cada um dos modelos metabonômicos de LDA e DD-SIMCA calculou-se a sensibilidade, especificidade, exatidão, valor preditivo positivo e valor preditivo negativo.

Foi utilizado o MetaboAnalyst 5.0 (<http://www.metaboanalyst.ca/>) para análise de vias metabólicas e as mais relevantes foram descritas, baseadas no valor de p . Foi escolhida a biblioteca *Homo sapiens* e utilizado o teste Exato de Fisher e o algoritmo de *relative-betweenness centrality* como as opções para a análise de sobre-representação e análise de topologia de via metabólica, respectivamente (PANG *et al.*, 2018).

5 RESULTADOS E DISCUSSÃO

5.1 DADOS DOS PACIENTES

A Tabela 3 apresenta os dados clínicos dos voluntários classificados neste estudo nos grupos de Sibilância Recorrente (n=14) e Sibilância não Recorrente (n=25). A Figura 14 mostra um espectro típico de RMN de ^1H do soro de um dos pacientes do grupo SNR com o processamento pós-instrumental e com alguns metabólitos identificados em seus respectivos deslocamentos químicos.

Tabela 3 – Características demográficas, clínicas e laboratoriais dos voluntários por grupo

	Sibilância não Recorrente	Sibilância Recorrente	Valor de p
Número de voluntários	25	14	
Idade em meses (Média ± desvio padrão)	22,3 ± 7,4	27,8 ± 6,2	0,024 ^a
Sexo (Masculino:Feminino)	20:5	8:6	0,156 ^b
Etnia (Afrodescendente:Caucasiano)	7:17*	3:10**	1,000 ^b
Teste cutâneo (Positivo:Negativo)	22:3	11:2**	1,000 ^b
Eosinófilos/mm³ (Média ± desvio padrão)	369,6 ± 297,9	585,6 ± 431,2	0,073 ^a
Eosinófilos (%) (Média ± desvio padrão)	3,4 ± 2,5	5,6 ± 4,0	0,042 ^a

^a Teste t de Student; ^b Teste exato de Fisher. *A etnia de um dos pacientes foi indicada como “outra”.

** Informação não preenchida por um dos voluntários.

Fonte: Autoria própria (2020).

A Tabela 4 apresenta a matriz de confusão com a predição feita usando o Índice Preditivo de Asma modificado (mAPI, do inglês *Modified Asthma Predictive Index*). A classificação para cada paciente está no APÊNDICE E. O mAPI foi desenvolvido como um indicador de ocorrência de asma, mas a título de comparação, classificamos todos os voluntários do estudo. Assim, o mAPI classificou corretamente as 14 amostras de pacientes que tiveram o diagnóstico de Sibilância Recorrente. Por outro lado, das 25 amostras fornecidas por pacientes que tiveram o diagnóstico de Sibilância Não-Recorrente, o mAPI classificou corretamente apenas 12 (48%) amostras.

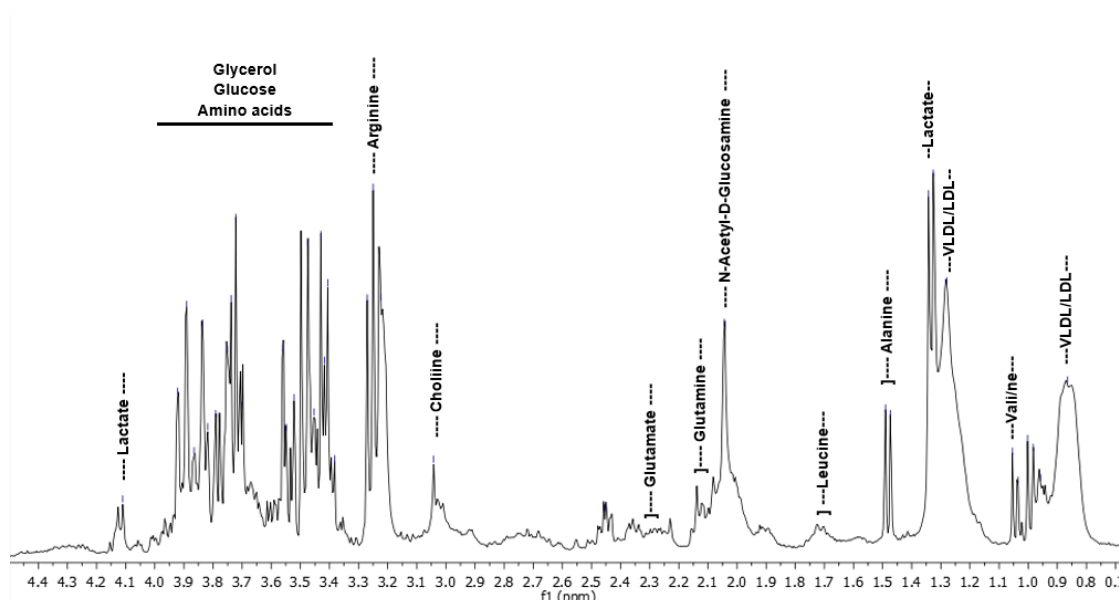
Tabela 4 – Matriz de confusão das classes previstas usando o Índice Preditivo de Asma modificado (mAPI)

		Diagnóstico Clínico	
		Sibilância Recorrente	Sibilância Não-Recorrente
Classe predita pelo mAPI	Positivo	14	13
	Negativo	0	12

Sensibilidade 100%; especificidade 48,0%; exatidão 66,6%; VPP 51,8% e VPN 100%.

Fonte: Autoria própria (2020).

Figura 14 – Típico espectro de RMN de ^1H (400 MHz, D_2O) de soro obtido nesse estudo, de um paciente do grupo de Sibilância não Recorrente. Alguns sinais foram atribuídos a metabólitos



Fonte: Autoria própria (2020).

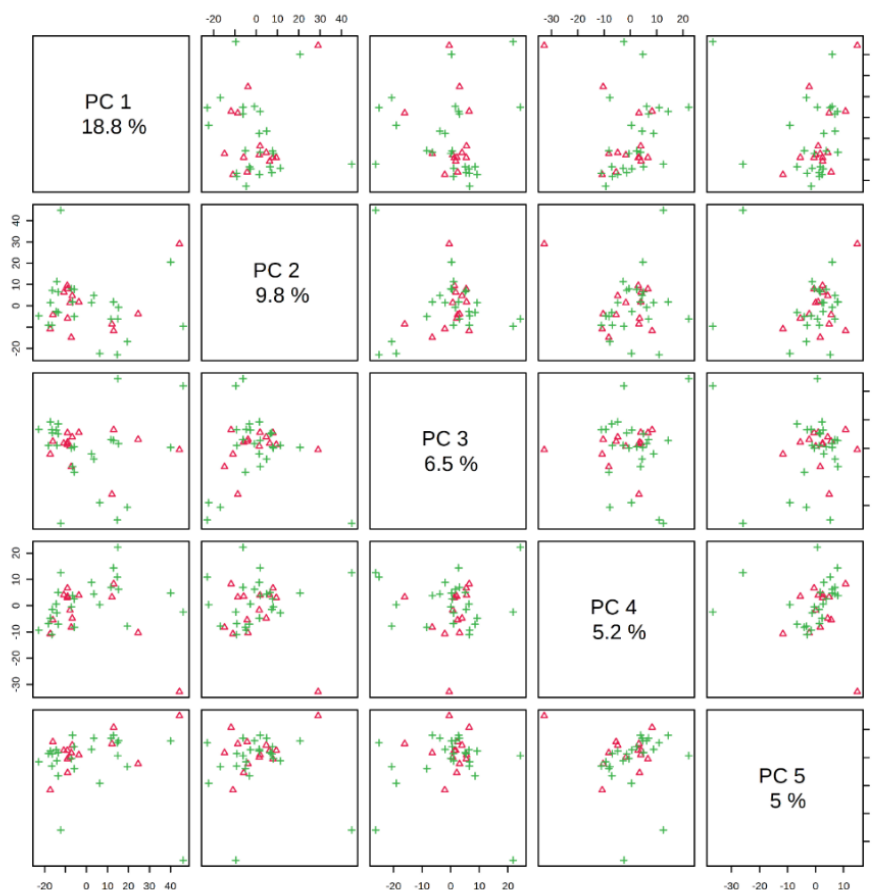
5.2 FORMALISMOS QUIMIOMÉTRICOS

A PCA do conjunto de dados analisado não mostrou uma diferença significativa entre os grupos em estudo (Figura 15). PC1 e PC2 explicaram apenas 28.6% da variância do conjunto de dados. Além disso, não foi observado agrupamento entre as amostras.

A primeira componente principal apresenta a máxima variância das variáveis. A segunda componente principal não é correlacionada com a primeira e representa o máximo da variância residual, e assim sucessivamente até que toda a variância seja explicada (IZQUIERDO-GARCÍA, 2011). Um dos motivos que justifique valores

baixos de variância explicada pelas componentes principais da PCA é a grande dimensionalidade dos dados analisados. Apesar da região que os espectros foram avaliados ter sido pequena, o número de bin utilizados gerou 963 variáveis para a matriz dos dados. Para esse estudo utilizou-se um número de bin menor visando diminuir a probabilidade de múltiplos picos residindo num único bin (ANDERSON, 2010). Como a estratégia de análise exploratória não atendeu aos propósitos deste estudo realizaram-se análises classificatórias supervisionadas.

Figura 15 – Análise de componentes principais do conjunto de dados analisado. Em vermelho as amostras de Sibilância Recorrente e em verde as amostras de Sibilância não Recorrente

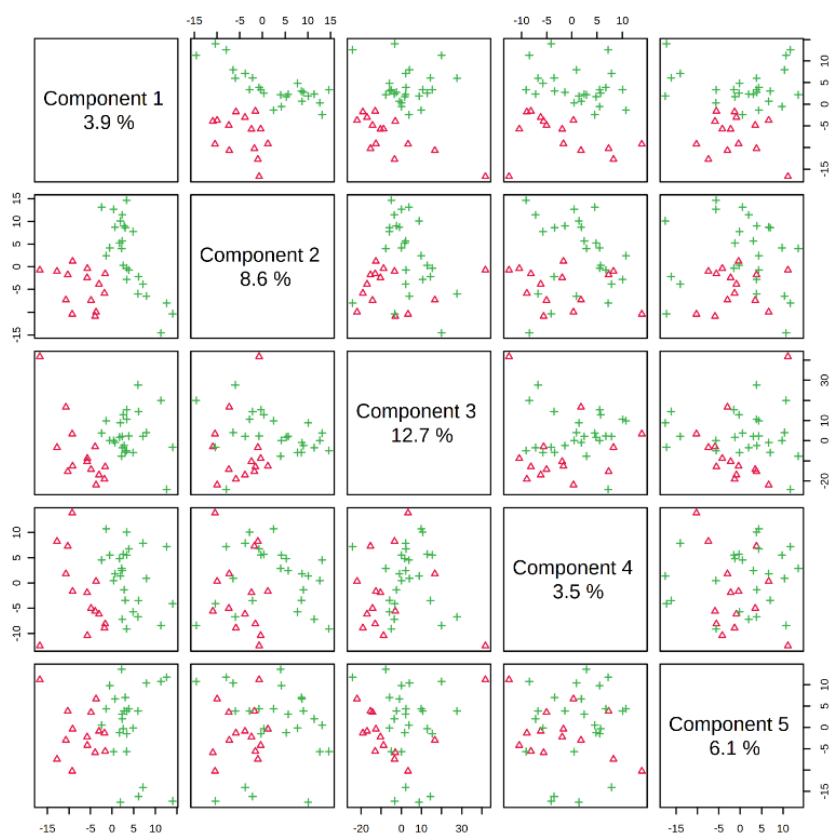


Fonte: Metaboanalyst 5.0 (2020).

Modelos de classificação são ferramentas importantes para fornecer medidas de correlação estatística (MADSEN, 2010). A Figura 16 revela como as amostras ficaram dispostas no modelo PLS-DA realizado, que apesar de um possível

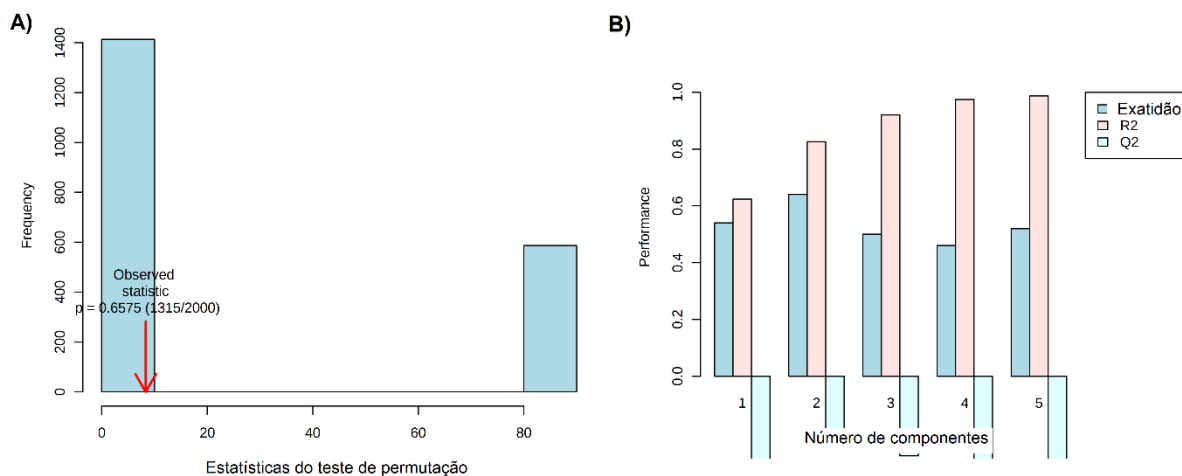
agrupamento de classes, apresenta baixa performance devido à baixa exatidão do modelo. As figuras de mérito do PLS-DA (Figura 17) não apresentaram resultados estatisticamente relevantes para a separação dos lactentes com Sibilância Recorrente e Sibilância não Recorrente. Uma possível causa desse resultado pode ter sido o grande número de variáveis, associadas às regiões espectrais que caracterizavam as amostras. Nessas condições, o modelo torna-se muito dependente do conjunto de dados usado para sua construção e, às vezes, eles apresentam baixa capacidade de generalização (RODRÍGUEZ-PÉREZ, 2018).

Figura 16 – Gráficos de escores da PLS-DA. Em vermelho, Sibilância Recorrente; e, em verde, Sibilância não Recorrente



Fonte: Metaboanalyst 5.0 (2020).

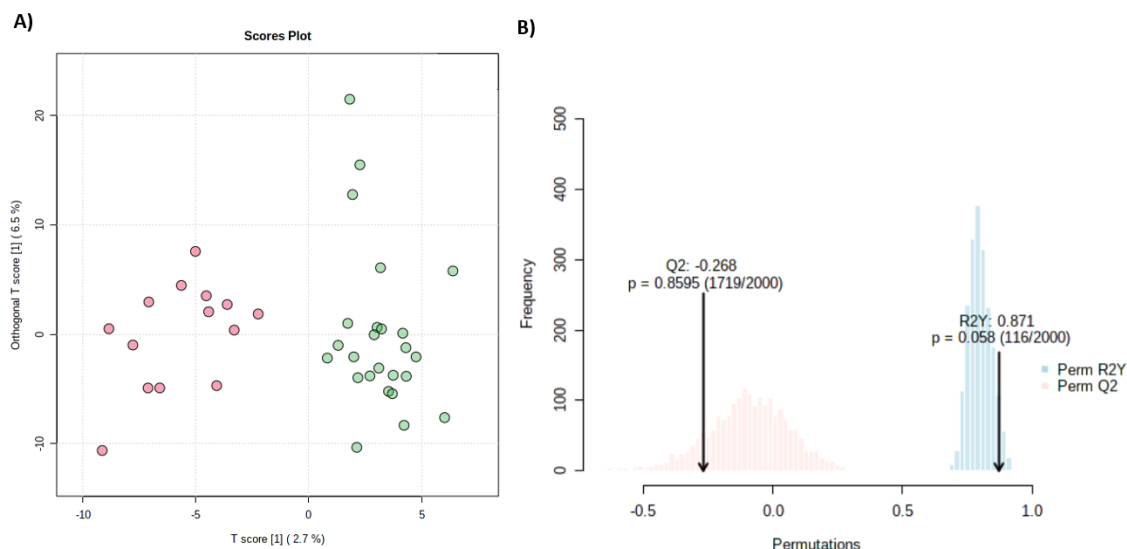
Figura 17 – Figuras de mérito do PLS-DA. A) Teste de permutação. B) Valores de performance do modelo



Fonte: Metaboanalyst 5.0 (2020).

O uso do OPLS-DA possibilitou uma melhor visualização da discriminação entre as amostras, apresentando um valor de R2Y igual a 0,87. O valor de R2Y representa a fração da variância explicada do modelo pela regressão (BORGONOVO,2006). Resultados semelhantes de R2Y foram observados em modelos que estudavam soro de pacientes com asma e doença pulmonar obstrutiva crônica, como apontam os estudos de Ghosh *et al.* (2019). A Figura 18 mostra o gráfico de escores e o teste de permutação do modelo metabonômico usando OPLS-DA. Tal ferramenta resultou num modelo sobreajustado, devido ao baixo valor de Q2 (Q2 igual a -0,268), resultando assim numa baixa capacidade preditiva.

Figura 18 – Modelo metabonômico de OPLS-DA calculado para discriminação entre os grupos de Sibilância Recorrente e não Recorrente (A). Resultado do teste de permutação do OPLS-DA (B)



Fonte: Metaboanalyst 5.0 (2020).

Devido a tendência a grupamento das amostras, os modelos construídos com PLS-DA e OPLS-DA indicam que a informação que discrimina os grupos de interesse está contida no conjunto de dados, mas a variação intra e intergrupo é muito intensa, dificultando a discriminação. Assim, optou-se por realizar duas abordagens distintas: construir funções discriminantes, usando LDA, a partir de um conjunto reduzido de variáveis; e uma modelagem usando DD-SIMCA, que utiliza um número limitado de componentes principais e considera apenas a variação intragrupo.

O resultado do formalismo GA-LDA, apresentado como matriz de contingência na Tabela 5, apresenta um modelo metabonômico com exatidão, especificidade e sensibilidade iguais a 75%, 62,5% e 100%, respectivamente. Outro resultado apresentado pelo GA-LDA foram as variáveis de interesse responsáveis pela separação das classes em estudo. Tais variáveis associadas ao deslocamento químico que apresentam podem ser associadas a metabólitos presentes nos quadros de sibilância recorrente e sibilância não recorrente.

Tabela 5 – Matriz de confusão dos modelos de GA-LDA para treinamento e teste

Modelo de Treinamendo		Diagnóstico Clínico	
		Sibilância Recorrente	Sibilância não Recorrente
Classe predita	Sibilância Recorrente	10	0
	Sibilância não Recorrente	0	17
Modelo de teste			
Classe predita	Sibilância Recorrente	4	3
	Sibilância não Recorrente	0	5

Sensibilidade 100%; especificidade 62,5%; exatidão 75,0%; VPP 57,1% e VPN 100%.

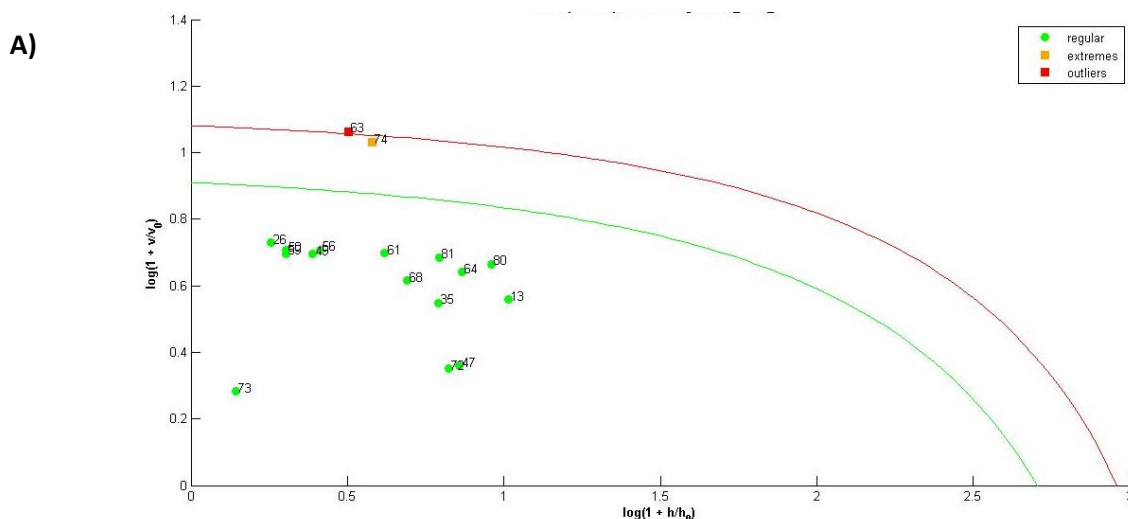
Fonte: Autoria própria (2020).

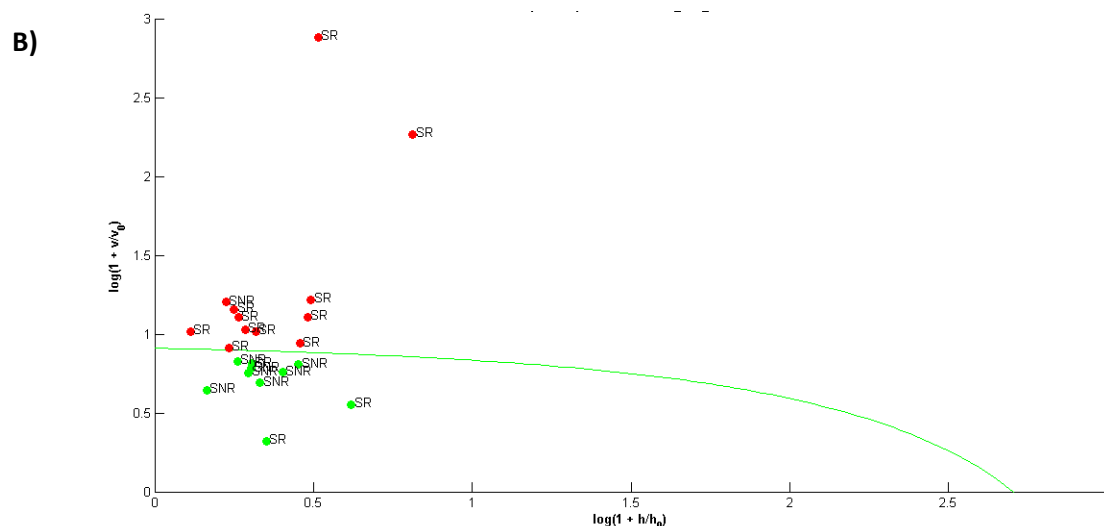
O uso do Algoritmo Genético possibilitou realizar a análise de LDA, contornando a restrição do cálculo de LDA quando existem mais variáveis que amostras (FERREIRA, 2016). Os resultados apresentados pelo modelo metabonômico de LDA tiveram exatidão de 75,0% ($\pm 24,5$), apresentando 42,9% de casos falsos-positivos. Isto implica numa baixa especificidade ($62,5\% \pm 27,4$) e também em baixo valor preditivo positivo ($57,1\% \pm 28,0$). Entretanto, este modelo poderia ser utilizado como ferramenta de triagem de novas amostras, uma vez que apresentou valor preditivo negativo igual a 100%, o que indica que todas as amostras do grupo de Sibilância não Recorrente foram classificadas corretamente. A performance do LDA apresenta-se então promissora, podendo-se melhorar o modelo inserindo mais amostras nos grupos de calibração e teste no futuro.

Modelos *one-class*, ou aqueles que apresentam modelagem de classe, consistem na descrição de uma classe alvo e na detecção se um objeto externo pertence ou não a essa classe (RODIONOVA *et al.*, 2016). Tais modelos também são úteis quando há baixa disponibilidade de amostras aceitas pelos critérios de inclusão (KHAN e MADDEN, 2014). Isso leva a um desbalanceamento de dados entre as classes, que faz com que os algoritmos de aprendizagem de métodos de aprendizagem de máquinas apresentem dados de baixo desempenho para a classe com menor número de amostras (MENA e GONZALEZ, 2009). Como a avaliação da capacidade preditiva de um modelo depende dos seus indicadores de exatidão, sensibilidade e especificidade (SOARES, 2020), optou-se por utilizar do formalismo

DD-SIMCA para construção de um modelo metabonômico que identificasse os pacientes que apresentam o quadro de sibilo não recorrente. Para o grupo de treinamento, das 17 amostras do grupo SNR selecionadas previamente pelo *Kennard-Stone* uma apresentou-se como *outlier* e uma como amostra de extremo, enquanto as demais estavam no intervalo de confiança estabelecido para o modelo. No grupo de teste foram selecionadas 8 amostras do grupo SNR e foram utilizadas todas as amostras do grupo SR para testar o modelo. Desta feita, o modelo pode prever que amostras apresentavam características semelhantes ao grupo inicial que serviu de treinamento para o formalismo empregado. Os resultados obtidos apresentaram sensibilidade, especificidade e exatidão de 78,6% ($\pm 17,0$), 87,5% ($\pm 14,0$) e 81,8% ($\pm 16,0$), respectivamente (Figura 18). Obteve-se para os valores preditivos positivo e negativo 91,7% ($\pm 12,0$) e 70% ($\pm 19,0$), respectivamente. A Tabela 6 apresenta as figuras de mérito do modelo.

Figura 19 – Modelo DD-SIMCA para o grupo de treinamento (A) e o grupo teste (B). Em verde, as amostras da classe de treinamento (SR); em laranja, amostras classificadas como extremos; e em vermelho, as amostras *outliers* preditas pelo modelo. A linha verde indica o intervalo de confiança (95%) do grupo de amostras modeladas, enquanto a linha vermelha indica o limite a partir do qual tem-se as amostras não pertencentes ao grupo modelado.





Fonte: Autoria própria (2020).

Tabela 6 – Matriz de Confusão do modelo metabonômico de DD-SIMCA utilizando o espectro de RMN de ^1H do soro de lactentes para discriminação entre Sibilância Recorrente e Sibilância não Recorrente

Grupo de teste		Diagnóstico Clínico	
		Sibilância Recorrente	Sibilância Não Recorrente
Classe Predita	Sibilância Recorrente	11	1
	Sibilância Não Recorrente	3	7

Sensibilidade 78,6%; especificidade 87,5%; exatidão 81,8%; VPP 91,7% e VPN 70,0%.

Fonte: Autoria própria (2020).

O baixo número de pacientes que apresenta Sibilância Recorrente pode não representar adequadamente todos os casos dessa condição fisiopatológica, mas ao optar por realizar uma metodologia com modelagem de classe para os casos negativos valida-se a descrição envolvendo os lactentes saudáveis (CABRAL e OLIVEIRA, 2014). Uma modelagem invertendo as classes de estudo foi realizada e seus achados foram incluídos no APÊNDICE D.

Comparando a razão de verossimilhança negativa entre os modelos de GALDA, DDSIMCA e a informação fornecida pelo mAPI, tem-se respectivamente os valores de 0; 0,24; 0,48. Este valor varia de 1 a 0 e quanto mais próximo a 0 menor a probabilidade de o voluntário apresentar-se doente se classificado como negativo

pelo modelo em estudo (FERREIRA e PATINO, 2018). Quando comparados (APÊNDICE F) com o mAPI, a performance dos modelos de DD-SIMCA e GA-LDA são melhores, o que indica que o uso de modelos metabonômicos podem levar a um diagnóstico diferencial entre lactentes que apresentem sibilância recorrente dos que não apresentam.

5.3 VIAS METABÓLICAS

Nove regiões espectrais mostraram-se importantes para discriminar amostras de sibilância recorrente de não recorrente, ao realizar a seleção de variáveis de interesse pelo Algoritmo Genético. Essas regiões foram atribuídas a oito metabólitos (Tabela 7) que podem ser agrupados em duas classes: lipídios (VLDL / LDL e colesterol) e aminoácidos e seus derivados (histidina, leucina, glutamina, glutamato e ácido N-acetil-L-aspartico).

Tabela 7 – Metabólitos endógenos mais importantes para discriminação entre Sibilância Recorrente e Sibilância não Recorrente e sua variação de concentração nestes grupos de acordo com o modelo metabonômico de GA-LDA

Componente	Deslocamento Químico (δ /ppm)	↑ Grupo	Referência
VLDL/LDL-CH ₃	0.791	SR	WHISHART <i>et al.</i> , 2009
Colesterol-C10	1.087	SR	LI <i>et al.</i> , 2017
Não atribuído	1.097	SNR	
Leucina	1.743	SNR	MAMTIMIN <i>et al.</i> , 2014
Glutamato	2.335	SNR	MAMTIMIN <i>et al.</i> , 2014
Ácido N-Acetil-L-Aspartico	2.519; 2.711	SR	WHISHART <i>et al.</i> , 2009
Glutamina	2.131	SR	WHISHART <i>et al.</i> , 2009
Histidina	3.976	SNR	MAMTIMIN <i>et al.</i> , 2014

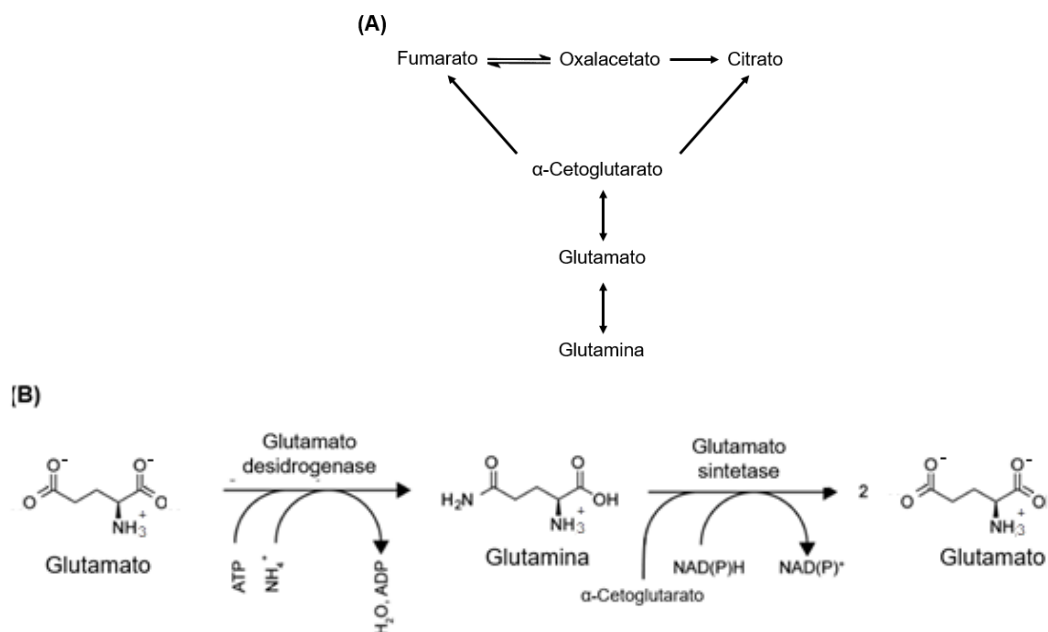
Fonte: Autoria própria (2020).

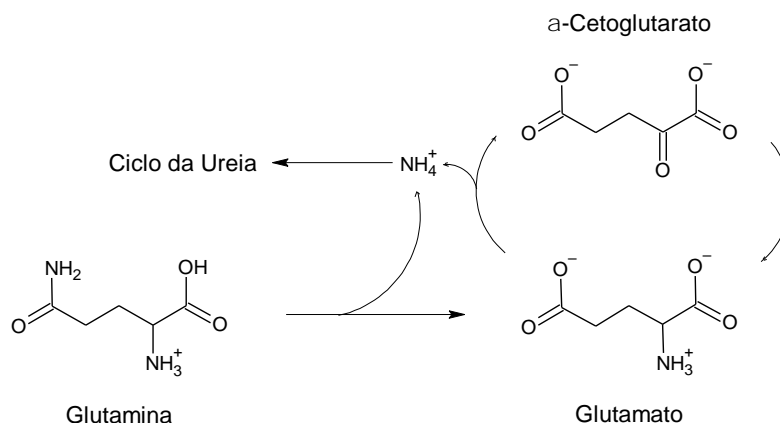
Os níveis séricos de VLDL e LDL aumentam a resposta imunológica, visto que essas moléculas induzem a produção de quimiocinas pró-inflamatórias (SARASWATHI e HASTY, 2006). O modelo metabonômico GA-LDA indica que os pacientes SR apresentaram níveis séricos mais elevados de VLDL e LDL e

colesterol. A mobilização de lipídeos relaciona-se com o metabolismo lipídico e estresse oxidativo (JUNG *et al.*, 2013; FARRAIA *et al.*, 2019). Moléculas oxidantes agem destruindo as funções ciliares do epitélio respiratório enquanto aumentam a secreção de muco, contribuem para a contração do músculo liso e servem como moléculas de quimiotaxia para células inflamatórias (SAGDIC *et al.*, 2010).

Muitas vezes, os aminoácidos são considerados mediadores da atividade imunológica na asma (MATYSIAK *et al.*, 2020). Níveis significativos de leucina em pacientes asmáticos relacionam-se à conversão para acetato, que posteriormente gera acetil-CoA e, por conseguinte, entra no ciclo do ácido tricarboxílico para atender à alta demanda metabólica desses pacientes (GHOSH *et al.*, 2019). Glutamato e glutamina foram observados com níveis séricos elevados nos grupos SR e SNR, respectivamente. São aminoácidos não essenciais e desempenham um papel importante em muitas funções metabólicas como o equilíbrio do *pH* sanguíneo pela aminação do glutamato e nas vias energéticas. O papel do glutamato e da glutamina nas vias metabólicas está relacionado tanto ao ciclo da ureia quanto ao ciclo de Krebs na geração de energia para a contração do músculo liso (NELSON e COX, 2014) e suas concentrações estão em constante equilíbrio (TAPIERO *et al.*, 2002) (Figura 19).

Figura 20 – (A) Balanço da conversão do glutamato em glutamina e a produção de intermediários de vias energéticas. (B) Vias metabólicas de biossíntese do glutamato



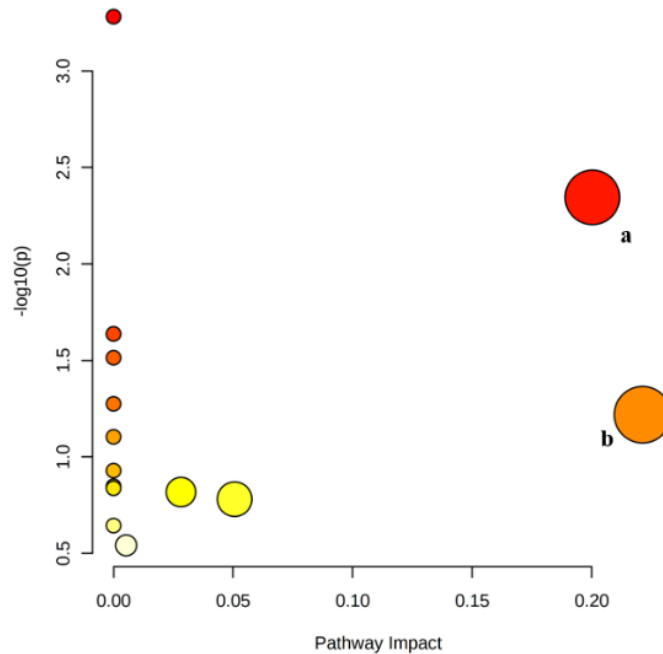


Fonte: (A) Autoria própria (2021); (B) Adaptado de WALKER e DONK (2016).

Com base nos dados obtidos, quando um paciente apresenta o quadro de sibilância recorrente, uma condição mais agravada considerando ambos grupos estudados, a reação tende a deslocar-se para a formação de novas moléculas de glutamina. Os níveis séricos elevados desses metabólitos podem estar relacionados à presença da sibilância recorrente, que possivelmente reflete a carga energética decorrente de inflamação e broncoconstrição (DE BENEDICTIS e ATTANASI, 2016) pela produção de energia e intermediários metabólicos do ciclo de Krebs e ciclo da ureia (Figura 20). O ácido N-acetil-L-aspartico sérico está relacionado ao ciclo glutamato-glutamina como um dos metabólitos intermediários (WISHART *et al.*, 2009). O Anexo A mostra as vias metabólicas da Alanina-Aspartato-Glutamato e D-Glutamina e D-Glutamato, as mais relevantes para a discriminação entre Sibilância Recorrente e não Recorrente, segundo nossos resultados.

A histidina e seus produtos metabólicos são potentes mediadores da inflamação aguda e já foram relatados achados de níveis séricos alterados de histidina em processos inflamatórios nas vias aéreas em pacientes asmáticos (MOTTA *et al.*, 2014). Este aminoácido foi encontrado com níveis elevados nos pacientes com sibilância não recorrente e por ser um aminoácido essencial não existem rotas de biossíntese que justifiquem seu aumento sérico. Todas as vias metabólicas discutidas estão de acordo com as mais relevantes indicadas na análise de vias metabólicas realizada (Figura 21).

Figura 21 – Vias metabólicas mais relevantes para a discriminação. (a) Alanina, aspartato e glutamato (p -value 0.0045); (b) D-glutamina e D-glutamato (p -value 0.0230)



Fonte: Metaboanalyst (2020).

A presença de diferentes arranjos bioquímicos-metabólicos reforça que a expressão de metabólitos está subjacente a perfis patológicos distintos já na idade pré-escolar (CARRARO *et al.*, 2018). Relacionar esses metabólitos com a condição dos pacientes fornece novos *insights* sobre o mecanismo patogênico da asma (JUNG, 2013) e pode levar a tratamentos adequados nos estágios iniciais da doença, melhorando a qualidade de vida do paciente.

6 PERSPECTIVAS

Espera-se que esse trabalho venha a ser útil para pesquisas e protocolos clínicos que busquem realizar um diagnóstico diferencial em lactentes que apresentem sibilo recorrentes daqueles que não apresentem. Ele também serve como linha de partida para diferentes abordagens metodológicas tais como testar se há diferença entre lactentes saudáveis e os grupos estudados, realização de modelos que englobem como variáveis IgE e valores de eosinófilos bem como que seus dados integrem bancos de dados que tratem acerca da condição clínica abordada ao longo do trabalho.

Além do que foi mencionado, espera-se que os modelos metabonômicos desenvolvidos possam ser utilizados na prática clínica, promovendo um diagnóstico precoce em lactentes e, conseqüentemente, seja possível realizar um tratamento adequado aos que apresentem um quadro de sibilo recorrente e evitar que aquelas crianças que não apresentam um quadro de recorrência não sejam submetidas a um tratamento indevido. Com isso, a academia integra-se à prática, aplicando pesquisa de ponta aos protocolos clínicos realizados no Sistema Único de Saúde brasileiro, estando de acordo com seu princípio básico de integralidade.

7 CONCLUSÃO

No presente trabalho foram desenvolvidos modelos metabonômicos capazes de identificar pacientes com Sibilância não Recorrente e diferenciá-los de lactentes que apresentam Sibilância Recorrente.

Com base nos achados do presente estudo, foi possível identificar metabólitos característicos de ambas condições fisiopatológicas e avaliar as vias metabólicas envolvidas, a partir dos espectros de RMN de ^1H adquiridos das amostras analisadas.

Os formalismos que resultaram em modelos com melhores performance foram GA-LDA (sensibilidade 100%; especificidade 62,5%; exatidão 75,0%; VPP 57,1% e VPN 100%) e DD-SIMCA (sensibilidade 78,6%; especificidade 87,5%; exatidão 81,8%; VPP 91,7% e VPN 70,0%). A análise multivariada de dados tornou possível desenvolver um modelo metabonômico com alta sensibilidade e especificidade. A integração de técnicas analíticas e algoritmos de análise multivariada pode fornecer uma nova ferramenta para o diagnóstico diferencial de Sibilância Recorrente e Sibilância não Recorrente.

As vias metabólicas da alanina- aspartato-glutamato e histidina mostram-se em congruência com as vias metabólicas reportadas na literatura de quadros precursores de asma.

A integração de modelos metabonômicos à prática clínica necessita de uma série de requisitos para ser implementada que incluem formação contínua de profissionais, avanços em pesquisa e adaptação de conteúdo. Espera-se que os resultados apresentados justifiquem o esforço de aumentar o número de pacientes para a obtenção de modelos mais robustos, podendo embasar pesquisas futuras no ramo das ciências ômicas e no diagnóstico de Sibilância Recorrente e Sibilância não Recorrente.

Os achados descritos neste trabalho não incluíram as dificuldades encontradas para a obtenção dos modelos metabonômicos mencionados no texto. Tais dificuldades foram poucas, entretanto responsáveis por traçarmos novos rumos para este trabalho. Eis aqui seu detalhamento: inicialmente havia um n de 81

voluntários no estudo com suas respectivas amostras de soro. Os pacientes do grupo “saudável”, não incluído neste estudo, tiveram suas amostras analisadas usando um espectrômetro de 7 T, operando a 300 MHz. No entanto, durante a execução do estudo, este equipamento quebrou e com isso as demais amostras, ou seja, todas aquelas do grupo de doentes (sibilantes recorrentes e sibilantes não recorrentes) foram analisadas no equipamento de 9 T, operando a 400 MHz, mencionado na metodologia. Ao realizar uma análise de PCA entre o grupo de saudáveis e doentes observou-se um agrupamento natural proveniente da diferença de dados obtidos entre os equipamentos, inerentes ao campo magnético aplicado, e não pelo perfil das amostras, o que levou ao descarte do uso de dados dos pacientes saudáveis. Também não foi possível continuar com a seleção de mais pacientes para o estudo para aumentar o n de pacientes com sibilância recorrente ou sibilância não recorrente devido a suspensão do atendimento no Ambulatório de Pediatria e Puericultura por conta da pandemia de COVID-19.

REFERÊNCIAS

- ALITH, M. B.; GAZZOTTI, M. R.; NASCIMENTO, O. A.; JARDIM, J. R. **Impact of asthma control on different age groups in five Latin American countries**. World Allergy Organization Journal, v. 13, n. 4, p. 100113, 2020.
- ÁLVAREZ-FERNÁNDEZ, J. A.; NÚÑEZ-REIZ, A. **Clinical ultrasound in the ICU: changing a medical paradigm**. Medicina Intensiva (English Edition), v. 40, n. 4, p. 246-249, 2016.
- ANDERSON, P. E.; MAHLE, D. A.; DOOM, T. E.; REO, N. V.; DELRASO, N. J.; RAYMER, M. L. **Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data**. Metabolomics, v. 7, n. 2, p. 179-190, 2011.
- ARANDA, C. S.; WANDALSEN, G.; FONZAR, L.; DELA BIANCA, A. C.; MALLOL, J.; SOLÉ, D. **Risk factors for recurrent wheezing—International Study of Wheezing in Infants (EISL) phase 3**. Allergologia et immunopathologia, v. 44, n. 1, p. 3-8, 2016.
- ATKINS, P. W.; DE PAULA, J. **Atkins Físico-Química**. LTC. 8ª ed, vol 8, 2008.
- BACHARIER, L. B.; BONER, A.; CARLSEN, K-H.; EIGENMANN, P. A.; FRISCHER, T.; GÖTZ, M.; HELMS, P. J.; HUNT, T.; LIU, A.; PAPADOPOULOS, N.; PLATTS-MILLS, T.; POHUNEK, P.; SIMINS, F. E. R.; VALOVIRTA, E.; WAHN, U.; WILDHABER, J. **Diagnosis and treatment of asthma in childhood: A PRACTALL consensus report**. Allergy: European Journal of Allergy and Clinical Immunology. 2008.
- BARCELLA, L.; LOPES, F. **Lute como uma Garota: 60 feministas que mudaram o mundo**. Cultrix. 2018.
- BARROS, C. J. P. **Metabonômica Baseada em RMN como Ferramenta para Discriminação de Grãos de Soja Irrradiados & Diagnóstico de Hepatites e Fibrose Hepática**. 2017. Tese de Doutorado. Universidade Federal de Pernambuco.
- BATISTA, A. D.; BARROS, C. J. P.; COSTA, T. B. B. C.; DE GODOY, M. M. G.; SILVA, R. D.; SANTOS, J. C.; LIRA, M. M. M.; JUCÁ, N. T.; LOPES, E. P. A.; SILVA, R. O. **Proton nuclear magnetic resonance-based metabonomic models for non-invasive diagnosis of liver fibrosis in chronic hepatitis C: Optimizing the classification of intermediate fibrosis**. World J Hepatol. 10: 105–115, 2018.
- BECKONERT, O.; KEUN, H. C.; EBBELS, T. M. D.; BUNDY, J.; HOLMES, E.; LINDON, J. C.; NICHOLSON, J. K. **Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts**. Nature protocols, v. 2, n. 11, p. 2692–2703, 2007.
- BOJA, E. S.; KINSINGER, C. R.; RODRIGUEZ, H.; SRINIVAS, P. **Integration of omics sciences to advance biology and medicine**. Clinical Proteomics, v. 11, n. 1, p. 1–12, 2014.

BLOEMBERGEN, N.; PURCELL, E. M.; POUND, R. V. **Relaxation effects in nuclear magnetic resonance absorption**. *Physical review*, v. 73, n. 7, p. 679, 1948.

BORGONOVO, E. **Measuring uncertainty importance: investigation and comparison of alternative approaches**. *Risk analysis*, v. 26, n. 5, p. 1349-1361, 2006.

BYLESJÖ, M.; RANTALAINEN, M.; CLOAREC, O.; NICHOLSON, J. K.; HOLMES, E.; TRYGG, J. **OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification**. *Journal of Chemometrics: A Journal of the Chemometrics Society*, v. 20, n. 8-10, p. 341-351, 2006.

BYLESJÖ, M.; RANTALAINEN, M. Model Based Preprocessing and Background Elimination: OSC, OPLS, and O2PLS. *In: BROWN, S. D.; TAULER, R.; WALCZAK, B. (ed.). Comprehensive chemometrics: chemical and biochemical data analysis*. Elsevier, 2009. v. 2, cap 8, p. 129-138

CABRAL, G. G.; OLIVEIRA, A. L. I. **One-class Classification for heart disease diagnosis**. *In: 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2014. p. 2551-2556.

CANUTO, G. A. B.; COSTA, J. L.; CRUZ, P. L. R.; SOUZA, A. R. L.; FACCIO, A. T.; KLASSEN, A.; RODRIGUES, K. T.; TAVARES, M. F. M. **Metabolomics: Definitions, state-of-the-art and representative applications**. *Quimica Nova*, v. 41, n. 1, p. 75–91, 2018.

CARRARO, S.; BOZZETTO, S.; GIORDANO, G.; EL MAZLOUM, D.; STOCCHERO, M.; PIRILLO, P.; ZANCONATO, S.; BARALDI, E. **Wheezing preschool children with early-onset asthma reveal a specific metabolomic profile**. *Pediatric Allergy and Immunology*, v. 29, n. 4, p. 375–382, 2018.

CASTRO-RODRIGUEZ, J. A.; CIFUENTES, L.; MARTINEZ, F. D. **Predicting asthma using clinical indexes**. *Frontiers in Pediatrics*, v. 7, n. JULY, p. 1–9, 2019.

CHONG NETO, H. J.; ROSARIO, N.; DELA BIANCA A. C.; SOLÉ, D.; MALLOL, J. **Validation of a questionnaire for epidemiologic studies of wheezing in infants**. *Pediatric allergy and immunology*, v. 18, n. 1, p. 86-87, 2007.

CLARIDGE, T. D. W. **High-Resolution NMR Techniques in Organic Chemistry**. Elsevier, 2009.

COLICINO, S.; MUNBLIT, D.; MINELLI, C.; CUSTOVIC, A.; CULLINAN, P. **Validation of childhood asthma predictive tools: a systematic review**. *Clinical & Experimental Allergy*, v. 49, n. 4, p. 410-418, 2019.

COLNAGO, L. A.; ANDRADE, F. D. **RMN no domínio do tempo: fundamentos e aplicações offline e online**. Embrapa Instrumentação-Capítulo em livro científico (ALICE), 2017.

COSTA, T. B. B. C. **Metabonomica aplicada ao diagnóstico e estadiamento de doenças hepáticas**. 2016. Dissertação de Mestrado. Universidade Federal de Pernambuco.

COSTA FILHO, P. A.; POPPI, R. J. **Algoritmo genético em química**. Química Nova, v. 22, n. 3, p. 405, 1999.

DA-COL, J. A.; DANTAS, W. F. C.; POPPI, R. J. **Experimento didático de quimiometria para o mapeamento de pellets farmacêuticos utilizando espectroscopia de imagem na região do infravermelho próximo e resolução multivariada de curvas com mínimos quadrados alternantes: Um tutorial, parte IV**. Química Nova, v. 41, n. 3, p. 345-354, 2018.

DE BENEDICTIS, F. M.; ATTANASI, M. **Asthma in childhood**. European respiratory review: an official journal of the European Respiratory Society, v. 25, n. 139, p. 41–47, 2016.

DE GRAAF, R. A.; PRINSEN, H.; GIANNINI, C.; CAPRIO, S.; HERZORG, R. I. **Quantification of ^1H NMR spectra from human plasma**. Metabolomics, v. 11, n. 6, p. 1702–1707, 2015.

DEEKS, J. J.; ALTMAN, D. G. **Diagnostic tests 4: likelihood ratios**. Bmj, v. 329, n. 7458, p. 168-169, 2004.

DELA BIANCA, A. C.; WANDALSEN, G. F.; MIYAGI, K.; CAMARGO, L.; CEZZARIN, D.; SOLÉ, D.; MALLOL, J. **Prevalência de sibilância em lactentes: proposta de protocolo internacional de estudo**. Rev. bras. alerg. imunopatol, v. 30, n. 3, p. 94-100, 2007.

DELIU, M.; BELGRAVE, D.; SPERRIN, M.; BUCHAN, I.; CUSTOVIC, A. **Asthma phenotypes in childhood**. Expert Review of Clinical Immunology, v. 13, n. 7, p. 705–713, 2017.

DUNN, W. B.; ELLIS, D. I. **Metabolomics: current analytical platforms and methodologies**. Trends in Analytical Chemistry, v. 24, n. 4, p. 285-294, 2005.

EMWAS, A. H.; SACCENTI, E.; GAO, X.; MCKAY, R. T.; SANTOS, V. A. M.; ROY, R.; WISHART, D. S. **Recommended strategies for spectral processing and post-processing of 1D ^1H -NMR data of biofluids with a particular focus on urine**. Metabolomics, v. 14, n. 3, p. 1-23, 2018.

FAN, Y.; LIU, S.; CHEN, X.; FENG, M.; SONG, F.; GAO, X. **Toxicological effects of Nux Vomica in rats urine and serum by means of clinical chemistry, histopathology and ^1H NMR-based metabonomics approach**. Journal of Ethnopharmacology, v. 210, n. January 2017, p. 242–253, 2018.

FARRAIA, M.; RUFO, C.J.; PACIÊNCIA, I.; MENDES, F.C.; DELGADO, L.; BOECHAT, J.L.; MOREIRA, A. **Metabolic interactions in asthma**. European Annals of Allergy and Clinical Immunology, v. 51, n. 5, p. 196–205, 2019.

FERREIRA, J. C.; PATINO, C. M. Entendendo os testes diagnósticos. Parte 3. **Jornal Brasileiro de Pneumologia**, v. 44, p. 04-04, 2018.

FERREIRA, M. M. C.. **Quimiometria: conceitos, métodos e aplicações**. Editora da Unicamp, 2015.

FIRMIDA, M. C. **Abordagem Clínica de Lactentes Sibilantes**. Pulmão RJ , v. 22, n. 3, p. 3–8, 2013.

GASTON, B. **The biochemistry of asthma**. Biochimica et Biophysica Acta - General Subjects, v. 1810, n. 11, p. 1017–1024, 2011.

GINA. **Global Strategy for Asthma Management and Prevention**. Disponível em: www.ginaasthma.org. Acesso em 30 Mar 2019.

GHINI, V.; QUAGLIO, D.; LUCHINAT, C.; TURANO, P. **NMR for sample quality assessment in metabolomics**. New Biotechnology, v. 52, p. 25–34, 2019.

GHOSH N, CHOUDHURY P, SUBRAMANI E, SAHA D, SENGUPTA S, JOSHI M, BANERJEE R, ROYCHOWDHURY S, BHATTACHARYYA P, CHAUDHURY K. **Metabolomic signatures of asthma-COPD overlap (ACO) are different from asthma and COPD**. Metabolomics 2019;15(6)1-11.

GILANY, K.; MOHAMADKHANI, A.; CHASMNIAH, S.; SHAHNAZARI, P.; AMINI, M.; ARJMAND, B.; MALEKZADEH, R.; GHOOCHANI, B. F. N. M. **Metabolomics analysis of the saliva in patients with chronic hepatitis b using nuclear magnetic resonance: A pilot study**. Iranian Journal of Basic Medical Sciences, v. 22, n. 9, p. 1044–1049, 2019.

GODOY, M. M. G.; LOPES, E. P. A.; SILVA, R. O.; HALLWASS, F.; KOURY, L. C. A.; MOURA, I. M.; GONÇALVES, S. M. C.; SIMAS, A. M. **Hepatitis C virus infection diagnosis using metabonomics**. Journal of Viral Hepatitis, v. 17, n. 12, p. 854–858, 2010.

GOETZ, L. H.; SCHORK, N. J. **Personalized medicine: motivation, challenges, and progress**. Fertility and sterility, v. 109, n. 6, p. 952-963, 2018.

GOUVEIA, L. R. **Metabonomica aplicada ao diagnóstico diferencial de doenças hepáticas**. 2017. Dissertação de Mestrado. Universidade Federal de Pernambuco.

GROMSKI, P. S.; MUHAMADALI, H.; ELLIS, D. I.; XU, Y.; CORREA, E.; TURNER, M. L.; GOODACRE, R. **A tutorial review: Metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding**. Analytica chimica acta, v. 879, p. 10-23, 2015.

GUILBERT, T. W.; MORGAN, W. J.; KRAWIEC, M.; LEMANSKE JR, R. F.; SORKNESS, C.; SZEFLER, S. J.; LARSEN, G.; SPAHN, J. D.; ZEIGER, R. S.; HELDT, G.; STRUNK, R. C.; BACHARIER, L. B.; BLOOMBERG, G. R.; CHINCHILLI, V. M.; BOEHMER, S. J.; MAUGER, E. A.; MAUGER, D. T.; TAUSSIG, L. M.; MARTINEZ, F. D. The Prevention of Early Asthma in Kids study: design, rationale

and methods for the Childhood Asthma Research and Education network. **Controlled clinical trials**, v. 25, n. 3, p. 286-310, 2004.

HOLMES, E.; TSANG, T. M.; TABRIZI, S. J. **The application of NMR-based metabonomics in neurological disorders**. *NeuroRx*, v. 3, n. 3, p. 358-372, 2006.

INGLE JR, J.D.; CROUCH, S. R. **Spectrochemical analysis**. Prentice-Hall. 1988.

IUPAC. **Compendium of Chemical Terminology**. Disponível em: <https://goldbook.iupac.org/terms/view/CT06948#:~:text=Chemometrics%20is%20the%20application%20of,of%20chemical%20experiments%20and%20simulations>. Acesso em: 10/03/2021.

IZQUIERDO-GARCÍA, J. L.; VILLA, P.; KYRIAZIS, A.; PUERTO-NEVADO, L.; PÉREZ-RIAL, S.; RODRIGUEZ, I.; HERNANDEZ, N.; RUIZ-CABELLO, J. **Descriptive review of current NMR-based metabolomic data analysis packages**. *Progress in Nuclear Magnetic Resonance Spectroscopy*, v. 59, n. 3, p. 263-270, 2011.

JUNG, J.; KIM, S-H.; LEE, H-S.; CHOI, G. S.; JUNG, Y-S.; RYU, D. H.; PARK, H-S.; HWANG, G-S. **Serum metabolomics reveals pathways and biomarkers associated with asthma pathogenesis**. *Clinical and Experimental Allergy*, v. 43, n. 4, p. 425–433, 2013.

KEGG. **Kyoto Encyclopedia of Genes and Genomes**. <https://www.genome.jp/kegg/pathway.html>. Acesso em: 26 de Fev 2021.

KELLY, R. S.; DAHLIN, A.; MCGEACHIE, M. J.; QIU, W.; SORDILLO, J.; WAN, E.; S.; WU, A. C.; LASKY-SU, J. **Asthma Metabolomics and the Potential for Integrative Omics in Research and the Clinic**. *Chest*, v. 151, n. 2, p. 262–277, 2017.

KHAN, S. S.; MADDEN, M. G. **One-class classification: taxonomy of study and review of techniques**. *The Knowledge Engineering Review*, v. 29, n. 3, p. 345-374, 2014.

KUMAR, V. **Robbins & Cotran - Patologia bases patológicas das doenças**. Elsevier Brasil, 2010.

LAPADULA, P.; MECCA, G.; SANTORO, D.; VELTRI, E. **Greg ML–Machine Learning for Healthcare at a Scale**. *Health and Technology*, p. 1-11, 2020.

LEHMAN-MCKEEMAN, L. D. **Metabonomics**. *Pharmacogenomics*, Second Edition, v. 455, n. October, p. 327–339, 2005.

LEVER, J.; KRZYWINSKI, M.; ALTMAN, N. **Classification evaluation**. *Nat Methods*, v. 13, p. 603-604, 2016.

LEVITT, M. H. **Symmetry in the design of NMR multiple-pulse sequences**. *The Journal of chemical physics*, v. 128, n. 5, p. 052205, 2008.

LI, X.; CAO, X.; GUO, M.; LIU, X. **Trends and risk factors of mortality and disability adjusted life years for chronic respiratory diseases from 1990 to 2017: systematic analysis for the Global Burden of Disease Study 2017**. *bmj*, v. 368, 2020.

LI, Y.; CAO, L.; YU, Q.; XUE, H.; LU, Y. **Association between peripheral blood mononuclear cell ORMDL3 expression and the asthma predictive index in preschool children**. *Journal of International Medical Research*, v. 47, n. 8, p. 3727–3736, 2019.

LINDON, J. C.; NICHOLSON, J. K. **Analytical technologies for metabonomics and metabolomics, and multi-omic information recovery**. *Trends in Analytical Chemistry*, v. 27, n. 3, p. 194–204, 2008.

LIRA NETO, F. T.; MARQUES, R. A.; CAVALCANTI FILHO, A. F.; ARAUJO, L. C. N.; LIMA, S. V. C.; PINTO, L.; SILVA, R. O. **¹H NMR-based metabonomics for infertility diagnosis in men with varicocele**. *Journal of Assisted Reproduction and Genetics*, 2020.

LIU, R.; YANG, B.; ZIO, E.; CHEN, X. **Artificial intelligence for fault diagnosis of rotating machinery: A review**. *Mechanical Systems and Signal Processing*, v. 108, p. 33-47, 2018.

LOPES, R. E. C. **Discriminação de madeiras similares por NIRS e PLS-DA considerando variações de temperatura e umidade**. 2015. Dissertação de Mestrado. Universidade de Brasília.

LUQUE, A.; CARRASCO, A.; MARTÍN, A.; HERAS, A. **The impact of class imbalance in classification performance metrics based on the binary confusion matrix**. *Pattern Recognition*, v. 91, p. 216-231, 2019.

LYRA, W. S.; SILVA, E. C.; ARAÚJO, M. C. U.; FRAGOSO, W. D. **Classificação periódica: um exemplo didático para ensinar análise de componentes principais**. *Química Nova*, v. 33, n. 7, p. 1594-1597, 2010.

MADSEN, R.; LUNDSTEDT, T.; TRYGG, J. **Chemometrics in metabolomics-A review in human disease diagnosis**. *Analytica Chimica Acta*, v. 659, n. 1–2, p. 23–33, 2010.

MALLOL, J.; GARCÍA-MARCOS, L.; AGUIRRE, V.; MARTINEZ-TORRES, A.; PEREZ-FERNÁNDEZ, V.; GALLARDO, A.; CALVO, M.; ROSARIO FILHO, N.; ROCHA, W.; FISCHER, G.; BAEZA-BACAB, M.; CHIARELLA, P.; PINTO, R.; BARRIA, C. **The international study of wheezing in infants: Questionnaire validation**. *International Archives of Allergy and Immunology*, v. 144, n. 1, p. 44–50, 2007.

MALLOL, J.; GARCÍA-MARCOS, L.; SOLÉ, D.; BRAND, P. **International prevalence of recurrent wheezing during the first year of life: variability,**

treatment patterns and use of health resources. Thorax, v. 65, n. 11, p. 1004–1009, Nov 2010.

MANZI, M.; RIQUELME, G.; ZABALEGUI, N.; MONGE, M. E. **Improving diagnosis of genitourinary cancers: Biomarker discovery strategies through mass spectrometry-based metabolomics.** Journal of pharmaceutical and biomedical analysis, v. 178, p. 112905, 2020.

MARKLEY, J. L.; BRÜSCHWEILER, R.; EDISON, A. S.; EGHBALNIA, H. R.; POWERS, R.; RAFTERY, D.; & WISHART, D. S. **The future of NMR-based metabolomics.** Current opinion in biotechnology, v. 43, p. 34-40, 2017.

MARTINEZ, F. D.; WRIGHT, A. L.; TAUSSIG, L. M.; HOLBERG, C. J.; HALONEN, M.; MORGAN, W. J. **Asthma and Wheezing in the First Six Years of Life.** N Engl J Med , v. 332, n. 3, p. 133–138, Jan 1995.

MATTIUZZI, C.; LIPPI, G. **Worldwide asthma epidemiology: insights from the Global Health Data Exchange database.** International Forum of Allergy and Rhinology, v. 10, n. 1, p. 75–80, 2020.

MATYSIAK, J.; KLUPCZYNSKA, A.; PACKI, K.; MACKOWIAK-JAKUBOWSKA, A.; BRĘBOROWICZ, A.; PAWLICKA, O.; OLEJNICZAK, K.; KOKOT, Z. J.; MATYSIAK, J. **Alterations in serum-free amino acid profiles in childhood asthma.** International Journal of Environmental Research and Public Health, v. 17, n. 13, p. 1–17, 2020.

MEDEIROS, D. SILVA, A. R.; RIZZO, J. A.; SARINHO, E.; MALLOL, J.; SOLÉ, D. **Prevalência de sibilância e fatores de risco associados em crianças no primeiro ano de vida, residentes no Município de Recife, Pernambuco, Brasil.** Cadernos de Saúde Pública , v. 27, n. 8, p. 1551–1559, Ago 2011.

MENA, L.; GONZALEZ, J. A. **Symbolic one-class learning from imbalanced datasets: application in medical diagnosis.** International Journal on Artificial Intelligence Tools, v. 18, n. 02, p. 273-309, 2009.

MICHÁLKOVÁ, L.; HORNÍK, Š.; SÝKORA, J.; HABARTOVÁ, L., & SETNIČKA, V. **Diagnosis of pancreatic cancer via ¹H NMR metabolomics of human plasma.** Analyst, v. 143, n. 24, p. 5974-5978, 2018.

MLYNÁRIK, V. **Introduction to nuclear magnetic resonance.** Analytical biochemistry, v. 529, p. 4-9, 2017.

MONARETTO, T. **Desenvolvimento e aplicação de sequências de pulsos CWFP uni e bidimensionais para RMN no domínio do tempo.** Tese de Doutorado. Universidade de São Paulo.

MOREIRA, E. D. T.; PONTES, M. J. C.; GALVÃO, R. K. H.; ARAÚJO, M. C. U. **Near infrared reflectance spectrometry classification of cigarettes using the successive projections algorithm for variable selection.** Talanta, v. 79, n. 5, p. 1260-1264, 2009.

MOTTA, A.; PARIS, D.; D'AMATO, M.; MELCK, D.; CALABRESE, C.; VITALE, C.; STANZIOLA, A. A.; CORSO, G.; SOFIA, M.; MANISCALCO, M. **NMR metabolomic analysis of exhaled breath condensate of asthmatic patients at two different temperatures.** Journal of proteome research, v. 13, n. 12, p. 6107-6120, 2014.

Nações Unidas. **Panorama Social de América Latina.** Comissão Econômica Para América Latina e Caribe. 2018.

NELSON, D.L.; COX, M.M. **Princípios de Bioquímica de Lehninger.** Artmed, 2014.

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, v. 29, n. 11, p. 1181-1189, 1999.

PAVIA, D. L. et al. **Introduction to spectroscopy.** Cengage Learn. 2009.

PANG, Z.; WANG, G.; WANG, C.; ZHANG, W.; LIU, J.; WANG, F. **Serum metabolomics analysis of asthma in different inflammatory phenotypes: a cross-sectional study in northeast China.** BioMed research international, v. 2018, 2018.

PEREIRA, A. M. M.; CASTRO, A. L. B.; OVIEDO, R. A. M.; BARBOSA, L. G.; GERASSI, C. D.; GIOVANELLA, L. **Atenção primária à saúde na América do Sul em perspectiva comparada: mudanças e tendências.** Saúde em Debate, v. 36, n. 94, p. 482-499, 2012.

PITCHON, R. R.; ALVIM, C. G.; ANDRADE, C. R.; LASMAR, L. M. B. F.; CRUZ, A. A.; REIS, A. P. **Mortalidade por asma em crianças e adolescentes: uma causa de morte quase sempre evitável.** Rev Med Minas Gerais, v. 28, n. Supl 6, p. S280607, 2018.

POMERANTSEV, A. L.; RODIONOVA, O. Y. **Concept and role of extreme objects in PCA/SIMCA.** Journal of Chemometrics, v. 28, n. 5, p. 429-438, 2014.

POMERANTSEV, A. L.; RODIONOVA, O. Y. **Popular decision rules in SIMCA: Critical review.** Journal of Chemometrics, p. e3250, 2020.

PONTES, T. A.; BARBOSA, A. D.; SILVA, R. D.; MELO-JUNIOR, M. R.; SILVA, R. O. **Osteopenia-osteoporosis discrimination in postmenopausal women by ¹H NMR-based metabonomics.** PLoS ONE, v. 14, n. 5, p. 1–10, 2019.

RABI, I. I.; ZACHARIAS, J. R.; MILLMAN, S.; KUSCH, P. **A new method of measuring nuclear magnetic moment.** Physical review, v. 53, n. 4, p. 318, 1938.

ROCHA, G. **Determinação do teor de biodiesel no petrodiesel através da RMN em baixo campo, aliada a calibração multivariada.** Dissertação de mestrado. Curitiba, 2013.

RODIONOVA, O. Y.; OLIVERI, P.; POMERANTSEV, A. L. **Rigorous and compliant approaches to one-class classification**. Chemometrics and Intelligent Laboratory Systems, v. 159, p. 89-96, 2016.

RODRIGUES, F. Z. **PANORAMA ATUAL DA ASMA: EPIDEMIOLOGIA, FISIOPATOLOGIA E TRATAMENTOS**. 2007. Trabalho de conclusão de curso. São Paulo, 2007.

RODRÍGUEZ-PÉREZ, R.; FERNÁNDEZ, L.; MARCO, S. **Overoptimism in cross-validation when using partial least squares-discriminant analysis for omics data: a systematic study**. Analytical and bioanalytical chemistry, v. 410, n. 23, p. 5981-5992, 2018.

RHODES, C. J. **Magnetic resonance spectroscopy**. Science Progress, v. 100, n. 3, p. 241-292, 2017.

ROSSI, R. G. **Análise de Componentes Principais em Data Warehouses**. Dissertação de mestrado. São Paulo, 2017.

SAVOIA, C.; VOLPE, M.; GRASSI, G.; BORGHI, C.; ROSEI, E. A.; TOUYZ, R. M. **Personalized medicine—a modern approach for the diagnosis and management of hypertension**. Clinical Science, v. 131, n. 22, p. 2671-2685, 2017.

SAGDIC, A.; SENER, O.; BULUCU, F.; KARADURMUS, N.; ÖZEL, H. E.; YAMANEL, L.; TASCI, C.; NAHARCI, I.; OCAL, R.; AYDIN, A. **Oxidative stress status and plasma trace elements in patients with asthma or allergic rhinitis**. Allergologia et immunopathologia, v. 39, n. 4, p. 200-205, 2011.

SANTANA, F. B.; SOUZA, A. M.; ALMEIDA, M. R.; BREITKREIZ, M. C.; FILGUEIRAS, P. R.; SENA, M. M.; POPPI, R. J. **Experimento Didático De Quimiometria para Classificação de Óleos Vegetais Comestíveis por Espectroscopia no Infravermelho Médio Combinado com Análise Discriminante por Mínimos Quadrados Parciais: Um Tutorial, Parte V**. Química Nova, v. 43, n. 3, p. 371-381, 2020.

SARASWATHI V.; HASTY A. H. **The role of lipolysis in mediating the proinflammatory effects of very low density lipoproteins in mouse peritoneal macrophages**. Journal of lipid research, v. 47, n. 7, p. 1406-1415, 2006.

SÉBÉDIO, J-L.; PUJOS-GUILLOT, E.; FERRARA, M. **Metabolomics in evaluation of glucose disorders**. Current Opinion in Clinical Nutrition & Metabolic Care, v. 12, n. 4, p. 412-418, 2009.

SHILLAN, D.; STEME, J. A.; CHAMPNEYS, A.; GIBBISON, B. **Use of machine learning to analyse routinely collected intensive care unit data: a systematic review**. Critical Care, v. 23, n. 1, p. 284, 2019.

SILVA, R. D. **Aplicações Metabonômicas usando Ressonância Magnética Nuclear de 1H: Diagnóstico Não-Invasivo de Câncer de Próstata e Urológico &**

Classificação de Azeite de Oliva Extra Virgem de Produção Orgânica. Tese de doutorado. Recife, 2017.

SOARES, S. F. C. **Um novo método para transferência de modelos de calibração NIR e uma nova estratégia para classificação de sementes de algodão usando imagem hiperespectral NIR.** Tese de doutorado. João Pessoa, 2016.

SOLÉ, D. **Sibilância na infância.** J Bras Pneumol , v. 34, n. 6, p. 337–339, Jun 2008.

SOLÉ, D.; ARANDA, C. S.; WANDALSEN, G. F. **Asthma: epidemiology of disease control in Latin America—short review.** Asthma research and practice, v. 3, n. 1, p. 4, 2017.

SOUSA, S. A. A.; MAGALHAES, A.; FERREIRA, M. M. C. **Optimized bucketing for NMR spectra: Three case studies.** Chemometrics and Intelligent Laboratory Systems, v. 122, p. 93-102, 2013.

THARWAT, A. **Classification assessment methods.** Applied Computing and Informatics, 2020.

TRYGG, J.; WOLD, S. **Orthogonal projections to latent structures (O-PLS).** Journal of Chemometrics: A Journal of the Chemometrics Society, v. 16, n. 3, p. 119-128, 2002.

TURI, K. N.; ROMICK-ROSENDALE, L.; RYCKMAN, K. K.; HARTERT, T. V. **A review of metabolomics approaches and their application in identifying causal pathways of childhood asthma.** Journal of Allergy and Clinical Immunology, v. 141, n. 4, p. 1191-1201, 2018.

VAN, Q. N.; CHMURNY, G. N.; VEENSTRA, T. D. **The depletion of protein signals in metabonomics analysis with the WET-CPMG pulse sequence.** Biochemical and biophysical research communications, v. 301, n. 4, p. 952-959, 2003.

WEIL, A. R. **Precision medicine.** Health Affairs, v. 37, n. 5, p. 687-687, 2018.

WISHART, D. S.; KNOX, C.; GUO, A. C.; EISNER, R.; YOUNG, N.; GAUTAM, B.; HAU, D. D.; PSYCHOGLIOS, N.; DONG, E.; BOUATRA, S.; MANDAL, R.; SINELNIKOV, I.; XIA, J.; JIA, L.; CRUZ, J. A.; LIM, E.; SOBSEY, C. A.; SHRIVASTAVA, S.; HUANG, P.; LIU, P.; FANG, L.; PENG, J.; FRADETTE, R.; CHENG, D.; TZUR, D.; CLEMENTS, M.; LEWIS, A.; SOUZA, A.; ZUNIGA, A.; DAWE, M.; XIONG, Y.; CLIVE, D.; GREINER, R.; NAZYROVA, A.; SHAYKHUTDINOV, R.; LI, L.; VOGEL, H. J.; FORSYTHE, I. **HMDB: A knowledgebase for the human metabolome.** Nucleic Acids Research, v. 37, n. SUPPL. 1, p. 603–610, 2009.

WISHART, D. S. **NMR metabolomics: A look ahead.** Journal of Magnetic Resonance, v. 306, p. 155-161, 2019.

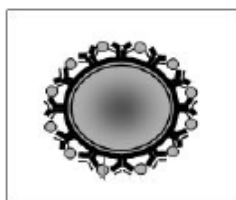
WOLD, S. **Chemometrics; what do we mean with it, and what do we want from it?**. Chemometrics and Intelligent Laboratory Systems, v. 30, n. 1, p. 109-115, 1995.

ZANG, X.; MONGE, M. E.; GAUL, D. A.; MCCARTY, N. A.; STECENKO, A.; FERNANDEZ, F. M. **Early Detection of Cystic Fibrosis Acute Pulmonary Exacerbations by Exhaled Breath Condensate Metabolomics**. Journal of Proteome Research, v. 19, n. 1, p. 144-152, 2019.

ZHANG, A.; SUN, H.; WANG, X. **Serum metabolomics as a novel diagnostic approach for disease: a systematic review**. Analytical and bioanalytical chemistry, v. 404, n. 4, p. 1239-1245, 2012.

ZONTOV, Y. V.; RODIONOVA, O. Y.; POMERANTSEV, A. L.; KUCHERYAVSKIY, S. V. **DD-SIMCA—a MATLAB GUI tool for data driven SIMCA approach**. Chemometrics and Intelligent Laboratory Systems, v. 167, p. 23-28, 2017.

APÊNDICE A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE) DO ESTUDO



EBSERH
HOSPITALS UNIVERSITÁRIOS FEDERAIS



**CENTRO DE PESQUISAS EM
ALERGIA E IMUNOLOGIA CLÍNICA
Serviço de Alergia e Imunologia Clínica
HOSPITAL DAS CLÍNICAS/UFPE**

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO RES. 466/12

Convidamos o (a) seu/sua filho (a) {ou menor de idade} que está sob sua responsabilidade para participar, como voluntário (a), da pesquisa: A METABONÔMICA NA AVALIAÇÃO DA SIBILÂNCIA EM LACTENTES. Esta pesquisa é da responsabilidade da pesquisadora Aline Duarte Maranhão, Fone: 982025172, endereço: Avenida Professor Moraes Rêgo S/N Hospital das Clínicas 2º andar – Ambulatório de Pré-natal, CEP: 50670-420, Fone: (81) 21263668, email: pesquisasibilancia@gmail.com e do pesquisador Décio Medeiros, Fone: 2126.3918, endereço: Avenida Professor Moraes Rêgo S/N Hospital das Clínicas 1º Andar – Ambulatório de Alergia Infantil, CEP: 50670-420, Fone: (81) 2126.3918; e-mail: daph@elogica.com.br.

Após ser esclarecido (a) sobre as informações a seguir, no caso de aceitar que seu/sua filho(a) faça parte do estudo, rubrique as folhas e assine ao final deste documento, que está em duas vias. Uma delas é sua e a outra é do pesquisador responsável. Em caso de recusa o (a) Sr.(a) ou o (a) seu/suas filho (a) NÃO serão penalizados (as) de forma alguma. O (a) Senhor (a) tem o direito de retirar o consentimento da participação dele (a) a qualquer tempo, sem qualquer penalidade.

Este termo de consentimento pode conter alguns tópicos que você não entenda. Caso haja alguma dúvida, pergunte a pessoa a quem está lhe entrevistando, para que você esteja bem esclarecido(a) sobre tudo que está respondendo.

O objetivo deste estudo é verificar se a análise de sangue pelo aparelho de ressonância magnética nuclear (semelhante ao aparelho de tomografia) é capaz de identificar em amostra de sangue de crianças de 6 a 36 meses atendidas no Hospital das Clínicas – UFPE e que apresentam diagnóstico de sibilância recorrente (cansaço, falta de ar, chiado no peito).

Como riscos, pode-se considerar o constrangimento em responder as perguntas dos questionários além do incômodo em realizar exames de rastreio de alergias (teste cutâneo – teste alérgico na pele,

podendo produzir uma pequena vermelhidão no local; coleta de sangue, com um pouco de dor e um pequeno hematoma local). Para minimizar os riscos, todos os testes serão realizados por profissionais treinados e que já têm experiência nos diversos procedimentos.

Como benefício direto poderão ser detectadas alterações que permitam o diagnóstico de sibilância e alergia na criança, o que favorecerá o seu tratamento. O benefício indireto será obter informação científica valiosa, ainda não disponível, relacionada à capacidade da análise do sangue pelo aparelho de ressonância magnética em identificar as crianças sibilantes e, posteriormente utilizar os resultados da pesquisa como subsídio para o melhoramento da assistência atualmente prestada à saúde dos pacientes portadores de doenças alérgicas.

As informações desta pesquisa serão confidenciais e serão divulgadas apenas em eventos ou publicações científicas, não havendo identificação dos voluntários, a não ser entre os responsáveis pelo estudo, sendo assegurado o sigilo sobre a participação dele (a). Os dados coletados serão armazenados em computador pessoal, sob a responsabilidade do pesquisador principal, no endereço acima informado, sigilosamente, pelo período de 5 anos.

O (a) senhor (a) não pagará nada para o (a) seu/sua filho (a) participar desta pesquisa. Se houver necessidade, as despesas para a sua participação serão assumidos pelos pesquisadores (ressarcimento de despesas).

Em caso de dúvidas relacionadas aos aspectos éticos deste estudo, você poderá consultar o Comitê de Ética em Pesquisa Envolvendo Seres Humanos da UFPE no endereço: (Avenida da Engenharia S/n - 1º Andar, sala 4 - Cidade Universitária, Recife-PE, CEP: 50740-600, Tel.: (81) 2126.8588 – e-mail: cepccs@ufpe.br).

Assinatura do pesquisador (a) _____

CONSENTIMENTO DO RESPONSÁVEL PARA A PARTICIPAÇÃO DA CRIANÇA

Eu, _____, CPF _____, abaixo assinado, responsável pelo (a) menor _____, autorizo a sua participação no estudo A METABONÔMICA NA AVALIAÇÃO DA SIBILÂNCIA EM LACTENTES como voluntário(a). Fui devidamente informado (a) e esclarecido (a) pelo pesquisador sobre a pesquisa, os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes da participação dele(a). Foi-me garantido que posso retirar o meu consentimento a qualquer momento, sem que isto leve a qualquer penalidade (ou interrupção de seu acompanhamento/ assistência/tratamento) para mim ou para meu/minha filho (a).

Recife, _____ de _____ de 20____.

Assinatura do (da) responsável: _____

Presenciamos a solicitação de consentimento, esclarecimentos sobre a pesquisa e aceite do sujeito em participar. Duas (02) testemunhas (não ligadas à equipe de pesquisadores):

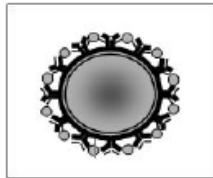
Testemunha : _____

Assinatura da Testemunha: _____

Testemunha : _____

Assinatura da Testemunha: _____

APÊNDICE B – FICHA CADASTRAL DE PACIENTES DO AMBULATÓRIO DE PEDIATRIA E DE PUERICULTURA DO HOSPITAL DAS CLÍNICAS DE PERNAMBUCO



EBSERH
HOSPITALS UNIVERSITÁRIOS FEDERAIS



**CENTRO DE PESQUISAS EM
ALERGIA E IMUNOLOGIA CLÍNICA**
Serviço de Alergia e Imunologia Clínica
HOSPITAL DAS CLÍNICAS/UFPE

PESQUISA: A METABONÔMICA NA AVALIAÇÃO DA SIBILÂNCIA EM LACTENTES

Data: ____/____/____

Nome do Paciente: _____

Data nascimento: ____/____/____ Prontuário: _____ Nº Registro Pesquisa: _____

Nome do responsável: _____

Idade: _____ Sexo: () Fem () Masc Peso: _____ (kg) Altura: _____ (cm)

Endereço: _____

Telefone: _____

Uso de medicações de urgência/resgate? () Não () Sim Há quanto tempo? _____

Grupo a encaminhar o paciente

() Grupo CASO () Grupo controle

Teste cutâneo:

() Positivo _____

() Negativo _____

Hemograma: Eosinófilos: _____ %

Dosagem de IgE total: Valor: _____ KU/L () normal () alterada

Metabolômica: _____

Critérios Maiores

1. História familiar de asma ()
2. Diagnóstico de dermatite atópica ()
3. Sensibilização a aeroalérgenos ()

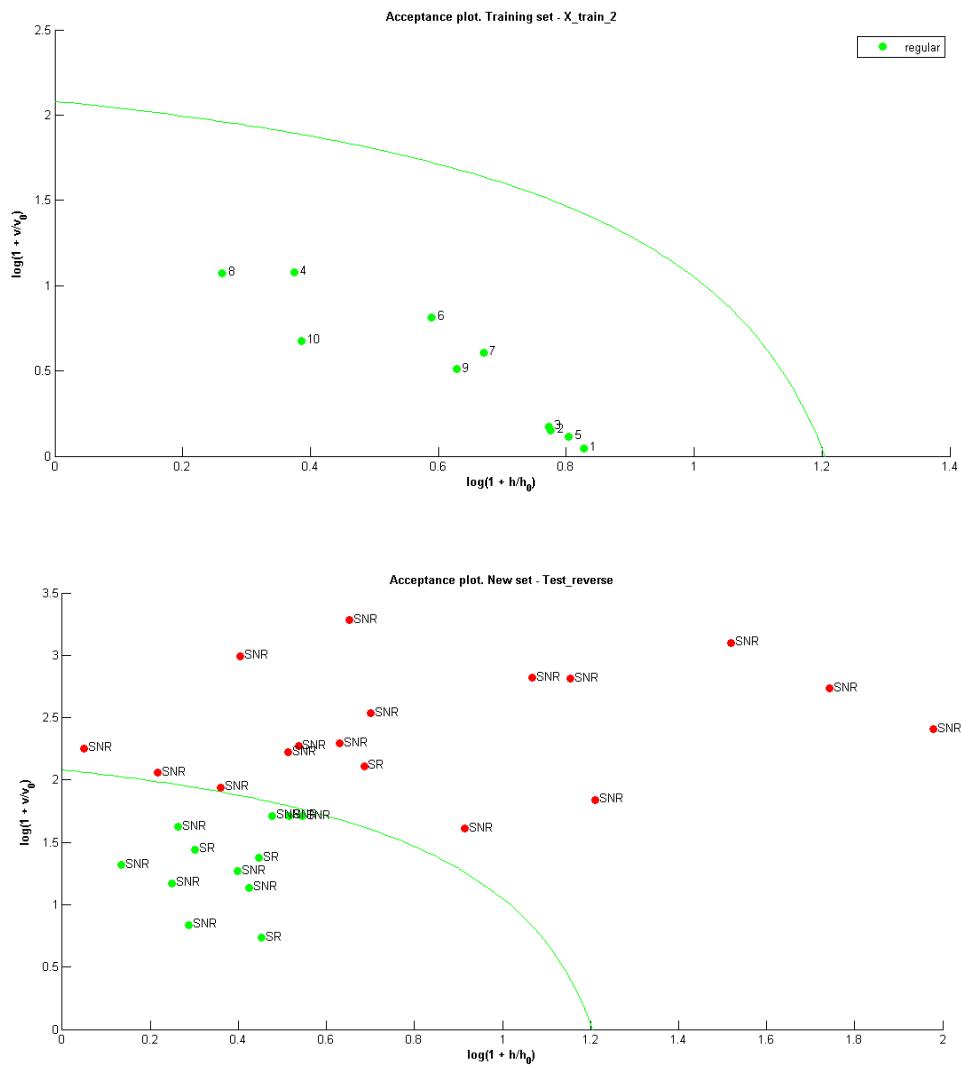
Critérios Menores

1. Sensibilização a alimentos ()
(leite de vaca, ovo, amendoim)
2. Eosinofilia $\geq 4\%$ ()
3. Sibilância não associada a resfriado ()

APÊNDICE C – QUESTIONÁRIO PADRÃO DE COLETA DE DADOS DE PACIENTES DO AMBULATÓRIO DE PEDIATRIA E DE PUERICULTURA DO HOSPITAL DAS CLÍNICAS DE PERNAMBUCO, BASEADO NO EISL

Estudo internacional sobre prevalência de sibilância recorrente em lactentes Questionário sobre sintomas respiratórios durante o primeiro ano de vida.			
Prezada(o) Mamãe (Pai ou responsável): Por favor, preencha (responda) o questionário abaixo sobre problemas respiratórios que o seu bebê teve no primeiro ano de vida (desde que nasceu até que tenha completado um ano de vida). Pedimos para que NÃO deixe quadradinhos em branco.			
Pessoa que irá preencher os dados: <input type="checkbox"/> Mãe <input type="checkbox"/> Pai <input type="checkbox"/> Outro.....			
Nome da criança.....			
IDADE (meses)	SEXO <input type="checkbox"/> Fem <input type="checkbox"/> Masc		
Endereço:.....			
Data de hoje	Telefone para contato		
Data de nascimento	(dia -mês -ano / exemplo 15-11-03)		
Peso ao nascer quilos	gramas (exemplo 3 kg 100 gramas)		
Quanto pesa agora quilos	gramas		
Quanto mediu ao nascer cm	Quanto mede agora cm		
Em que país seu bebê nasceu?			
Por favor, coloque um X no quadradinho que corresponda a sua resposta correta.			
1. Seu bebê teve chiado no peito ou bronquite ou sibilâncias nos seus primeiros 12 meses de vida : <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
Se você respondeu NÃO, por favor, vá à pergunta de número 12			
2. Quantos episódios de chiado no peito (bronquite ou sibilâncias) ele teve no primeiro ano de vida? <input type="checkbox"/> Nenhum <input type="checkbox"/> Menos de 3 episódios <input type="checkbox"/> 3 a 6 episódios <input type="checkbox"/> Mais de 6 episódios			
3. Com que idade seu bebê teve o primeiro episódio de chiado no peito (primeira bronquite)? Aos meses			
4. Seu bebê recebeu tratamento com medicamentos inalados para aliviar o chiado no peito (broncodilatadores) por nebulizadores ou inaladores (bombinhas), por exemplo: Salbutamol, Aerolin®, Berotec®, Brycanil®? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO <input type="checkbox"/> NÃO SEI			
5. Seu bebê recebeu tratamento com corticóides (cortisonas) inalados (bombinhas) (Por exemplo, Symbicort®, Flixotide®, Seretide®, Clenil®, Beclosol®, Budesonida, Busonid®, Pulmicort®, Beclometasona, Fluticasona etc.). <input type="checkbox"/> SIM <input type="checkbox"/> NÃO <input type="checkbox"/> NÃO SEI			
6. Seu bebê recebeu tratamento com: Antileucotrienos (Singulair®) - <input type="checkbox"/> SIM <input type="checkbox"/> NÃO <input type="checkbox"/> NÃO SEI			
7. Nestes últimos doze meses quantas vezes você acordou durante a noite devido à tosse, ou chiado no peito do seu bebê? <input type="checkbox"/> Nunca <input type="checkbox"/> Raras vezes (menos de uma vez ao mês) <input type="checkbox"/> Algumas vezes (algumas semanas em alguns meses) <input type="checkbox"/> Frequentemente (duas ou mais noites por semana, quase todos os meses)			
8. Nestes últimos doze meses o chiado no peito (sibilâncias) do seu bebê foi tão forte a ponto de ser necessário levá-lo a um serviço de emergência (Hospital, Clínica ou Posto de Saúde)? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
9. Nestes últimos doze meses o chiado no peito (sibilância) do seu bebê foi tão intenso a ponto de você vê-lo com muita dificuldade para respirar (com falta de ar)? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
10. Seu bebê já foi hospitalizado (internado em hospital) por bronquite? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
11. Algum médico lhe disse alguma vez que seu bebê tem asma? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
12. Seu bebê já teve pneumonia? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
13. Seu bebê já foi hospitalizado por pneumonia? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
14. Alguma pessoa fuma dentro da sua casa (pai, mãe, avós, tios)? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
15. Você fuma? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
16. A mãe do bebê fumou durante a gravidez? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
17. Seu bebê tem familiares com asma? <input type="checkbox"/> SIM <input type="checkbox"/> mãe <input type="checkbox"/> pai <input type="checkbox"/> Irmãos <input type="checkbox"/> NÃO			
18. Seu bebê tem familiares com alergia no nariz ou rinite alérgica? <input type="checkbox"/> SIM <input type="checkbox"/> mãe <input type="checkbox"/> pai <input type="checkbox"/> irmãos <input type="checkbox"/> NÃO			
19. Seu bebê tem familiares com alergia de pele (dermatite alérgica)? <input type="checkbox"/> SIM <input type="checkbox"/> mãe <input type="checkbox"/> pai <input type="checkbox"/> irmãos <input type="checkbox"/> NÃO			
20. Seu bebê nasceu por cesariana (parto cesárea)? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
21. Seu bebê foi à creche este ano? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
22. Com que idade seu bebê entrou na creche? Meses			
23. Com que frequência você dá ao seu bebê algum dos seguintes produtos (não feitos em casa): iogurte, pudim, salgadinhos (chips), chocolate, refrigerantes, suco de frutas de caixinha ou de garrafa, geléia artificial. <input type="checkbox"/> Nunca <input type="checkbox"/> Uma vez por semana <input type="checkbox"/> Uma vez ao mês <input type="checkbox"/> Todos os dias da semana			
24. Que tipo de aquecedor ou calefação você utiliza na sua casa? <input type="checkbox"/> Não uso <input type="checkbox"/> Estufa a gás <input type="checkbox"/> estufa elétrica <input type="checkbox"/> lenha <input type="checkbox"/> carvão <input type="checkbox"/> parafina <input type="checkbox"/> outra			
25. Que tipo de combustível você usa para cozinhar na sua casa? <input type="checkbox"/> Gás encanado <input type="checkbox"/> gás de bujão <input type="checkbox"/> eletricidade <input type="checkbox"/> carvão <input type="checkbox"/> madeira <input type="checkbox"/> outro			
26. Você tem ar condicionado em sua casa? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
27. Você tinha algum animal de estimação (cachorro, gato, passarinho, coelho) em sua casa quando seu filho nasceu? <input type="checkbox"/> SIM <input type="checkbox"/> Cachorro <input type="checkbox"/> gato <input type="checkbox"/> outros <input type="checkbox"/> NÃO			
28. Você tem algum bicho de estimação na sua casa atualmente? (cachorro, gato, passarinho, coelho)? <input type="checkbox"/> SIM <input type="checkbox"/> cachorro <input type="checkbox"/> gato <input type="checkbox"/> outros <input type="checkbox"/> NÃO			
29. Você tem carpete na sua casa? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
30. Você tem banheiro com pia, chuveiro e vaso sanitário dentro de casa? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
31. A cozinha da sua casa (ou o lugar onde a comida é preparada) é dentro da casa? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
32. Você tem telefone (fixo ou celular)? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
33. Marque qual o seu grau de escolaridade <input type="checkbox"/> Educação básica, primária ou nenhuma (8 anos ou menos). <input type="checkbox"/> Educação média ou secundária incompleta (nove a 11 anos). <input type="checkbox"/> Educação média ou secundária completa e nível superior (doze ou mais anos)			
34. Por quantos meses você deu somente leite do peito para o seu bebê (sem dar sucos, papinhas ou outro tipo de leite)? meses			
35. Quantos resfriados (episódios de espirros, tosse e secreção nasal com ou sem febre) seu bebê teve no primeiro ano de vida (escreva o número no quadradinho) episódios			
36. Com que idade seu bebê se resfriou pela primeira vez? (escreva a idade em meses no quadradinho). meses			
37. Seu bebê tem ou teve alguma alergia de pele durante o primeiro ano de vida? (manchas vermelhas na pele com coceira, alergia à fralda, alergia à picada de mosquito, comida, metais, etc.). <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
38. Você considera que o lugar onde você vive é um lugar com poluição atmosférica (fumaça de fábricas, alto tráfego de veículos, etc.). <input type="checkbox"/> SIM <input type="checkbox"/> muito <input type="checkbox"/> moderado <input type="checkbox"/> pouco <input type="checkbox"/> NÃO			
39. Existe mofo (bolor) ou manchas de umidade em sua casa? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
40. Seu bebê tem as vacinas em dia (as que correspondem ao primeiro ano)? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
41. Quantos irmãos ou irmãs o seu bebê tem?			
42. Quantas pessoas (adultos e crianças) vivem atualmente na sua casa?			
43. Você (a mãe) tem atualmente um trabalho remunerado? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
44. Qual é a raça do seu bebê? <input type="checkbox"/> Branca <input type="checkbox"/> Negra (mulato, pardo) <input type="checkbox"/> Asiática (japonês e chinês.) <input type="checkbox"/> outra			
45. Seu bebê recebeu tratamento com corticóides orais (Predsim®, Prelone®, Decadron®)? <input type="checkbox"/> SIM <input type="checkbox"/> NÃO <input type="checkbox"/> NÃO SEI			

APÊNDICE D – MODELOS DE TREINAMENTO E TESTE, RESPECTIVAMENTE,
DA MODELAGEM COM O GRUPO DE SIBILÂNCIA RECORRENTE COMO
GRUPO ALVO E MATRIZ DE CONTIGÊNCIA



Diagnóstico Clínico			
Grupo de teste		Sibilância Recorrente	Sibilância Não Recorrente
Classe Predita	Sibilância Recorrente	3	9
	Sibilância Não Recorrente	1	16

Sensibilidade 75%; Especificidade 64%; Exatidão 65,51%; VPP 25% e VPN 94,12%.

APÊNDICE E – RESULTADO DO ÍNDICE PREDITOR DE ASMA MODIFICADO DOS VOLUNTÁRIOS DO ESTUDO

Numeração dos voluntários	Índice Predictor de Asma Modificado (mAPI)						Resultado
	Critérios Maiores			Critérios Menores			
	1	2	3	1	2	3	
10	S	N	N	N	N	N	Positivo
13	N	N	N	N	S	N	Negativo
26	N	N	N	N	N	N	Negativo
34	N	N	N	N	N	N	Negativo
35	S	N	N	N	N	N	Positivo
36	S	N	N	N	N	N	Positivo
40	N	N	S	N	N	N	Positivo
43	S	N	N	N	N	N	Positivo
45	S	N	N	N	*	N	Positivo
46	N	S	N	N	*	N	Negativo
47	N	N	N	N	S	N	Negativo
48	S	N	N	N	N	N	Positivo
49	N	N	N	N	N	S	Negativo
50	N	N	S	N	N	N	Positivo
51	S	S	N	N	N	S	Positivo
52	S	N	N	N	S	N	Positivo
54	S	N	N	N	N	N	Positivo
55	S	N	N	N	S	N	Positivo
56	S	N	N	S	N	N	Positivo
58	S	N	N	N	N	N	Positivo
59	N	N	S	N	N	N	Positivo
60	S	N	N	N	N	N	Positivo
61	N	N	N	N	N	N	Negativo
62	S	N	N	N	S	N	Positivo
63	N	N	N	N	S	N	Negativo
64	N	N	N	N	N	N	Negativo
65	S	N	N	N	S	N	Positivo
66	S	N	N	N	N	N	Positivo
67	S	N	N	N	N	N	Positivo
68	S	N	N	N	N	N	Positivo
71	S	N	N	N	N	N	Positivo
72	N	N	N	N	S	N	Negativo
73	N	N	N	N	S	N	Negativo
74	S	N	N	N	N	N	Positivo
75	S	N	N	N	N	N	Positivo
76	S	N	N	N	S	N	Positivo
78	N	N	N	N	N	N	Negativo
80	S	N	N	N	S	N	Positivo
81	S	N	N	N	N	N	Positivo

*Não informado

APÊNDICE F – CLASSIFICAÇÃO DOS PACIENTES VOLUNTÁRIOS DO ESTUDO SEGUNDO O mAPI, GA-LDA E DDSIMCA

Numeração dos voluntários	Diagnóstico Clínico	Classificação segundo <i>mAPI</i>	Modelo GA-LDA	Modelo DD-SIMCA
10	SR	Positivo	SR	SR
13	SNR	Negativo	SNR	SNR
26	SNR	Negativo	SNR	SNR
34	SNR	Negativo	SNR	SNR
35	SNR	Positivo	SNR	SNR
36	SR	Positivo	SR	SR
40	SNR	Positivo	SR	SNR
43	SR	Positivo	SR	SR
45	SR	Positivo	SR	SR
46	SNR	Negativo	SR	SNR
47	SNR	Negativo	SNR	SNR
48	SNR	Positivo	SNR	SNR
49	SNR	Negativo	SNR	SNR
50	SNR	Positivo	SNR	SNR
51	SR	Positivo	SR	SR
52	SR	Positivo	SR	SR
54	SNR	Positivo	SNR	SNR
55	SR	Positivo	SR	SR
56	SNR	Positivo	SNR	SNR
58	SNR	Positivo	SR	SNR
59	SNR	Positivo	SNR	SNR
60	SR	Positivo	SR	SNR
61	SNR	Negativo	SNR	SNR
62	SR	Positivo	SR	SR
63	SNR	Negativo	SNR	Amostra outlier do grupo SNR
64	SNR	Negativo	SNR	SNR
65	SR	Positivo	SR	SR
66	SR	Positivo	SR	SR
67	SNR	Positivo	SNR	SNR
68	SNR	Positivo	SNR	SNR
71	SR	Positivo	SR	SNR
72	SNR	Negativo	SNR	SNR
73	SNR	Negativo	SNR	SNR
74	SNR	Positivo	SNR	Amostra de extremo do grupo SNR
75	SR	Positivo	SR	SNR
76	SR	Positivo	SR	SR
78	SNR	Negativo	SNR	SR
80	SNR	Positivo	SNR	SNR
81	SNR	Positivo	SNR	SNR
Métrica de performance				
Exatidão	–	66,6%	75%	81,8%
Sensibilidade	–	100%	100%	78,6%
Especificidade	–	48%	62,5%	87,5%
Valor Preditivo Positivo	–	51,8%	57,1%	91,7%
Valor Preditivo Negativo	–	100%	100%	70,0%
Em azul as amostras de treinamento e em verde as amostras de teste				

ANEXO A – VIAS METABÓLICAS DA ALANINA-ASPARTATO-GLUTAMATO E D-GLUTAMINA E D-GLUTAMATO

